RESOURCE SATELLITES AND REMOTE AIRBORNE SENSING FOR CANADA

This document was produced
by scanning the original publication.

Ce document est le produit d'une
numérisation par balayage
de la publication originale.

PROCEEDINGS OF THE FIRST CANADIAN SYMPOSIUM ON REMOTE SENSING OTTAWA, FEBRUARY 1972

VOLUME 2

GEOMATICS BIBLIOGRAPHIC DATABASE
3 6503 00001432 5

PROCEEDINGS OF
THE FIRST CANADIAN SYMPOSIUM ON REMOTE SENSING
OTTAWA, FEBRUARY 1972

VOLUME 2

Published by The First Canadian Synposium on Remote Sensing
for the Canada Centre for Remote Sensing, Department of
Energy, Mines and Resources, Ottawa

Price $10 (2 volumes)

Edited and prepared by
Dennis White
Public Relations and Information Services
Department of Energy, Mines and Resources
Ottawa

AUGUST 1972

Lithographed in Canada by
CAMPBELL PRINTING

C O N T E N T S

VOLUME 2

SESSION 1 (Cont)

TERRAIN ANALYSES (B)

Use of stereoscopic orthophotos for soil-terrain analyses
L.S. Crosson and R. Protz 347

Supplementary reconnaissance air photography as an aid for detailed geomorphological mapping
A. Kesik 353

Terrain analysis from small-scale aerial photographs
P. Gimbarzevsky 367

Film and filter combinations for the study of the periglacial landscape of the high Arctic
P.J. Howarth 379

Thermal infrared imagery at the St. Jean Vianney landslide
Jean-Yves Chagnon and Marc G. Tanguay 387

SENSORS AND INSTRUMENTATION (B)

A poor man's remote sensing system
W.B. McCoy 405

Effects of spectral filtration and atmospheric conditions on aerial photography obtained in 1970 and 1971
U. Nielsen 411

A qualitative study of Kodak Aerochrome infrared film, Type 2443, and the effect produced by Kodak colour compensating filters, at high altitudes
R.D. Worsfold 417

Colour photography from the Yellowknife district
V.R. Slaney 429

Camera mounts for 35 mm mono-and multi-spectral photography
V.G. Zsilinszky 441

TERRAIN ANALYSES (C)

Etalonnage au sol sur Parcelles Experimentales de Radiometre Barnes PRT-5 dans la fenetre 9.5 - 11.5 microns
F. Bonn, P. Clément, B. Bareilhe, C. Jalbert, Y. Leblanc, C. Trudel 453

Use of ground test array in assessing infrared line scan performance
G.A. Morley and D. McKinnon 461

A colour laser beam image recorder for Canada's remote sensing program
J.W. Locke 471

Principles of SLAR systems
G.E. Haslim and D.F. Page 479

SLAR imagery interpretation techniques and procedures for reconnaissance forestry and soils mapping
P.M. Addison 483

SESSION II

METHODOLOGY (A)

Perceptual principles related to remote sensing
M.M. Taylor 497

Photogrammetric aspects of precision processing of ERTS imagery
V. Kratky 505

A quick-look readout for ERTS multispectral scanner imagery
R. Barrington and R. Hutchinson 527

Problems in relating user requirements to quantitative parameters of imagery quality
M.A. Underhill 533

Geometric consideration in remote sensing
Eugene E. Derenyi 547

Contents (Cont)

METHODOLOGY (B)

The interpretation of Arctic imagery
I.H.S. Henderson and R.J. Brown 553

Can we teach computers to see?
T. Kasvand 557

The block adjustment of colour in high-altitude photography
S.H. Collins 569

Minimum distance classification in remote sensing
A.G. Wacker 577

Multispectral-multitemporal photography and automatic terrain recognition
D. Steiner 601

SENSORS AND INSTRUMENTATION (A)

Microwave radiometry for remote sensing from aircraft and spacecraft
A.W. Adey 613

New directions in microwave radiometry for remote sensing
M.A.K. Hamid 627

The measurement of snow water equivalent using natural gamma radiation
R.L. Grasty and P.B. Holman 633

Airborne remote sensing of resistivity through the use of E-Phase techniques
J.H. Davies and J.D. McNeill 647

Detection of atmospheric constituents
J.H. Davies and J.D. McNeill 647

ECONOMIC ANALYSES, PROGRAMS AND PLANS (A)

Decisions on combining data from several sensors
A.H. Aldred 681

The economics of remote sensing of forest land
H. Rae Grinnell 691

Data handling facility of the Canada Centre for Remote Sensing
W. Murray Strome 697

The ERTS experiments of the Canadian Forestry Service
L. Sayn-Wittgenstein and W.C. Moore .. 705

Experiments in aerial remote sensing for highway engineering
C.D. Bricker 713

ECONOMIC ANALYSES, PROGRAM AND PLANS (B)

A proposed organization for the efficient interpretation of remote sensing data
H.W. Thiessen 719

The Canadian Forces Airborne Sensors Familiarization Program
L.M. Sugimoto 723

The application of multispectral remote sensing to the study of soil properties affecting erosion
G.F. Mills 731

High flying hardware and the common man
R.W. Higgs 745

Remote sensing evaluation of environmental factors affecting the developmental capacity of inland lakes
S.J. Glenn Bird 755

SESSION 1 (Cont)

TERRAIN ANALYSES (B)

USE OF STEREOSCOPIC ORTHOPHOTOS FOR SOIL-TERRAIN ANALYSES

L. S. Crosson and R. Protz,
Department of Land Resource Science,
University of Guelph,
Guelph, Ontario.

ABSTRACT

Stereoscopic orthophotos facilitate accurate and precise horizontal and vertical measurements of any terrain feature imaged on them. They are particularly suited to terrain analyses because of their geometric fidelity and the ease with which automation could be implemented. In a study of the use of 1:1000 scale stereoscopic orthophotos for soil-terrain analyses, elevation, slope and aspect were measured on stereoscopic orthophotos and were found to have no significant differences from the same measurements made in the field. When the terrain measurements were related to soil properties it was found that a high correlation existed between distance, elevation and slope measurements and the particle size distribution in the surface soil. Use of a linear regression program enabled the description of 75% of the variability in the sand content, 74% of the variability in the silt content and 71% of the variability in the clay content of the surface soils using only distance from the highest point in the sampling area, elevation and slope as independent variables. The results of this study indicate the value of stereoscopic orthophotos as a tool and a data base for land inventory studies.

INTRODUCTION

In the last several years great strides have been taken in the field of remote sensing in providing up-to-date information on the earth's resources. The development of stereoscopic orthophotos (Collins, 1968) shows promise of meeting the increased demand for precision and accuracy in order to keep pace with other technological advances in compiling land resource inventories. It is now possible for anyone interested in the study of land resources to reap the benefits of stereophotographs that have the planimetric accuracy of a line map and require only the simplest of photogrammetric equipment for analysis.

TERRAIN ANALYSES BY AERIAL PHOTOGRAPHS

Buringh (1966) described two types of stereo aerial photograph analyses which may be used for terrain or landform studies : (1) the analytical analysis in which landform elements are studied separately or in combinations and (2) the physiographic analysis in which physiographic units of the terrain are studied. In general the analytical elements including distance, slopes, aspects, relief and photographic tone may be accurately measured on aerial photographs by photogrammetric or other methods. The analytical analysis can be practised with little knowledge of the terrain. The physiographic analysis requires considerable knowledge of geology, geomorphology and soils. This study is concerned with the former method of terrain analysis.

THE ACCURACY OF STEREO AERIAL PHOTOGRAPH MEASUREMENTS

A combination of relief and tilt displacements make measurements of terrain features on conventional stereo aerial photographs unreliable unless expensive photogrammetric equipment is used to rectify the image before measurements are made and then there is no permanent record of the rectified image.

Most aerial photographs contain some degree of tilt displacements due to pitch, roll and yaw of the aircraft at the time of film exposure. In addition there are relief displacments of variable magnitude which depend upon the relative relief of the terrain, the height of the aircraft above the terrain and the distance of the terrain image from the principal point of the photograph.

Figure 1 (Moffit, 1967) illustrates the difference between perspective projection as obtained when viewing conventional stereo aerial photographs and orthographic projection as obtained when viewing stereoscopic orthophotos and emphasizes the effects of relief displacements caused by perspective projection. The orthographic or map projection of elevation contours

is shown beneath an imaginary terrain. The perspective contours, as shown on the left hand photograph, have been enlarged so that the lowest contour line is at the same scale as that of the map projection, and then superimposed over the map projection. The remainder of the perspective contours are shown as dashed lines. Two important facts are immediately apparent: (1) the scale of the perspective contours increase from the bottom contour line to the top contour line- i.e. each of the contour lines has a different scale; (2) each closed perspective contour line is displaced outwards from the center of the photograph because of the relief displacements- i.e. relief displacements are radial about the principal point.

STEREOSCOPIC ORTHOPHOTOS

The recent introduction of stereoscopic orthophotos facilitates accurate landform measurements from photographic images without the use of expensive equipment and provides a permanent record of the rectified image. Within the limits of accuracy of the production method and the reproduction processes stereoscopic orthophotos offer a simple and direct means of making any desired measurements over the terrain represented by the images (Collins, 1968). Horizontal distances, angles and areas may be measured directly on the orthophoto and terrain elevations can be measured on the stereoscopic orthophotos.

USE OF STEREOSCOPIC ORTHOPHOTOS FOR TERRAIN ANALYSES

A hypothesis was made that if terrain properties can be adequately described by measurements made on stereoscopic orthophotos and if there is a strong relationship between the measured terrain properties and soil properties then stereoscopic orthophotos can be used as a valuable tool for soil-terrain analyses. A study was set up to evaluate this hypothesis.

An 80 by 190 metre rectangular grid with 10 metre spacings was located on the edge of a dissected lacustrine plain immediately south of the city of Brantford, Ontario (Figure 2). The soils in the area are stratified silty clays and silty clay loams with variable depths of sandy overburden. Preliminary investigations indicated that the greater the erosion, the less sandy overburden present, therefore, the grid was located such that various degrees of surface erosion would be represented.

The grid sampling points were numbered according to their X and Y coordinates and the elevation of the surface at each sampling point was measured to the nearest one-hundredth of a metre.

Terrain slopes at each grid sampling point were calculated from the surface elevations and distances between points as explained in Figure 3. The maximum slope at each sampling point was recorded along with the direction in which the slope was facing (aspect).

The surface soils were sampled at each grid point and the percent sand, silt and clay determined by the pipette method.

A set of 1:4000 scale glass diapositives were used to prepare a set of 1:1000 scale stereoscopic orthophotos of the grid by the Guelph Airphoto laboratory. Contact prints were made of the orthophoto and its stereomate. The sampling grid was then carefully located on the orthophoto and elevations, slopes, aspects and photo densities were measured or calculated for each of the grid points. A mirror stereoscope with a 3X magnification and a floating dot parallax measuring attachment was used to measure the elevations. A reflection densitometer with a 4 mm aperature was used to measure the photo densities.

All data was recorded on computer cards. A computer program was used to determine the highest point in the grid (apex) and the distance from the apex to all other points in the grid. The distances were then added to the data.

A simple correlation program was used to relate the field measured surface elevations to the elevations measured on the stereoscopic orthophotos and to relate the particle size distribution in the surface horizon of the grid to the terrain properties measured on the stereoscopic orthophotos.

A "F"-test was used to determine if there was a significant difference between field measured elevations and stereoscopic orthophoto measured elevations.

The residuals between the field measured elevations and the stereoscopic orthophoto measured elevations were plotted in grid form.

A stepwise multiple linear regression program

was used to estimate the particle size distribution of the surface horizon using the terrain data obtained from the stereoscopic orthophotos as independent variables.

RESULTS AND DISCUSSION

The accuracy of the elevations measured on the stereoscopic orthophotos was tested in three ways:

- The calculated simple correlation coefficient was 0.998 which indicated that the field and orthophoto measured elevations behaved in much the same manner over the entire grid.

- The "F"-test gave a calculated "F" value of <0.01 which indicated that statistically there were no significant differences between the two methods of measuring elevation.

- The residual plot (Figure 4) indicated that over 95% of the elevations measured on the stereoscopic orthophotos were within $\pm$ 15 cm of the field measurements and the remainder of the orthophoto elevations were within $\pm$ 20 cm of the field measurements.

Table 1 lists the simple correlation coefficients between the stereoscopic orthophoto measurements and the particle size distribution of the surface soil.

Table 1: Simple correlations between terrain measurements and particle size distribution of the surface horizon.

	% Sand	% Silt	% Clay
Distance from apex	- 0.82	0.83	0.75
Elevation	0.72	- 0.71	- 0.73
Slope	0.52	- 0.51	- 0.55
Aspect	- 0.27	0.29	0.17
Photo density	0.23	- 0.23	- 0.24

The correlation coefficients indicated that there was a relatively high correlation between distance from the apex and particle size distribution and between elevation and particle size distribution as was indicated by the preliminary examinations of the area- i.e. the sand content of the surface horizon tended to decrease and the silt and clay contents tended to increase with distance from the least eroded or highest point in the grid. There was a moderate correlation between terrain slope and particle size distribution and relatively low correlations between aspect and particle size distribution and photo density and particle size distribution. The relatively low correlation with photo density was suspected because of the many factors which affect photo density.

The stepwise multiple linear regression analysis gave the following regression equations for the particle size distribution of the surface soils:

- % Sand = -830.551 - 0.267 (distance) + 3.996 (elevation) + 0.446 (slope)
  $r^2 = 0.75$

- % Silt = 710.236 + 0.215 (distance) - 3.042 (elevation) - 0.281 (slope)
  $r^2 = 0.74$

- % Clay = 227.501 + 0.051 (distance) - 0.956 (elevation) - 0.155 (slope)
  $r^2 = 0.71$

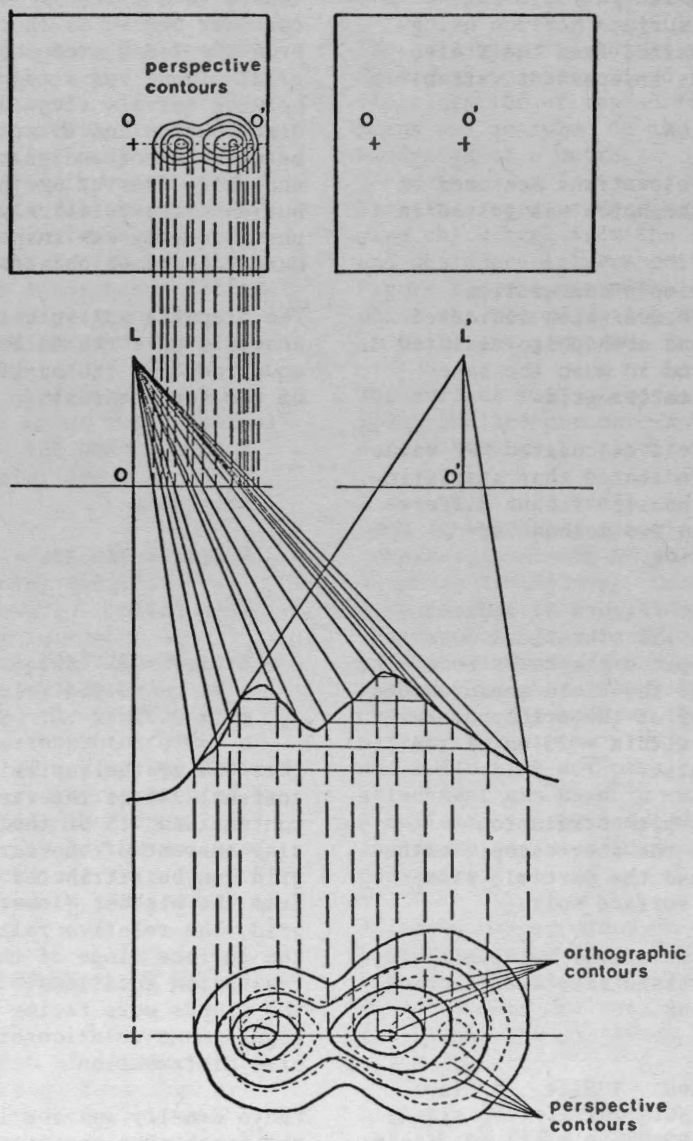
i.e. 75% of the variability in the sand content, 74% of the variability in the silt content and 71% of the variability in the clay content of the surface horizon of the grid can be attributed to the linear distance from the highest (least eroded) point in the grid, the relative relief of the terrain and the surface slope of the terrain by means of regression equations. The direction in which the slopes were facing (aspect) had no significant relationship with the particle size distribution.

Photo density was the last variable to enter the regression equations and contributed very little to the reduction in sums of squares, it was therefore left out of the equations.

CONCLUSIONS

Since there were no significant differences between the field measured elevations and the stereoscopic orthophoto elevations it must also be true that there are no significant differences between slopes calculated from the field and stereoscopic orthophoto elevations. Similarly there must be no significant differences in aspects of the slopes. Therefore, it can be concluded that stereoscopic orthophotos may be used as a valuable tool for accurately measuring terrain properties.

From the results of the correlation and regres-



O = principal point of the left hand photograph
O^1 = principal point of the right hand photograph
L = exposure station for the left hand photograph
L^1 = exposure station for the right hand photograph

Figure 1. Perspective and orthographic projection (taken from Moffit, 1967)

sion analyses it can be concluded that at a given level of abstraction there is a strong relationship between terrain properties and soil properties. Since stereoscopic orthophotos can be used to accurately measure terrain properties it may also be concluded that stereoscopic orthophotos are a valuable tool for soil-terrain analyses.

APPLICATIONS

The results of this study have indicated that stereoscopic orthophotos can be an invaluable tool for soil studies. It would be necessary to conduct detailed soil-terrain studies in selected areas only and then extrapolate to other areas by deciphering soils data from landform measurements made on stereoscopic orthophotos. In addition stereoscopic orthophotos are well adapted to automatic terrain analyses which would greatly improve the efficiency of soil-terrain studies. Loelkes (1969) has proposed the value of orthophotographs as data base for many uses. The advent of the stereoscopic orthophoto has increased the utility of orthophotos and they should be seriously considered as the data base for the majority of land resource studies.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to the Canada Department of Agriculture for grant No. 9079 which enabled this research to be carried out.

The authors are indebted to Professor S. H. Collins, Mrs. J. Law and Mr. R. Shillum of the Guelph Airphoto Laboratory at the University of Guelph for the use of equipment and the production of the stereoscopic orthophotos.

REFERENCES

Buringh, P. 1966. In (Manual of Photographic Interpretation. American Society of Photogrammetry, Publishers). pp. 633-666.

Collins, S. H. 1968. Stereoscopic Orthophoto Maps. The Canadian Surveyor, Vol. 23, No. 1: 167-176.

Loelkes, G. L. 1969. Orthophotography an optimum data base. Proceedings of the ASP-ACSM Fall National Meeting, Portland, Oregon, September 23-26, 1969. pp. 149-218.

Moffit, F. H. 1967. Photogrammetry. International Textbook Company Publishers. pp. 216-218.

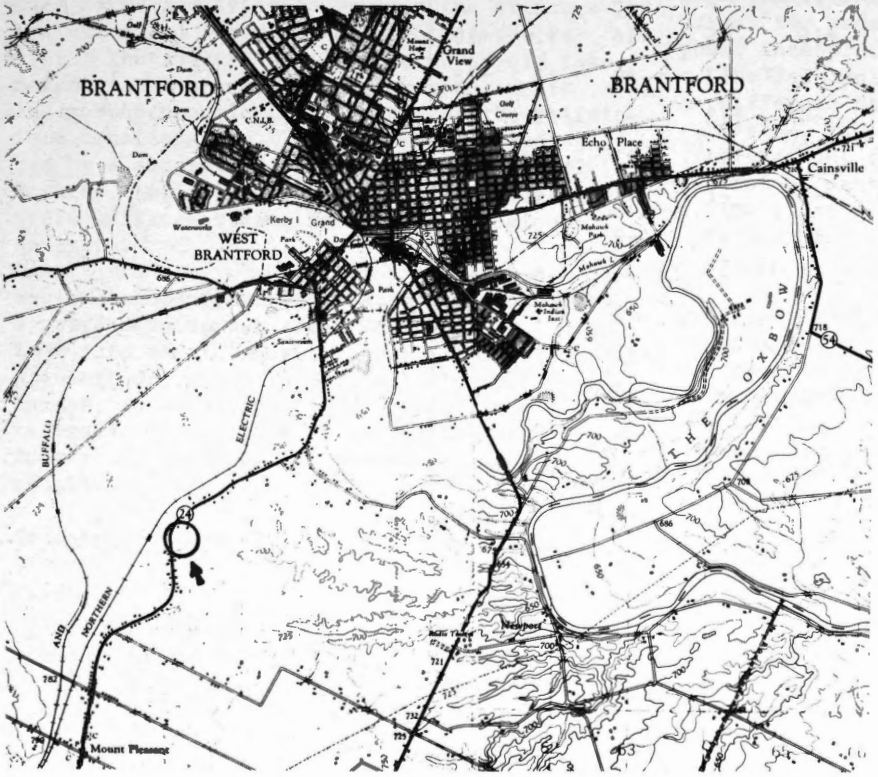
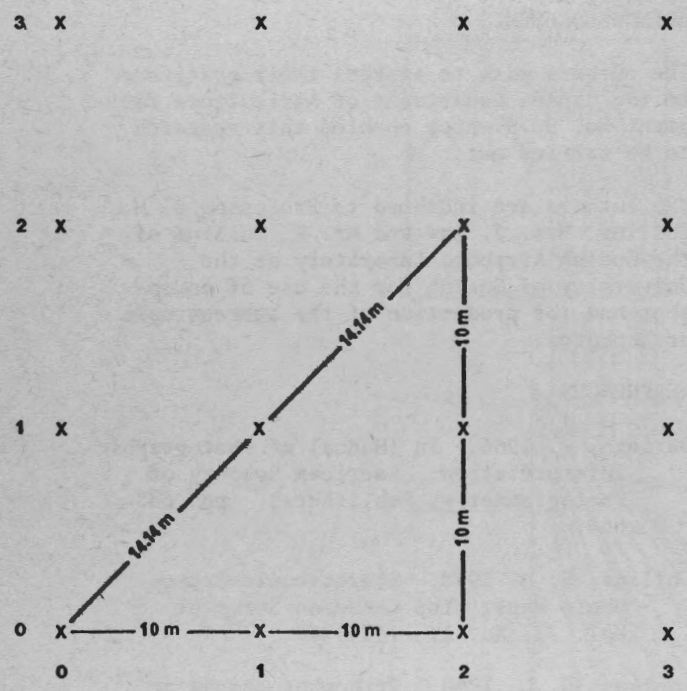


Figure 2. Topographic map showing location of research site.



$$\% \text{ Slope at } (1,1) = (1)\ \frac{h_{(0,0)} - h_{(2,2)}}{28.28} \times 100$$

$$(2)\ \frac{h_{(o,1)} - h_{(2,1)}}{20.00} \times 100$$

$$(3)\ \frac{h_{(1,2)} - h_{(2,0)}}{28.28} \times 100$$

$$(4)\ \frac{h_{(1,2)} - h_{(1,0)}}{20.00} \times 100$$

where: h = elevation in metres

Figure 3. Calculation of surface slope

Y									
19	+1	-1	+14	-6	-6	+6	-10	+4	-2
18	-1	-2	- 8	+15	+2	-14	+18	+6	+9
17	+2	-13	+15	+8	+4	-3	+13	+7	+8
16	+15	-1	+6	+2	+7	+5	+12	-6	-2
15	+4	+5	+14	+6	+15	+5	+14	+2	+10
14	-2	+1	+3	+7	-5	-7	+9	+5	+18
13	+20	-6	+1	+5	-4	+13	-1	+3	-1
12	-12	-1	+15	+8	0	+14	+1	-7	+8
11	+6	-4	+20	0	-3	-1	-1	+14	+9
10	-8	+12	-13	-13	-2	-7	-9	+10	+5
9	+4	+2	-4	+9	+9	-2	-4	-6	+2
8	-9	+5	+3	-6	+11	-11	-5	+11	-12
7	0	-15	+6	-7	+14	+7	-14	-2	-4
6	+1	+17	+5	-3	-1	+3	+3	+6	+3
5	+1	-5	-7	-7	+5	-6	-16	0	+7
4	-14	+3	+8	-8	-1	-5	0	+5	+3
3	-5	-2	-5	+2	-4	-5	-5	-14	+11
2	-1	+6	-3	-10	+2	-1	-5	-12	-1
1	+11	+6	-4	-3	-2	-6	+5	-4	+1
0	+2	-2	-15	+2	-3	+12	+12	-3	-4
	0	1	2	3	4	5	6	7	8

X

Figure 4. Residuals (in cm) between field measurements of surface elevation measurements made on stereo orthophotos

SUPPLEMENTARY RECONNAISSANCE AIR PHOTOGRAPHY AS AN AID FOR DETAILED GEOMORPHOLOGICAL MAPPING

A. Kesik
Associate Professor
Department of Geography
University of Waterloo, Waterloo, Ontario

ABSTRACT

Geomorphological mapping today represents an important method for regional geomorphological surveying. It makes use of data concerning land forms, including existing air photographs. The more widespread application of air photography and photo interpretation becomes feasible through the use of cheap, simple and flexible systems of small frame cameras (35 or 70 mm). This is referred to as Supplementary Air Photography (SAP).

During a geomorphological mapping, project in the area covered by the map, sheet Galt 1:50.000, (SW Ontario), in summer 1971, four Hasselblad cameras were used for this type of photography.

INTRODUCTION

Geomorphology as an earth science deals with land forms and tries to explain the origin, evolution and future development of the relief. The analysis of land forms should take morphographic, morphometric, morphogenetic and morphochronological aspects of the relief into consideration. Each of these aspects can be a separate topic of geomorphological studies, but only comprehensive investigations yield sufficient data for valuable regional analysis.

Regional geomorphological surveying underwent a rapid development after World War II, both in Europe and in America. One of the principal methods, which is particularly applied in Europe, is geomorphological mapping which can be regarded as a form of surveying with the construction of special thematical maps as an objective.

Geomorphological maps at large scale should present a proper balance of four main geomorphological aspects: morphography, morphometry, morphogenesis and morphochronology. The legend for such a map was discussed by the Subcommission on Geomorphological Mapping of the International Geographical Union (11) and a proposal has been published together with a recommendation for use in detailed mapping (3, 9, 12).

Very often the lack of sufficient geomorphological data causes difficulties in the map compilation phase. Many recently published geomorphological maps have a limited legend which only partially corresponds to the comprehensive proposals for detailed maps by IGU Subcommission.

Today the importance of air photography and and photo interpretation in geomorphological surveying is universally accepted. Examples of analytical interpretation can be found in the well known book by Lueder (7) or in the French publication "Photo Interprétation" (10). A review of the application of air photographs to dynamic and regional geomorphological studies was recently published by F. Fezew (5).

Publications dealing with the methodology of geomorpholgical surveying and mapping (9,13) stress two aspects of air photography and photo interpretation in particular:

a. The far applying of air photographs at different stages of geomorphological surveying, including preliminary photo interpretation, field check and final compilation of data.
b. The great usefulness of sequential photographs for obtaining adequate morpho-dynamic data.

In spite of the acknowledged wide applicability of air photographs to geomorphological surveying, the material used, is very often rather incidental with respect to time, scale and technique. Frequently, the photos have been taken for other purposes, e.g., topographic mapping or forest inventory. It is a rather exceptional situation when the photo mission is organized especially for geomorphological purposes, with the time of photography, scale and film-filter combination being optimal from the geomorphological point of view.

Optimal air photographs can be briefly described as those which are taken in the season

when the morphographical and morphodynamical characteristics of the land forms are most visible. This problem is important, because a quantified geomorphology should have a greater input from the photo interpretation techniques.

Geomorphologists will apply air photographs more often when the photographic system is cheaper, simpler in operation and available at the time it is needed. A complete system of SAP should consist of:

1. A set of reconnaissance cameras (35 or 70 mm) fitted into a frame that can be mounted easily in light aircraft. The cameras should be connected to an intervalometer, so that they can be triggered at desired regular intervals.
2. A light aircraft with a cabin spacious enough for the camera mount and for convenient operation during the mission.
3. A photographic darkroom equipped for the controlled processing of, at least black and white air photo films.
4. A photo interpretation laboratory with the necessary instrumentation for visual photo examination and photometric measurements.

At present, many educational institutions in Canada have elements of such a system. For example, at the University of Waterloo we have a set of Hasselblad cameras, whereas there is an airphotographic darkroom at the University of Guelph. It would be beneficial, not only to geomorphology, but to other sciences as well, to collect information about existing equipment and to organise working groups which then could make joint use of such equipment and carry out cooperative projects.

SUPPLEMENTARY AIR PHOTOGRAPHY IN GEOMORPHOLOGICAL SURVEYS

Supplementary air photography should be taken for the purpose of acquiring additional information on land forms and morpho-dynamic processes. To attain greatest profit possible from SAP, the following aspects should be taken into consideration:

a. The selection of objects and terrains to be photographed.
b. Time of photography.
c. Scale of photography.
d. The selection of cameras and photographic material.

The following is a brief discussion of these points.

a. The Selection of Objects and Terrains to be Photographed

Morpho-structural and morpho-climatic factors control the intensity as well as the character of morpho-dynamic processes, which consequently show a spatial differentiation. Zones of slow morphological evolution alternate with zones of intensive morphological activity, where the action of such agents as water or ice, brings about positive or negative changes in the existing relief.

In each morphological landscape or group of land forms (micro, meso, macro), it is possible to distinguish zones, or elements of more active land forms, from those with more stable features. The proper recognition and the qualitative as well as quantitative characterisation of such morphological "catenas" is essential for morphological regionalisation and for predicting the morphological evolution of the terrain.

The dynamic approach to morphological landscape analysis requires the concentration of observations and measurements on the dynamic elements of landforms. Clearly there is a need for more versatile data. These should be collected at time intervals appropriate for the quantitative description of morpho-dynamic evolution.

It is here that geomorphologists can gain additional information, exceeding that available from conventional photo coverages, but taking supplementary photography in morpho-dynamic key areas.

Such areas should be carefully selected on the basis of an analysis of topographic, geologic and soil maps as well as a field or air reconnaissance of the terrain. The examination of photo mosaics and existing conventional photographs is also helpful.

If, for the selection of sample areas, such a supporting data are not available, then the researcher should use his general knowledge of geomorphology and air photo interpretation.

b. Time of Photography

What is the best season for geomorphological air photography? The problem is important, but there is no universal and precise answer. In general, the season of photography should correlate with the season of maximum activity of morpho-dynamic processes. These processes are geographically differentiated and depend upon such factors as climatic conditions, topographic and geographical location, morpho-structural characteristics and man's activity.

With the use of climatic data the best season of photography can be determined in broad terms only; e.g., beginning of spring, or period of summer rains. A more precise timing is complicated by climatic fluctuations which can affect the duration of the best season as well as the frequency of favorable conditions.

It is also important to remember that some microforms of relief whose origin and development are related to man's activity, are short-lived and can be recorded on air photographs at particular seasons and during short periods of time only. The capability of a supplementary air photography system to detect such anthropogenic land forms, especially in areas with a great impact of man on environment is another advantage.

c. Scale of Supplementary Air Photography

The scale of SAP for detailed geomorphological mapping should be large enough to permit the recognition of micro and meso forms of relief, as well as of indicators of dynamic processes. The range of suitable scales is probably from 1:5,000 to 1:40,000 with the optimum between 1:5,000 and 1:15,000. Scales larger than 1:5,000 may cause problems for the photographic mission, due to the limitation of the camera cycle, when 60% or more overlap is required. For Hasselblad cameras the practical limit is an interval time of 2 sec.

Photographs at a scale smaller than 1:40,000 are useful for geomorphological reconnaissance but are more difficult to obtain from small aircraft.

d. Selection of Cameras and Photographic Material

Supplementary air photography can be obtained by using commercially available cameras equipped with motor for film rewind. Several publications (1, 2, 4, 8, 14) describe applications of 35 or 70 mm cameras. New models of Nikon F-2 Photomic and Cannon F-1 (35 mm cameras) have a maximum shutter speed of 1/2000 sec. and can expose to 5-7 frames per sec. During the selection of cameras for a SAP system the size of imagery, the capacity of the film magazine, the most desirable scale of photography and the category of aircraft have to be considered.

SAP with more than one camera makes the simultaneous use of different combinations of film and filters possible. This leads to multispectral photography which usually employs one or two black and white films (panchromatic or infrared) with a combination of various filters.

If conventional colour, as well as false colour photography is desirable, in addition to black and white films, the number of cameras needed increases to four. A set of four reconnaissance cameras, 35 or 70 mm, seems to be the present standard.

The interpretation of morphological features is based on both direct and indirect indicators. Indirect indicators such as hydrography or vegetation are important keys for geomorphological mapping. For this reason, the use of infrared films in combination with true colour or panchromatic film is preferable.

If only two cameras can be accomodated, the following order of preference applies:

1. Conventional colour - Colour infrared (false colour)
2. Conventional colour - Black and white infrared
3. Colour infrared - Panchromatic
4. Panchromatic - Black and white infrared

Regional morphological differences, local hydrographical conditions, as well as land use or natural vegetation patterns play an important part in the selection of the best combination of films.

SUPPLEMENTARY AIR PHOTOGRAPHS FOR THE MAP, SHEET GALT, 1:50,000

Technical Characteristics

Supplementary air photographs for the Galt area were made in connection with the geomorphological mapping, project for sheet Galt (40P/8 West), carried out by the author in summer 1971 (Research grant from NRC). Geological characteristic of the Galt terrain is presented in Karrow's report (6).

During the presentation phase a preliminary photo interpretation was carried out on the basis of conventional air photographs at a scale of 1:24,000 (Brant County, Wild RC5, May, 1965) and 1:15,000 (Waterloo County, Wild RC5, April, 1968). The preliminary photo interpretation included an examination of the air photographs under mirror stereoscope and a delineation of the principal morphological forms and boundaries.

During the mapping season it was possible to organise three short photo missions using four Hasselblad cameras set up in a special fram (fig. 1) and connected to a Robot

intervalometer. The Hasselblad camera system was originally developed at the Department of Geography, University of Waterloo, for another project (see paper on "Multispectral - Multitemporal Air Photography and Automatic Terrain Recognition" by Dieter Steiner). A Beaver-Havilland aircraft was available due to the courtesy of the Ontario Department of Land and Forests.

The purpose of the project was to investigate the applicability of SAP for geomorphological surveying and mapping. One particular goal was to examine photo scale, films and season of photography as parameters. Two perpendicular flight lines (N-S and E-W) crossing principal morphological units were chosen for these missions. Due to some technical difficulties the program could be realized only in part. The greatest drawback is the fact that no photographs could be taken during spring season. The timing for the three coverages, acquired later in the year, was not entirely ideal and was partly dictated by aircraft availability and weather condition. Photo mission data are presented in Table 1.

The black and white films were processed under sensitimetric control at the Air Photo Laboratory, University of Guelph, whereas the colour films were processed by the National Air Photo Library in Ottawa. So far the quality of the processed films have been evaluated visually by means of a light table and a Bausch and Lomb Stereo Zoom 70 instrument.

It should be noted also that, once the cameras are in operation during a flight, lens aperture and/or exposure time cannot be varied without interrupting the coverage. This may have an adverse affect on film exposure particularly if the illumination of the terrain changes rapidly during the mission.

The visual examination of the films shows that the greatest differences in optical densities appear on the Infrared Aerographic film. One lesson learned was that care should be exercised when working with black-and-white film. It is considerably thinner than regular film and a light leakage in a cassette can affect it more readily. As a matter of fact, the exchange of cassettes in the aircraft caused an exposure of approximately the first two feet of film. Consequently, the use of spare magazines preloaded in a darkroom becomes a necessity.

False colour film appeared to be less sensitive to the variations in terrain illumination that Infrared Aerographic. Some transparencies have a low degree of colour saturation (overexposure) but, in general the well known advantages of this film were confirmed.

The conventional colour (Ektachrome) and the panchromatic Tri-X films were more adequately exposed as a result of their greater exposure tolerance. This is an important aspect for SAP. More precise comparisons can be done only on the basis of sensitometric measurements.

Photo Interpretation Procedure

The original colour reversal films and black-and-white diapositive copies were examined under a William's viewer with a Bausch and Lomb Stereo-Zoom 70. All photographs were analised using direct and indirect criteria for the identification of geomorphological features. Previously existing data as well as field observations and knowledge of the terrain were compared during the photo interpretation process with the information derived from the SAP. This procedure is illustrated by some examples.

Pinehurst Morainic Ridge

The Pinehurst Morainic Ridge lies south of Galt, between HW. 24A and the Grand River Valley. It is 2½ km long and oriented NW-SE; it has relative elevations of 80-100 feet and is dissected by three parallel subglacial channels which are visible on the stereogram in the form of elongated depressions partly occupied by lakes. (fig.2). The lakes are in different stages of hydrographical evolution due to processes of sedimentation and succession of vegetation.

The morainic ridge changes through 90° the direction of the main channel of the Grand River. A steep valley slope with relative elevations of approximately 150 feet is strongly modulated by cut-back erosion and by mass movements. Northeast and southwest from the morainic ridge there are gentle slopes modulated by fluvial erosion. The drainage pattern is visible on all photographs. The slopes change gradually into outwash plains, which, in NE corner of the photo, are dissected by the Grand River Valley. (Fig.3)

A comparison of a conventional, 1965 air photographs with the SAP of July 1971 shows the following differences:

a. Panchromatic photographs of 1965 and 1971 exhibit a different content of morphologically important information. This is mainly due to a difference in season of

photography. On the photographs taken in May 1965 the topographic surface in the morainic ridge zone is much better visible since dense tree foliage is absent. After the spring thaw lakes show a higher water level and no sign of the off-shore vegetation which occurs in shallow water during the summer. The erosional bluff of the Grand River Valley as well as land slides on the slope also are better visible on the May photographs.

b. Infrared Aerographic film, in comparison with panchromatic Tri-X film is clearly better in the protrayal of water surfaces and drainage patterns but there are only small differences in photographic density between hardwoods and pastures (Fig. 4).

c. Of the two colour films, the false colour one is better with respect to all hydrographical features, drainage patterns and boundaries between different plant formations.

A suitability ranking of the films for geomorphological mapping in the Pinehurst Morainic Ridge area is the following: 1. False colour; 2. True colour; 3. Panchromatic Tri-X; 4. Infrared Aerographic.

Hummocky Moraine Landscape (Galt Moraine)

The photographs shown in Fig. 5 illustrate the topographic surface of the Galt Moraine, approximately 9 km south of Galt and 5 km ENE of Glen Morris. The topographic surface of the hummocky moraine consists of irregular hills, ridges and hollows with relative elevations of 30-60 feet. The drainage pattern is in general, irregular. Some depressions do not have a superficial outflow, and some are connected by ditches. The east slope and part of the lacustrine accumulation surface are made of younger elements of morphology connected with the postglacial period. (Fig.6)

A comparison of the conventional photographs taken in May 1965 with the SAP of August 1971 demonstrates that the summer photographs in spite of their smaller scale contain more information concerning land form than the May photographs. Differences in vegetation are correlated with geomorphological features more clearly. Hills and depressions as well as drainage pattern elements can be detected easily. Special attention should be paid to the interesting rectangular ground pattern which is clearly visible in the upper left corner of the photograph. (Fig.7)

An analysis of the colour films* shows that, in the hummocky morainic terrain with a predominance of pasture land relatively small elevations, correlated with differences in moisture contents, are more readily apparent on false colour film.

It would appear that the suitability ranking of the films is close to the one derived from the first example.

The next two examples are at a different scale, approximately 1:7,600. The photographs show small elements of meso or microforms of the relief.

Nith River Valley (1 km West of Ayr)

Conventional photographs at a scale of 1:15,000 were taken in April 1967. The following features are clearly visible on them: river channel, fresh sediments accumulated along the channel, and the steep slope which is visible in spite of the forest cover.

The four photographs of October 1971 show interesting differences. The panchromatic film is a poor differentiator of vegetation and this effectively masks the presence of an old river channel. Partly, this may also be a result of the season. This channel is visible on both infrared photographs (Fig.8). The infrared photographs, particularly the false colour bring out differences in vegetation and also give more information in shadows. Geomorphological sketch is presented on Fig. 9.
The best single film from a geomorphological point of view is false colour one.

Example of Soil Erosion (Slope of Nith River Valley, 3 km W of Ayr

The SAP of October 1971 show a complex of microforms developed on the Nith River Valley slope, which has a gradient of approximately 15°. The following microforms can be distinguished: 1. Erosional rills; 2. Accumulative venners (small fans); 3. Zones of degraded soil (Fig. 10,11)

The relationship between the distribution of microforms and the orientation of principal man-made features such as roads and small scarps is interesting. In the upper part of

*Infrared Aerographic film was not obtained due to a camera failure.

the photo a small scarp intercepts running water. Accumulated material increase the relative elevation of the scarp. The main road also stops the development of rills, but unconsolidated excess material accumulated on the road has been washed out onto the neighbouring field.

A comparison of the four different images shows that microforms are best detectable on the false colour film, followed by the true colour and panchromatic films. The poorest image in this case is provided by the Infrared Aerographic film.

CONCLUSIONS FROM THE EXAMINATION OF SUPPLEMENTARY AIR PHOTOGRAPHS

The results of the visual examination of SAP can be summarized as follows:

1. The photographs taken in July and August at a scale of 1:30,500 demonstrate their general applicability for geomorphological surveying and mapping, particularly in nonforested areas. For forest-covered terrains it is necessary to take photographs early in spring or late in fall.

2. The photographs at a scale of 1:7,600 show the variability and abundance of microform and they can be used for qualitative and quantitative analysis of particular elements of relief.

3. Out of the four films used during the experimental SAP project the most suitable for geomorphological purposes was the false colour film, followed by the conventional colour and the panchromatic Tri-X films.

For future applications of SAP for geomorphological mapping it is necessary to strive for optimal conditions. It is advisable to consider the following problems and suggestions:

1. There should be a better synchronization of photography with the season of morphodynamic activity. In our geographical zone the majority of photographs should be obtained at the beginning of spring.

2. To avoid a duplication of conventional photographs, in terms of scale, it is recommended to take SAP at scales larger than 1:20,000.

3. Supplementary air photography can be restricted, if necessary to two cameras loaded with false colour and true colour films. Even if only two cameras are required, having four cameras has the advantage, that banks of two can be used alternately, then enabling a continuos operation.

4. For each project sample areas should be selected carefully. This work should be done during the preliminary photo interpretation phase, and, if possible in connection with a field reconnaissance.

It is also suggested that further advances in the application of SAP to geomorphological surveying could be brought about by an improved technical and scientific cooperation between persons and institutions involved in the development of earth sciences and aeromethods.

REFERENCES

1. Aldrich, R.C., Bailey, W.F., Heller, R.S., 1959. Large-scale 70 mm colour photography techniques and equipment and their application to a forest sampling problem. Photogramm. Eng., 25(5): 747-754.

2. Aldrich, R.C., 1966. Forestry application of 70 mm colour. Photogramm. Eng., 32(5): 802-810.

3. Bashenina, N.V., Gellert, J.F., Jolly, F., and others. The unified key to the detail geomorphological map. IGU. Commission on Applied Geomorphology. Krakow. 1968.

4. Corneggie, D.M., 1968. Large Scale 70 mm photography - a potential tool for improving Range Resources Inventories. Proceedings of 34th Ann. Meeting of Am. Soc. of Photogrammetry.

5. Fezer, F., 1971. Photo interpretation applied to geomorphology - a review. Photogrammetria, 27; 7-53.

6. Karrow, P.F., 1963. Pleistocene Geology of the Hamilton-Galt Area. Ontario Department of Mines. Geological Report No. 16. Toronto.

7. Lueder, D.R., 1959. Aerial Photographic Interpretation. McGraw-Hill Co. N. York.

8. Lyons, E.H., 1967. Forest sampling with 70 mm fixed airbase photography from helicopters. Photogrammetria, 22: 213-231.

9. Manual of Detailed Geomorphological Mapping (Preliminary Edition), 1971. IGU, Commission on Geomorphological Survey and Mapping, Brno.

10. Photo Interprétation. Édition Technip. Paris.

11. Problems of Geomorphological Mapping. 1963. Institute of Geography, Polish Academy of Science, Geographical Studies, No. 46. Warsaw.

12. St.-Onge, D. Geomorphic Maps, 1968. The Encyclopedia of Geomorphology. Reinhold Book Corp. p. 388-403.

13. Verstappen, H.Th., Van Zuidam, R.A. 1968. ITC. System of Geomorphological Survey. ITC Textbook of Photo-Interpretation. Vol. VII. Delft.

14. Zsilinszky, V.G. 1970. Supplementary Aerial Photography with miniature cameras. Photogrammetria. 25: 27-38.

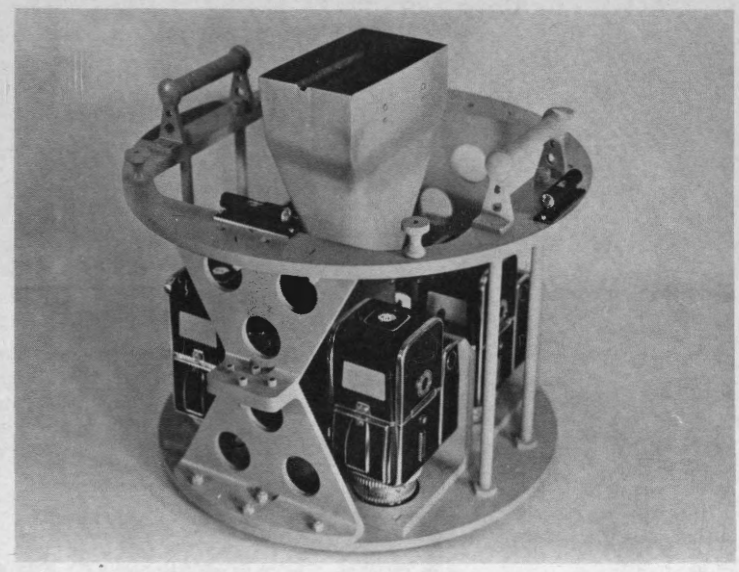


Figure 1. Supplementary air photographic system (SAP) developed at the Department of Geography University of Waterloo. Frame with four Hasselblad cameras.

Table 1. Characteristic of Photo Missions.

Date & Time of A.P.	Cameras Lenses	Films	Filters	Relative Aperture ( )	Exposure (sec.)	Flying Height Above Ground	Approx. Scale	No. of Pictures	Weather Conditions
15 June 1971 T=11:30 a.m.	4 Hasselblads, model EL, with Biotars f = 80mm. Remote control with intervalometer Robot	Tri-X	Wratten 12	16½	1/500	8,000 ft.	1:30,500	64	Sunny, few clouds, haze
		IR-Aerographic	Wratten 87	5.6½				30	
		Ektachrome	HF (3 + 4)	5.6½				64	
		IR-Aerochrome	Wratten 12	5.6½				64	
12 August 1971 T=11:30 a.m.	3 Hasselblads model EL (as above)	Tri-X	Wratten 12	22	1/500	8,000 ft.	1:30,500	72	Sunny, few clouds
		Ektachrome	HF (3 + 4)	8				72	
		IR-Aerochrome	Wratten 12	5.6½				72	
7 October 1971 T=10:50 a.m. 01:10 p.m.	4 Hasselblads model EL (as above)	Tri-X	Wratten 12	11-16	1/500	2,000 ft.	1:7,600	136	Cloudy, short periods of sun, haze
		IR-Aerographic	Wratten 87	4- 5.6				120	
		Ektachrome	HF (3 + 4)	4- 5.6				136	
		IR-Aerochrome	Wratten 12	2.8½-4½				136	

Additional Information: Aircraft speed during all missions was 100 m.p.h. Exposure time intervals calculated for 60% overlap was 15 sec. during the first two missions, and 4.4 sec. during the October mission.

Figure 2. Pinehurst Morainic Ridge. Stereotriplet. SAP, Hasselblad EL, with Biotar 80mm lens, Panchromatic Tri-X film, Wratten 12, F-16½, 1/500 sec. Flying height 8000 ft. above terrain, approx. scale 1:30,500, 15 July, 1971, h - 11,30 a.m.

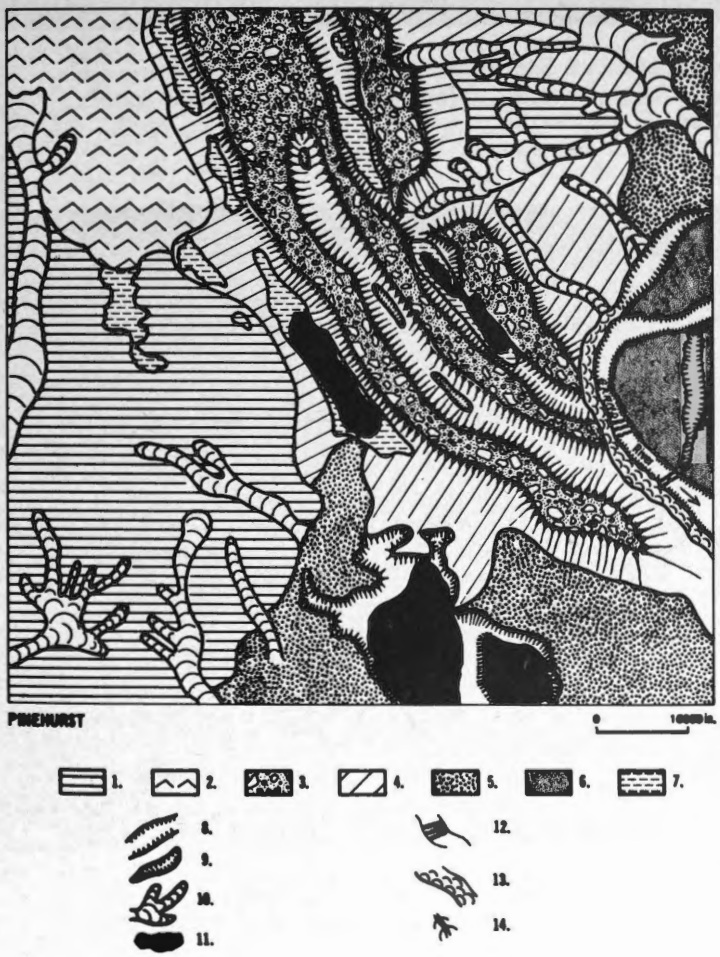


Figure 3. Geomorphological sketch of Pinehurst Morainic Ridge.
Legend: 1. Till plain 2. Hummocky till plain 3. Morainic ridges 4. Slopes of morainic ridges and subglacial channels. 5. Glacio-lacustrine plain. 6. Holocene Grand River terraces 7. Holocene lacustrine and organogenous plains 8. River bed cut in Quaternary deposits 9. a) Abandoned loops (cut-offs) b) Depressions in subglacial channels. 10. Dry, denudation valleys. 11. Lakes 12. Steps in river bed. 13 Undercutted bank with land-slides and slumps 14. Gulliess

A B

Figure 4. Pinehurst Morainic Ridge . Comparison of panchromatic and infrared photographs. SAP,Hasselblads EL, 15 July,1971,h=11,30 a.m., 1/500 sec.,Flying hight 8000ft.
a) Panchromatic,Tri-X film, Wratten 12, F-16½
b) Infrared Aerographic film, Wratten 87, F-5,6½

Figure 5. Hummocky morain landscape (Galt Morain). Stereotriplet. SAP, Hasselblad EL with Biotar 80mm, Panchromatic Tri-X film, Wratten 12, F-22, 1/500 sec. Flying height 8000 ft. above terrain, approx. scale of orginal 1:30.500, 12 August,1971 h= 11,30 a.m.

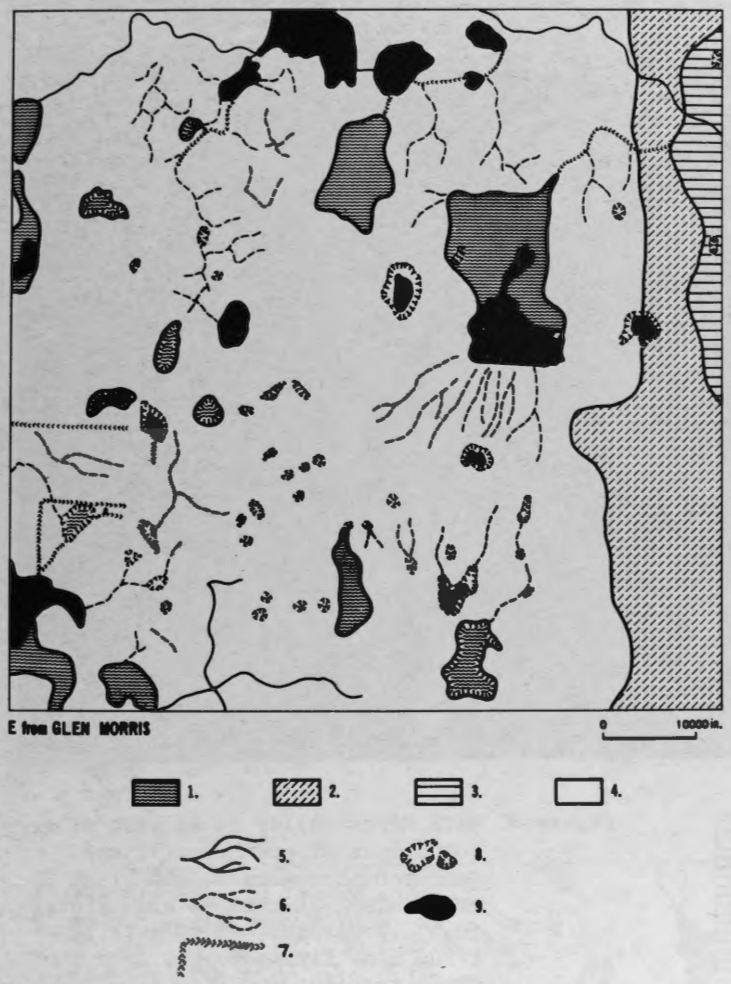
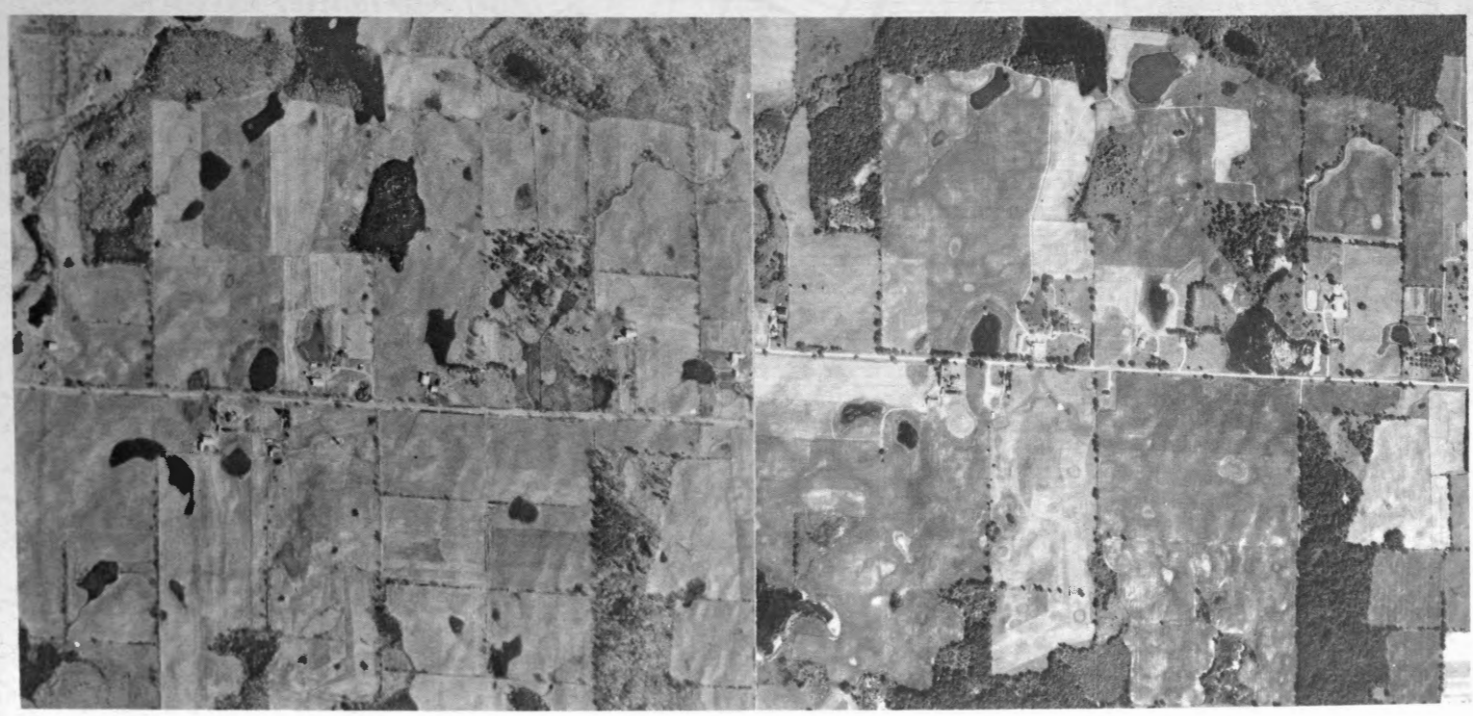


Figure 6. Geomorphological sketch of hummocky morain landscape (Galt Morain) Legend: 1. Kettles 2. Slope of till plain 3. Lacustrine plain. 4. Hummocky till plain. 5. River bed cut in quaternary deposits 6. Rills and small denudation valleys 7. Ditches 8. Small depresions without permanent lakes. 9. Lakes in morainic depressions.

Figure 7. Hummocky morainic landscape (Galt Morain) Comparison of conventional and SAP photographs. a) Conventional air photo, Wild RC5, lens 6', May 1965, scale of orginal 1:24,000 b) SAP, Hasselblad EL, Biotar 80mm, Panchromatic Tri-X film, Wratten 12, 1/500 sec. 12 August, 1971, h-11,30 a.m.



A B

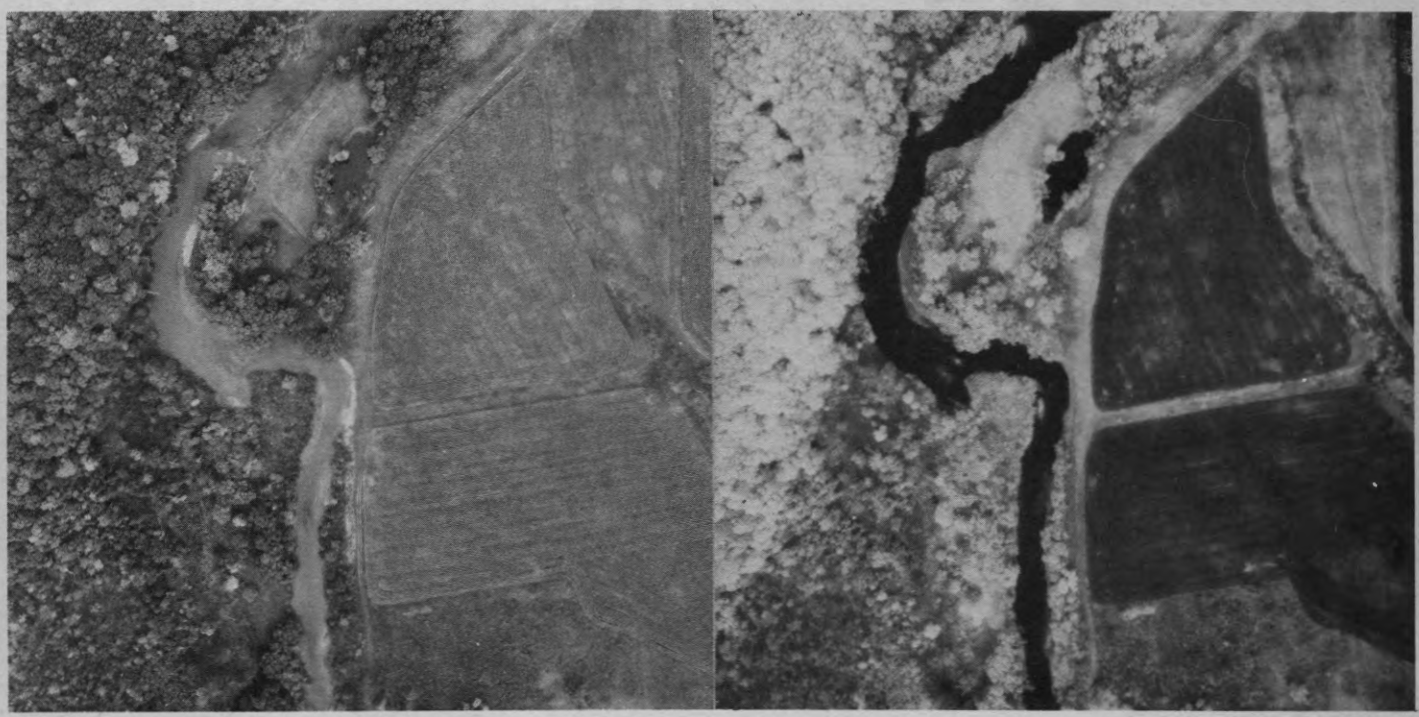


Figure 8. Nith River valley (1 km west of Ayr) Comparison of panchromatic and infrared photographs. SAP, Hasselblad EL cameras with Biotar 80mm, 7 October, 1971, h-12,10 1/500 sec. flying height 2000 ft. above terrain.
a) Panchromatic, Tri-X film, Wratten 12, F-11
b) Infrared Aeorgraphic film, Wratten 87, F-5,6

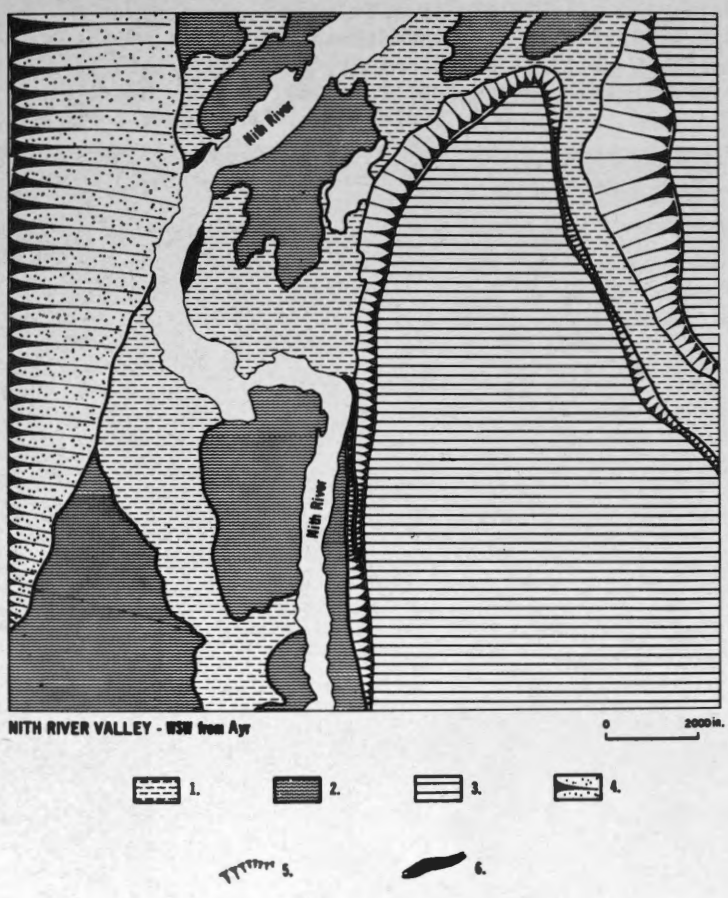


Figure 9. Geomorphological sketch of Nith River valley (1 km west of Ayr) Legend: 1 Seasonal channels, abandoned loops. 2. Low holocene terrace. 3. Holocene cultivated terrace. 4. Slopes of valley cut in the quaternary deposits. 5. Slopes of holocene terraces. 6. River cut scar.

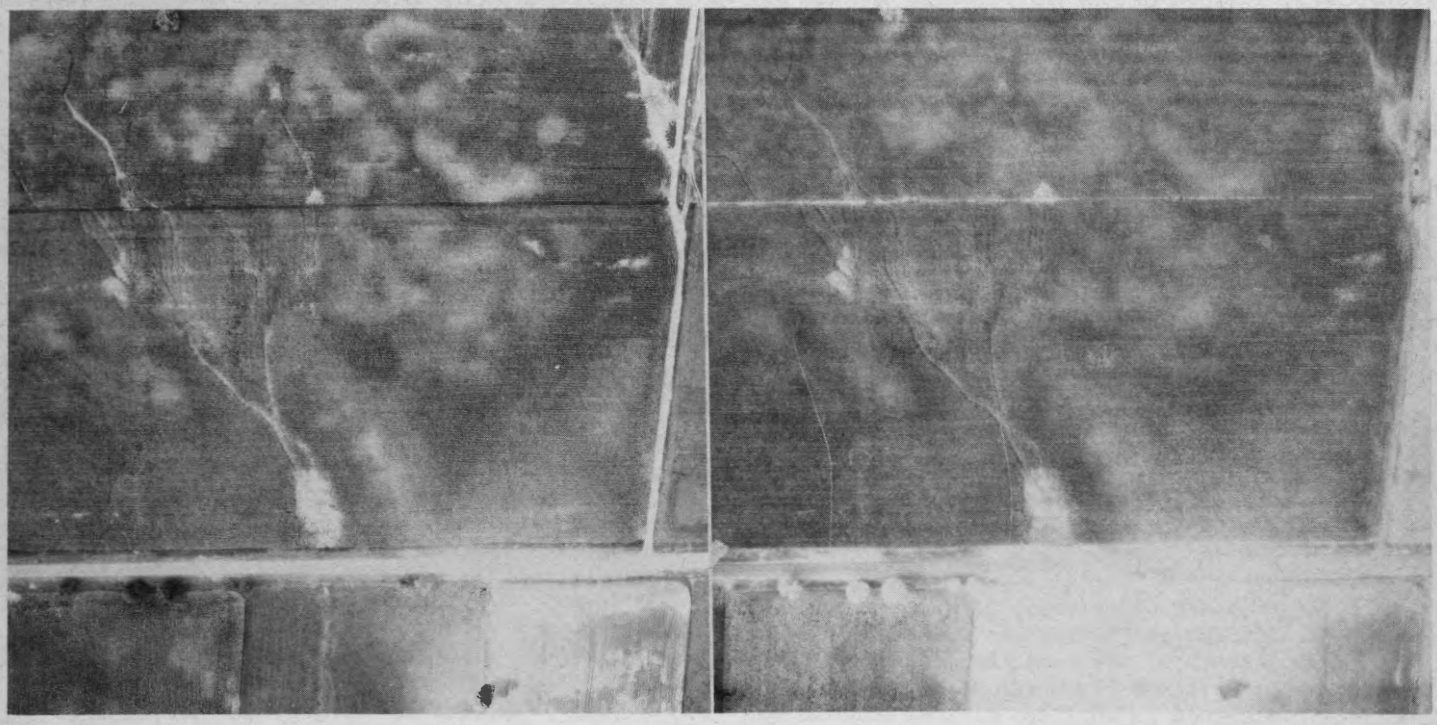


Figure 10 Soil erosion (slope of the Nith River valley, 3 km west of Ayr) Comparison of panchromatic and infrared photographs. SAP, Hasselblad EL cameras with Biotar 80mm, 7 October 1971, h-12,15 p.m. 1/500 sec. Flying height 2000 ft. above terrain.
a) Panchromatic, Tri-X film, Wratten 12, F-11
b) Infrared Aerographic film, Wratten 87, f-5,6

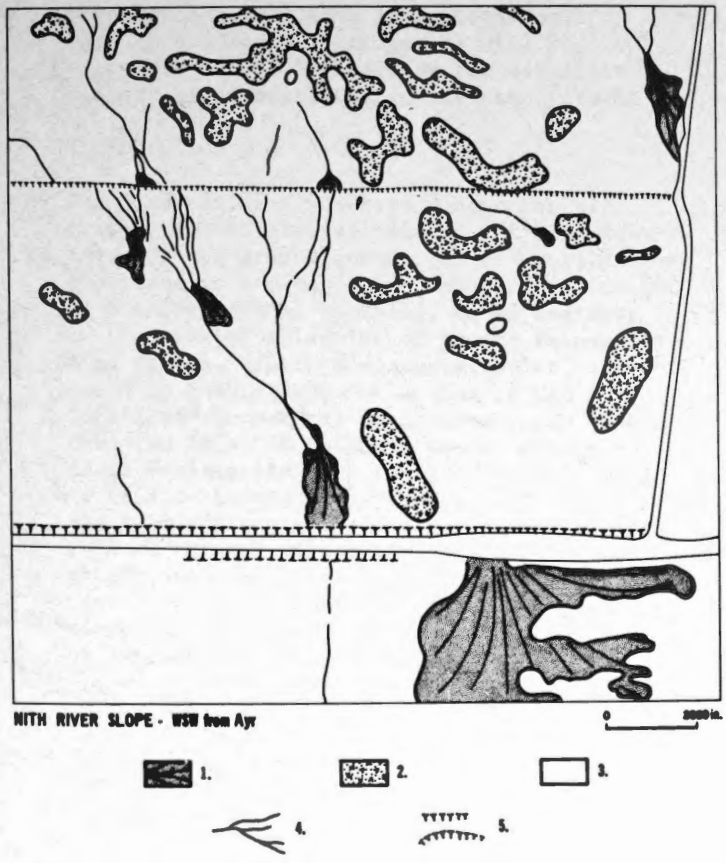


Figure 11 Geomorphological sketch. Example of soil erosion
Legend: 1. Recent accumulation cover (fans) 2. Culminations of small hills with washed out (degradated) A Horison of soil profile. 3. Slope of the valley (gradient 15°) 4. Rills 5. Man-made scarps and road cuts.

TERRAIN ANALYSIS FROM SMALL-SCALE AERIAL PHOTOGRAPHS

Philip Gimbarzevsky,
Research Scientist,
Forest Management Institute,
Canadian Forestry Service,
Ottawa, Ontario, K1A 0H3.

ABSTRACT

A stereoscopic analysis of vertical aerial photographs provides a firm base for evaluation of main environmental factors such as regional and local topography, internal and external drainage, vegetative cover, physical land characteristics and their interrelationships. Such an analysis may be carried out at a desired level of intensity.

Examples of terrain analysis are described and illustrated with 1:80,000, 1:100,000 and 1:160,000 scale aerial photographs. The scales used in high altitude photography are much smaller than those employed in conventional aerial photography but larger than those expected from satellite imagery. Thus experience gained from interpretation of small-scale aerial photographs will be extremely useful in handling the satellite imagery to be available in the near future.

INTRODUCTION

The land surface is a common base for all earth-oriented disciplines and intimate knowledge of the ground conditions is equally important to a manager of renewable resources in his operational planning, to an engineer in his work on selection of proper transportation routes, and to a conservationist, interested in the rational balancing of man's impact on the natural environment. In each case the final decision is based on an adequate familiarity with the land's surface or terrain features: topography, kind of surface material, moisture conditions, structure and type of natural vegetation, presence of open waters, man-made features, etc. During the past three decades aerial photography has played an important role in the acquisition of the required information and some types or scales of photography became "conventional" in the stereoscopic studies or mapping. The large and medium scales (1:10,000 - 1:31,680), for example, are preferred in photo interpretation, smaller scales (1:40,000 - 1:63,360) in topographic mapping.

Recent advances in remote sensing technology, and the introduction of new optics, films and processing techniques have produced a considerably improved image quality and generated a special interest for the small-scale photography among photo interpreters, particularly those involved in a rapid evaluation of extensive land areas.

The purpose of this paper is to describe briefly the basic concept of air-photo analysis of terrain conditions and to demonstrate some advantages of the small-scale aerial photography in this qualitative analysis.

A "terrain" is a part of the earth's surface not covered by the ocean. It may be dominated by land or fresh water, by vegetative cover or man-made features.

A "small" scale in this paper refers to the photographs obtained from altitudes greater than 30,000 feet above the ground with a 6-inch or shorter focal length camera.

A SYSTEMATIC APPROACH TO AIR-PHOTO ANALYSIS OF LAND CONDITIONS

A "terrain analysis" may be simply defined as a process of stereoscopic examination of aerial photographs or other imagery to extract basic data on physical characteristics of the area under observation and to determine the effect the natural and man-made features may have on the specific study, or operations. Depending upon the purpose of the study the results of terrain analysis may be presented at a desired intensity level — as a *broad evaluation of landscape pattern* showing ecologically significant land-vegetation units, or as an *intensive survey*, where the segments of the landscape pattern are individually identified and classified.

Figures 1 and 2 are examples of the 1970 small-scale (1:80,000) photography, where a single 9 x 9-inch print covers about 124 square miles. Using a broad classification, the land area may be simply subdivided into

several land-vegetation units dominated by similar features: a mineral-organic complex composed of sand dunes and organic accumulations in depressions between the dunes, a well-defined river valley, a rough, broken bedrock-controlled topography with very steep slopes, or an outwash plain composed of coarse-textured sand and gravel, with some alluvial deposits along the stream valley and coarse colluvial accumulations bordering the lower slopes of steep hills.

Occasionally such general analysis may be sufficient to provide *reconnaissance-type information*, or to select land units having specific characteristics for a more *intensive study*. In that case the individual sand dunes, portions of the flood plain or particular slope positions may be viewed under a proper magnification and mapped as separate units.

The main purpose of the analysis is to provide fundamental information on the area under observation and regardless of the intensity level, the systematic terrain analysis is basically a stratification of the land surface into relatively homogeneous *terrain units* having common topographic features, moisture conditions and kind of surface material. This is accomplished through a stereoscopic evaluation of main *components* or *elements* of the image pattern formed on the photograph: landform, drainage, erosion, vegetation, land use and photo tone (Belcher, 1948).

IMAGE COMPONENTS

Landform

Landform is a topographic arrangement of surface features and is indicated by relief, relative size, shape and orientation of hills, ridges, plains and depressions, their degree of accordancy and boundary characteristics. Small-scale photography is an excellent tool for interpretation of regional relief, as shown on Figures 1 and 2. A proportion of hilly, flat and depressional topography, for example, may be estimated quite easily. Also, the relative length, width and height, as well as the shape of upland on lowland areas, their topographic forms, type of slopes, surface configuration and their boundary characteristics are quite clear and well defined.

Drainage

Drainage is any natural or artificial channel that carries runoff. It may be continuous or intermittent, external or internal. External drainage which is controlled mainly by the permeability of surface material and ground slope produces typical drainage patterns directly related to physical properties of the underlying material. A dendritic pattern, for example, usually indicates a uniform surface material, a rectangular pattern with all its variations indicates a bedrock-controlled landscape, while a deranged or disordered pattern with its irregular channels, numerous lakes, ponds and wetlands is usually associated with glacial till landforms.

Erosion

Erosion on the land surface may be caused by the wind, running water, gravitation or a combination of these forces. The shape of erosional features, density of occurrence, and steepness of gradient usually indicates the cohesive properties and the texture of surface material. Many gullies develop for example, on the surface of fine-grained silts and clays, in contrast to coarse-textured sands and gravels, having very few gullies, because of their permeability.

Vegetation

Natural vegetation within a climatic region is very closely related to physical land characteristics and recognition of vegetation types is frequently an excellent indicator of terrain conditions: kind of surficial material, moisture content, depth to underlying bedrock, etc. In the Boreal Region, for example, pure pine stands are usually associated with well-drained, medium- to coarse-textured materials, while black spruce may be used as an indicator of poorly drained, wet conditions. In addition to ecological characteristics of vegetation types and regional climate, the effect of forest fires, logging, insect damage and other factors should be considered in relating vegetative cover as an indicator of specific land conditions.

Land Use

There is frequently a direct correlation between physical land properties and dominant land utilization within a particular region. Irrigation or drainage ditches for example, will indicate moisture conditions. Restriction of cultivated fields to special soil types or topographic positions, shape and size of cultivated fields, gravel and borrow pits, road pattern, rock piles, rock fences or any other detail on the use of the land

may provide a valuable clue for a successful evaluation of local conditions.

Photo Tone

The earth surface features appear on aerial photographs as a pattern composed of various shades of grey (black and white photography) or as hues of principal colors (color photography). Because the tonal differences are affected by many factors, such as vegetative cover, time of photography, type of film and filter, processing, and many others, the photo tone should be evaluated with other elements of terrain analysis. Tonal differences between the vegetation types in Figure 1, for example, also reflect different land features: the dark grey tone of stabilized sand dunes indicates a vegetation pattern and local relief. In comparison with the light grey tones of organic depressions between the dunes the boundary characteristics are very clear. On infrared or color photographs the land-surface features will have tonal patterns related to the particular spectral region. Generally speaking the light tones may indicate sand, recently exposed mineral soil, bare bedrock, grass, shrubs, deciduous forest types or high concentration of calcium carbonates. Dark tones may indicate a higher moisture content, dark colored rocks, clays, high organic content or coniferous tree species.

Terrain Factors

A systematic approach to terrain analysis is based on a stereoscopic study of a stereo model and subdivision of the land surface into relatively homogeneous units composed of similar terrain factors: *local topography*, *moisture conditions* and *kind of surface material*. The vegetative cover, open waters and man-made features may be included in the original analysis or added at a later date, if required.

Local Topography

A stereo model is an exact replica of relief features of the area under observation and local topography of a land unit may be expressed in a descriptive form or by a set of conventional symbols (Miles, 1962). A graphical presentation (Table 1) of topographical features resembling sections of the land-surface profile, has been found to be a convenient and simple way to show local topography in terrain analyses (Fig. 3 and Fig. 4).

TABLE 1. Local topography

Symbol	Topography
	Knob or high hill.
	Hill.
	Knoll.
	Plateau or elevated flat.
	Flat.
	Depression.
	Trough.
	Depressed flat.
	Scarp.
	Rolling.
	Rough.
	Slopes 5%, 5-20%, 20% plus.

Moisture Conditions

The moisture conditions may be expressed as moisture classes, based on a combination of internal drainage or permeability and external drainage or runoff. Within a climatic region the permeability or downward drainage is controlled by the texture of surface material and depth of the underlying bedrock. The external drainage is a function of topography, texture and vegetative cover (Miller, 1968). Considering a scale of six moisture classes with moisture class 1 (dry, excessively drained) and class 6 (very poorly drained) as two extremes, the remaining four classes (2-5) provide a transition between extremely dry and extremely wet conditions:

Class 1. Dry, rapidly drained land units, having an excessive internal or external drainage, or both, as a result of coarse texture, slope, or both. Usually in this class eroding steep slopes, loose sand and gravel, stratified medium sands, sharp morainic ridges, exposed bedrock or crests of bedrock-controlled hills.

2. This class is a standard for a

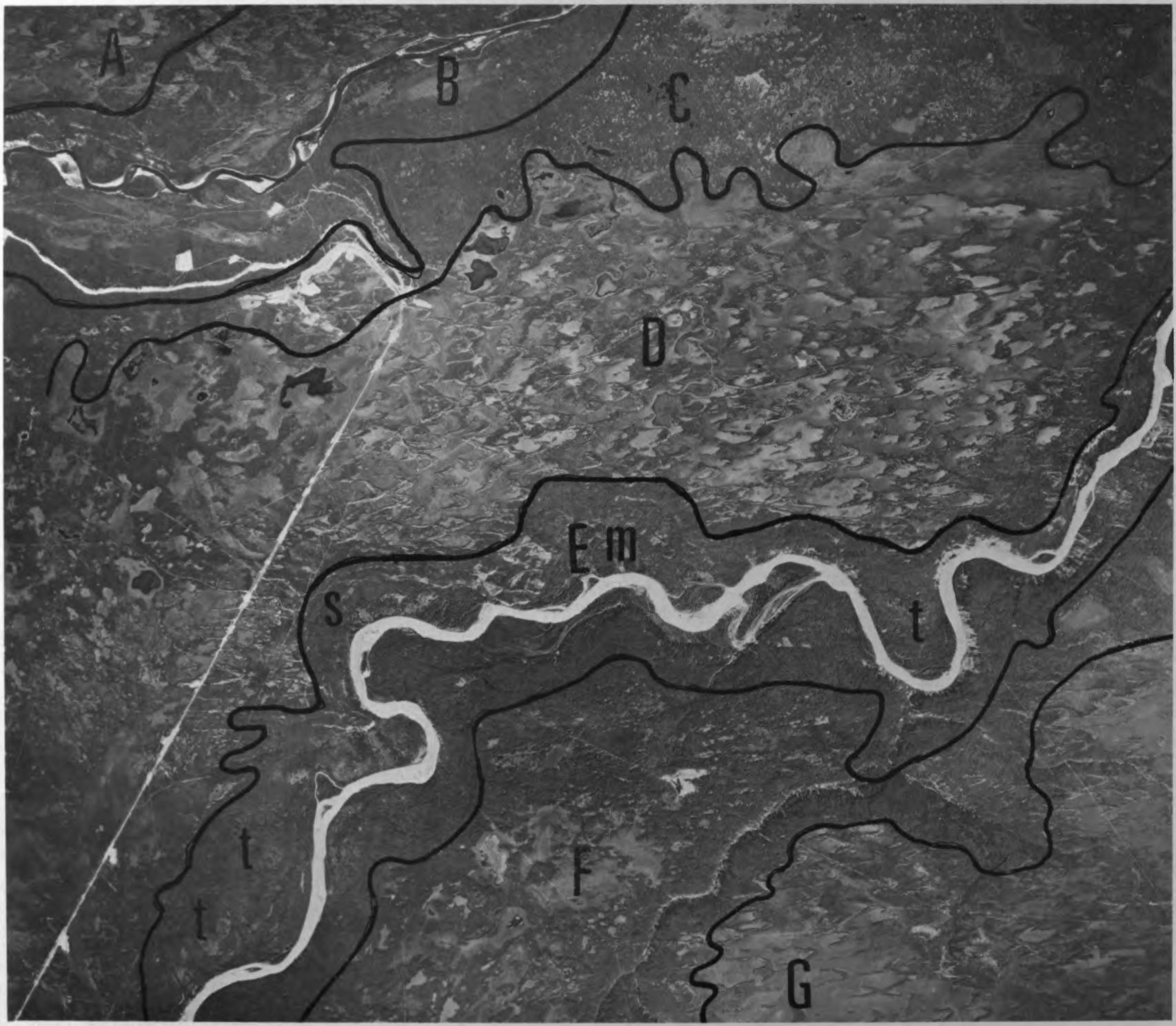

A
B
C
D
Em
s
t
t
t
F
G

Figure 1. Delineation of landscape patterns. Scale 1:80,000 (NAPL).

Unit A is a mineral-organic complex composed of partly stabilized sand dunes 15-20 feet high, surrounded by wet organic depressions. Distribution of natural vegetation is closely related to landforms, drainage conditions and the texture of the surface material. Sand dunes are supporting open-growing and medium-stocked stands of aspen, pine and spruce, while depressions between the dunes are poorly drained organic accumulations. They may support a growth of scattered black spruce or locally occur as "non-forested", with some tamarack, dwarf birch and sedges.

Unit B The Wapity River Valley, with very steep, eroded slopes 250 to 300 feet high and a narrow flood plain. The south-facing slopes support an open growth of a pine-aspen stand with numerous bare spots indicating recent wind erosion or blowouts. The north-facing slopes are well forested by a spruce-aspen-pine stand about 50 feet high. The narrow flood plain supports locally, fairly good white spruce stands on recent alluvial deposits.

Unit C Sandy-silty plain with numerous small depressions. Medium-stocked 50-60 foot high aspen-pine-spruce stand.

Unit D Mineral-organic complex, similar to Unit A.

Unit E Smoky River Valley about 300 feet deep with very steep eroded slopes. Some recent slumping at point S indicates unstable slopes. The flood plain shows several meander scars (m), abandoned channels and alluvial terraces (T). The lower portion of the flood plain is a very narrow valley confined to steep eroded banks. Vegetation is similar to Unit B.

Unit F Sandy plain similar to Unit C.

Unit G Organic-mineral complex - similar to A and D but with a higher proportion of poorly drained organic accumulations.

given climatic region; soil is usually slightly moist during normal summers, except where exposed after removal of vegetation. This class may include moderate slopes composed of sandy material, compacted silty and loamy sands, stratified and compact fine sands.

3. Moderately drained. Soils in this class are wet in spring with internal high water table up to 30 cm from surface, or water perched in upper layers in spring if lower layer impermeable. Occurs usually on gentle slopes and flats composed of silty material.

4. Imperfectly drained. Occurs in depressions or flats composed of massive clays and silts. Organic cap may be present up to 25 cm deep.

5. Wet, poorly drained. Occurs in small depressions and on fringes of large swamps. Mineral soil is wet all year, may flood in spring. Organic accumulation from 30 to 100 cm. In northern latitudes the ice pan may support perched water.

6. Very wet, very poorly drained, occurs in broad organic areas. Deep peat accumulation (100 cm +) and permanent saturation most of the year.

Surface Material

The kind of surface material comprising the landscape unit is directly related to the geomorphological process of deposition and erosion performed by the glacier ice, water and wind, or a combination of these forces. As the landscapes resulting from the action of the same forces are generally composed of similar materials, the air-photo analysis of the surface pattern of an area under observation is based on a recognition of landforms and grouping them by their origin into major categories: glacial, alluvial, aeolian, organic, gravity and bedrock (Gimbarzevsky, 1964, 1965).

The kind of surface material may be expressed in general terms by origin (glacial, alluvial, etc.) or as specific classes of texture. For most practical purposes the mineral surface material may be placed in three main classes: coarse, medium and fine, in addition to a separate class for bedrock and organic deposits (Table 2).

TABLE 2. Surface material

Texture Class	Material	Origin
Coarse (c)	Sand, gravel	Glacio-fluvial, lacustrine, aeolian, alluvial
Medium (m)	Sandy loam, silt loam	Glacial till, some alluvial and lacustrine
Fine (f)	Sandy clay, silty clay, clay	Lacustrine, alluvial, glacial till
Organic (o)	Peat, muck	Organic
Bedrock (r)	Bedrock	Sedimentary, igneous, metamorphic

Final results of the analysis provide a stratification of the study or project area into smaller terrain units separated on the basis of their physical properties. They may be presented on a topographic map, on a mosaic, or left on the original working set of photographic prints used for the interpretation. The physical properties may be briefly described (*see* description under Fig. 1 and 2) or expressed by conventional symbols, as shown on the following examples.

Example 1.

Terrain symbol: ↘ 1cW

Interpretation:

The arrow (↘) indicates a steep, southeast facing slope.

The numeral (1) indicates moisture conditions — dry, rapidly drained land as a result of a relatively steep slope and coarse texture of surface material (cW) — (c) coarse, (W) wind deposit.

Example 2.

Terrain symbol: ~ 2-4 $\frac{mT}{B}$

Interpretation:

A rolling till plain composed of sandy textured till material (mT) — (m) medium texture, (T) glacial till. The unit has a range of moisture condition from well drained (2) on elevated till to imperfectly drained (4) conditions in small depressions. The double symbol for material $\frac{mT}{B}$ has been used to indicate that sandy till is underlain by bedrock (B) less than 100 cm below the surface.

Example 3.

Terrain symbol: ⌊⌋ 5-6 pO

Interpretation:

Wet (5) to very wet (6) very poorly drained depressed flat (⌊⌋) filled with organic peat (pO) accumulations.

CONCLUSIONS

A completed terrain analysis map or mosaic provides in a comprehensive form the essential background data on environmental features of the area under observation. On some engineering projects these maps or mosaics may be used directly in feasibility studies, for example, in the selection of proper sites, location of granular materials, or pinpointing problem areas for special investigations.

The results of the initial terrain analysis may be easily modified and adapted for use in special studies, such as, land classification, hydrology, or bio-physical land classification.

The land-capability mapping is based on photo-interpretation of physical land characteristics and a capacity of the land within a given climatic region (1) to produce a specified amount of common field crops, forest products, (2) to support a wildlife habitat or (3) to serve for recreation purposes. The photo interpretation procedure is very similar to the one described for terrain analysis. A glacio-fluvial outwash, for example, will be outlined by the land expert as a separate land unit, but instead of calling it by name ("flat, excessively drained, coarse sand and gravel") he will assign to it a capability class symbol and "limiting factor", which in this case will be, moisture deficiency, because of the coarse texture. The land capability classes are actually a result of a deductive terrain analysis.

Delineation of drainage basins, evaluation of infiltration components, estimation of runoff, sediment yield, selection of gauging station sites, classification of open waters, and other land-surface factors related to hydrologic studies may also be interpreted from the terrain-analysis data.

The bio-physical land classification system is based on the survey of main environmental components — land, vegetation and water, and the evaluation of their interrelationship within a climatic region. The hierarchical structure of this classification (four levels: land type, land system, land district and land region) permits an intensive use of aerial photographs for mapping the recurring pattern of landforms, soil and vegetation. The photo-interpretation procedure in bio-physical land classification is essentially a terrain analysis carried out at one of the four intensity levels. The knowledge of physical land characteristics may be supplemented with additional information on open waters, vegetation, productive capacity, cultural features, to provide a firm "land base" for an integrated survey of natural resources.

The past developments of photo-interpretation techniques and their continuous improvement were possible only through the joint efforts of the professionals working in various fields. As in the past, the interpreter should always consider aerial photography as one of many remote sensing tools, and he, the user, is best qualified to select the proper scale for his job, to judge the usefulness and limitations of this tool, and to recommend improvements or modifications. The small-scale photography is a new tool and also a new challenge to the interpreters. It is a forerunner of even smaller scale satellite imagery, which no doubt will be more complex to interpret than any small scale available today. Thus experience gained from the interpretation of the small-scale photographs will be essential in handling satellite imagery in the future.

REFERENCES

American Society of Photogrammetry — Manual of Photographic Interpretation, 1960. Washington, D.C.

BELCHER, D.J. 1948. The Engineering Significance of Landforms. Highway Research Board, Bull. No. 13, Washington, D.C.

GIMBARZEVSKY, P. 1964. The Significance of Landforms in the Evaluation of Forest Land. Woodlands Rev., Can. Pulp Pap. Mag.

GIMBARZEVSKY, P. 1965. Land Inventory Interpretation. Photogramm. Eng.

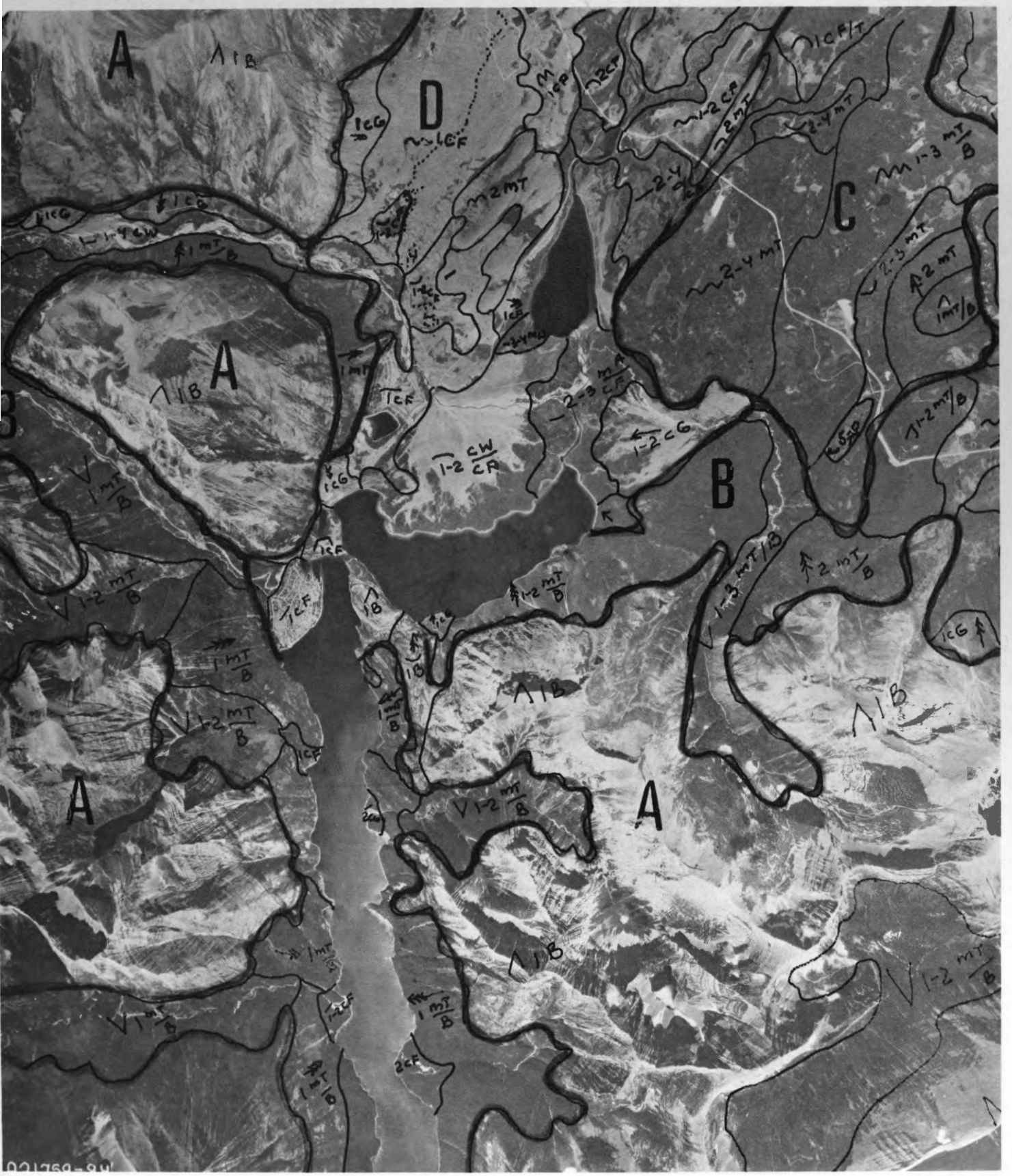
MILLER, V.C. and S.A. SCHUMM. 1968. Aerial Photographs and Surface Features in "Aerial Surveys and Integrated Studies". Proc. Toulouse Conf., UNESCO, Paris.

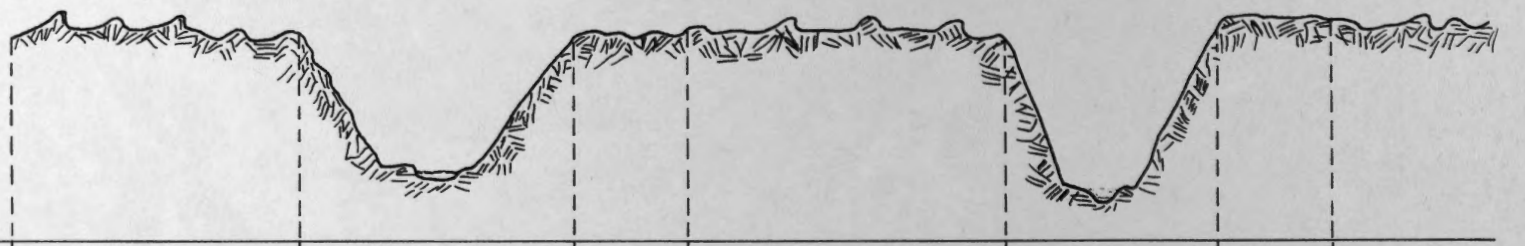
MILES, R.D. 1962. A Concept of Landforms, Parent Materials, and Soils in Airphoto Interpretation Studies for Engineering Purposes. Trans. Symp. Photo Interpretation, Delft.

Figure 2. Two levels of terrain analysis. Scale 1:80,000 (NAPL).

1. Delineation of broad landscape patterns, (A) very rugged bedrock hills, with sharp crests and steep slopes; (B) shallow till over bedrock occupying lower positions of steep slopes. Includes some deep eroded valleys and coarse colluvial accumulations on the valley bottom; (C) till plain, partly drumlinized, with local relief from 5 to 25 feet; (D) glacio-fluvial plain - outwash, deltas, eskers and colluvial fans, mainly coarse, stratified well-drained deposits with some recent alluvial silt and fine sand along the modern stream.

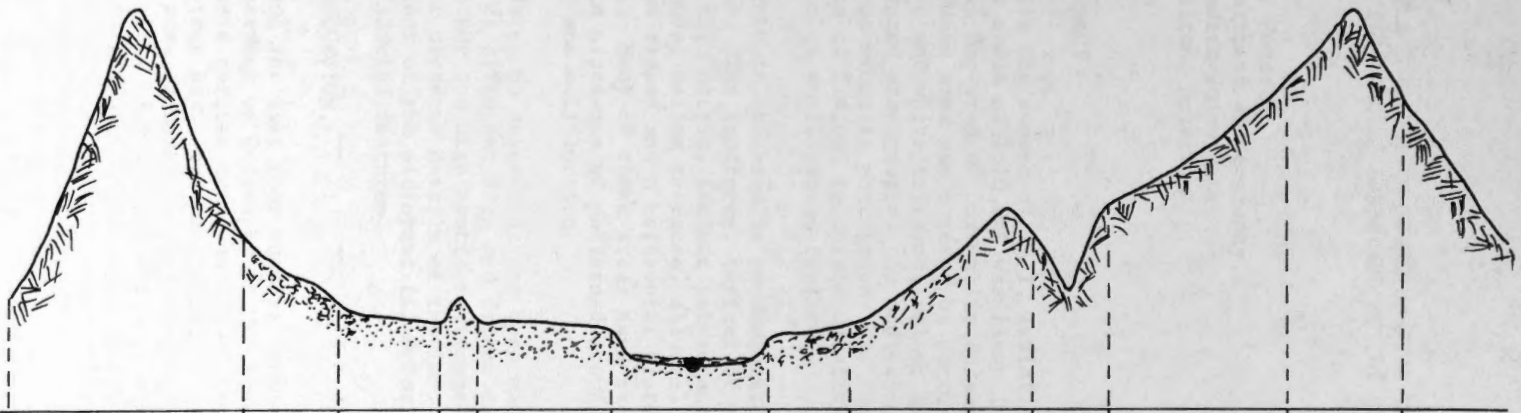
2. Detailed terrain analysis - delineation of terrain units on the basis of topographic features, moisture conditions and surface material.


A
1B
D
1CG
1CF
2CF
M 1-2 CR
2-4 MT
1CF/T
M 1-3 MT/B
C
2-4 MT
2-3 MT
2 MT
1 MT/B
M2MT
1-2 CW/CR
2-3 MT/CF
1-2 CG
B
1-2 MT/B
1 MT/B
2 MT/B
V 1-3 MT/B
ICG
V 1-2 MT/B
1 MT/B
1CF
1B
A
1B
V 1-2 MT/B
A
1B
V 1-2 MT/B
1 MT/B
2CF



Land Unit	A Mineral-organic complex	B River valley	C Plain	D Mineral-organic complex	E River valley	F Plain	G Organic-mineral complex
Topography	Sand dune ridges and organic depressions	Level flood plain & very steep slopes	Level	Sand dune ridges and organic depressions	Level flood plain and very steep slopes	Level	Depressions and sand dune ridges
Moisture	Dry (1) on dunes wet to very wet (5-6) in depressions	3-5 on flood plain 1-2 on slopes	2-4	1-6	3-5 flood plain 1-2 slopes	2-4	1-6
Surface material	Wind-blown sand and organic accumulations (peat)	Alluvial silty sand, sand and sandy till on eroded slopes	Sandy or silty loam (till)	Wind-blown sand and organic peat	Alluvial sand and sandy loam	Sandy loam, sand	Peat and sand

Figure 3. General terrain analysis - diagram of landscape units outlined in Figure No. 1.



Topography	High Hill	Slope	Flat	Hill	Flat	Depressed Flat	Flat	Steep Slope	Hill	Trough	Steep Slope	High Hill	Very Steep Slope
Moisture	1	1-2	1	1	1	2-3	1	1-2	1	1-2	2	1	1-2
Surface material	Bedrock (sandstone)	Rock fragments, till (gravity fan)	Coarse sand and gravel (outwash plain)	Sand and gravel (esker)	Sand and gravel (outwash plain)	Silty sand over sand and gravel (flood plain)	Sand and gravel (outwash plain)	Sandy loam (shallow till over bedrock)	Exposed bedrock	Sandy loam (till) over bedrock	Sandy loam (till) over bedrock	Exposed bedrock	Sandy loam (till) over bedrock

Figure 4. Detailed terrain analysis - diagram of terrain units outlined in Figure No. 2

FILM AND FILTER COMBINATIONS FOR THE STUDY OF THE PERIGLACIAL LANDSCAPE OF THE HIGH ARCTIC

P.J. Howarth,
Department of Geography,
McMaster University,
Hamilton, Ontario.

ABSTRACT

During the summer of 1971, aerial photography at a scale of 1:10,000 was flown in the Resolute Bay area of Cornwallis Island, N.W.T. The same area was covered by panchromatic, black and white infrared, colour and colour infrared photography. In addition, one strip of panchromatic photography was flown at a scale of 1:4,000 to obtain detailed information on small ground features.

Bedrock in the area is predominantly limestone. The landforms, typical of much of the High Arctic, include raised and modern beaches, marine terraces, alluvial fans, talus slopes and a horizontal plateau surface. Many of these areas have been subjected to processes of patterned ground development and solifluction.

To date, no assessment has been made of the use of different film and filter combinations to study the High Arctic environment. This paper presents details of the information content of the different films for studying periglacial features.

INTRODUCTION

During the last four summers, members of the Department of Geography at McMaster University have carried out a series of studies at Radstock Bay, south-west Devon Island and in the area of Resolute Bay, Cornwallis Island (Fig. 1). The program has been concerned with investigations of the physical landscape and research work has been concentrated on studies of Arctic beaches, talus slope characteristics, drainage basin studies and soil characteristics. The use of aerial photographs has provided an important source of information both for mapping and for interpretation in many of these studies. At the same time, the studies have provided a wealth of ground truth information for photo-interpretation of the periglacial landscape of this area.

To date, the photography used has been panchromatic coverage obtained from the National Air Photo Library. In order to determine if other types of film and filter combinations aid the interpretation of landscape detail and provide additional information on ground characteristics vertical aerial photography using four film and filter combinations was flown in the Resolute Bay area of Cornwallis Island during the summer of 1971. This paper presents a first assessment of this photography and indicates the type of information that may be obtained concerning landscape characteristics within the periglacial environment of the High Arctic.

Later work will involve more detailed interpretation of selected frames of photography from the four types of films, in particular the possible differentiation of soil types and terrain units that have been mapped in the field by Cruickshank (1971). Studies of image densities and spectroradiometer measurements over various types of terrain will also be undertaken.

REMOTE SENSING IN ARCTIC AREAS

With such vast areas to be studied in the Canadian Arctic, it is apparent that aerial photography and other remotely sensed data have an important role to play in the initial stages of geological and geomorphological investigations, resource inventory or site and route selection. The majority of photo-interpretation studies in the higher latitudes of North America have been concerned with the more vegetated areas of northwestern Canada, the sub-Arctic and also Alaska. For example, Sager (1951) and Frost (1960) have discussed the appearance on panchromatic aerial photographs of landforms unique to permafrost areas. Frost (1960) and Pressman (1963) have been concerned with identification of surface materials and site and route selection for various engineering projects. To date, studies involving thermal infrared imagery and radar imagery have been concentrated on investigations of sea ice and glacier ice (e.g. Poulin and Harwood, 1966; Leighty, 1966), though in the latter article reference is also made to terrain characteristics. Apart from the studies of Dunbar and

and Greenaway (1956) and Pressman (1963), it is difficult to obtain information on the interpretation from aerial photographs of the characteristics of the landscape of the High Arctic, an area that is becoming of increasing interest with the growth of oil exploration.

PHOTOGRAPHY

The study being undertaken at the present time has involved the flying of panchromatic, black and white infrared, colour and colour infrared photography at a scale of 1:10,000 over an area near Resolute Bay. Details of the films, filters, exposures and times of flying are given in Table 1. The photography was flown by the ice reconnaissance Beechcraft of the Polar Continental Shelf Project carrying a Wild RC-10 camera (with a 6 inches focal length lens) belonging to the Glaciology Subdivision of the Inland Waters Branch.

For each type of film in turn, photography covering the same area was flown from an altitude of 5,000 feet along flight lines running approximately north-south. In each case six lines were flown, except for the colour infrared photography for which film was available for only five flight lines. In addition, four extra lines were flown using the panchromatic photography to provide complete coverage of the drainage area of the Mecham River, which is being studied by members of the Department of Geography at McMaster University as part of the program for the International Hydrological Decade. One line of panchromatic photography was also flown from an altitude of 2000 feet (1:4,000 scale) at an angle to the main flight lines to be able to study small scale features in detail and to record on the photography a series of panels of different densities which are to be used in further studies related to this research program.

Film processing was arranged through the Remote Sensing Division of the National Air Photo Library and studies of the photography are being made on positive film transparencies.

AREA OF STUDY

Resolute Bay (74°40'N and 95°00'W) is situated on the southern coast of Cornwallis Island (Fig. 1). Bedrock in the area is predominantly Ordovician and Silurian in age, and within the area covered by the photography rocks of the Allen Bay Formation and the Read Bay Formation predominate (Thorsteinsson, 1958). They consist of argillaceous, crystalline, crinoidal and dolomitic limestones with occasional thin layers of shale. These rocks have been gently folded into a series of anticlines and synclines with their axes running approximately north-south. These rock types and formation are typical of much of the Queen Elizabeth Islands.

The major physiographic feature is a nearly horizontal plateau surface reaching elevations of about 200 m. in the Resolute Bay area. Similar plateau surfaces up to elevations of 400 m. can be found in other parts of the Queen Elizabeth Islands.

The coastal zone consists of raised beach and lagoonal deposits reaching elevations of over 100 m. (Bird, 1967, p. 130). These deposits have been produced and exposed during isostatic uplift, following retreat of the ice after the last glaciation about 9000 years ago.

Separating the plateau and the coastal zone is a section of scree and talus slopes which in many areas have been developed beneath the edge of the plateau. Also included within this category are the valley side slopes produced by the entrenching of rivers into the plateau surface.

Permafrost underlies the whole area and the active layer at the end of the summer period averages about 0.5 m. in depth. The summer period begins in mid-June with most of the snow melting by the beginning of July. Freeze-up starts at the beginning of September. Within the whole area there are only isolated sections of vegetation. The presence of the permafrost and its prevention of subsurface drainage lead to the formation of periglacial landscape features that are not encountered in lower latitudes.

INTERPRETATION

Within this section, the appearance on the different types of films of various rock types, and also landscape features typical of periglacial areas are discussed. A number of papers have already been written comparing the information content of different types of aerial films for studying the physical landscape. For example, Welch (1966) has studied a glacial area in Iceland, Fischer (1958, 1962) has investigated rock types and Anson (1968, 1970) has considered the interpretation of geological formations and soil. With this general background of information, the emphasis in this study is on the special characteristics of the periglacial environment.

Rock Types

Although bedrock is infrequently exposed, mechanical breakdown of the rock and the small degree of soil development often permits the identification of the types of bedrock at shallow depths beneath the surface. Even though the rocks are almost predominantly limestones, they do show frequent lithologic variations which can sometimes be differentiated on the films. The majority are various shades of grey in colour, but yellow, brown and orange-pink colouration does occasionally occur in some strata (Thorsteinsson, 1958).

On the panchromatic and black and white infrared films it is impossible to distinguish any tonal variation that may be related to distinct colouration of the rock type, because the wide range of grey rocks image in numerous grey tones on these films. On the colour photography, however, it is possible to provide a much clearer distinction and to trace certain rock types for some distance. This agrees with Thorsteinsson's (1958, p.29) field observations that "it is possible to walk for several miles between streams without observing an exposure, although ubiquitous frost-shattered debris (felsenmeer) of bedrock generally permits the tracing of formational contacts."

Of particular interest in this respect is a rock type producing a distinctive pale red colouration on the colour film. Even on the colour infrared film it appears as a green-blue colour, but placing a boundary between this and adjacent rock types is extremely difficult on the colour infrared film. This is because there is little visual distinction between the green-blue and the predominantly blue images of the remainder of the rock types. The colour infrared film is in fact of less benefit than the panchromatic or black and white infrared film for differentiating rock types, due to the narrow exposure latitude of the film. On some frames of photography, certain areas are over-exposed, while other areas are under-exposed, thereby adding to the problems of interpretation.

In the area covered by the photography to the north of Resolute, where relative relief is comparatively subdued, many of the rocks image in dark tones on the panchromatic film. On a number of occasions, just using the panchromatic film, the dark tones have been mistaken for areas of vegetation. Although parts of this area consist of the pale red rocks, referred to above, others are dark-grey on the colour film.

Where rock structure is clearly defined, the patterns exhibited can be readily identified on all types of film, but where contacts are not clearly displayed and continuous, colour photography is of definite benefit for geological interpretation.

Beaches

Lowland areas have been modified by beach development. Raised beach gravel and deposits of finer particles occupy many areas to altitudes of over 100 metres. The raised beaches are readily identified on all four films as a series of strips extending across the landscape, with each strip maintaining a more or less constant altitude. In some cases the strips represent low ridges and intervening hollows; in other cases they are produced by different particle sizes of material producing slightly different image characteristics. Thus the pattern produced is one of alternating lighter and darker tones or colours.

Along the shoreline, it is possible to distinguish the present active beach zone. At the time the photographs were taken, sea ice still occupied Resolute Bay and the active beach zone was buried beneath beach-fast ice. Experience with panchromatic photographs, however, has indicated that the active beach zone images in darker tones than the raised beaches.

Identification of cracks in the near-shore sea ice is also easily achieved on all films and can aid the interpretation of near-shore bottom topography. At the time the photography was taken, "puddling" of the ice surface due to melting was well developed. Only on the colour and the colour infrared films was it possible to differentiate areas of surface "puddling" from areas where the ice had actually cracked apart to form holes through the ice. In the latter case, images were recorded as darker blues.

In some beach locations, an indication of major directions of ice push can be identified by small ridges formed as the ice is forced on shore. Such features can be readily identified on all four types of film.

Solifluction

Solifluction of surface deposits occurs in many areas in this part of the High Arctic. It is a landform modifying process which can effect various types of deposits, not only on the plateau surface, but also in areas of raised beaches. In some cases, solifluction produces

a sheet of saturated surface deposits, sometimes it forms individual or inter-connected lobes, and in other instances there is just a general downslope movement of surface materials producing stripes. Frequently there is a sorting of the material across the slope into coarse and fine particles which aids concentration of water flow down the slope.

Areas of solifluction can be identified on all four types of film by the pattern of alternating wet and dry deposits produced by the stream and rill flow. The patterns, in this case, can be more easily identified on the panchromatic film. This, however, is probably due to the fact that parts of the black and white infrared film copy have been slightly over-exposed and areas of water do not show any really dense images.

On both the colour and the colour infrared films, the darker patterns of the solifluction stand out clearly against the surrounding landscape. On the colour film it is difficult to obtain an assessment of the relative amounts of saturation of the surface materials in different areas. The colour infrared film, however, shows a greater range of density variations associated with the presence of varying amounts of moisture. Frequently, vegetation is associated with areas of solifluction. It helps to emphasize the occurrence of solifluction on the colour infrared film and appears, to some extent, to be correlated with the amount of saturation of the surface deposits.

Patterned Ground

Within this classification, features of a variety of different sizes may be considered. First, are small frost-heaved hummocks averaging about 1 m. in diameter. On the aerial photographs they produce a textured image. They occur predominantly on near level surfaces, particularly on the tops of small mounds. As the slope steepens, they give way to a striped surface. The important factor in their identification is correct exposure of the film. Particularly if the density of the film is too light, they cannot be detected. They appear to be most easily seen on the colour infrared film, but as the exposure of this type of film is most critical, they are frequently not identifiable in many areas where they occur.

The second feature associated with patterned ground is the polygon. Many of these, particularly in areas of raised beaches, are approximately the same size as the frost-heaved mounds and difficult to differentiate from them. While these polygons are usually not sorted, larger polygons, averaging about 3 m. in diameter normally have coarse material forming the borders with finer particles in the centre. Such polygons, although limited in occurrence at Resolute Bay, can be identified on all four types of photography. The presence of vegetation helps to emphasize their pattern.

The third and largest category of patterned ground is the ice wedge polygon. The majority are 10 m. to 20 m. in diameter, but the larger ones are in the range of 75 m. to 100 m. in diameter. They develop in many areas, but can be most clearly identified on deposits of finer-grained sediments. In such areas, the pattern of the ice wedges is imaged in darker tones and colours on the photography, due to increased moisture content. They stand out clearly against the generally light colouration of the deposits in which they are developed, particularly on the colour and colour infrared films. In some areas of raised beaches, ice wedge polygons have developed a more regular, rectangular pattern. No one film is of particular advantage for identifying these polygons as it is just their pattern which is recorded on the film.

Other Features

Brief mention should also be made of the characteristics of the vegetation and water bodies in the area. In the limited areas where low vegetation occurs (primarily grasses and mosses) it is usually discontinuous. This and the light tones of many of the rocks makes it difficult to identify the vegetation on the black and white infrared film. Even on the colour film, the vegetation is difficult to distinguish from the surrounding browns and greys of the surface deposits. Only on the colour infrared film can the vegetation be clearly identified by its magenta colouration.

At the time of photography, the larger lakes were still covered by ice. Many shallow pools also occur, some of which dry up in the later parts of summer. On the colour photography it is possible to see vegetation differences at the bottom of the lakes, but the pools can be more readily identified by their bright blue colouration on the colour infrared film.

DISCUSSION

It is apparent that within the periglacial environment of the High Arctic, the pattern of different landscape features is one of considerable importance in their identification.

For example, rock structures may be identified from the pattern produced by rocks of different lithologies; areas of raised beach are characterised by the patterns of individual ridges or variations in particle size; large scale ice wedge polygons have distinct geometric patterns, and the downslope movement of solifluctedmaterial produces a lined pattern on the photography. The patterns can be identified on all four types of photography. In cases where one film and filter combination is better than another for identification and delineation of a feature, it is because the characteristics of that combination of film and filter produce a greater tonal or colour contrast in the final image. Thus, for example, the occurrence of moisture variations in some areas of patterned ground and in zones of solifluction leads to easier recognition on the colour infrared films.

In situations where the pattern of the landscape features is irregular and uncontrolled, colour contrast becomesmuch more important than pattern. Thus identification of isolated areas of individual rock types or areas of vegetation requires the film and filter combinations to maximise the colour differentiation. In the case of vegetation, for example, colour differentiation on the normal colour film is not sufficient to provide easy identification and it becomes preferable to record the near infrared reflectivity of the vegetation on the colour infrared film. It is suggested that considerations of pattern and colour differentiation should be borne in mind when planning aerial photography for specific purposes in High Arctic regions.

In summary, for aerial photography at scales of approximately 1:10,000 flown during the summer period in areas of the High Arctic similar to Resolute Bay, the following points can be made:

- Panchromatic film is adequate for most purposes, except for geological mapping in areas where structure is poorly defined and for vegetation studies.

- Black and white infrared photography has no particular advantages over panchromatic photography. If the exposure on the film copy had been better, identification of features where the presence of water is important, may have been improved.

- Colour photography is especially useful for geological mapping, sea ice studies, underwater identification and patterned ground studies.

- Colour infrared photography is of benefit for solifluction and vegetation studies, and for identification of some aspects of patterned ground.

ACKNOWLEDGEMENTS

Financial support for this project has been received from the National Research Council of Canada, Grant No.A5586 and the Defence Research Board of Canada, Grant No.3480-01. Assistance from Dr. E.F. Roots of the Polar Continental Shelf Project, Mr. K.C. Arnold and Mr. A.D. Terroux of the Glaciology Subdivision, Inland Waters Branch, the Remote Sensing Division of the National Air Photo Library and the Interdepartmental Committee on Aerial Surveys is also gratefully acknowledged.

REFERENCES

(1) A. Anson. Developments in Aerial Color Photography for Terrain Analysis. Photogrammetric Engineering, Vol. 34, No.10, 1968, pp. 1048-1057.

(2) A. Anson. Color Aerial Photographs in the Reconnaissance of Soils and Rocks. Photogrammetric Engineering, Vol.36, No.4, 1970, pp. 343-354.

(3) J.B. Bird. The Physiography of Arctic Canada. The Johns Hopkins Press, 1967.

(4) J.G. Cruickshank. Soils and Terrain Units around Resolute, Cornwallis Island. Arctic, Vol. 24, No.3, 1971, pp. 195-209.

(5) M. Dunbar and K. Greenaway. Arctic Canada from the Air. Canada Defence Research Board, 1956.

(6) W.A. Fischer. Color Aerial Photography in Photogeologic Interpretation. Photogrammetric Engineering, Vol. 24, No.4, 1958, pp. 545-549.

(7) W.A. Fischer. Color Aerial Photography in Geologic Investigations. Photogrammetric Engineering, Vol. 28, No.1, 1962, pp. 133-139.

(8) R.E. Frost. Aerial Photography in Arctic and Subarctic Engineering. Journal of the Air Transport Division,

Proceedings of the American Society of Civil Engineers, Vol. 86, No.AT1, 1960, pp. 27-56.

(9) R.D. Leighty. Terrain information from High Altitude Side-Looking Radar Imagery of an Arctic Area. Proceedings Fourth Symposium on Remote Sensing of the Environment, 1966, pp. 575-597.

(10) A.O. Poulin and T.A. Harwood. Infrared imagery in the Arctic under Daylight Conditions. Proceedings Fourth Symposium on Remote Sensing of the Environment, 1966, pp. 231-241.

(11) A.E. Pressman. Comparison of Aerial Photographic Terrain Analysis with Investigation in Arctic Canada. Photogrammetric Engineering, Vol. 29, No.2, 1963, 245-252.

(12) R.C. Sager. Aerial Analysis of Permanently Frozen Ground. Photogrammetric Engineering, Vol.17, No.4, 1951, pp. 551-571.

(13) R. Thorsteinsson. Cornwallis and Little Cornwallis Islands, District of Franklin, Northwest Territories. Geological Survey of Canada, Memoir 294, 1958.

(14) R. Welch. A Comparison of Aerial Films in the Study of the Breidamerkur Glacier Area, Iceland. Photogrammetric Record, Vol. 9, No.28, 1966, pp. 289-306.

TABLE I

Details of Aerial Photography flown at Resolute Bay, July 15, 1971

	FILM	FILTERS	EXPOSURE	TIME (GMT)	N.A.P.L. ROLL AND PHOTO NUMBERS	
1.	Panchromatic (Kodak Aero XX, Type 2405)	500 mm., Pan 2x. A.V. 1·4x	$\frac{1}{500}$ sec at f8	22.37 - 23.36	A22365	1 - 230
2.	Black & White Infrared (Kodak Infrared Aero, Type 2424)	700 mm., A.V. 1·4x	$\frac{1}{300}$ sec at f8	21.56 - 22.28	A22353	1 - 115
3.	Colour (Kodak Aero-colour Neg., Type 2445)	none	$\frac{1}{400}$ sec at f8	21.15 - 21.47	A30343	1 - 111
4.	Colour Infrared (Kodak Ektachrome Infrared, Type 2443)	500 mm., Pan 2x. A.V. 1·4x	$\frac{1}{250}$ sec at f5.6	20.34 - 21.04	A30344	1 - 98

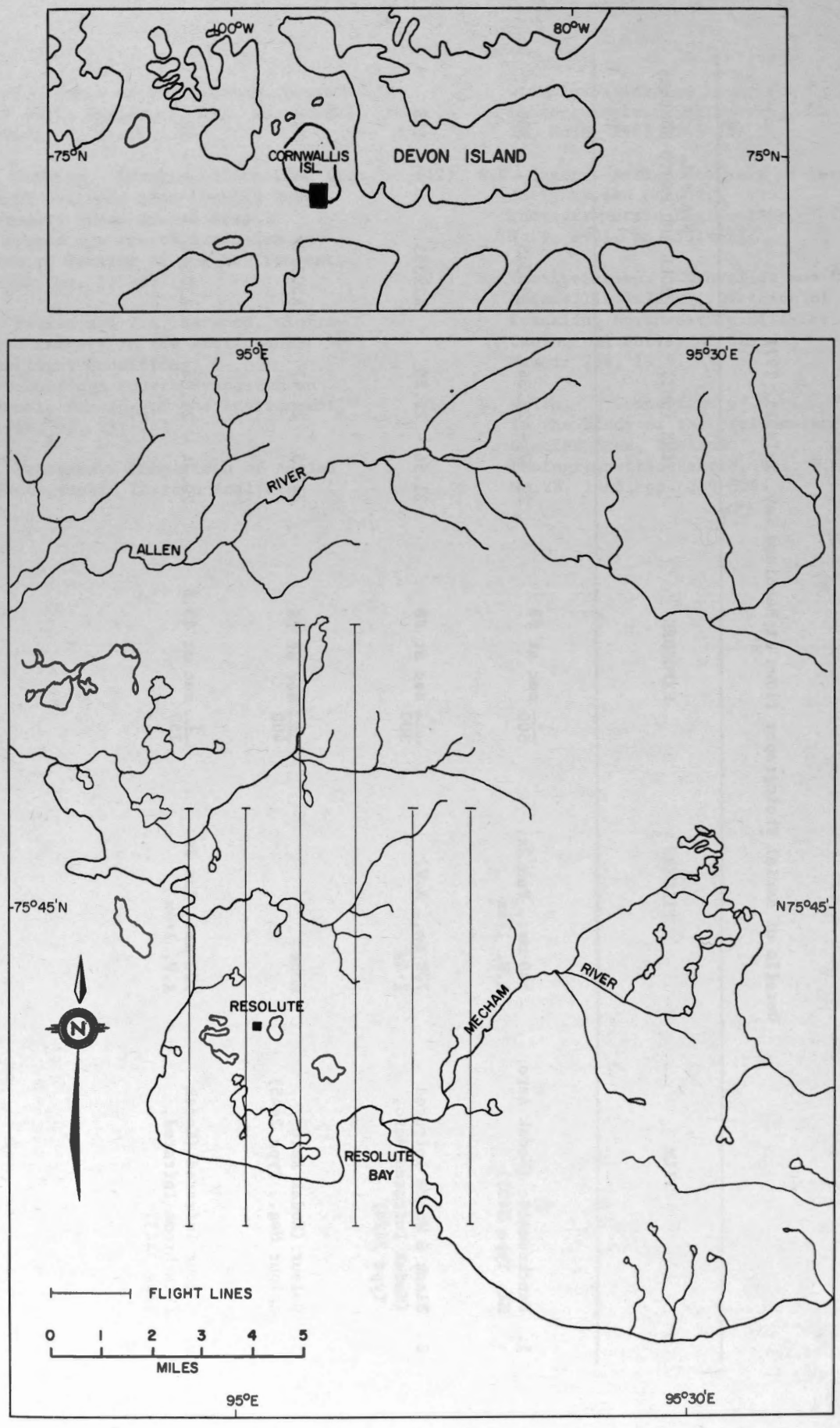


Figure 1. Location of Resolute Bay and planned flight lines for photography

THERMAL INFRARED IMAGERY AT THE ST-JEAN-VIANNEY LANDSLIDE

Marc G. Tanguay
Département de Génie Géologique
Ecole Polytechnique, Montréal

Jean-Yves Chagnon *
Division de Géotechnique
Ministère des Richesses Naturelles, Québec

INTRODUCTION

A catastrophic landslide of the mud-flow type, or a sensitive clay slide, occurred in the vicinity of St-Jean-Vianney and Shipshaw municipalities on May 4th, 1971. The event took place at about 10:30 p.m. and the major part of slide occurred within a few minutes after the initial failure at the bank of Petit-Bras stream. Thirty-one lives were lost and 40 houses were destroyed. It is believed that the sliding was almost over and that the final configuration of the crater was finalized before midnight. However, the area was not fully stabilized until a few days after the slide.

In order to obtain as much information as possible about the soil conditions and the groundwater flow in the area adjacent to the rim of the crater so as to evaluate the possibility of further large scale movements, the Department of Natural Resources asked the Canadian Armed Forces to launch a program of photographic reconnaissance including thermal imagery, black and white, black and white infrared, colour and colour infrared films.

The infrared imagery was obtained by one of the Canadian Armed Forces remote sensing CF-100 aircraft, for the Quebec Department of Natural Resources, using a Reconofax XIII-A scanner. Most of the photographic films were obtained using 70 mm. Vinten cameras. Table I gives further information on the types of imagery, dates, time of day, bandwidth and weather conditions.

REGIONAL SETTING

In order to understand the May 4th slide, it is necessary to present a brief outline of the regional geological features.

Saint-Jean-Vianney is located 1.5 miles north of the Saguenay river and approximately six miles west of Chicoutimi.

An extension of the Champlain sea, sometimes referred to as the Laflamme sea, flooded this area some 9,000 years ago. The sea receded leaving behind a thick sequence of marine clay layers covered by a thin veneer of sand. These deposits were then uplifted to the present level of about 650 feet, and were subjected to weathering and erosional processes.

The stratigraphy of the surficial deposits in the area is: a thin cover of sand and gravel of glacial origin overlying the bedrock. Thick layers of stratified clay, the thickness varying from a few feet up to 300 feet. Thin layers of silt are interstratified within the clay. The strata are nearly horizontal with minor local variations. Layers of sand up to 30 feet thick cover the clay.

The mechanical properties of the clay are generally similar to those of the Champlain sea deposits in the St-Lawrence Lowlands; the susceptibility of this clay to sliding is now well known. Fourteen slides have been recognized in the Chicoutimi area and the largest slide ever to occur in this type of marine sediments is located at St-Jean-Vianney and was described by Lasalle and Chagnon (1968). Geochronological studies (C^{14} method) indicate that it occurred some 500 years ago. It covered an area of approximately 8 square miles and affected an estimated volume of 270 million cubic yards. More recent studies (Tavenas and al., 1971) show that it probably occurred in more than one stage. The material now found in the crater was displaced during these events. The municipality of St-Jean-Vianney is located almost entirely within the limits of the ancient landslide, and the May 4th slide is entirely within these limits.

Therefore the stratigraphy has been disturbed by these various movements and is complex, at least near the surface. The first slide (shown on map of Figure I) deranged the stratigraphy and was followed by the deposition of sand layers along the scarp.

*Paper approved for publication by the Minister, Dept. of Natural Resources of Quebec.

These deposits were disturbed then and covered up by the second slide. Material from minor sliding along Rivière aux Vases to the east also covered these deposits.

It has been generally considered that the remoulded zone of a landslide was stable, mainly because of the destruction of the original flocculated structure of the clay and the removal of the excess water. The St-Jean-Vianney area, within the limits of the old slides, was not considered likely to be affected by further mud-flow slides. Therefore the May 4th catastrophe was very surprising not only because of its extent but simply because it took place completely within the limits of a disturbed zone. This precisely is one of the reasons that necessitated the use of remote sensing monitoring in order to see the extent of the endangered zone.

LOCAL SETTING

A complete geotechnical and hydrogeological investigation of the area has been initiated by the Department of Natural Resources, after the slide. Although the results are not fully interpreted at this stage, a picture of the soil conditions can be outlined on the basis of the available data.

As mentioned above, the stratigraphy is complex over the floors of the previous slides and has not been worked out in detail yet. Further geochronological data is needed to outline the sequence of events. The material within this zone consists of an heterogeneous mixture of sand and clay. At a few different locations, the clay is homogeneous and fairly massive, as if huge blocks had been moved bodily, or the clay is massive but it has been reworked. The sand deposits are located mostly from Harvey street to the west (section A - A, Figure 2). The sand seems to extend from the scarp of the previous slides as if it had been deposited along the base of the perimeter of the ancient crater as talus deposits, much as some sand is now being deposited at the bottom of the May 4th slide, along the rim. Some sand pockets probably result from the burial of the upper horizons by the ancient slides. Under the failure plane of the old slides, the original stratified clay is found to a depth of up to 200 feet. Near the banks of Petit-Bras stream, the strata dip 6° to the east. Under the clay a thin layer of sand and gravel overlies the bedrock. This layer however is not continuous.

Section A - A shows the general aspect of the regional and local stratigraphy as indicated by a few borings.

THE MAY 4TH SLIDE

This event took place during the spring thaw and was preceded by heavy rainfall, amounting to 1.44 inch (Tavenas et al., 1971). A few days before the slide the residential development zone was partly flooded and some sectors were covered by 2 to 3 feet of water.

Around April 20th, 1971 (the actual date is not known), a major landslide occurred along the bank of Petit-Bras stream. It was located at the base of the May 4th slide, as shown on section A-A, and involved an area 200 feet wide and 500 feet long. From the evidence now available, it is apparent that this slide was related to the one that followed it.

The May 4th slide took place at 10:30 hrs. p.m. and extended from Petit-Bras stream well into the new residential development area. A volume of approximately 9 million cubic yards of sand and clay was removed, 1.5 million cubic yards of which is believed to have remained in the crater. Most of this material was remoulded and flowed down Petit-Bras stream, into Rivière Aux Vases over a distance of 1.8 miles into the Saguenay river. The speed of the flow was estimated to be around 16 mph. and the average height of the flow, as measured from the mud cake on the slopes of the river banks, was 60 feet. At the mouth of Rivière Aux Vases, the height of the flow was about 35 feet. Most of the houses that were taken by the slide were destroyed in the rivers and the debris accumulated near the site of the bridge downstream which was removed bodily into the Saguenay river.

Following the cessation of movement, much water was removed from adjacent areas, thus contributing to the stabilization of the crater slopes.

An extensive program of study was launched soon after the event, including a geotechnical study of the areas adjacent to the slide and the Petit-Bras stream and Aux Vases river in order to determine the stability of the slopes; an hydrogeological study of the old slide in and around the zone involved in the ancient slides; an hydrological study of the Petit-Bras river and an hydrometeorological study in the vicinity of Shipshaw and St-Jean-Vianney.

While these various programs were underway, the slopes of the crater regressed somewhat

due to slope trimming by water run-off and seepage and the rupture of unstable parts. More important was the lowering of the bed of Petit-Bras stream. The swiftly flowing Petit-Bras eroded the banks and its bed, lowering its channel by as much as 20 feet. This was an unfavorable development which aggravated the stability conditions of the slopes and had to be corrected as soon as possible. Therefore it was decided to stabilize the hydraulic conditions by channeling the Petit-Bras over a major part of its course. This was accomplished rapidly by erecting aprons on the bottom of the river and retaining walls. These structure were made up of gabions, wire baskets filled with rocks.

The results of the studies are not completely analysed but they allow us to outline the geological, geotechnical and hydrogeological conditions prevailing prior to the May 4th slide. Section A - A' (Figure 2) is a generalized concept of these conditions. Huge sand pockets were held back by a wall of clay extending to the Petit-Bras stream. The pockets are fairly deep, up to 90 feet, and act as water reservoirs where strong hydrostatic pressures are built up. These pressures overcame the clay mass which moved down into the valley.

Some soil parameters, as measured on site, came as a complete surprise. The shear strength of the clay is around 7,000 p.s.f., which is very high. The sensitivity on the other hand varies from moderate to infinite.

WOULD THE SLIDE EXTEND FURTHER?

These technical data now being released as the different studies are progressing evidently were not available at the time of the catastrophy. An emergency situation was created and the question was raised immediately as to the possibility of other slides to occur. The Canadian Armed Forces in their assistance program also offered to use their reconnaissance and remote sensing equipment Aircraft equipped with standard cameras and film were dispatched to the site for monitoring. Of all the data thus collected the black and white infrared photos and the thermal infrared imagery became most useful.

A CF-100 aircraft equipped with basic remote sensing instrumentation for civilian uses under the Canadian Center for Remote Sensing was called to monitor the area. The thermal infrared imagery was collected by the Reconofax XIII-A scanner. Other equipment in that aircraft included a set of Vinten cameras which collected the other types of photos.

Upon delivery of the first imagery available, a rapid and detailed interpretation was conducted in order to detect seepage zones, wet areas or any other information that could determine the extent of the danger zone. One must remember that several houses were still inhabited in the area adjacent to the housing complex that had suffered the disaster.

The first rapid interpretation of the data on May the 11th, indicated that seepage zones and wet areas could be detected on the combined thermal imagery and black and white infrared photos. These seepage zones and wet areas were thought to be potentially dangerous areas of saturated wet and weak clays and sands possibly related to some artesian conditions or major subsurface infiltration paths. An emergency drilling program was undertaken immediately with the cooperation and assistance of the Quebec Department of Natural Resources Hydrogeological Branch, and evacuation of the threatened houses was ordered at once.

Examples on the following figures (3,4 and 5) show the location of seepage zones that were subjected to a close analysis in the field and were drilled. Figure 3 shows both the thermal infrared imagery and black and white infrared photograph for the ground just downstream from the Shipshaw embankment and dam. This dam which at one time was thought to be related to the slide later was proven to be in excellent condition. A drill hole located at the seepage zone indicated by arrow (a) proved that no connection existed with the reservoir; only a small amount of seepage was occurring which added little water to the surface run-off. Arrows (b) indicate some ponds due to snow melting which substantially added to the run-off (arrows c) flowing towards the slide area (arrow d). Topographic shadows (arrows e) and anomalies due to vegetation (arrows f) could be sorted out from the seepage, wet ground and run-off areas with the help of the combined thermal IR imagery and black and white IR photography. Part of the landslide is shown in (g) and the residential area that suffered the disaster is shown in (h).

Figure 4 shows other seepage and saturated areas which were tested by drilling. Arrow (a) indicates a seepage zone east of the school at St-Jean-Vianney. Arrow (b) shows a seepage zone just north of the school; the drill test hole indicated very deep and soft

clay starting from the surface at that point. Arrows (c) indicate both seepage and run-off in a farm field. These zones were drilled also. Arrows (d) indicate standing water coming from the snow melt and from the surface drainage of the rain water (some not shown by arrows). Arrows (e) show the May 4th slide boundary and arrow (f) indicates the uppermost boundary of the older slide (about 500 years ago).

Figure 5 shows an interesting case of run-off and stagnant waters. Arrows (a) point at ponds of stagnant water much warmer than the water seeping and running off as indicated by arrows (b). The infrared black and white photographs alone could not help determine this situation. The thermal infrared imagery as in the two previous cases was the decisive input for interpretation. On Figure 5 the center of the black and white infrared flight line is indicated by lines c - c. Full coverage was not possible because of some underlap on both types of images. The outflow of the seepage and run-off as indicated by arrows (d) passed the overlap zone. Arrow (e) indicates some intensity level change in the scanner, not a terrain anomaly. The thermal infrared portion of this figure is in the 4 to 5 micron band as compared to the previous two which were in the 8 - 14 micron band.

AN EVALUATION OF THE DIFFERENT REMOTE SENSING TYPES OF DATA

The previous section by means of a few examples established the usefulness of both the thermal IR imagery and the infrared black and white photographs. This section deals with the total content of interpreted information for each type of film and imagery with a discussion of the respective advantages and the different problems.

THE THERMAL INFRARED IMAGERY

The first interpretation of the May 10 and May 12 thermal infrared imagery was conducted on the negative 70 mm. strip films on a small light table in order to provide a fast evaluation of the landslide area and adjacent grounds. This first level interpretation was given all the attention and care and with the assistance of the infrared black and white photos the following anomalous zones could be located and identified: 1) seepage zones, 2) water run-off, 3) wet soil areas, 4) standing waters, 5) topographic shadows, 6) dense vegetation, brush and forested areas. Of these, the seepage zones, water run-off and wet soil areas were retained as a direct input for guiding the first drilling program. Some anomalies due to shadows, vegetation, water bodies not related to the problem or due to the weather conditions were discarded.

In order to make the interpretation most effective, no single anomaly could be left aside unless it was identified. Once it was identified and its cause could be inferred, each anomaly was then checked in the field and decision was taken on the need of a drill hole, based on the terrain conditions and the inferred data from the imagery.

The following figures show some of the interpretation problems raised by the thermal infrared imagery. One problem of course is due to the scale distortions both lateral and longitudinal but with some eye and brain gymnastics and a topographic base map or regular aerial photographs, the interpretation of the thermal imagery can be conducted relatively fast. Two aspects that are most difficult to deal with are the time of the day effect on the imagery and the final decision as to what is the actual cause of the anomaly.

On Figure 6, some of these anomalies can be sorted out quite readily. The anomalies (dark areas in this case) due to vegetation (arrows A) and topographic shadows (arrows B and white line B) were identified readily. The most difficult task was to evaluate the origin, type and influence of the conspicuous anomalies (C to H) to the upper right of the water conduit (K) and the embankment (L). The question raised by the C to H anomalies as to their possible connection to the water body had to be answered as fast as possible and the nature of these anomalies determined. The C anomalies proved to be standing water mainly due to the melting of snow. The D anomaly was due to water saturated soils (wet soils) and anomaly E was due to both standing water and wet soils, partly drained by the system of ditches as in F. The G and H anomalies were slightly less intense; G was due principally to wet soils and H was due to a combined effect of short vegetation (shrubs and grass) and of surface drainage and wet soils.

The question thus remained as to the relationships between these anomalies with themselves and most particularly with the substratum. The dynamic thermal changes in time for these anomalies were not known and were necessitating some additional flights. In addition, the evidence of the anomalies alignment in a direction towards the landslide crater, indicative of some possible underground weakness zone connecting with it, was responsible for the decision of evacuating a few more houses on

May 12th in the area between letters D G E F and on down the street to the intersection at Harvey St. At that moment all of the residential development in M had been evacuated already. The following day the rotary drilling program was undertaken and the doubts were lifted although extensive thicknesses of weak clays were encountered in some of the drill holes.

The thermal infrared imagery thus proved most useful (1) in locating the wet soils, standing water conditions and the possible seepage zones; (2) in providing a powerful guide to the drilling program; (3) in discarding the presence of zones of weakness susceptible to slide.

Figure 7 is an additional example of the thermal IR imagery being used for locating seeps and run-off in a region down-stream from the landslide near the limits of Shipshaw. Points A show medium size seepage zones partly due to the melting of snow and to underground water seeping off the tabular topographic rise (M) at the water tower E. Anomaly B was checked to be a second rate seepage and anomaly C corresponded to some minor seepage and to some run-off. Anomaly D was similar to B, but smaller yet. The fears of an eventual slide in that area were due to the seepage zones in the clay and to the high local relief. There is an actual difference of elevation of some 270 feet between the water tower base and the bottom of the valley at point K. The valley itself at K is 130 feet deep.

Figure 8 is another outstanding example of the thermal imagery usability. It is a repetition of part of Figure 4 so as to emphasize its importance. Cross-section A - A' from Figure 2 is transferred here and the seepage zones, the water run-off and the highly remoulded soils of the old slide are much better understood in the light of the thermal imagery. The seepage zones at B readily were detected and the run-off and system of ditches (C) could be evaluated easily. The highly remoulded zone D covered up by a great many small ponds like E could be evaluated. Area under D actually corresponds to the grounds below the 1st old slide's escarpment (see also Figure 2). Area under F is a part of the remolded zone D and the forest cover emphasizes the anomaly due to very wet soils.

The figure 8 thermal imagery also shows a highly saturated zone of sands over clay indicated by arrows G. This was believed to be highly susceptible to sliding although the relief produced by the May 4th slide temporarily decreased the likelihood of further sliding. As the picture shows this zone G is now connected to the tip of May 4th slide through which it drains itself. Area under H is in a somewhat similar situation except that it is forested and it drains itself over the scarp of the May 4th slide, with the water running off on the slide floor.

THE BLACK AND WHITE INFRARED PHOTOGRAPHS

These photographs were obtained by a Vinten camera on 70 mm. film and printed to produce 9" x 9" enlargments with the regular 60% overlap. There was little or no sidelap as the flight lines were not planned for a full coverage. However these prints sensible to the 0.7 to 0.9 micron band were most useful because of (1) the geometric properties (little or no distortion), (2) a workable mapping size for data transfer, (3) and the sensing of data in a convenient band allowing some discrimination for water, soil moisture and vegetation.

The method used during the interpretation was to interpret the 9" x 9" black and white infrared photos using a pocket stereoscope and color pens to delineate all the zones of possible interest for the present purpose. The thermal infrared imagery on the strip negatives were then interpreted and the informative data transferred onto the 9" x 9" prints in a different color.

The black and white infrared photos thus were most useful. This spectral band however was limited in that it did not help differentiate the wet soil zones from burnt grass, vegetation with high evaporation and cloud shadows.

Besides these limitations these photos proved most adequate. With the thermal imagery, they provided the best combination of remote sensing data to the solution of this particular problem.

The color films: this film type generally produces the most useful information on soils. It would have been the case most probably here also but the cloud cover was almost continuous for all day on May 7th (see Table 1). So from the two 70 mm. film rolls that were inspected, only about a dozen or more frames were acceptable for interpretation purposes and yet the scale was not convenient and they were not on target. Therefore, this imagery was not subjected to any further examination.

The color infrared films: this film type is well known to be more sensitive to shadows than regular color films. For this reason, only a few frames were available for interpretation.

As most of these frames were off the target, this imagery again was abandoned. However, one frame was taken just over the landslide crater. The color balance was good but the film slightly over-exposed. It could be seen from this frame that the photo scale was not convenient as in the color film case and the color contrasts between soils, water and vegetation were too sharp and too high to produce a set of data of much value.

The standard black and white photos: in addition to the 70 mm. Vinten enlarged black and white prints, the normal aerial photographs come under this category. These photos were not particularly useful in terms of interpreted data. They were used mostly for location purposes and they were found useful to this project for this reason only.

DYNAMIC THERMAL INTENSITY CHANGES AS A FUNCTION OF TIME OF THE DAY

The collection of thermal infrared scanner imagery is known to be affected by the thermal behavior of the features being sensed. The dynamic thermal intensity changes or the emitted energy as a function of time of the day generally follows a curved pattern fairly similar to a sine function. However, the thermal inertia of certain features can produce a time lag to appear in the cooling-off or in the warm-up of these features. This produces a reversal of gray tones between the daytime anomalies and the night time anomalies. For instance, the water bodies at night are relatively warmer than the soils and vice-versa for the day.

This became most important when interpreting the St-Jean-Vianney thermal imagery. Numerous water bodies were of concern in this study and the soils conditions as well. It became evident early during the imagery interpretation that the daytime imagery procured only part of the answers that were looked for. For this reason additional IR trace flights were requested as early in the day as possible and preferably before the sunrise. An additional aspect to the thermal behavior of things is the changes as a function of time over a longer period like a week, or even a month. Some of the anomalies as it was expected would decrease in intensity or even disappear as the amount of run-off due to the thawing would decrease, leaving the anomalies really due to water infiltrations and seepage most evident. Additional thermal IR coverages thus were necessary.

The sensor wavelength is another parameter of major importance. In the early surveys at St-Jean-Vianney, an 8-14 micron band detector was used. For the last survey in August, the 4 - 5 micron band was used. When comparing the different images, it appears that the 8 - 14 micron imagery taken before the sunrise gave the maximum of information desired. The 8 - 14 micron imagery taken after the sunrise procured some information but it had to be supplemented by the subsequent thermal imagery. The 4 - 5 micron was much useful but it was affected by the weather conditions.

Figure 9 shows four imagery strips taken at different times of the day strips 1 to 3 were collected by the Reconofax XIII-A in the 8 - 14 micron band. Strip 4 was collected by the Reconofax IV in the 4 - 5 micron band. Points A and B are for location and orientation purposes. Points C and D on strip no. 2 show the standing water (C) in the crater and the running water (D). This was not evident on the other strips. On strip 4 running water in the streams is most evident and an extensive area of moist soils in the remoulded zone of the old slide is apparent (arrows E). This image also was taken at a time of the day past the temperature reversals. This zone delineated by the E arrows might correspond to either a moisture condition or to a zone of higher pore pressures in the soils. The monitoring program at this stage is not sufficiently advanced to answer this question. Strip no. 3 is closely similar to strip no. 2, except for the lack of discrimination between standing and running water.

Figure 10 is a second example of the time of the day influence for imagery taken under the same conditions as in figure 9. Here again the imagery taken before sunrise on May 12 was most relevant and the flow pattern shown in D is not detected elsewhere except on strip no. 1, where it appears in part at point E. Strip no. 3 is not as relevant and there happens to be either some fog or a gain level change on the scanner (see point H). Strip no. 4 is not very relevant here but it still has an input. The tip of lake K is shown in part as a very faint change in tone from the background, as a possible result of a water reflectivity changes. The light tone for water indicates a partially stagnant condition and relatively warmer temperature. The darker tone water bodies are indicative of a running condition as for stream (S) or possibly related to underground seepage to the surface thus giving a much colder temperature as for the water bodies (W).

GROUND CONTROL

In the study of a problem with so many facets, it obviously becomes necessary to collect as much ground control (ground truth) as possible. And so, for a more complete knowledge of the area and for a better understanding of the anomalous zones, extensive field checking was done by walking. In addition, a program of some rotary drill holes as mentioned was undertaken rapidly to try and lift the remaining doubts of subsequent sliding. Once this first program was completed, additional monitoring was undertaken which is to be reported by the competent people in reports to be published at a later date.

CIVIL ENGINEERING APPLICATIONS OF THERMAL INFRARED IMAGERY

This study of the St-Jean-Vianney landslide using different remote sensing imagery types showed the numerous uses of the thermal scanner imagery. Besides its capability of locating seepage zones, wet soils, run-off conditions and infiltration, it can be used as a powerful means of checking the efficiency of ditches and drainage systems. It can locate stagnant waters, running waters and soil saturation conditions. This type of imagery thus can become a powerful and reliable source of data for the civil engineer dealing particularly with soils and hydrology problems.

CONCLUSIONS

In the St-Jean-Vianney landslide study using five different imagery types as a source of data, the following conclusions were reached:

- the thermal infrared scanner data on 70 mm. strip negatives provided most reliable information in the guidance of an early drilling program designed to lift the remaining doubts of additional sliding
- the thermal infrared scanner data allowed to detect the seepage zones, the water saturated soils (wet soils), the run-off conditions, the standing waters
- the thermal anomalous zones had to be sorted out in order to reject the ones due to the vegetation, the topographic shadows, the water bodies not related to the landslide and the climatic conditions
- the optimum combination of remote sensing data was provided by both the thermal scanner imagery and the black and white infrared photographs, the latter being used as a base free of distortion for correlation purposes
- the color and color infrared films did not yield sufficient data for interpretation because of the cloud conditions that prevailed at the time these films were exposed
- the thermal infrared imagery collected before the sunrise yielded the maximum useful data in the study of the present problem
- the thermal infrared imagery combined with black and white infrared aerial photographs can provide a useful monitoring method of landslide areas, particularly if repetitive coverages can be provided in the early period following a slide, and it can be used as a fast data collection means to assist further field studies.

ACKNOWLEDGMENTS

The authors are very grateful to the Department of Natural Resources of the Province of Quebec and to the Canadian Armed Forces personnel who helped procure the imagery. The authors wish to express their deepest appreciation to the CFB Bagotville Alouette personnel who permitted the use of their interpretation facilities at the time of the sinister and to the PI unit at the CFB Rockcliff for providing copies of the imagery.

REFERENCES

(1) P. Lasalle, J.Y. Chagnon. An Ancient Landslide along the Saguenay River, Quebec. Can. Journ. of Earth Sci., Vol. 5, no. 3, pp. 548-549 1968.

(2) F. Tavenas et al. The St-Jean-Vianney Landslide: Observations and Eyewitnesses Account. Can. Geotech. Journ., Vol. 8, no. 3, pp. 463-478 1971.

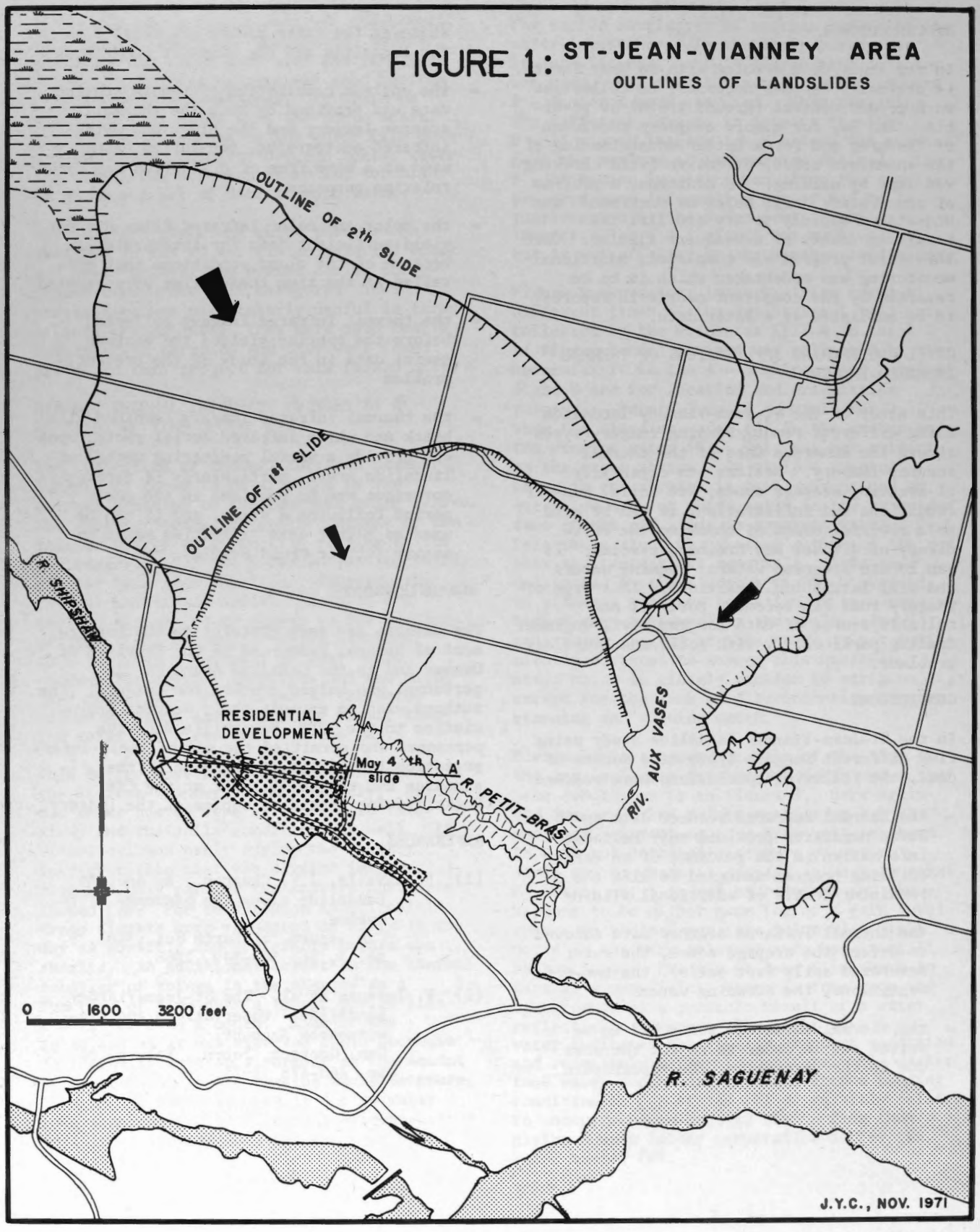


FIGURE 1: ST-JEAN-VIANNEY AREA
OUTLINES OF LANDSLIDES
OUTLINE OF 2th SLIDE
OUTLINE OF 1st SLIDE
R. SHIPSHAW
RESIDENTIAL DEVELOPMENT
A
May 4 th slide
A'
R. PETIT-BRAS
RIV. AUX VASES
0 1600 3200 feet
R. SAGUENAY
J.Y.C., NOV. 1971

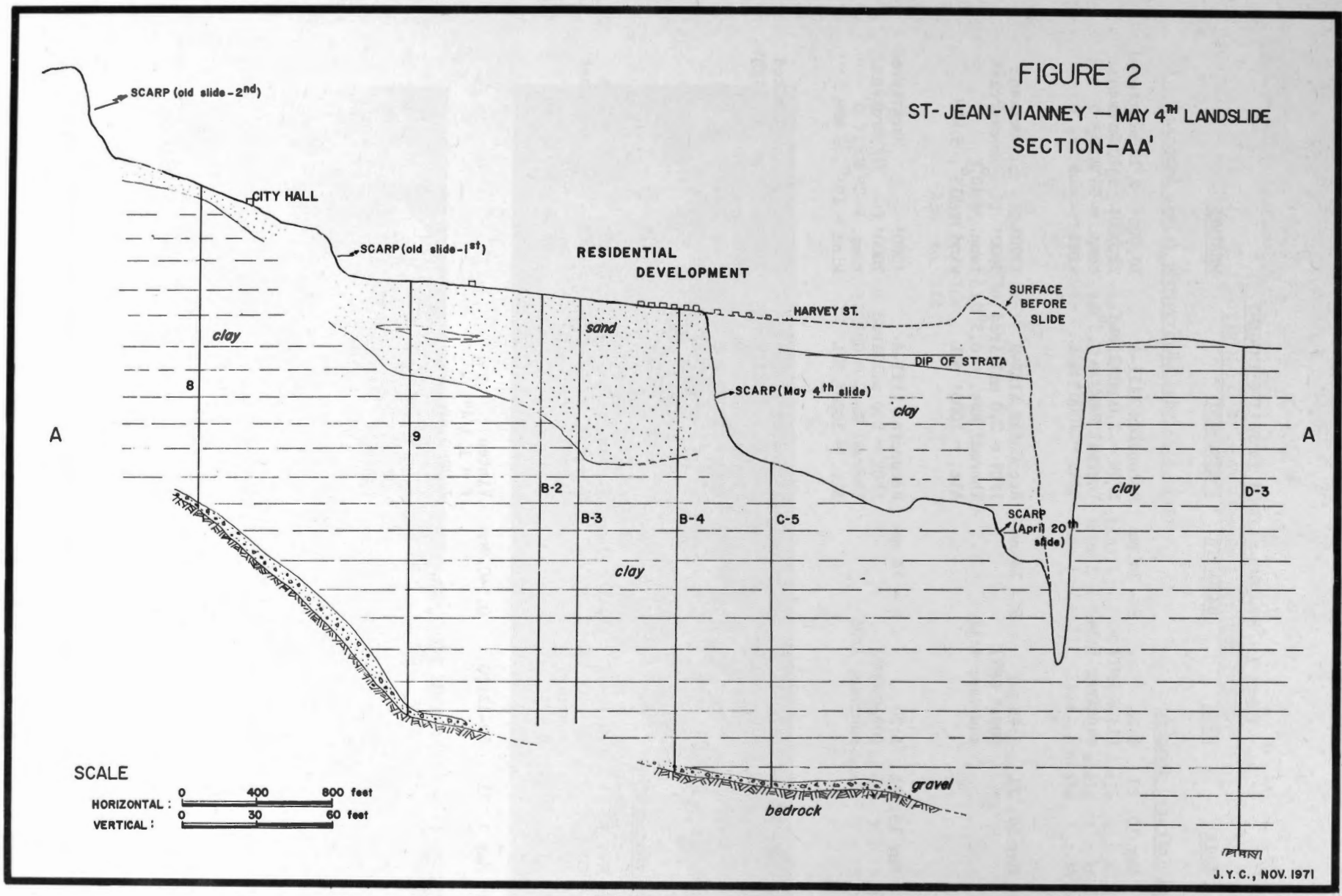


FIGURE 2
ST-JEAN-VIANNEY — MAY 4TH LANDSLIDE
SECTION-AA'
SCARP (old slide-2nd)
CITY HALL
SCARP (old slide-1st)
RESIDENTIAL
DEVELOPMENT
HARVEY ST.
SURFACE
BEFORE
SLIDE
DIP OF STRATA
SCARP (May 4th slide)
SCARP
(April 20th
slide)
clay
sand
clay
clay
clay
A
A
8
9
B-2
B-3
B-4
C-5
D-3
gravel
bedrock
SCALE
HORIZONTAL : 0 400 800 feet
VERTICAL : 0 30 60 feet
J.Y.C., NOV. 1971

TABLE 1 ST-JEAN-VIANNEY LANDSLIDE IMAGERY

IMAGERY TYPE	DATE	TIME	BANDWITH	INSTRUMENT	WEATHER
1. Thermal infrared imagery					
1T007	May 10, 71	9:26 (1326 GMT) sunrise; 5:14	8 - 14 nm	Reconofax XIII-A IFOV = 2.0 millirad Thermal Res. = 0.3°C Alt. = 1000' AGL	10,000' 2/10 scattered 22,000' 5/10 broken temp. = 55°F wind = calm
IT008	May 12, 71	5:45 (0945 GMT) sunrise; 5:11	8 - 14 nm	Reconofax XIII-A IFOV = 2.0 millirad Thermal Res. = 0.3°C Alt. = 1000' AGL	1700' broken 3600' overcast temp. = 40°F wind = 115°, 5 MPH
1T009	May 15, 71	5:20 (0920 GMT) sunrise; 5:05	8 - 14 nm	Reconofax XIII-A IFOV = 2.0 millirad Thermal Res. = 0.3°C Alt. = 1000' AGL	4500' scattered 7000' broken temp. = 28°F wind = 195°, 3 MPH
3T001	Aug. 20, 71	6-7:00 (10-1100 GMT) sunrise; N.A.	4 - 5 nm	Reconofax IV IFOV = 3.0 millirad Thermal Res. = 0.7°C Alt. = 1000' AGL	Overcast, viz. limited 7 mi., 1-3 mi. in light rain and some fog. temp. = 61°F wind = calm
2. 70 - mm photography					
a- colour IR, 2443 wratten 12	May 7, 71	10-12:00 14-1600 GMT sunrise, 5:17	0.5-0.9 nm	Vinten f = 3.0" 1/2000 sec, f stop 2.8	1700' broken 3600' overcast temp. = 40°F wind = 270°, 3 MPH
b- colour IR 2443 wratten 12	May 7, 71	10-12:00	0.5-0.9nm	Vinten f = 1 3/4" 1/2000 sec., f stop 4	-
c- colour 8442 no filter	May 7, 71	10-12:00	0.4-0.7nm	Vinten f = 3.0" 1/2000 sec., f stop 4	-
d- black and white infrared 5424 wratten 89B	May 7, 71	10-12:00	0.7-0.9nm	Vinten f = 2" - 4" 1/2000 sec., f stop -	-
e- black and white Plus-X wratten 12	May 7, 71	10-12:00	0.5-0.7nm	Vinten f = 2" - 4" 1/2000 sec. f stop -	-

TABLE 1 ST-JEAN-VIANNEY LANDSLIDE IMAGERY

3. Standard aerial photographs (abridged list)

a.	Q 64174 - 77 to Q 64174 - 79,	line L - 4870,	1964,	scale	1 : 15,840
	Q 64174 - 131 to Q 64174 - 134,	line L - 4821,	1964,	scale	1 : 15,840
	Q 64176 - 255 to Q 64176 - 260,	line L - 4822,	1964,	scale	1 : 15,840
b.	Contract 71-705, roll 6920		1969,	scale	1 : 24,000
	Prints 32 to 38				
	56 to 61				
	159 to 164				
	238 to 244				
c.	Q 71301 - 1 to Q 71301 - 8		1971,	scale	1 : 5,000
	Q 71302 - 1 to Q 71302 - 17		1971,	scale	1 : 5,000

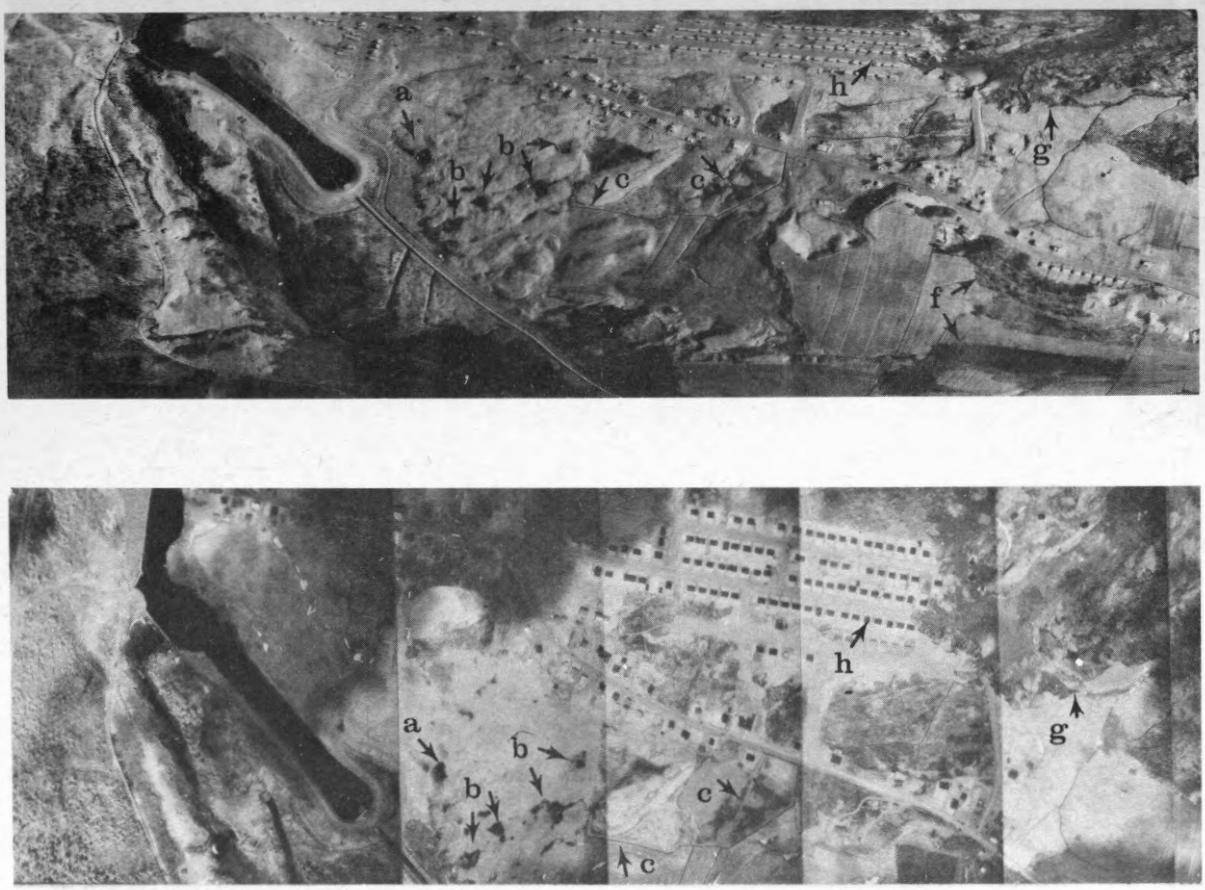


Figure 3: Thermal infrared imagery (above), May 10, 1971, 09:26 hr., 8-14 microns. Mosaic (below) of black and white infrared photos, May 7, 1971.

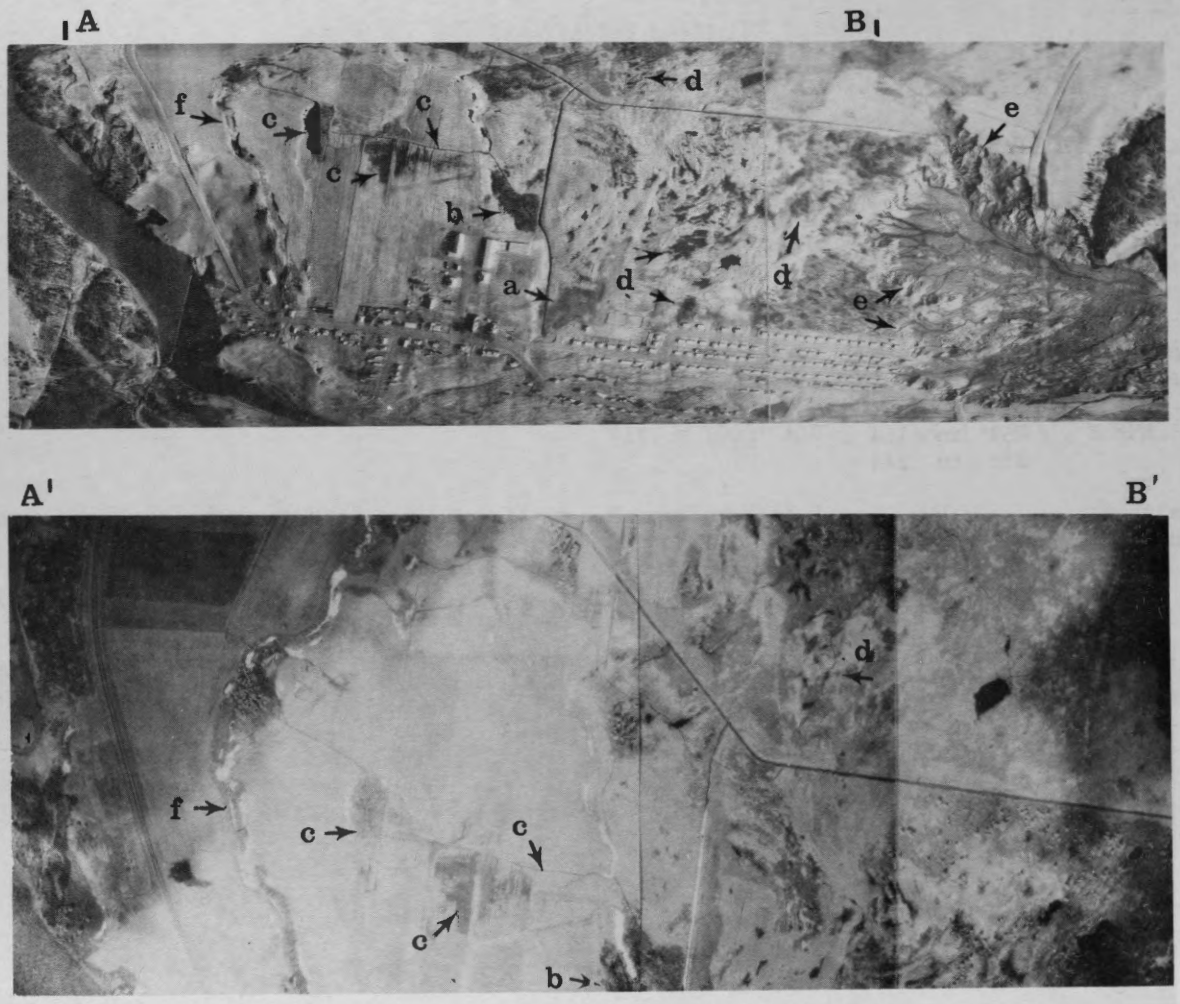


Figure 4: Thermal infrared imagery (above) May 10, 1971, 09:26 hr., 8-14 microns. Mosaic (below) of black and white infrared photos, May 7, 1971.

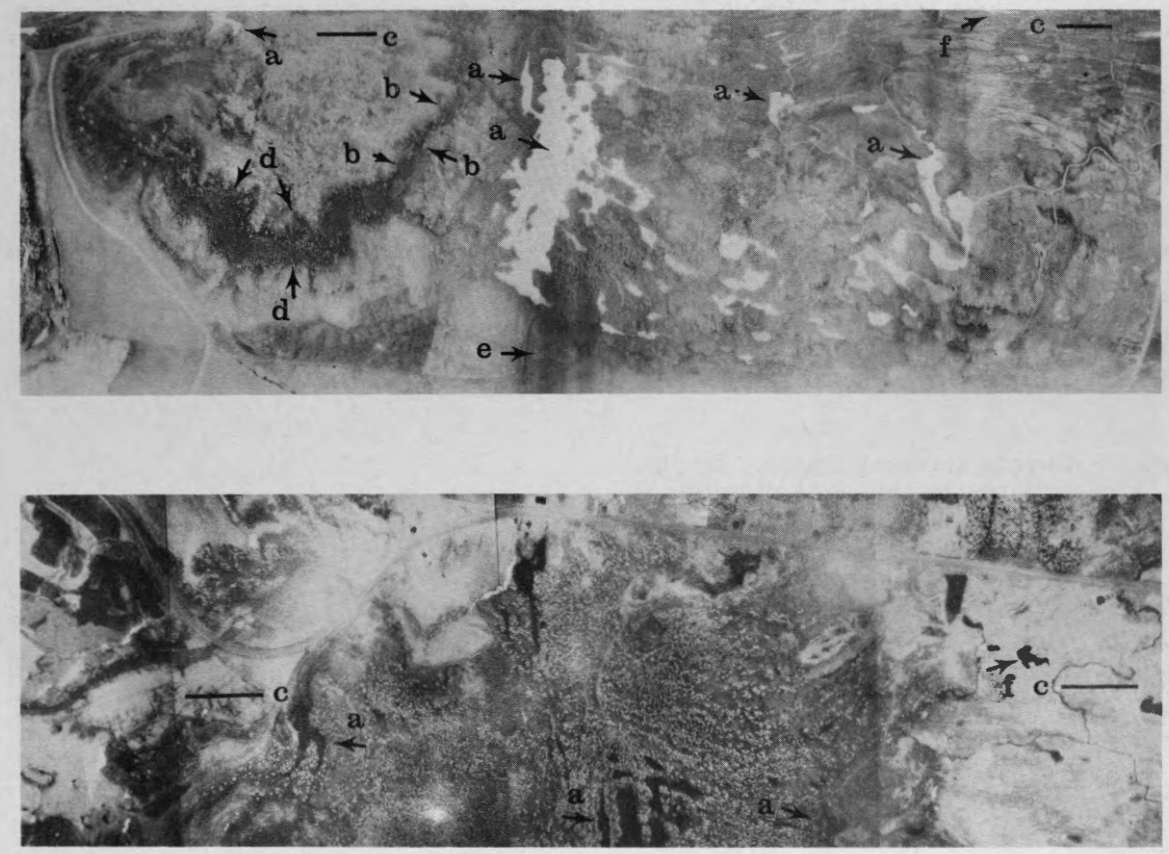


Figure 5: Thermal infrared imagery (above) May 12, 1971, 05:45 hr., 8-14 microns. Mosaic (below) of black and white infrared photos, May 7, 1971. Scale and flight line axis were tried to be made compatible.

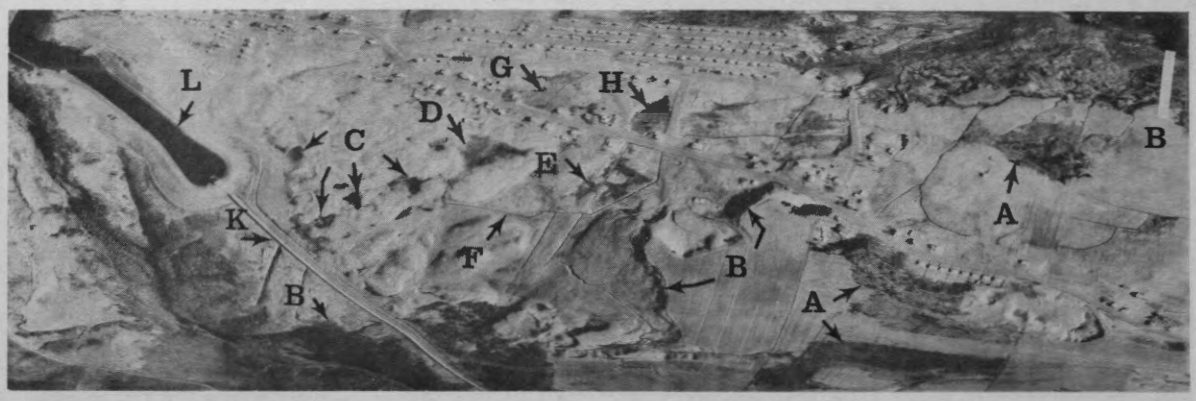


Figure 6: Thermal infrared imagery, May 10, 1971, 09:26 hr., 8-14 microns.

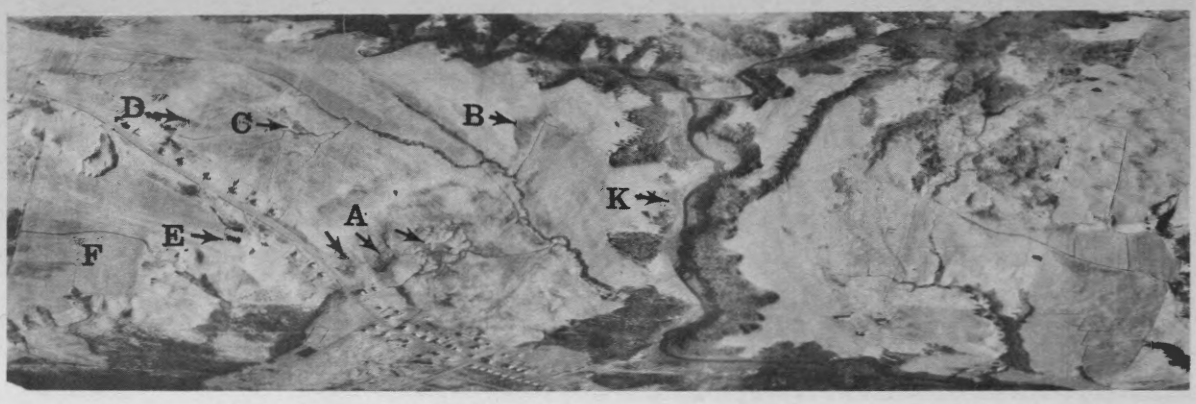


Figure 7: Thermal infrared imagery, May 10, 1971, 09:26 hr., 8-14 microns.

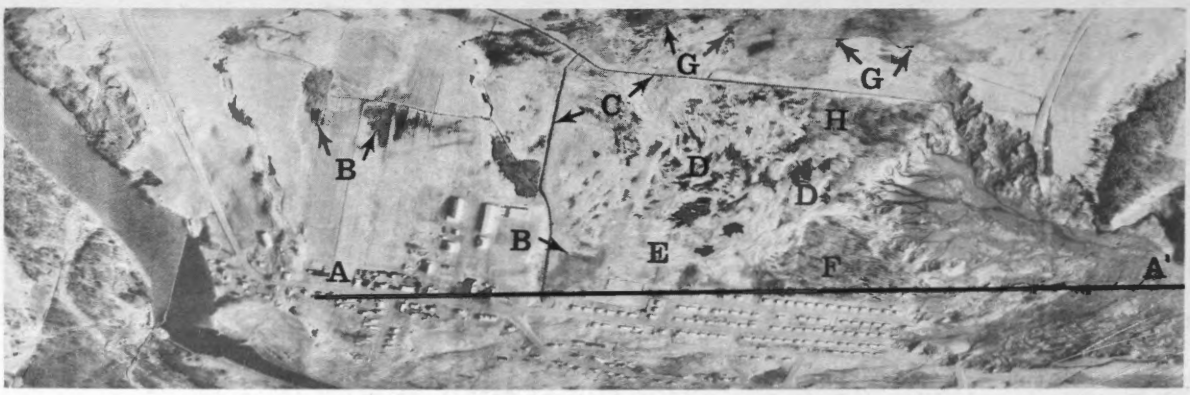


Figure 8: Thermal infrared imagery, May 10, 1971, 09:26 hr., 8-14 microns.

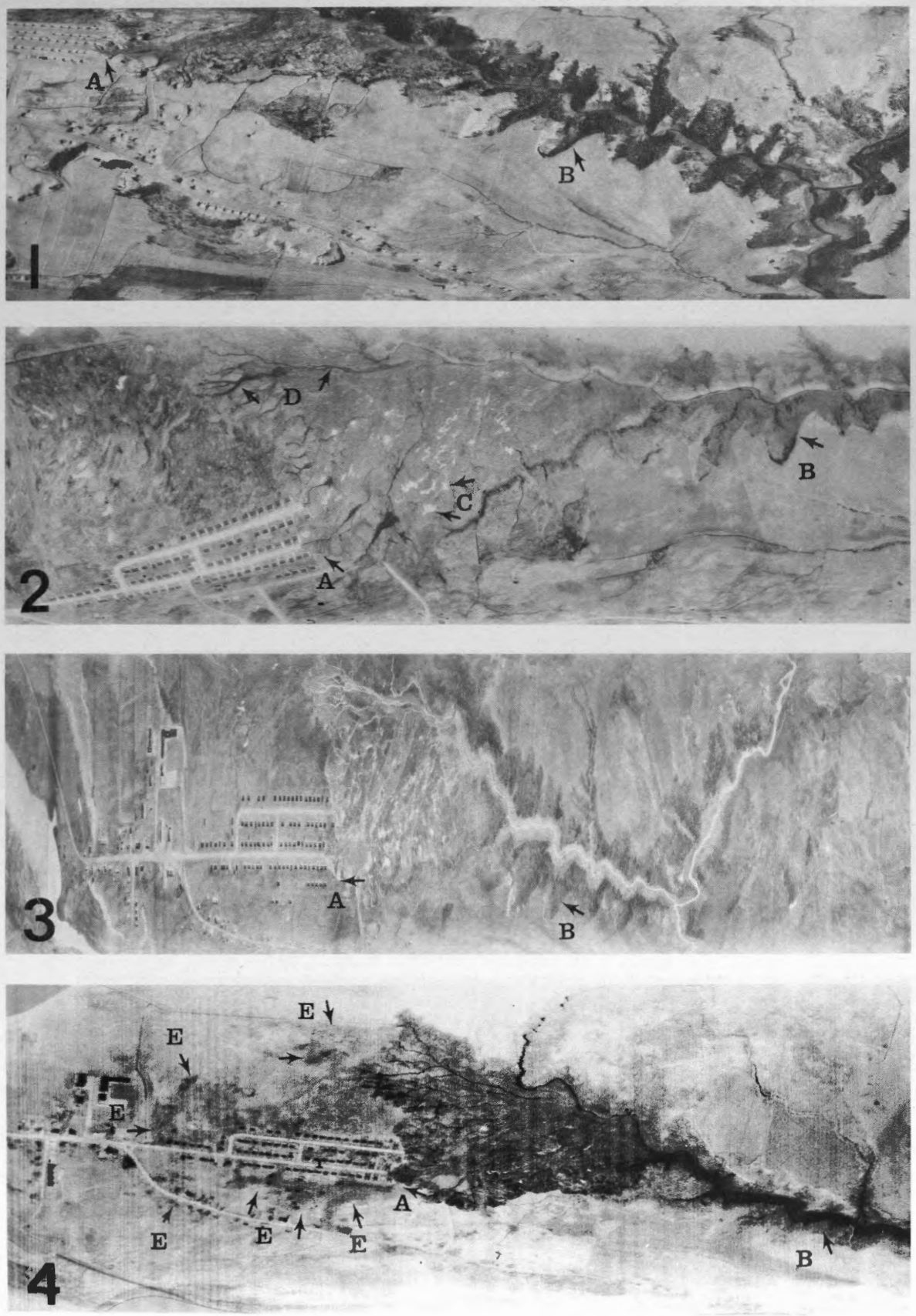


Figure 9: (1) 8-14 microns, May 10, 71, 9:26 AM
(2) 8-14 microns, May 12, 71, 5:45 AM
(3) 8-14 microns, May 15, 71, 5:20 AM
(4) 4-5 microns, Aug. 20, 71, 10:-11:00 AM

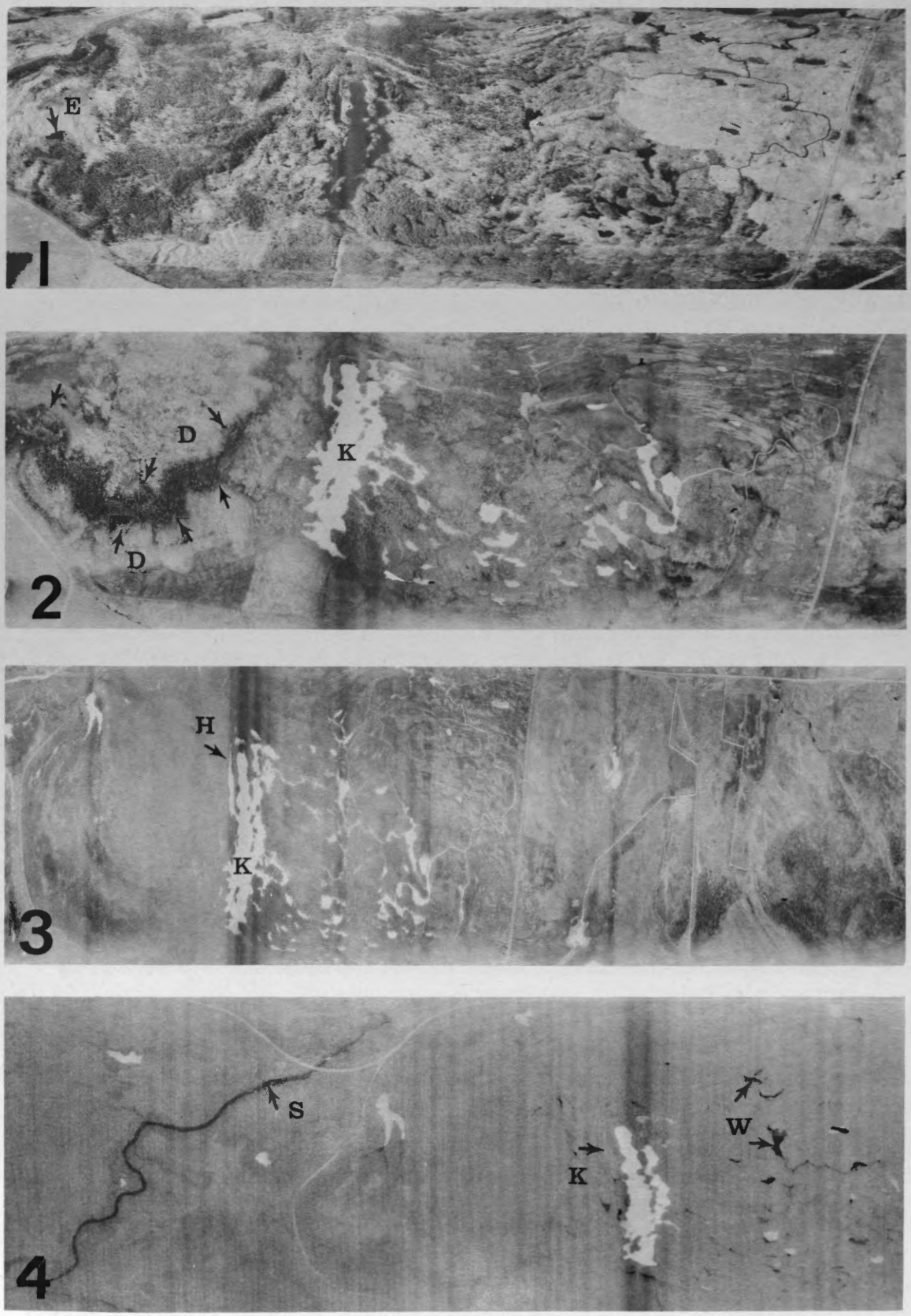


Figure 10: (1) 8-14 microns, May 10, 71, 9:26 AM
(2) 8-14 microns, May 12, 71, 5:45 AM
(3) 8-14 microns, May 15, 71, 5:20 AM
(4) 4-5 microns, Aug. 20, 71, 10:11:00 AM.

A POOR MAN'S REMOTE SENSING SYSTEM

W.B. McCoy
Civil Engineering Department
University of Saskatchewan
Saskatoon, Saskatchewan

ABSTRACT

The University of Saskatchewan in co-operation with the Canadian Wildlife Service and Ray's Flying Service of Saskatoon, Saskatchewan, has developed an inexpensive, flexible aerial camera system. This system is capable of providing standard black/white and colour imagery in the visible and near infra-red spectra. The system provides a convenient link between ground truthing and the type of regional imagery which will become available with the CF 100 and ERTS Systems.

Professor I.W. Tweddell and the author, in considering many of the research projects at the University, came to the realization that these projects could be assisted by special aerial photography. Many of the University research projects are well serviced by the conventional aerial photography supplied by the National Air Photo Library in Ottawa; the NAPL provides an excellent and inexpensive source of aerial photography. A large number of the research projects at the University require special photography — that is, photography which is not available through the National Air Photo Library. The research problem can be stated as one involving time or special viewing. A number of research projects are time dependent; the researcher must obtain his information at a specific time of the day, week, month, or year. A number of researchers wish to obtain information in specific portions of the spectrum (i.e. infra-red, blue-green, etc.); this requires special purpose photography. Time dependent and special purpose photography is not a service of the National Air Photo Library.

The University of Saskatchewan suffers from the disadvantage that air survey cameras and aircraft are not based in the province of Saskatchewan; this means that the cost of "deadheading" the aircraft to Saskatoon is prohibitive and may deplete the research budget.

Attempts to solve the problem of special purpose photography by using a Cessna 172 (high wing aircraft) for oblique photography were carried out in 1968. It was immediately apparent that oblique photography was not suitable for many of these projects.

During 1968, Don Dabbs, a graduate student who was employed by the Canadian Wildlife Service, discovered that removing one panel in the floor of a Piper Twin Apache presented an opening large enough to mount a camera in the vertical configuration. Mr. Dabbs used this system over the Cumberland Marsh Project for Canadian Wildlife during the summer of 1968.

The Surveys and Mapping Branch of the Federal Department of Energy, Mines and Resources made a research grant available to the Civil Engineering Department, University of Saskatchewan which permitted the purchase of the necessary camera equipment for this system. The camera mount was designed and built in the winter of 1968/69. The camera and mount have been used extensively since that time; the various projects are listed in an appendix.

The camera system is based on the Hasselblad camera. The Hasselblad is a Swedish built, 70 mm camera of high quality; in fact, this was the type of camera chosen for the astronauts on their moon missions. The Hasselblad has interchangeable lens from 40 mm to 500 mm, giving the photographer great flexibility. The 80 mm is the standard angle lens and was used for most of the University research. There are several magazines available; the twelve exposure magazine is standard with the camera and is used with ordinary 120 film. The University purchased seventy exposure magazines; these proved to be most suitable for the work. Hasselblad recently announced a

500 exposure magazine; this magazine would be used on large projects. The camera can be purchased as a manually or as a battery operated instrument. The battery operated camera cycles in less than two seconds and is more easily operated; this type of configuration is well worth the additional purchase price. The hand operated camera requires about three seconds per cycle.

The camera must be mounted in the aircraft; this involves a custom-built mount. The user must design his own mount to fit the camera and aircraft which is available to him. The mount designed for the Piper Twin Apache is not necessarily suitable for other aircraft. The mount should provide a safe, stable platform for the camera and should permit the camera operator to see the ground over which he is flying. The University camera mount has a separate horizontal head which supports the camera; this head is mounted on nylon clamps which allow the operator to rotate the camera about the "Z" axis to compensate for drift. The camera head has horizontal pivots on the camera mount; these pivots enable the operator to use a bull's-eye level bubble to bring the camera lens axis vertical.

The researchers at the University have found that generally a photographic scale of 1:12000 is suitable; on some of the projects the scales have differed from this. They have generally used transparency type film rather than the print type for colour and colour infra-red work. Transparencies have two advantages over prints. Transparencies can be projected to a variety of scales; this is especially useful when photos are to be used by several researchers — each wishing to use a different scale. Transparencies show better resolution and detail than do the colour prints; this is an important consideration when coverage must be limited and details are important.

The University has found that it is advantageous to charter the aircraft from a local flying service. The researcher who uses a local charter service has the advantage of flexibility; for example,it requires less than half an hour to modify the Apache; this means that photography can be flown at very short notice. Researchers often wish to study ephemeral or unexpected phenomena; the flexibility of a local charter makes this possible. The researcher will find that the local charter owner and his pilots take a personal, friendly interest in his work and will often be able to provide special service for him. The aircraft is a fairly light one; this means that the charter cost is moderate and the significance of this for small research budgets need not be belaboured.

The camera system does have a major disadvantage. The Hasselblad camera has a 70 mm format and area coverage does not compare with that which can be obtained on the 9-inch square format of the normal survey camera. The University has not found this to be a serious limitation to the present; the projects have involved research studies over limited areas requiring modest line lengths.

The University has been doing the development of the film; this is not recommended. The demand for darkroom facilities in research projects will fluctuate. The National Air Photo Library provides 70 mm development service for government and University research projects; the quality of the work is excellent; the charge is modest and "turn around" time is very reasonable. In the future, the University of Saskatchewan will probably use this service exclusively.

The University has been using a single camera type of mount to the present. Most researchers wish to have their projects flown with more than one film/filter combination. The requirement of multi-spectral studies has led to repetitive flights over the same line and this has two disadvantages. The obvious disadvantage is the greater cost and time of the project if additional flights are required The major disadvantage of multiple flights over the same area lies in the fact that the coverage on the various flights never do exactly correlate and this increases the difficulty of cross correlation somewhat. The University and Canadian Wildlife Service have often co-operated in the past and plan to carry this further in the future by "pooling" their cameras and building a multi-camera mount.

The new camera mount will accommodate up to four Hasselblad Electric Cameras operated synchronously. The cameras will be independently attached to the top plate of the mount so that the lens collimation can be adjusted for photo coverage coincidence. The top plate of the mount will be adjustable for level and for drift. A Government surplus drift sight and an intervalometer will be incorporated into the mount. At present, the details have not been completed. It is hoped that the mount will be built in the near future.

In conclusion, a researcher who is willing to display a bit of ingenuity and to co-operate

with others, can build an inexpensive, special purpose air camera system to further his research studies. The researcher will be surprised at the diversity of use which will develop for the system.

AIR PHOTO PROJECTS APPENDIX

Projects carried out by Canadian Wildlife Service:

1. Cumberland Marsh Studies - summer 1968; this was a study of waterfowl habitat photographed in black and white and in colour; in part, it was used to test the system.

2. Matador Project - spring 1969; this was a flight photographed in colour and colour infrared film over the International Biological study area at Matador, Saskatchewan. This flight combined further testing of the system with service to the Matadór Research Group.

3. The Athabasca Delta Project - summer 1969; continuing. This project is the prime raison d'etre for the Wildlife photographic research. The project is an "in depth" continuing study of the impact of the W.A.C. Bennett Dam upon the ecology of the Athabasca Delta. Photography is one of the important tools in this study.

4. The South Athabasca Sandhill Study - 1971; this is a photo study of the ecology of a sandhill area in co-operation with Professor J.S. Rowe, University of Saskatchewan. This study points up the excellent co-operation which the Canadian Wildlife fosters with our University.

Projects carried out by the University of Saskatchewan:

1. Pollution Studies on the South Saskatchewan River at Saskatoon - 1968 - continuing. In 1968, a few oblique photos were flown in colour and black and white — in subsequent years vertical photography missions using various film/filter combinations were flown. During the summer of 1971 we hired two students under the Opportunities for Youth Programme; a report upon our work until that time was submitted to the Resources Study Group (University).

2. Pollution Studies on the North Saskatchewan River from Prince Albert - East - 1970 - continuing. This study has paralleled the South Saskatchewan study in most of its details; a report for this project was submitted to the Resources Study Group.

3. The Langham Study - 1970/71 - terminated. This project was a major geotechnical study of a portion of the North Saskatchewan River near Langham, Saskatchewan. Mr. Ed Wilson, a graduate student in Transportation Engineering, wrote his Master's Thesis on this project. Mr. Wilson carried out a very thorough study of the inter-relationships of plants, soils, geology and groundwater units.

4. The Potash Salt Project - 1970/71 - terminated. This was a study conducted in the vicinity of a local potash mine by Messrs. Olson and Wilkie for their fourth year Civil Engineering Graduand Thesis. The study was instituted to test the practicability of measuring salt fallout from the mine stack using plants under stress as the indicator. Certain extraneous elements prevented these men from reaching a solid set of conclusions and the work was terminated with inconclusive results.

5. The Mercury Pollution Study - 1971 - continuing. The Chemical Engineering Dept., University of Saskatchewan, have made limited use of our photos on the South Saskatchewan River study of possible mercury pollution from a chemical plant.

6. Arctic Land Use Research - 1971 - continuing. Our group provided a service facility for ALUR in that we flew photographic missions for them in the Watson Lake, Y.T., project. We supplied photography only; we did not carry out an interpretation role.

7. Horticultural Studies - 1971 - continuing. The Department of Horticulture, University of Saskatchewan, requested a number of flights in the Saskatoon area, and at the Outlook Irrigation area; we supplied photography for them and they are presently studying this photography.

8. South Saskatchewan River - 1971 - continuing. Messrs. Rosseker and McAvoy, fourth year Civil Engineering students, are writing their Graduand Thesis on a

geotechnical and sedimentation study of the South Saskatchewan River from Cutbank to Pike Lake, Saskatchewan. We flew a photographic mission in support of this thesis in the autumn of 1970. The area is under study at present and no firm conclusions have been reached at present.

9. Saskatoon Collegiates - 1969 - continuing. We have been co-operating with Geography teachers in two Saskatoon Collegiates since 1969 and have flown several small missions for them. The missions we have flown to the present have been large scale urban photography in black and white and colour. The photos are used for a number of urban planning studies by the collegiate students.

10. South Saskatchewan River Ice Study - 1971 - continuing. This is a study of the open water patterns in the ice cover of the river in the vicinity of Saskatoon. A number of sources of warmed water (powerhouse, sewage plant and others) maintain certain open water areas in the vicinity and we are studying the patterns of this open water.

11. Moon Lake Study - 1971 - terminated. An area on the South Saskatchewan River in the vicinity of Moon Lake near Saskatoon was flown; R. Hall, a graduate student, used these near infrared photos for a plant ecology study. In this project, our only contribution was the photography; Mr. Hall carried out the remainder of the work.

Figure 2

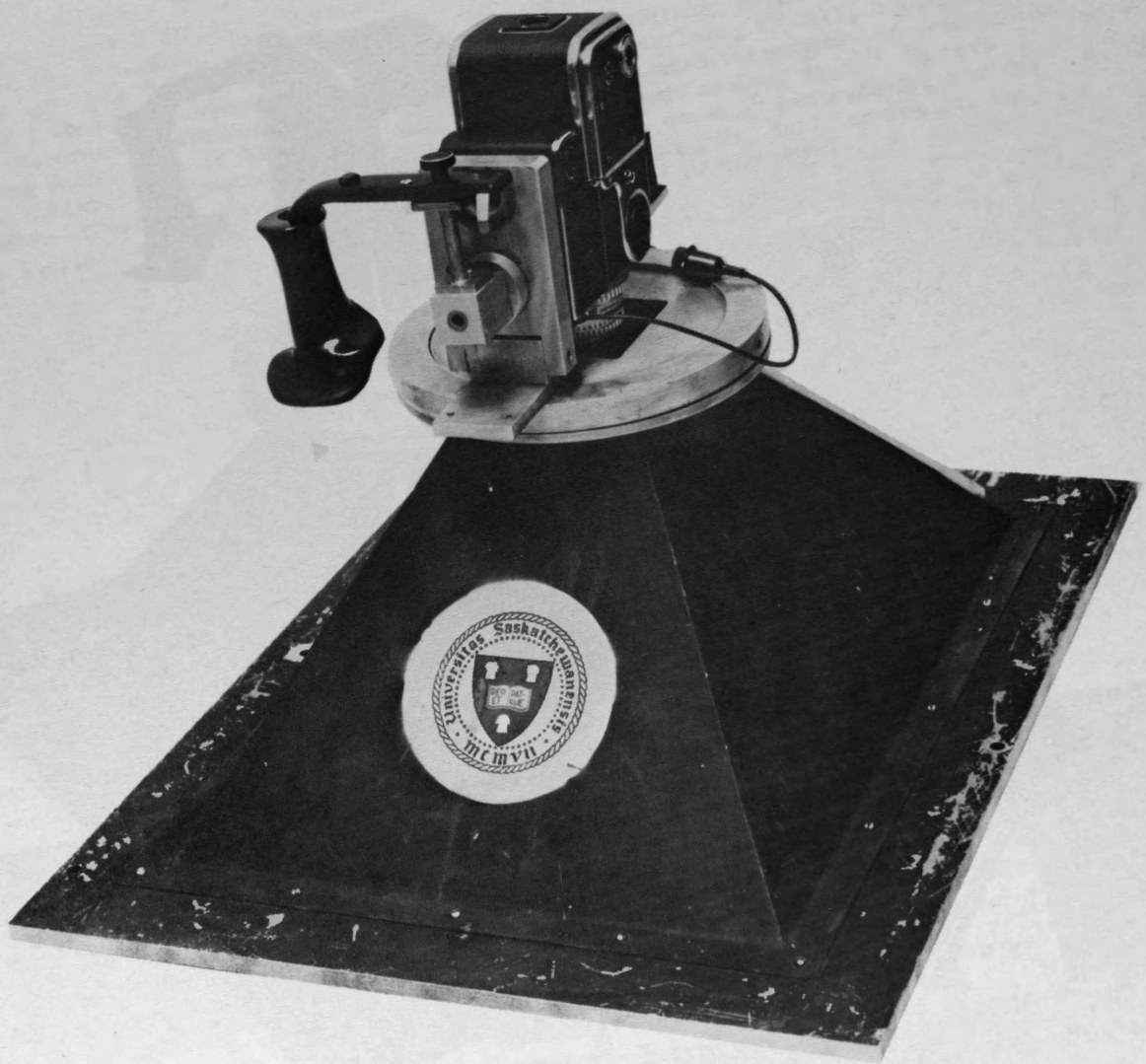


Figure 3

EFFECTS OF SPECTRAL FILTRATION AND ATMOSPHERIC CONDITIONS ON AERIAL PHOTOGRAPHY OBTAINED IN 1970 and 1971.

U. Nielsen,
Research Scientist,
Forest Management Institute,
Canadian Forestry Service,
Ottawa, Ontario, K1A 0H3.

ABSTRACT

Photography obtained by the Airborne Sensing Unit provided material which illustrates many of the basic concepts of spectral reflectance, atmospheric attenuation and spectral filtration discussed in the paper. These concepts have to be understood and the presently very limited amount of information concerning spectral reflectance and effect of atmospheric factors has to be vastly increased if specifications for aerial photography are to be optimized.

INTRODUCTION

Over the last two years remote sensing has become a major research tool for Canadian environmental scientists. Although great advances have been made in the development of sensors aimed at recording electromagnetic energy far beyond the visible range of the spectrum, photography will probably remain the primal technique in monitoring many of the earth surface's resources. The capabilities of photographic techniques have been greatly increased with the development of multiband and multidate, and high- and orbital-altitude photography. Related to this increase in capabilities has been the increase in complexity and costs of the photographic missions, most of which cannot be repeated when failure occurs. Optimum quality of the material obtained has therefore become extremely important. Photography obtained by the Canadian Centre for Remote Sensing during 1970 and most of 1971 provided material with a wide range of quality. The purpose of this paper is to discuss some basic concepts which, one hopes, will provide a better understanding of these quality differences.

SPECTRAL FILTRATION

Tonal contrast is one of the most important criteria for discrimination of vegetation types or, generally, features of the same nature on aerial photographs, especially when scales are too small to recognize discriminating morphological characteristics. Recording of differences in spectral reflectance is therefore one of the main tasks of aerial photography for interpretation purposes. These differences in spectral reflectance are usually limited to narrow bands of the photographic spectrum and therefore can only be recorded with the proper film-filter combination. Spectral reflectance of vegetation, especially in the near-infrared, changes with the season. Since the timing of phenological events is characteristic of species, their separation is much easier on photography obtained early in the growing season; phenological variation within species must be considered. In the summer, after leaves have reached maximum areal development, spectral reflectance of different species tends to equalize making a separation more difficult. In the fall, differentiation again increases, but near-infrared reflectance is, in general, considerably lower. Unfortunately data on airborne spectral reflectance measurements are very scarce and consequently, photographic specifications are seldom optimized.

Film and filters presently available permit the recording of almost any portion of the photographic spectrum. Simultaneous recording of several bands (multiband photography) normally increases the interpretation capabilities. It has become almost standard procedure to approximately record the three bands to be recorded by the return beam vidicon sensors aboard NASA's Earth Resources and Technology Satellite. The film-filter combinations normally used by the Airborne Sensing Unit (ASU) are:

- Film No. 5063 with a Wratten 12 and a Wratten 44 filter (green band),
- Film No. 5063 with a Wratten 25A filter (red band), and
- Film No. 2424 with a Wratten 89B filter (near-infrared band).

There is no apparent reason for the use of two filters in the first combination. The ERTS band 475-575 nanometers (green) is best

covered with the Wratten 58 filter. If a slightly narrower band is desired the Wratten 61 filter could be used. The advantage of these two filters as opposed to the combination of the 12 and 44 is their higher transmittance (Figure 1) and stability. In addition, the requirement for filter-flatness to avoid loss in image sharpness is easier to satisfy when one filter is used. Since Film 5063 has no extended red sensitivity, the second combination covers a somewhat narrower band than the corresponding ERTS return beam vidicon camera (580-640 as opposed to 580-680 nm.)

Simultaneously with the ERTS bands the ASU often exposed a color-infrared film (Kodak Ektachrome Infrared Aero Film 8443 and later, Kodak Aerochrome Infrared Film 2443) which, in most cases, has been of better quality than the black and white films. Since there has been some confusion about the use of filters with the color-infrared film, a summary of the film's characteristics seems justified.

The color-infrared film has the same three dye-forming layers (yellow, magenta, and cyan) as the normal-color films, but they are sensitized for longer wavelengths (Fritz, 1967). Since all three layers are also sensitive to blue light (haze), a minus-blue filter (Wratten 12 or similar) is always used to limit the exposure to wavelengths longer than 500 nanometers.

It is important to note, that an increase in exposure decreases the density of the dyes formed by the reversal development. Colors are therefore not formed by the exposed, but rather by the unexposed dyes. Blue light will expose the yellow-forming dye of the normal-color film; the remaining magenta- and cyan-forming dyes will, when developed, combine to a blue color. Blue light will not expose the color-infrared film (minus-blue filter). Green light will expose the magenta-forming dye of the normal-color film (leaving the yellow- and cyan-forming dyes unexposed) and the yellow-forming dye of the color-infrared film (leaving the magenta- and cyan-forming dyes unexposed). Resulting colors will be green and blue respectively. Red light will expose the cyan-forming dye of the normal-color film and the magenta-forming dye of the color-infrared film. The unexposed dyes will combine to red and green colors respectively. Near-infrared energy will not expose the normal-color film but will expose the cyan-forming dye of the color-infrared film. Remaining yellow and magenta combine to a red color.

The sequence is simple. Both films have yellow, magenta- and cyan-forming dye layers, exposure of which results in blue, green and red colors respectively. The difference is the sensitivity, which is to blue, green and red energy respectively for the normal-color film and green, red and near-infrared energy for the color-infrared film with a Wratten 12 filter.

Another difference between both color films is that the cyan-forming layer (near-infrared sensitive) of the infrared film is slower than the other two. The reason for this is the high near-infrared reflectance of vegetation, which is one of the primary subjects of interest. If the speed of the three layers is equalized, the reflectance from vegetation may yield very low cyan densities (toe of the curve in Figure 2). The images would be excessively red and small differences in near-infrared reflectance would not be detectable. The balance between the three layers can be manipulated with the help of additional filters (Fritz, 1967). Visible-transmitting and infrared-absorbing filters, such as the Corning 3966, would decrease the amount of energy exposing the cyan-forming layer (shift the cyan curve further to the right). Visible-absorbing and infrared-transmitting filters such as many of the Kodak color-compensating (CC), light-balancing and photometric filters, would decrease the amount of visible energy reaching the film (shift the curves of the yellow- and magenta-forming layers to the right).

The optimum filter selection will depend on film condition, the thickness and condition of the atmosphere through which the reflected light must pass, and, of course, the purpose of the photography. While film aging and low temperatures decrease the sensitivity of the film, particularly that of the cyan-forming layer, water vapour in the atmosphere and increased thickness of the air mass (high altitudes) decrease the near-infrared energy reaching the film. The scattering of short wavelengths which increases the amount of visible energy also weakens the infrared record. In any of these cases, visible-absorbing and infrared-transmitting filters will restore the balance between the three film layers. If the opposite is the case, i.e., if the amount of near-infrared energy reaching the film is too great compared with the visible energy (low-altitude aerial photography), an infrared-absorbing filter should be used.

The balance between the three film layers can also be manipulated by using different

processing techniques. But this approach is difficult to follow in practice.

A properly exposed infrared film will yield good quality color separations. These can be obtained by copying the color-infrared film onto black and white copying film through the proper filters. A Wratten 47B filter would transmit blue, which is the record of the green reflectance of the original scene. A Wratten 58 filter would transmit green (record of original red) and a Wratten 25A filter would transmit red (record of original near-infrared). So far, there is little reported on this approach. A test made resulted in color separations, which had a much higher information content and were of far better quality than the simultaneously exposed black and white films (Figure 3). If this is consistent, separate exposure of the three bands seems redundant. It is stressed, that these findings are based on the material investigated. Better black and white photography can be obtained with properly calibrated systems.

Since the three dye-forming layers of the normal-color films have approximately the same speed, and since their spectral response is limited to the lesser reflected visible energy, they are, perhaps, easier to use than color-infrared films. But these characteristics limit their use, at least for vegetation studies, to low altitudes. Their sensitivity to blue light causes an extreme contrast reduction. Contrast ratios, which can be 1000:1 at ground level, may be reduced to 10:1 from high altitudes. Moreover, short wavelengths seem to overlap and are difficult to separate. Blue and green (vegetation) objects will appear in a similar color on the photographic record. Haze-cutting filters may improve the picture slightly, but do not solve the problem. Cutting too much blue light will yield an overly yellow image, especially if the photograph is exposed at low altitudes or in clear atmospheric conditions. In this case, a filter similar to the Wratten 1A or no filter at all will give the best results. The often used combination of the HF3 and HF4 or HF5 filters has no apparent advantage over the use of the HF3 filter alone.

ATMOSPHERIC CONDITIONS

The light reaching the earth's surface is composed of direct sunlight and skylight. The latter is solar energy scattered and reflected by particles suspended in the atmosphere. The proportion of these two components, which is strongly affected by atmospheric conditions, has a direct influence on the spectral composition of the light reaching the ground. Besides smoke, dust and other solid particles, condensed water vapour is the most light-disturbing factor. It increases scattering of short wavelengths and absorbs long wavelengths. Since scattering of reflected light reaching the camera increases luminance (but decreases contrast), the exposure of films recording shorter wavelengths has to be decreased. If color-infrared film is used, this shift in exposure, added to the absorption of long wavelengths, will further weaken the infrared record. At solar altitudes of 50° and more and for a clear sky condition the proportion of visible sunlight to skylight is about 3 to 1. With light haze this proportion might change to 1 to 30 in favour of skylight. Jones and Condit (1948) report that the proportion of the two light components show considerable variation, even when the conditions as appraised visually are the same. Undoubtedly there is a need for more research in this field to quantify the effect of these factors upon photographic exposure.

The poor quality of some of the photography obtained by the Airborne Sensing Unit was mostly due to the condition of the atmosphere at time of exposure. Quality could be correlated with water content as indicated by high dew point or high temperature and relative humidity.

Vignetting or light falloff towards the edges of the photographs is also a serious problem which particularly affects densitometric measurements. In addition to lens characteristics and lens aperture used, other factors that increase vignetting are under-exposure, filters and filter mounts (reflections). Vignetting is serious on most photography obtained under the remote sensing program.

OTHER FACTORS

In addition to the two variables discussed above, there are many others which affect quality of photographic records. For example, the chromatic correction of lenses is important. A lens corrected for wavelengths of 500 to 600 nanometers will not yield a sharp image resulting from light of shorter or longer wavelengths. This consideration is very important when using band-pass filters (ERTS bands). Ideally, a camera system should be optimized for the specific band for which it is to be used.

Another important factor is processing. While exposure controls density of a negative, degree of development controls density

differences or contrast. A numerical measure for the degree of development is gamma. In theory, a gamma smaller than one decreases, a gamma higher than one enhances the contrast ratios of a scene. But due to many factors, partly discussed above, contrast ratios are decreased and a gamma higher than one is needed to retain the brightness range of the scene. Processing should therefore be more specific. High-altitude photography should be processed to higher gammas than low-altitude photography. Also films exposed within the three ERTS bands should be processed differently, the record of the shortest wavelengths being processed to the highest gamma.

Processing of color films is much more critical since, in addition to contrast, color balance has to be considered. The color-negative film number 2445 can only be processed to a negative, while the color-reversal film number 2448 may be processed either to positive or negative. If the original roll of film is available to the user, a reversal-type film processed to positive will yield the most "truthful" record. By further processing and printing, the user has to rely upon the processor's judgement on the truthful color rendition of the scene at time of exposure. Color-infrared film should be processed to positive only. Negative processing techniques available to date produce extremely poor results.

CONCLUSIONS

One of the main tasks of small-scale aerial photography is to record differences or changes in spectral reflectance. To optimize this recording, a knowledge of the reflectance characteristics of the objects of interest and the effect of environmental factors on the intensity and composition of the reflected energy reaching the sensor is mandatory. Spectral regions, where differences in reflectance are either greatest or lowest could be selected. This would eliminate the recording of superfluous information and would involve reductions in the time and costs of interpretation. Unfortunately, there is an almost complete lack of reliable airborne data on spectral reflectance obtained under similar conditions as photography. Data obtained from laboratory and field measurements often conflict with measurements taken from a flying aircraft (Steiner and Gutermann, 1966). Also, little is known on the effect of environmental variables such as illumination and atmospheric attenuation. This absence of information leads to superfluous duplications of research experiments, and little progress in the improvement of photographic specifications for specific purposes.

Color-infrared film has been exposed in almost every mission flown by the Airborne Sensing Unit. For most applications, at least those involving vegetation studies, it is the most useful film, provided it is properly used. The balance of its three dye-forming layers can be manipulated with the aid of additional filters. Some will attenuate the visible energy without reducing energy in the near-infrared. This requires an increase in exposure, which will better expose the cyan-forming layer. This equalizing of the speed of the three layers is often necessary to compensate for haze, low near-infrared reflectance (late summer and winter for vegetation or any season for non-vegetation targets) and improper film storage. On the other hand, if near-infrared energy reaching the film is very high (vegetation targets from low altitudes), the density of the cyan-forming layer can be increased by using an infrared-attenuating filter.

At present, there is insufficient information to recommend optimum specifications for given conditions. The activities of the Interpretative Physics Section to be included within the Canadian Remote Sensing Centre will be most welcome.

REFERENCES

JONES, L.A. and H.R. CONDIT. 1948. Sunlight and skylight as determinants of photographic exposure. J. Opt. Soc. Am. 38(2): 123-178.

FRITZ, N.L. 1967. Optimum methods for using infrared sensitive color films. Photogramm. Eng. 33(10): 1128-1138.

_____ 1971. New color films for aerial photography. Proc. Bienn. Workshop Color Aerial Photogr. Plant Sci. Relat. Fields. March 2-4, 1971, Gainesville, Fla.

STEINER, D. and T. GUTERMANN. 1966. Russian data on spectral reflectance of vegetation, soil and rock types. Dep. Geogr. Univ. Zurich. 232 p.

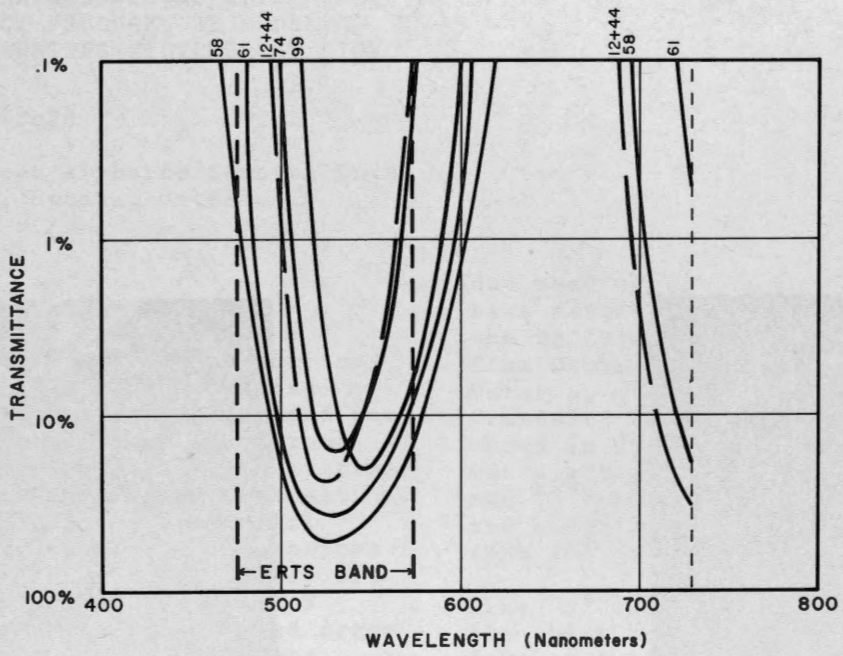


Figure 1. Transmittance of several Wratten band-pass filters within the sensitivity range of a panchromatic film with extended red.

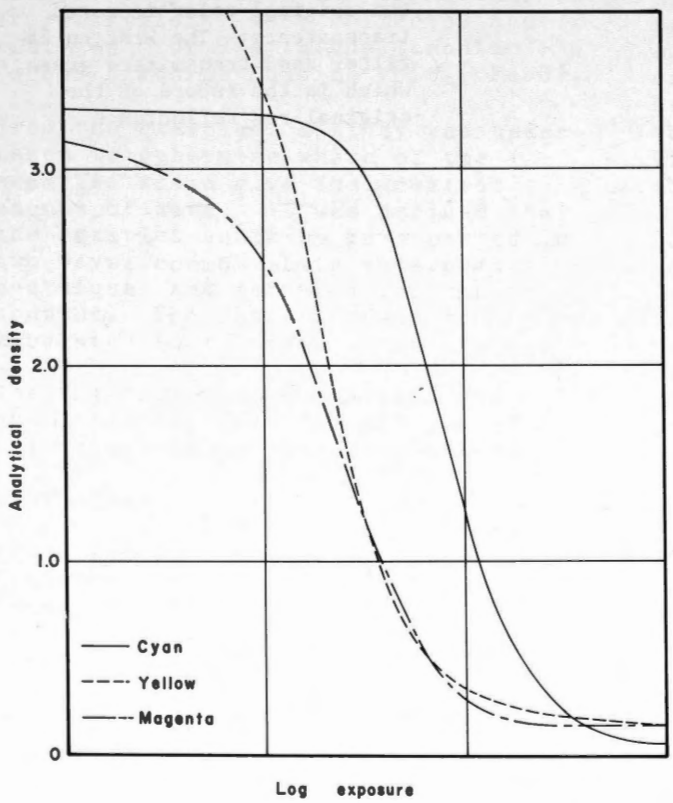


Figure 2. Sensitometric curves for Aerochrome Infrared Film 2443 (modified from Fritz, 1971).

Figure 3. A: Photograph obtained on panchromatic film (Tri-X, Type 5063) with a Wratten 25A filter (red transmitting).

B: Color separation obtained from the original color-infrared transparency. The Wratten 58 filter used transmitted green, which is the record of the original red reflectance.

A QUALITATIVE STUDY OF KODAK AEROCHROME INFRARED FILM, TYPE 2443, AND THE EFFECT PRODUCED BY KODAK COLOUR COMPENSATING FILTERS, AT HIGH ALTITUDES

Lt. R.D. Worsfold
RSEO
Canadian Forces Airborne Sensing Unit
CFB Uplands, Ottawa, Ontario

INTRODUCTION

AETE Uplands, Photo Development (now Canadian Forces Airborne Sensing Unit), on August 18, 1970 carried out film/filter tests using Aerochrome Infrared Film, Type 2443, at altitudes of 20,000 feet and 40,000 feet with Vinten 492 and Vinten 547 aerial reconnaissance cameras. The purpose of the test was to contribute film/filter samples to a catalogue of examples of forested land, land forms, geologic formations, agricultural land, and industrial land being compiled for mission planning. Utilizing the various film/filter combinations and evaluating each filter with respect to the film, selected positive transparencies, with seventy millimeter format and in stereotriplets, were collected and analyzed for use in the Canadian Forces Airborne Sensing Unit (CFASU).

From the examples, a study was undertaken to determine which of the examples would give the greatest amount of data. It was decided that the examples could be interpreted in two ways; common photo interpretation techniques and infrared information content. The results could be tabulated in relation to each other.

The film/filter combinations are available at CFASU for the use of investigators in mission planning.

TEST AREA

Three flight lines were flown near Kingston, Ontario, Canada (map sheet 31D). Line "A" followed a path from Simcoe Island located in the Lower Gap of the St. Lawrence River northwest over the Amherst Bar to the C.I.L. Ammonia plant east of Milhaven, Ontario. Line "B" began due west of Inverary at Loughborough Lake northwest over Sydenham Lake to the Holleford Crater. Line "C" ran from South Lake, south of Seeleys Bay, Ontario, due west over Leo Lake to Cranberry Lake. The flight lines are shown in Figure 1. Flight line "A" was eight miles long, "B" nine miles and "C" three miles. The cloud cover and cloud shadow combined was less than 10% on the day of the flight.

Line "A" includes Amherst Island and the shore near Milhaven, Ontario. Amherst Island is mainly glacial material underlain by Paleozoic limestone. The shoreline consists of rocky limestone outcrops, glacial gravels and sands deposited by offshore currents. Amherst Island is made up of mixed farms with a few areas of deciduous bush. The mainland area near Milhaven is similar to Amherst Island, low flat glacial material underlain by moderately flat Paleozoic limestones. The streams on the mainland have cut through the limestone forming small gorges up to thirty or forty feet in depth.

Line "B" extends from Loughborough Lake to the Holleford Crater passing from Paleozoic limestones covered with thin topsoil and glacial material to Precambrian material. The area near Loughborough Lake is characterized by limestone cliffs and glacial soils and areas of deciduous bush. Sydenham Lake in the centre of the flight line lies on the boundary between the Paleozoic and the Precambrian. Both deciduous and coniferous vegetation are found in the area. The area near the Holleford Crater also lies on the Paleozoic/Precambrian boundary and is famous for the Holleford Crater formed when a meteorite struck the earth over 600 million years ago during the Precambrian.

Line "C" over Leo Lake is an area of folded precambrian rocks of the Grenville Province. There are a variety of rock types and geologic structure in the area with pockets of glacial soils covered mainly with coniferous vegetation.

SENSOR SYSTEM

For purposes of the test, a CF100 aircraft serial number 100767 was used with seven Vinten aerial reconnaissance cameras. The gunbay contained three Vinten Type 547 cameras and one Vinten Type 492. The nose section of the aircraft contained two Vinten 547's and a Vinten 492. The shutter speed of the Vinten 547's was 1/2000 of a second and the shutter speed of the Vinten 492's was 1/800 of a second. The bay cameras were controlled by one intervalometer and the nose cameras a second. Therefore, bay cameras exposed simultaneously and the nose cameras exposed simultaneously but the two systems were not electrically synchronized so there was a slight difference in the principal point between nose and bay imagery.

The CF100 is a two seat jet aircraft. The cameras are located in unpressurized areas of the aircraft; therefore, they must be remotely controlled from the navigators section of the cockpit. The intervalometers are located on the forward navigation panel. Because some navigation aids had to be removed so the intervalometers could be used, a closed-circuit television system was installed inclined fifteen degrees forward from the vertical so accurate tracking could be carried out.

The original system design was for simulation of ERTS imagery by the bay cameras and for colour, false colour and track recovery films in the nose area. Although the system was designed for a special purpose all cameras were versatile. Instead of the standard three inch lens, six inch and one and three quarter inch lenses could be used and all lenses had screw on filter holders for any combination of filters. Exposure was controlled by shutter speed, f-stop and filtration. The films were equally exposed so the evaluation could be carried out by comparing the films to each other.

FILM/FILTER DATA

Kodak Type 2443 was the film used for the film/filter tests in the Kingston area. The film is sensitive to the visible and near infrared portions of the electromagnetic spectrum (Figure 2). The visible spectrum ranges from 400 nanometers to about 700 nanometers and the near infrared portion from about 700 nanometers to 900 nanometers.

Type 2443 is a special colour film that has three image layers sensitive to green, red, and infrared instead of the blue, green, and red of ordinary colour film. A yellow filter is used to "cut out" blue radiation because all three layers are sensitive to blue radiation (Figure 3). When the film is processed, a blue image is recorded for the green sensitive region, a green image for the red region and a red image for the infrared region. The resulting colours of blue, green and red are formed from the yellow, magenta, and cyan dyes with densities depending upon the proportions of green, red and infrared reflected by the subject.

Type 2443 film is normally used with a Wratten 12 (W12). The W12 prevents radiation shorter than 500 nanometers from being recorded. The sensitivity of the yellow forming layer is from 500 nanometers to 610 nanometers, the sensitivity of the magenta forming layer is from 500 nanometers to 690 nanometers, and the sensitivity of the cyan forming layer is from 500 nanometers to 900 nanometers (Figures 4 and 5).

Because the film is sensitive to aging and because of differences in emulsion batches it is often necessary to apply colour adjustments. To carry out the colour adjustments, colour compensating filters are used. The filters control portions of the spectrum by attenuating the electromagnetic radiation. Each filter controls one colour region, transmitting one or both of the other two colours.

The colour compensating (CC) filter absorption curves use diffuse density on the vertical scale and wavelength

on the horizontal scale. The diffuse density is the common logarith of the reciprocal of transmittance. The diffuse density scale is a linear scale and cannot be compared directly to the logarithmic transmittance scale, therefore, tabular data must be used in conjunction with absorption curves.

SENSOR PLANNING

Because of the flying costs and the flying time involved, it was decided that only one flight could be carried out. The CC20 series of filters was chosen because the peak diffuse densities and hence the transmittance minima were approximately one half the values of the CC50 series (greatest diffuse density). Each CC20 filter has its transmittance minimum at a certain wavelength. The tables correlate the diffuse density with transmittance and indicate what film layers will receive the least exposure when the transmittance curves of the CC filters are correlated with the spectral sensitivity curves of the film.

Once the wavelengths were characterized and the areas of interest chosen the mission was flown and the film rolls were tabulated as per Table 1.

ASSESSMENT

Two methods of assessment must be used to interpret properly the various film/filter combinations. The first were the standard photo interpretation methods of assessing the data as to visual information content. Because the tests were with false colour film, the second method was an assessment for infrared content. The method involved the study of the samples to ascertain the amount of information available in the recorded infrared region.

Results Using Standard Methods

Each frame on the test films were analyzed with respect to the major features in that area. The interpretation was made so as to determine what film/filter combination resulted in the best qualitative data with the intent to classify the combinations as to their relative usefulness according to standard interpretation techniques. The following results are the author's opinion based on his interpretation.

Amherst Island -- The water surrounding Amherst Island was analyzed with respect to the water pattern and water penetration. The combinations that gave the best results were the W12 and the W12-CC20M with the W12-CC20G and W12-CC20R also giving usable results. For agricultural and forested areas the W12-CC20M gave the best results and the W12 almost as good. The glacial patterns were best observed with the W12 and the W12-CC20M. The W12-CC20R and W12-CC20G showed better than average results. The beaches around the island were shown equally as well by the W12, W12-CC20M and W12-CC20R with the W12-CC20G also proving adequate. Overall the W12-CC20M gave the best results followed by the W12, with the W12-CC20G and W12-CC20R being almost the same. The W12-CC20B, W12-CC20C, and W12-CC20Y were dark and visual interpretation proved difficult to impossible.

Milhaven Industrial Complex (C.I.L. Plant) -- There was no stereotriplet with a W12 available for the area. The W12-CC20M and W12-CC20R showed the industrial areas the best with the W12-CC20B and W12-CC20C giving adequate results. Water pattern and water penetration were best observed with the W12-CC20M and W12-CC20R. Agricultural and forested areas showed best with the W12-CC20M with the W12-CC20R and W12-CC20G giving good results. The glacial pattern and drainage patterns were best shown with the W12-CC20M and the W12-CC20R. The W12-CC20M and W12-CC20R showed the best overall results for the area with the W12-CC20G and W12-CC20B giving better than average results.

Loughborough Lake -- The most important feature of the area was its geology. Strike and dip, outcrop, structure and pattern were most easily discerned with the W12-CC20M and W12-CC20B. The remaining filters gave adequate results except for the W12-CC20Y. With respect to water pattern and water penetration, the

W12-CC20M and the W12-CC20R were the best. Agriculture and forested areas showed best with the W12-CC20M; the W12-CC20R and W12-CC20G gave adequate results. The best overall results were with the W12-CC20M, followed by the W12-CC20R and W12-CC20G. There was no stereotriplet with a W12 available.

Sydenham Lake -- The geologic data mentioned above for Loughborough Lake showed best with the W12-CC20M, with good results also obtained with the W12, W12-CC20R and W12-CC20G. Water pattern and water penetration showed best with the W12-CC20M and W12-CC20R, with all other combinations showing reasonable results except the W12-CC20Y. The agricultural and forested areas also showed well with all combinations except the W12-CC20Y. Overall, the W12-CC20M gave the best results followed by the W12-CC20R and the W12.

Holleford Crater -- The geologic data was shown equally well by the W12, W12-CC20M and W12-CC20R, as were the water pattern, water penetration and the agricultural and forested areas. The W12-CC20M, W12-CC20R and W12 gave the best overall results, with the W12-CC20Y showing a complete loss of data.

Leo Lake -- The geologic data showed equally well with the W12-CC20M and W12-CC20R. Agricultural and forested areas showed the same as did water penetration and water pattern. The W12-CC20M gave the best overall results, and the W12-CC20Y showed the worst.

From the above results, the various film/filter combinations were classified according to Table 2 with the best combination at the top decending to the worst.

Results Using Infrared Content -- Using the second method the samples were analyzed for the infrared enhancement produced by the compensating filters. As with the first method each frame of each stereotriplet was assessed relative to each combination showing the same scene. It was not necessary to use the parameters of Method 1 as guidelines for Method 2. Based on the infrared assessment, Table 3 was established showing the order of infrared enhancement from the best to the worst. The results are valid for imagery at 40,000 feet and 20,000 feet.

DISCUSSION

Both methods can be used for information content. The first method can be used when it is necessary to obtain sharp results as well as infrared information from the same phtotgraphs. If infrared information is of prime importance and sharpness not the second method should be used.

Each table can be subdivided into:

a) useful film/filter combinations

b) combinations that do not improve upon the W12

c) those that are not usable.

From Table 2, the W12-CC20M, W12-CC20R and W12 are usable combinations. The W12, the standard filtration, gave the usual pinkish-red appearance indicating a high infrared return. The sharpness of the combination with the W12 must be considered normal because it is the "standard" for near infrared studies. The W12-CC20R gave sharper results than the W12 and the infrared content was almost as good. The W12-CC20M had the sharpest results and the infrared record was almost as good as the infrared record of the W12.

The second section from Table 2 was the W12-CC20G, W12-CC20B and W12-CC20C. The sharpness of these combinations was of a poorer quality than the first group. The third section was the W12-CC20Y combination. It was considered useless for interpretation.

The second method could also be divided into three sections. Section one was the W12-CC20B, W12-CC20C and W12 combinations. Again the W12 was taken as the standard. The infrared content was improved with the W12-CC20C and very much improved with the W12-CC20B.

Section two was the W12-CC20G, W12-CC20R and W12-CC20M combinations. The infrared content of these samples approached the infrared content of the W12-CC20Y combination, and as in the

first method no useful information was available.

From the above discussion, Figure 6 can be drawn. The length of the joining lines has no significant meaning but the off-centering indicates that the infrared content is much better than the sharpness for the W12, W12-CC20G and W12-CC20Y combinations.

The reasons for the differences in colour balance can best be understood from Figure 7. The characteristic curves represent a typical roll of Type 2443 film with a W12 filter. The red curve represents the infrared radiation, the green the red radiation and the blue the green radiation. The CC filters cannot affect the red curve because they do not attenuate infrared radiation; therefore, the colour shifts occur when the blue and green curves are shifted.

When the W12-CC20M combination is used the blue curve shifts changing the colour balance by removing the blue colour from the positive. With the W12-CC20R the green curve shifts removing the red radiation from the positive record. The W12-CC20B results in enhancement of the infrared radiation because the green and blue curves are both shifted resulting in decrease in exposure of the green and red radiation giving a relative increase in exposure of the infrared radiation. With the W12-CC20C the green curve shifts removing the red record from the positive giving a relative increase in exposure of the infrared and green radiation. In the case of the W12-CC20G the green curve shift slightly giving a small relative enhancement of the infrared region. The W12 does not change the colour balance and the W12-CC20Y does not cause a relative colour shift.

The results shown by the W12-CC20Y imagery indicate that that combination results in underexposure of film caused by the increase in yellow filtration when the yellow W12 and yellow colour compensating filter are combined.

CONCLUSIONS

It is apparent from the tests that the colour balance, can easily be changed by using colour compensating filters but the results are not always as useful as the "standard W12 combination". Therefore, unless the chosen combination improves upon "the standard" it is of no use. The W12-CC20B and W12-CC20M greatly improve upon the standard and the W12-CC20C and W12-CC20R moderately improve upon it while the W12-CC20G is poorer than the standard, and the W12-CC20Y causes underexposure.

It is concluded that the W12-CC20B is best suited for enhancement of the infrared, and W12-CC20M is best for sharpness and good infrared return combined. Both are improvements on the standard and neither obscures the recorded data.

Based on the tests it is recommended that if investigators only have one camera available to them they must decide between high infrared return or a combination of sharpness and good infrared record. If more than one similar format camera is available then both the W12-CC20B and W12-CC20M should be used. If more than one camera is available but they have different formats, the combination that is of most use to the investigator should be used with the larger format camera and the other combination with the smaller format camera.

It is recommended that the "standard W12" should be avoided in any study concerned with the near infrared because the resulting red overpowers the blue and green colours giving the whole scene a pinkish-red colour.

With many sensor configurations, four seventy-millimeter format cameras are available. Three of the cameras are used for what is defined as ERTS band photography. The choice of film and filter for the fourth camera is normally left to the investigator. Usually a colour film is used because the three ERTS bands can be combined using monochromatic light sources to form a colour image. Many investigators have made the mistake of using normal colour film with the fourth camera when Type 2443 with a W12-CC combination would give them a closer approximation to the combined ERTS bands. Therefore, it is

recommended that the fourth camera of any ERTS configuration should utilize false colour film with an appropriate colour compensating filter.

It must be noted, that the results of the tests are concerned with the CC20 series of compensating filters. The next stage of the experiment will be testing of the complete CC Magenta and Blue filter series to ascertain the differences in exposure that can result with CC10, CC20, CC30, CC40 and CC50 filters.

All results are the author's opinion based on his interpretation of the test film. Individual investigators will have to determine from their own needs what filtration is best suited for their studies.

TABLE 1

Unit Roll Number	NAPL Roll Number	Film	Filter
70/61-66	CP-1100 IR	Type 2443	W12
70/61-77	CP-1106 IR	Type 2443	W12-CC20M
70/61-78	CP-1107 IR	Type 2443	W12-CC20R
70/61-79	CP-1108 IR	Type 2443	W12-CC20B
70/61-81	CP-1110 IR	Type 2443	W12-CC20C
70/61-82	CP-1111 IR	Type 2443	W12-CC20Y
70/61-83	CP-1112 IR	Type 2443	W12-CC20G

TABLE 2

(Method 1)

1. W12-CC20M
2. W12-CC20R
3. W12
4. W12-CC20G
5. W12-CC20B
6. W12-CC20C
7. W12-CC20Y

TABLE 3

(Method 2)

1. W12-CC20B
2. W12-CC20C
3. W12
4. W12-CC20G
5. W12-CC20R
6. W12-CC20M
7. W12-CC20Y

REFERENCES

1) Optimum Methods for Using Infrared Sensitive Colour Film, Norman L. Fritz, Photogrammetric Engineering, p. 1128, Oct. 1967

2) Making Colour Infrared Film a More Effective High-Altitude Remote Sensor, Robt. W. Pease and Leonard W. Bowden, Remote Sensing of Environment, Vol. 1, No. 1, p. 23, March 1969

3) More Information Relating to the High-Altitude Use of Colour Infrared Film, Robt. W. Pease, Remote Sensing Environment, Vol. 1, No. 2, p. 123, March 1970

4) Color Aerial Photos in the Reconnaissance of Soils and Rocks, Abraham Anson, Photogrammetric Engineering, Vol. XXXVI, No. 4, p. 343, April 1970

5) Color Photos, Cotton Leaves and Soil Salinity, H.W. Gausman, W.A. Allen, R. Cardenas R.L. Bowen, Photogrammetric Engineering, Vol. XXXVI, No. 5, p. 454, May 1970

6) Color and IR Photos for Soils, A.D. Kuhl, Photogrammetric Engineering, Vol. XXXVI, No. 5, p. 475, May 1970

7) Colour Infrared Film as a Negative Material, Robt. W. Pease, Remote Sensing of Environment, Vol. 1, No. 4, p. 195, Dec. 1970

8) Applied Infrared Photography, Kodak Technical Publication, M-28, 1970

9) Kodak Filters for Scientific and Technical Users, Kodak Technical Publication, B-3, 1970

10) Photographic Remote Sensing H. Ross Jackson, Industrial Photography, p. 20, Aug. 1971

11) Kodak Data for Aerial Photography, Kodak Technical Publication, M-29, 1971

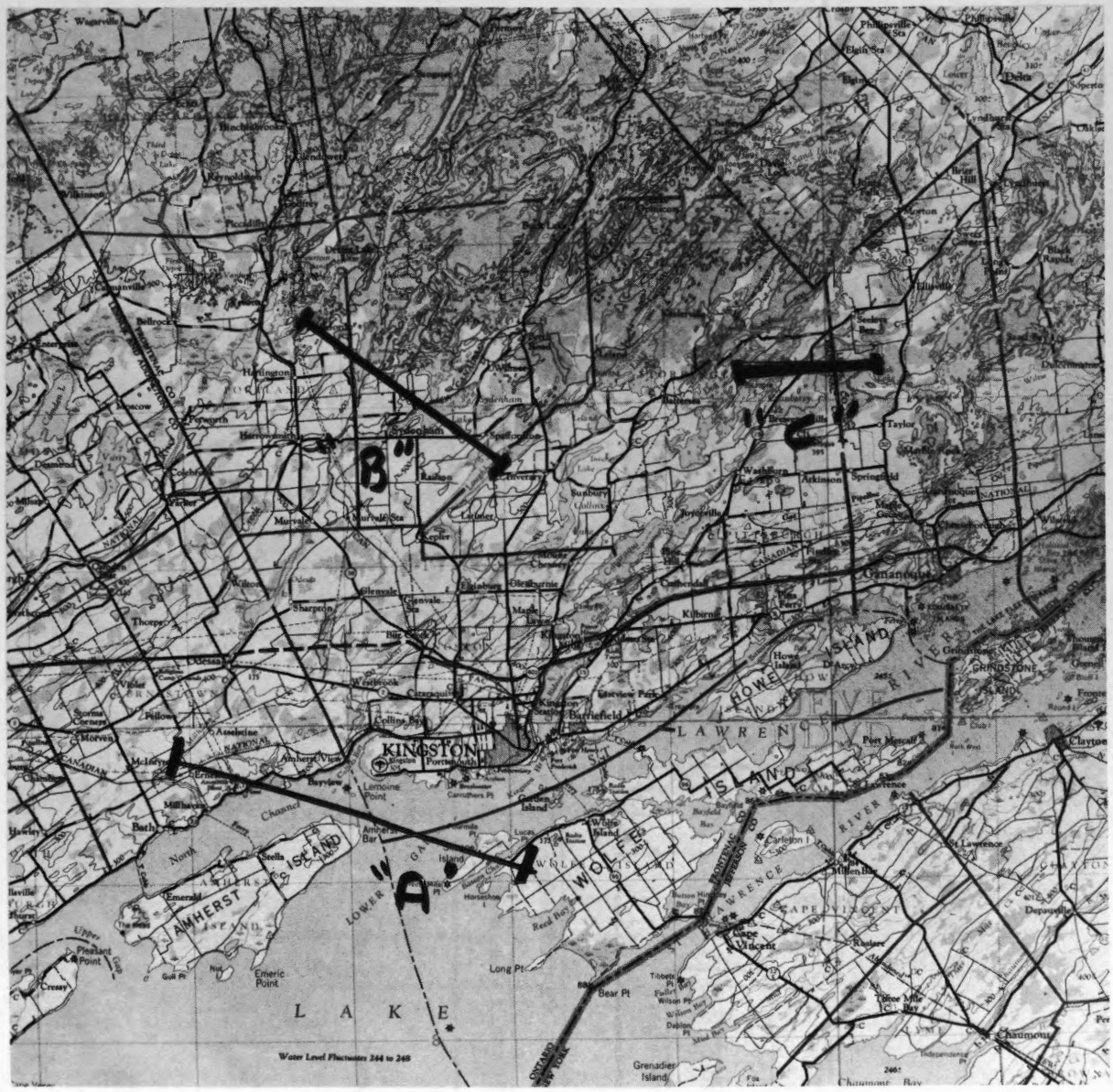


Figure 1. Flight lines (Map sheet 31 D Kingston)

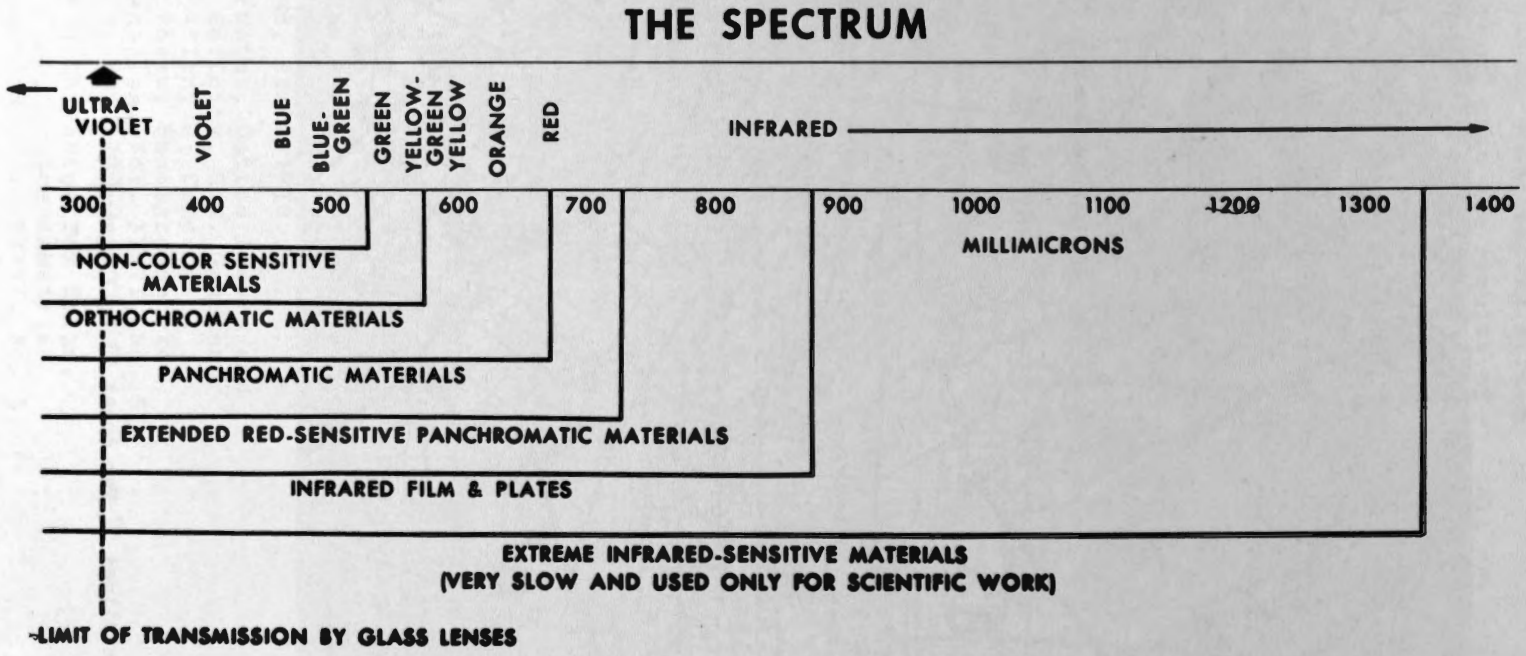


Figure 2. The photographically actinic regions of the electromagnetic spectrum.
(After Kodak Technical Publication M-28, Applied Infrared Photography, p2)

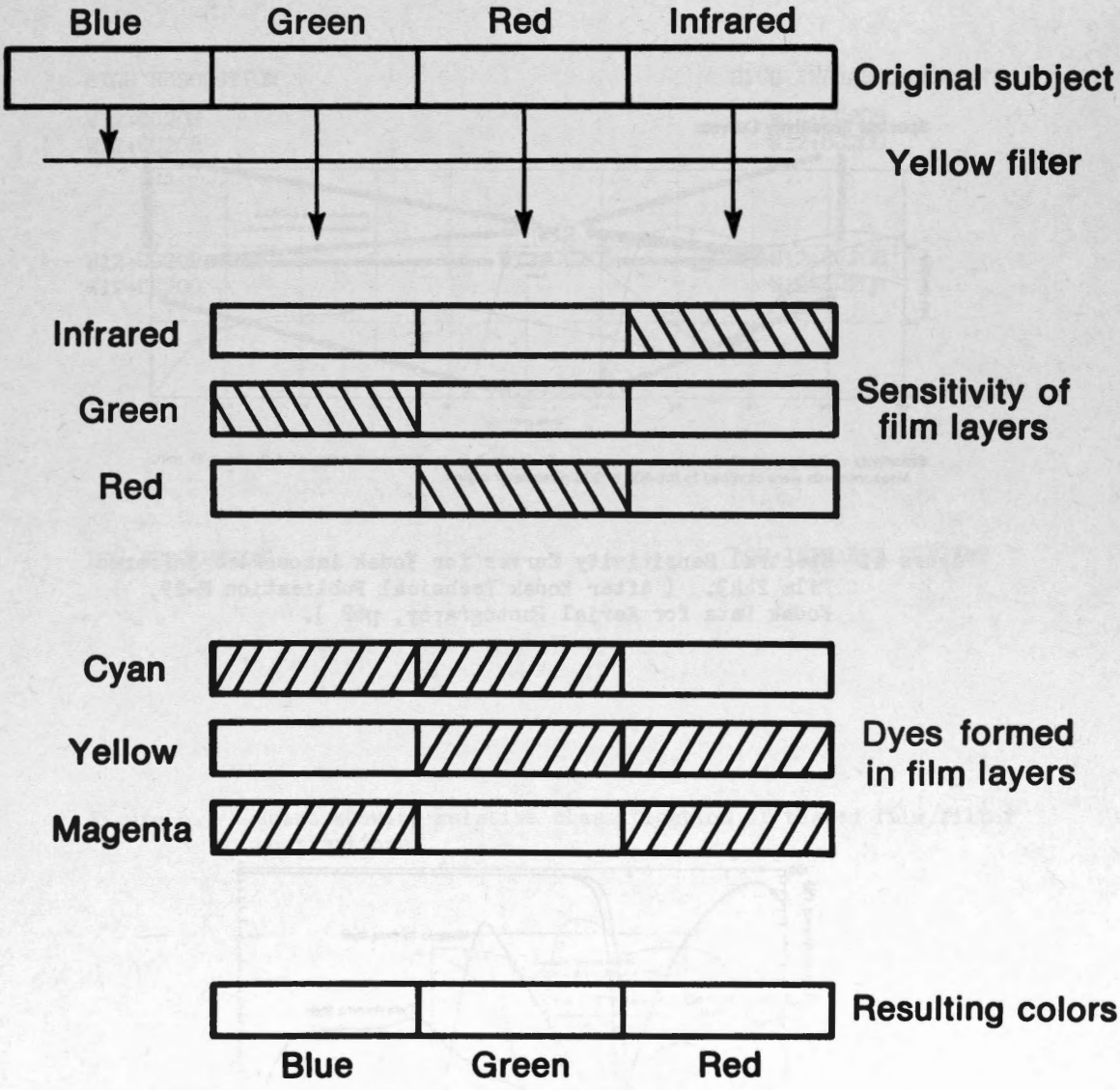
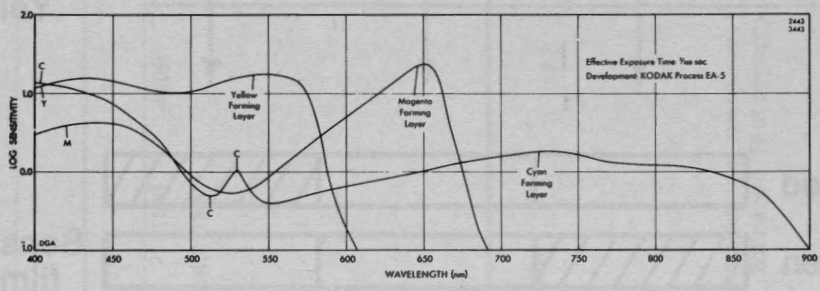


Figure 3. Colour formation with infrared colour film. (After Kodak Technical Publication M-28, Applied Infrared Photography, p5)

Spectral Sensitivity Curves:



Sensitivity = Reciprocal of the exposure (ergs/cm²) required to produce a density of 1.0 above D min. Measurements were confined to the 400 to 900 nanometer region.

Figure 4. Spectral Sensitivity Curves for Kodak Aerochrome Infrared Film 2443. (After Kodak Technical Publication M-29, Kodak Data for Aerial Photography, p62).

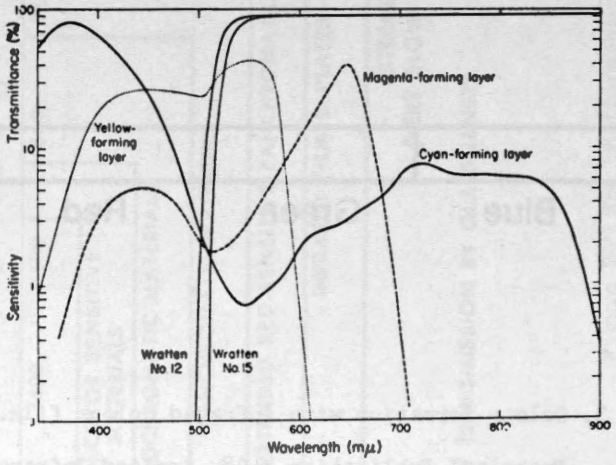


Figure 5. Spectral sensitivities of the three film layers. Spectral transmittance of Kodak Wratten filters as indicated. (After N.L. Fritz, Optimum Methods for Using Infrared Sensitive Colour Films, Photogrammetric Engineering , October 1967, p1128-1138).

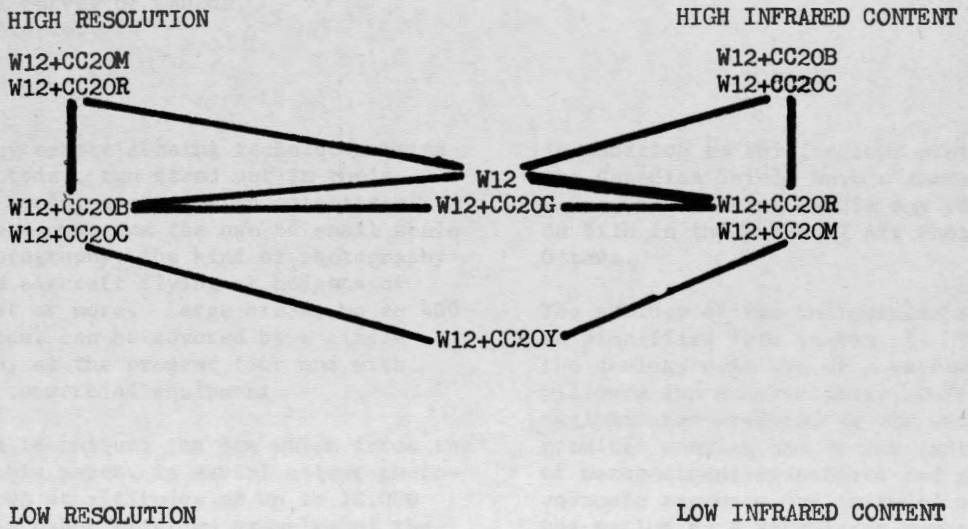


Figure 6. Chart showing relative classification of tested film/filter combinations

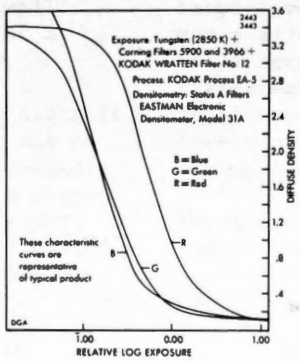
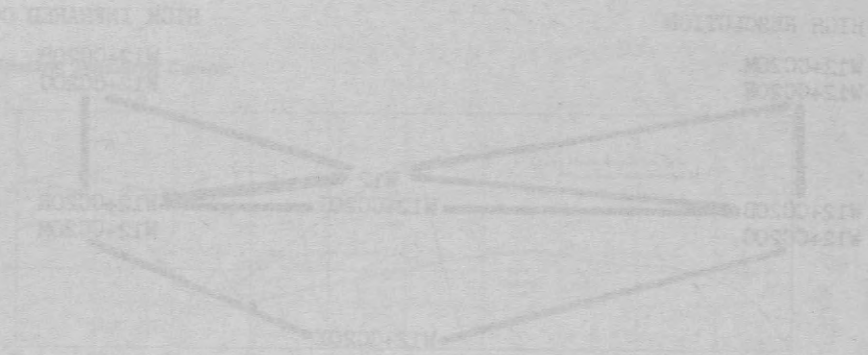


Figure 7. Characteristic Curves (After Kodak Technical Publication M-29, Kodak Data for Aerial Photography)



COLOUR PHOTOGRAPHY FROM THE YELLOWKNIFE DISTRICT

V. R. Slaney,
Geological Survey of Canada,
Ottawa, Ontario.

Of the many remote sensing techniques being practiced today, two stand out in their practical usefulness to earth scientists. One of these concerns the use of small scale aerial photography, the kind of photography taken from aircraft flying at heights of 40,000 feet or more. Large areas, up to 400 square miles, can be covered by a single photograph, at the present time and with available commercial equipment.

The second technique, the one which forms the basis of this paper, is aerial colour photography flown at altitudes of up to 10,000 feet. This paper provides examples of the kinds of information that colour photographs can show. The cost of obtaining aerial photography in Canada is also discussed.

Of course colour photography is not new. It is, however, a fast developing technology. Colour corrected lenses are becoming standard for all cartographic cameras. Faster films permit photography to be carried out at other than maximum sun angles and so the flying day is extended. And the introduction of automatic processing machines, and automatic dodging colour printers, will provide more even quality of output and hopefully will maintain the price of these services at a reasonable economic level.

In the summer of 1970, the GSC Skyvan aircraft flew a number of experimental surveys out of Fort Smith and Yellowknife in the Northwest Territories. Equipment consisted of a gamma-ray spectrometer, thermal infrared scanner, and a number of cameras which included a Wild RC8 cartographic camera. The colour photographs included in this paper were taken with this camera during the stay at Yellowknife.

Yellowknife is particularly suitable for photogeological studies because of the sparse nature of the soil and vegetation cover and because of the availability of geological maps of the area at scales up to 1:12,000 which provide very detailed ground control.

In addition to this, colour photographs of the Canadian Shield have a certain rarity value, there being hardly any other examples on file in the National Air Photo Library in Ottawa.

The geology of the Yellowknife area is shown in simplified form in Fig. 1. Very simply the geology consists of a succession of pillowed and massive lavas, tuffs and agglomerates bordered to the west by a granitic complex and to the east by a series of metasedimentary schists and gneisses. The volcanic sequence was intruded at more than one period by a very large number of metagabbro and metadiorite dykes, sills and irregular shaped bodies. A detailed account of the geology is contained in the publication by Henderson and Brown, 1966.

The photographs used in this study cover a rectangular area approximately fifteen miles long and 3.5 miles wide elongated along the north-south trending granodiorite-volcanic boundary. Photographs were taken from three heights, 3,000 feet, 6,000 feet and 7,500 feet above average ground level providing average photoscales of 1:6,000, 1:12,000 and 1:15,000 respectively. As mentioned above, the camera was a Wild RC8 (9" x 9" format) with a 6" focal length lens. The film used was Kodak Ektachrome MS type 2448 developed to a negative. From the negatives, colour transparencies were prepared for stereoscopic examination.

The negatives will shortly be handed over to the National Air Photo Library in Ottawa. Prints of the illustrations can then be ordered directly from the library. The reproductions shown here are not of the same quality as the original transparencies but are included to demonstrate the features as they are described in the text.

Figure 2 is taken four miles north of Yellowknife. The white, cream and pink colours represent the western granodiorite

* Presented at the Symposium under the title "Getting Better Photography for Geologists".

and the dark green and buff colours indicate the area underlain by volcanic rocks. Point 1 locates the West Bay fault which in this area has a sinistral displacement of about 3 miles. The Giant Yellowknife Mine at point 2 is one of the two gold mines still operating in this area. The margin of the granodiorite rocks is in places a fault and elsewhere is intrusive (point 3) into the volcanic assemblage. The granodiorite itself is variable in appearance. One form of the rock (5) is distinctly paler in colour and is ridged indicating the presence of a banded or strongly gneissic type of foliation. A second form of granodiorite (4) is often more pink in colour, does not have visible signs of foliation, and is evidently more resistant to erosion since it forms bosses of rather higher ground. The small nearly circular lake point 7, has a 300-foot wide rim of more resistant granodiorite. The origin of the lake has not been explained. A number of prominent gabbro dykes (6) are readily distinguished from the granodiorite even where individual outcrops are small as in the area immediately north of the circular lake. There is considerable iron staining in the host rock near the margins of the gabbro dykes and along the faulted edge of the granodiorite. This staining is not readily detectible on the black and white photography and because of the gradational nature of the staining is very difficult to map in the field. At point 8 a minor cross fault offsets a basic dyke. The banding in the volcanics shows particularly well at point 9 where cherty tuffs are interbedded with basalt flows.

Figure 3 is located immediately southeast of figure 2. The West Bay Fault (1) is again the dominent structural feature. The ore bodies of the Giant Yellowknife mine are located along shear planes (2) which can be recognized as areas of no outcrop. The sandy beach at 3 and the slime pond sediment at 4 are reminders of the ease with which non-vegetated surfaces can be recognized on colour photographs. The dark green volcanics west of West Bay Fault are lithologically identical to the buff coloured rocks east of the fault. The marked contrast in colour and resistance to erosion are due to differences in metamorphic grade. West of the fault the volcanics lie close to the intrusive margin of the granodiorite and are metamorphosed to epidote-amphibolite grade. East of the fault the volcanics lie about 3 miles from the eastward extension of the intrusive contact and are metamorphosed to green schist grade. The buff colouring of the volcanics is superficial but is enough to obscure many of the cherty tuff bands (5) and dacite dykes (6) that can be mapped so easily in the higher grade metamorphic rocks. Outcrops of porphyritic quartz dacite can usually be recognized by the distinctive pale yellow colour. Against the fault, however, the rock has a ferruginous stain and is difficult to separate from the volcanics. Quartzites of the Jackson formation (9) can be recognized by the scattered nature of the outcrop pattern by the fine laminations (bedding traces) visible on the outcrop and by the neutral grey colour of the rock. All of the rocks within a half mile radius of the Giant Yellowknife Mine are stained a darker shade than the rocks further away. The difference in colour is presumed to be due to the effect of gasses and particulate matter originating from the mine workings.

Figure 4 shows a portion of the volcanic sequence approximately two miles east of Ryan Lake at the northern limit of the photographed area. The main feature of interest here is the complex of quartz feldspar leucodacite dykes and irregular shaped bodies (1) which intrude the flows. Many shear zones (2) lie close to the strike of the volcanic layering. The volcanic formation contains numerous dykes and sills of metagabbro and metadiorite, most of which are practically identical in mineral content with the host rocks. The intrusions are very difficult to recognize in the field so it is hardly surprising that they can only rarely be recognized on photographs. Point 3 shows dark striped areas which represent metagabbro dykes. It has not yet been determined why the dykes can be recognized here and not elsewhere. Point 4 locates the workings of the abandoned Crestaurum Gold Mine. At point 5 there is an irregularly shaped outcrop of cream coloured cherty tuff which cannot be distinguished from the dacite intrusions.

Figure 5 is centered over Ryan Lake (1). Low domes of more massive and resistant pink-tinted granodiorite are shown at 2 and more foliated varieties at 3. At point 4 the rocks have a colour intermediate between that of the volcanics and that of the granodiorite. The published maps indicate that the area is predominantly granodiorite. A close inspection of the transparency shows many inclusions of volcanic origin. The rock here is clearly of mixed origin and quite different to the granodiorite west of Ryan Lake. The lesson to be learned is that it may be easier to map gradational boundaries on photographs than in the field. Point 5 indicates large diabase gabbro dykes clearly recognizable within the granodiorite because of the colour

difference and the wall rock staining. Within the volcanics, the gabbro can often be recognized by its red-brown tink and by the occurrence of scattered cream-coloured patches of lichen. At point 6 a pair of gabbro dykes each two to four feet wide can be recognized quite easily because of the contrast in colour with the granodiorite. Such dykes cannot be recognized where they intrude volcanic rocks.

Figure 6 is taken from a good quality 1:20,000 scale black and white negative from the National Air Photo Library and covers almost the same area as figure 3. The West Bay Fault, point 1, can be seen just as clearly here as on the colour photographs. The granodiorite too has a distinctive high reflectance and can be separated easily from the volcanics. Colour changes within the granodiorite cannot be recognized, nor can the ferruginous staining associated with the West Bay Fault (2) or with the margins of the diabase gabbro dykes (3). The gabbro dykes can be recognized by their lower reflectance and by the presence of cross joints but not nearly so surely as with the colour photographs. When outcrops are small and separated, recognition becomes uncertain. Point 4 indicates the area of darker rocks around the mine. The darkening effect on panchromatic film is rather subtle and would probably be missed on black and white film alone.

Figure 7 is centered on Lake Ryan indicated by point 1. North of Ryan Lake the granodiorite-volcanic boundary is clearly defined (2). South of the lake the boundary is far more difficult to recognize (3) where the difference in reflectance between granodiorite and volcanic material is obscured by a thin soil cover. Point 4 indicates an area with well defined contact against the volcanics and with a reflectance less than but fairly close to that of the granodiorite. From the panchromatic photography this area would probably be interpreted as granodiorite. From the colour photography (figure 5) the area can be recognized as a mixture of granodiorite and volcanic material. At point 5 a large diabase gabbro dyke is easily recognized by its darker tone and by the prominent jointing. At point 6 near the edge of the granodiorite intrusion, the dyke is reduced by local fracturing to a series of scattered outcrops. The cross jointing is no longer obvious and the normally low reflectance of the rock is obscured by soil so that the location of the dyke here is very difficult to determine. At points 7 and 8 diabase gabbro dykes lie in volcanic rocks and are difficult to recognize on this photograph. On the colour photograph of this area (figure 5) both dykes can be recognized.

ROCK RECOGNITION BY COLOUR

On fresh surfaces, Yellowknife rocks have a variety of colours ranging from white through pinks, browns and greens to black. Weathering effects are largely superficial in character but when combined with surface dust and rock fragments and organic material, particularly lichens and mosses, the range of rock colour exposed to the aerial camera is reduced to a significant degree. In turn, the range of colour actually recorded on aerial film depends not only on the camera system (lens plus filter plus film plus development) but also on the subject view angle, the elevation and azimuth of the sun and on the nature and path length of the atmosphere at the moment of exposure.

The effect of these many variables is that relative differences on the aerial photograph assume greater significance than absolute colour differences particularly with photographs taken at altitudes close to or higher than 10,000 feet. For this reason, Figure 8 is presented in tabular form indicating which of the rocks present in the area photographed can be distinguished from the surrounding lithology by differences in colour.

The remarkable feature of this table is the large number of rock units which in fact can be recognized in this way. The X's or failures in the main occur when gabbro intrudes gabbro, a difficult subject to map even in the field.

THE COST OF AERIAL PHOTOGRAPHY

Once the decision has been made to obtain new photography of an area, the question of the cost of colour relative to that of panchromatic photography becomes important. Comparative cost studies should be made regularly to monitor the advances being made in rapid processing and printing techniques. An exercise of this kind was undertaken in 1971 to cost colour and panchromatic photography for a reasonably large area, in this case the area covered by a 1:250,000 map sheet in the Yellowknife district of the Northwest Territories. This represents an area of 70 x 70 miles, or 4,900 square miles. A 9 x 9" format cartographic camera would be flown at 10,000 feet above average ground level

providing an average negative scale of 1:20,000. A flying height of 10,000 feet was chosen because this represents a reasonable upper limit for flying optimum quality colour. Unless the atmosphere is exceptionally clear, moisture and particulates in the atmosphere usually affect the colour balance of colour film at altitudes above 10,000 feet. The effect of light scattering by the atmosphere is to produce a yellow tint on the colour negative (blue on the print). This can be filtered out at the printing stage but only by reducing the range and brightness of colours reproducible by the print or transparency.

The main item is the cost of flying which is the same for colour as for monochrome photography. The cost of flying is estimated at $11.00 per line mile, a figure taken from the records of the Interdepartmental Committee on Air Surveys (ICAS), the group responsible for most of the Federal air survey contracts. Line mile costs could be higher for smaller projects or projects located in less accessible areas or for areas with less favourable weather conditions. Basic cost of the film, film development and reproduction is three to five times that for panchromatic material. For this exercise, the photography is needed for interpretation. The output is, in the case of panchromatic film, two sets of black and white paper prints. In the case of colour photography, the output is one set of colour transparencies for analysis in an office and one set of black and white paper prints (printed from the colour negatives) for use in the field. The result of the analysis is shown in Figure 9 which indicates that in these particular circumstances it would be entirely reasonable for colour photography to cost 34% more than panchromatic photography.

It is worthwhile comparing the cost of aerial photography with the cost of flying other airborne geophysical methods. Table I lists the average cost per line mile of a number of established airborne geophysical systems. Most of the survey systems quoted here are flown at low altitudes along closely spaced flight lines. Two to four line miles of survey are needed to provide adequate coverage for one square mile. Whichever way the figures are arranged, the cost of new photography is low compared with any other type of geophysical survey. One would not suggest that photography should or indeed could replace any of these geophysical methods but for the cost involved relative to the amount of information obtained it is surely not unreasonable for a geologist or mining engineer to insist on the provision of suitable aerial photographs to back up geological and airborne geophysical surveys.

CONCLUSIONS

1. Nine by nine inch format colour photography is in information value the cheapest and most effective remote sensing system available today.

2. By using colour negative film, colour transparencies may be printed for analysis in the office and black and white paper prints produced for use in the field.

3. With some notable exceptions, most of the features recognized on colour film can also be recognized on black and white photographs. Even so, coloured film is preferred because it permits the interpreter to recognize and understand information more rapidly and with greater confidence.

4. The cost of colour is still significantly greater than the cost of black and white photography. A particularly strong case for colour photography can be advanced in areas where there is a high percentage of outcrop and where photo scales of 1:20,000 or larger are acceptable.

5. Research work must be undertaken to improve the quality of colour photography taken at heights above 10,000 feet where moisture and particulates in the atmosphere affect the colour balance of aerial film.

REFERENCES

Henderson, J.F. and I.C. Brown:
Geology and Structure of the Yellowknife Greenstone Belt, District of Mackenzie. Geological Survey of Canada, Bulletin 141, 1966.

Henderson, J. B.
Stratigraphy of the Archaean Yellowknife Supergroup, Yellowknife Bay - Prosperous Bay Area, District of Mackenzie. Geological Survey of Canada, Paper 70-26, 1970.

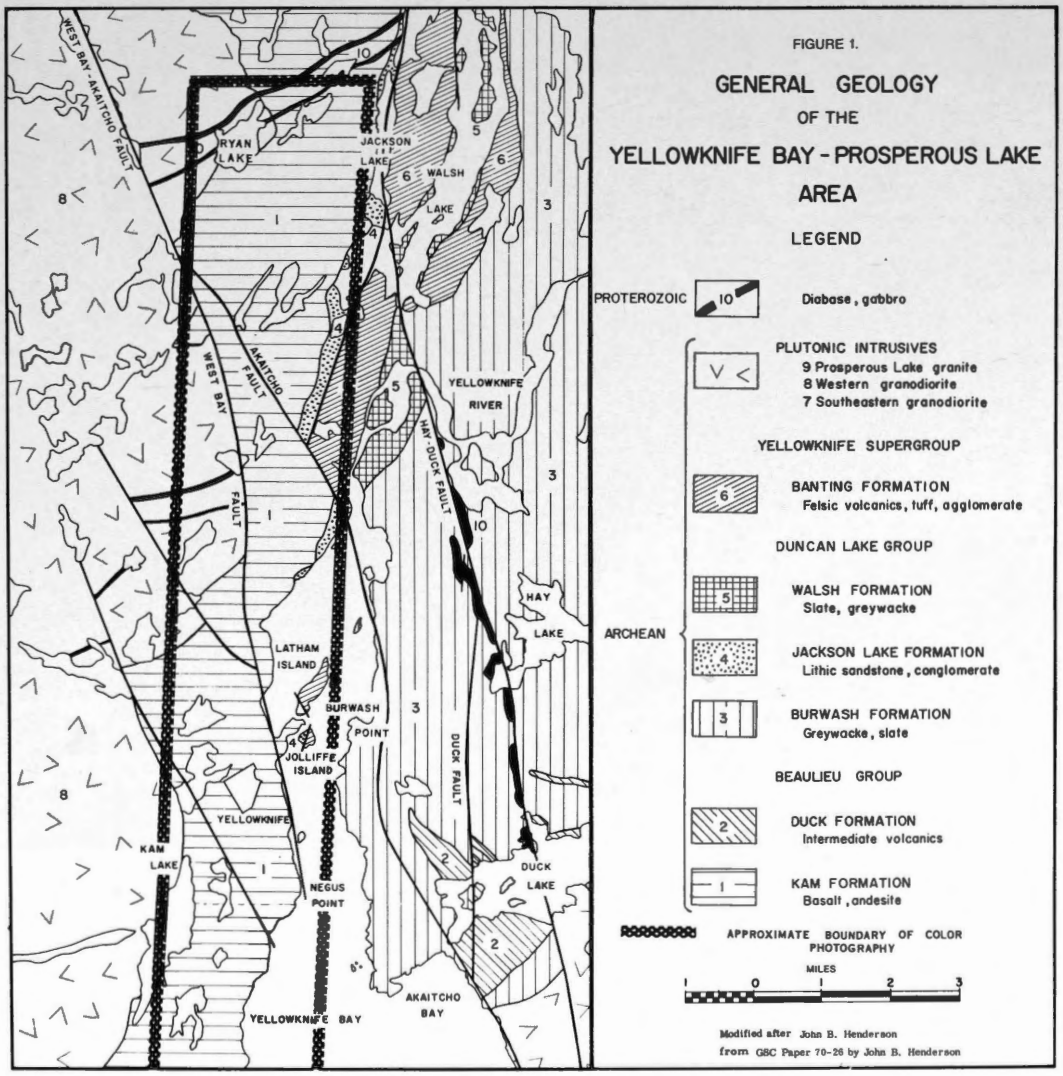
Allen, S. J.
Geophysical Activity in 1969. Geophysics, Vol. 36, No. 1., February 1971.

TABLE 1. COST OF AIRBORNE GEOPHYSICAL SURVEYS IN CANADA

E.M.	*	$13.90/line mile
Combined E.M./Mag.	*	17.87 "
Magnetics	*	6.34 "
Radioactivity	*	7.15 "
AFMAG	*	15.00 "
High Sensitivity Gamma Spectrometer	#	17.50 "
Monochrome Photography		5.18/square mile
Colour Photography		6.86/square mile

*Taken from Geophysical Activity in 1969, S. J. Allen, Geophysics, Vol. 36, No.1. Feb. 1971.

#Pers. Comm.



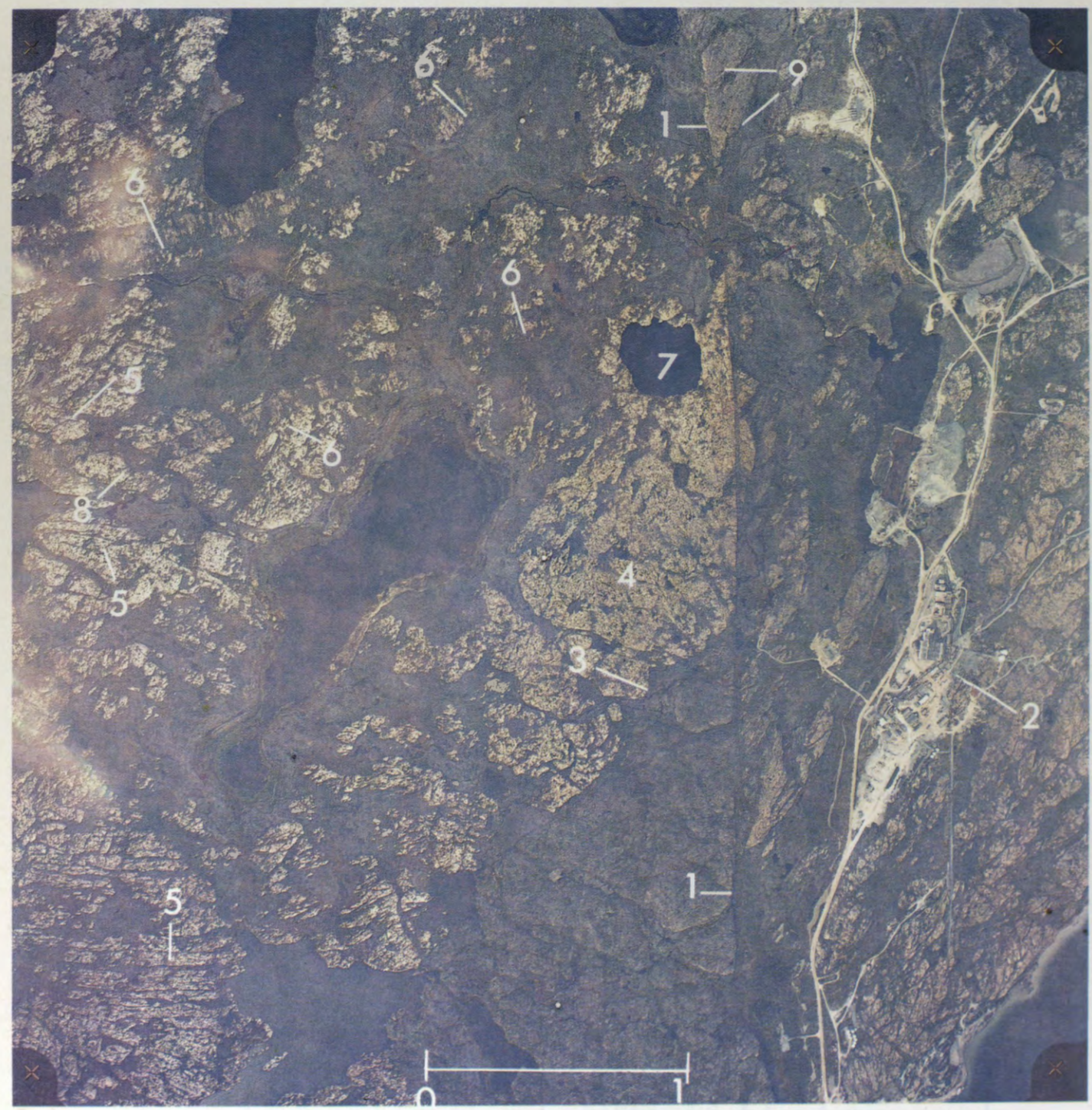


Figure 2.

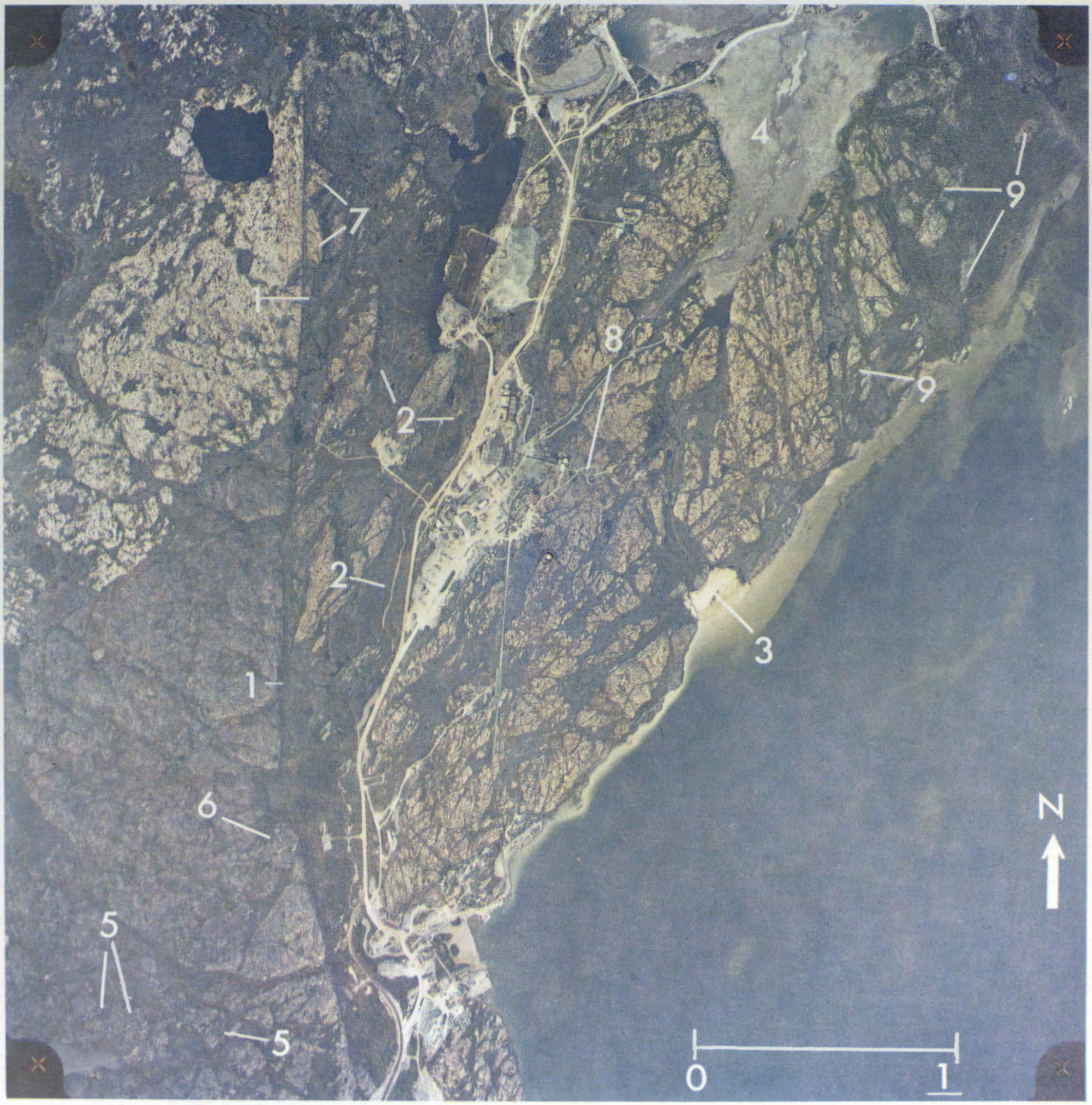


Figure 3.

Figure 4.

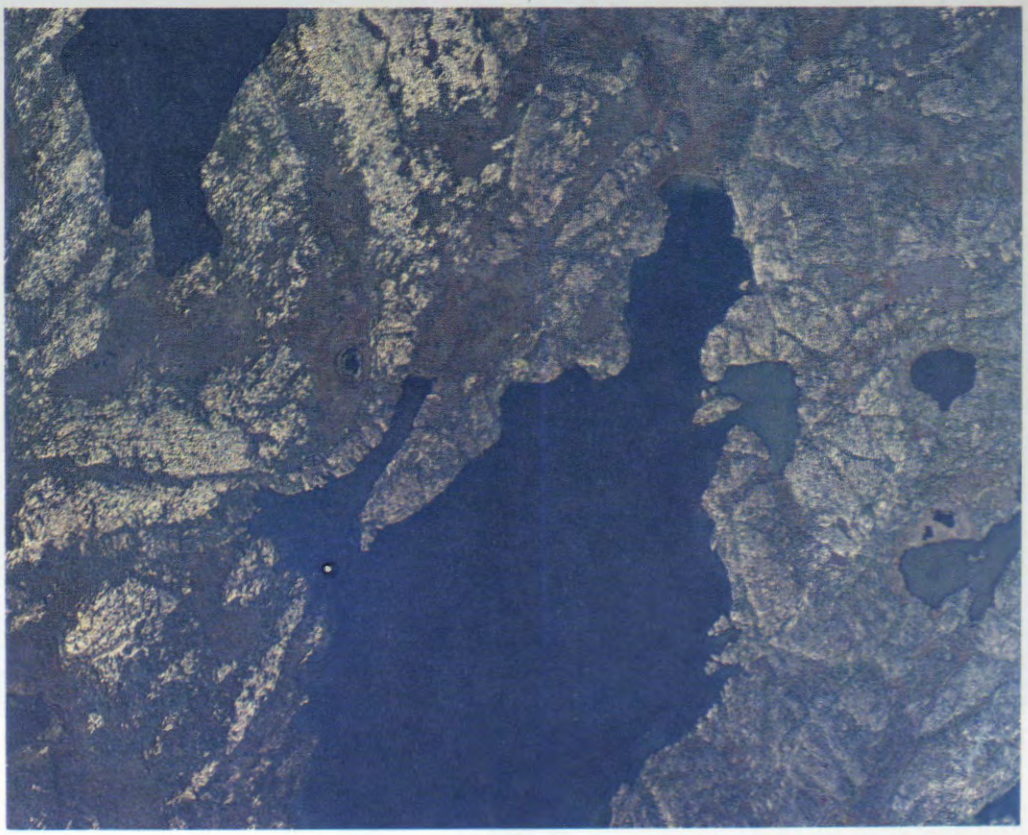


Figure 5.

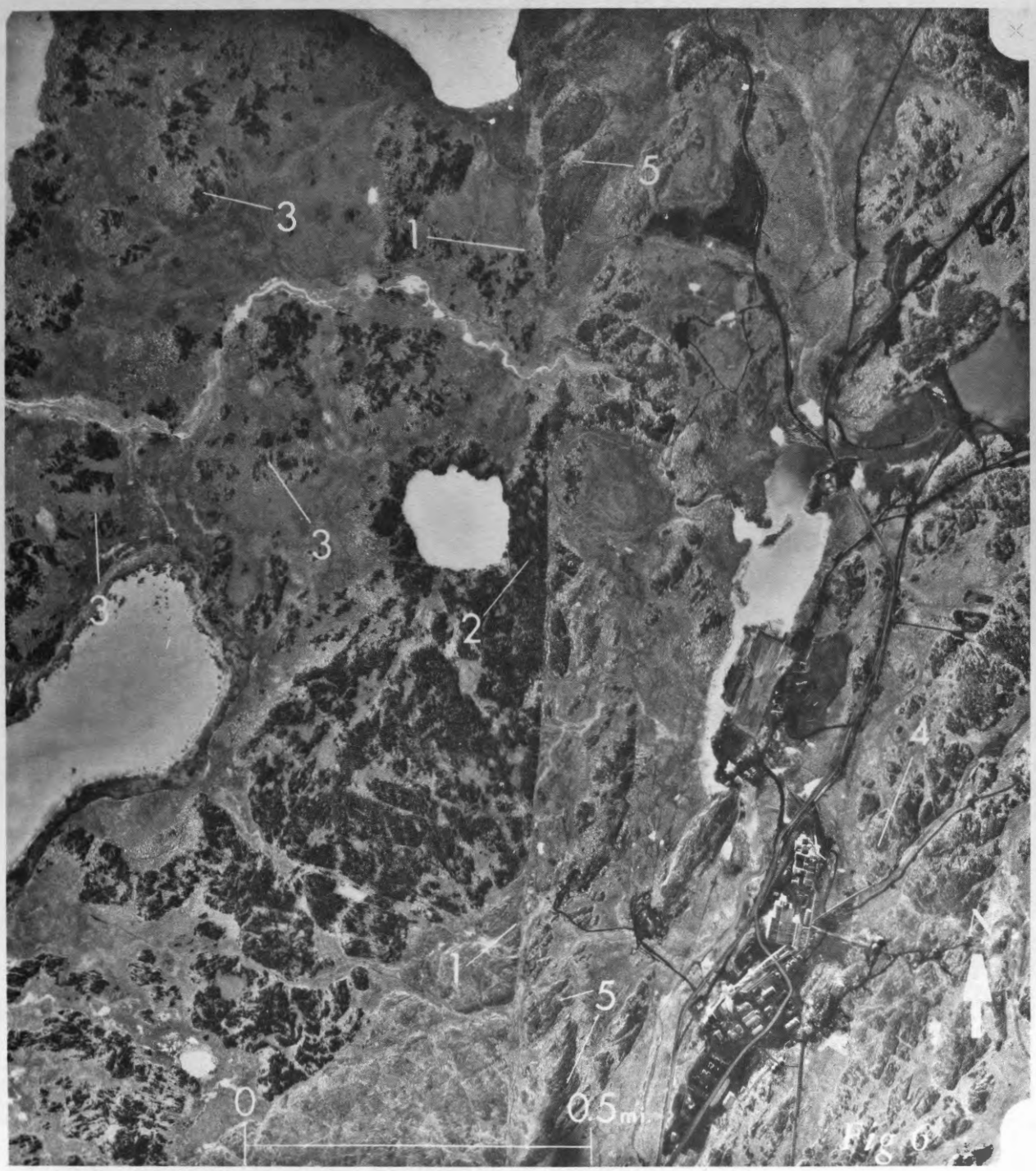


Figure 6.

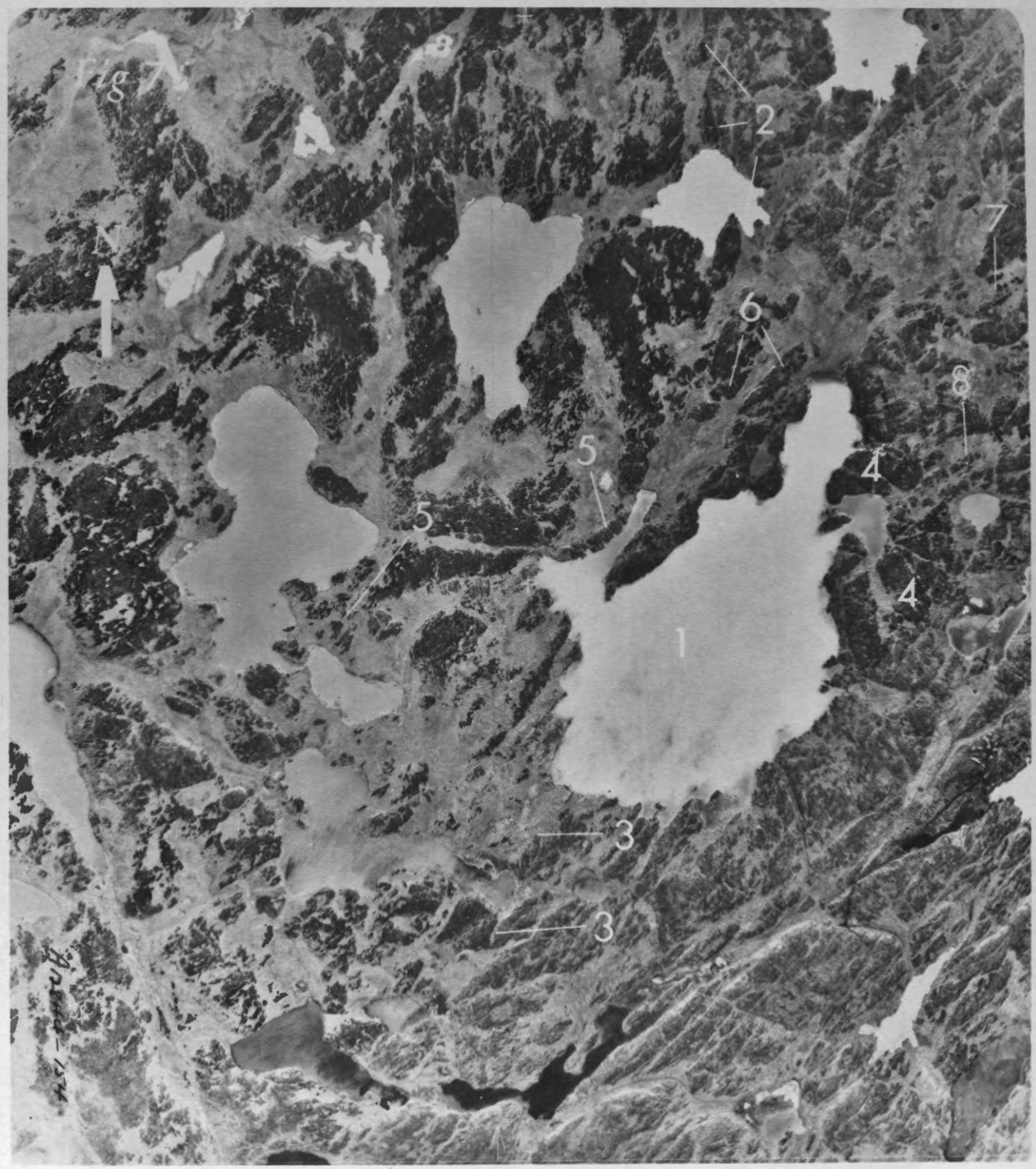
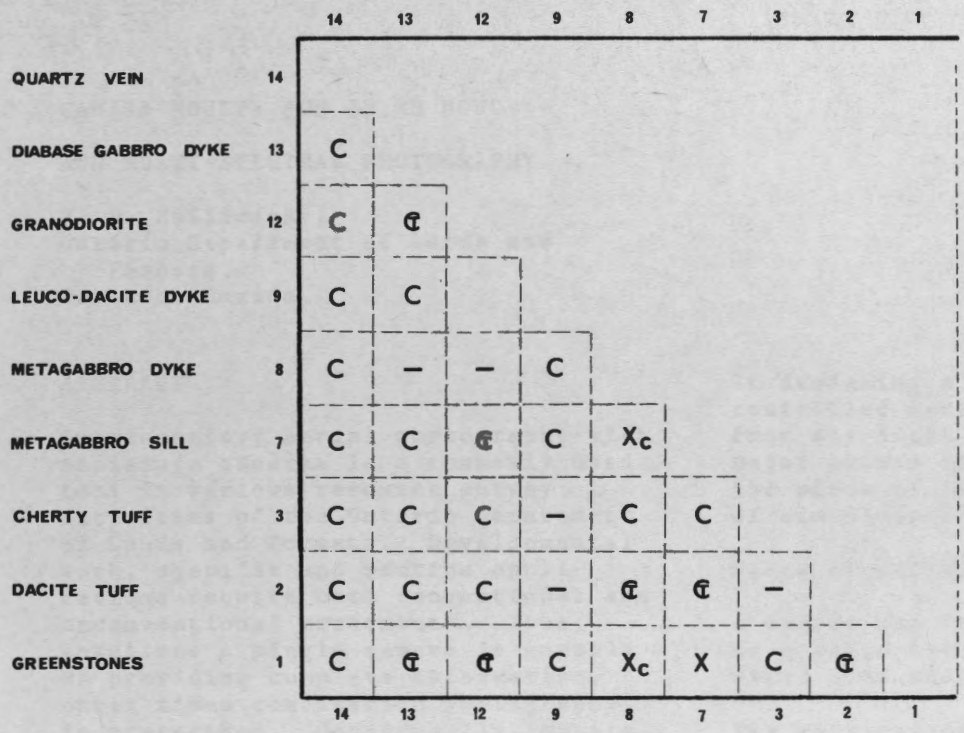


Figure 7.

YELLOWKNIFE COLOUR PHOTOGRAPHY

FIGURE 8



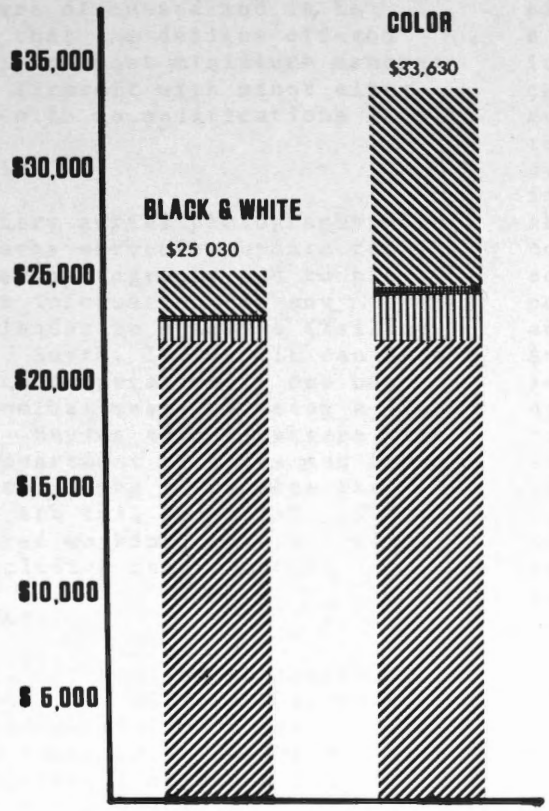
ROCK RECOGNITION

COLOUR	C
TOPOGRAPHY	T
ROCKS NOT SEPARATED	X
ROCKS NOT ASSOCIATED	–
COLOUR DIFFERENCE IN LOWER METAMORPHIC GRADE ROCKS	X_C

GROUND RESOLUTION
1:15,000 SCALE PHOTOGRAPHY
HIGH CONTRAST < 2 FEET
LOW CONTRAST 2-4 FEET

FIGURE 9

THE COST OF AERIAL PHOTOGRAPHY



AREA = 70 X 70 miles
= 4900 square miles
PHOTO SCALE 1:20,000
NO. OF LINE MILES = 2170
NO. OF PHOTOS = 1922

B.+W. PHOTOGRAPHY = $ 5.18 sq. mile
(2 sets paper prints)

COLOUR PHOTOGRAPHY = $ 6.86 sq. mile
(1 set Colour transparency)
(1 set B.+W. paper prints)

Colour costs 34 per cent more than B.+W.

CAMERA MOUNTS FOR 35 MM MONO- AND MULTI-SPECTRAL PHOTOGRAPHY

V. G. Zsilinszky,
Ontario Department of Lands and Forests,
Toronto, Canada.

ABSTRACT

Supplementary aerial photography with miniature cameras is a commonly used tool in various resource survey activities of the Ontario Department of Lands and Forests. Developmental work, specific and routine applications require both conventional and unconventional treatments. Thus, sometimes a single camera is capable of providing complete information, other times combination photography is prescribed. Consequently, mounts of various capacity have been engineered for single, two, three and four cameras. Each working model is described and illustrated, principles involved are discussed and it is suggested that the designs offered could apply to most miniature cameras and light aircraft with minor alterations or with no modifications at all.

INTRODUCTION

Supplementary aerial photography with 35 mm cameras serves to update regular survey photography and to provide quick information for any specific landscape analysis (Zsilinszky 1969, Smyth, 1972). It can be a very low cost operation if one has some economical means of using a light aircraft. Having this advantage, the Ontario Department of Lands and Forests is attempting to develop the system to its full potential. This has required working out numerous details, including camera mounts.

REQUIREMENTS

To accommodate varying photographic requirements for different purposes, single- and multi-camera installations have been designed. Thus both routine and developmental missions can be specified and handled.

In designing a 35 mm camera mount for controlled vertical aerial photography from any light aircraft, there are two major points to consider: these are the place of installation and the mode of stabilization.

Place of Installation

A camera may be installed *internally* in a cargo hatch or *externally* provided some manual control is possible.

The working condition of an aerial photographer due to atmospheric turbulence can be difficult at any time. Therefore, it is important that the operation of the camera(s) be as simple and secure as possible. Unless a fixed stereo-base is required, the internal installation, which is discussed below, is advantageous because even if the operation is controllable remotely, manual access to the camera system is necessary for changing lens, filter, film magazine, aperture, shutter and camera position and for correcting mechanical failure. In addition, the cargo hatch is usually positioned so that the exhaust heat and fumes are cleared away from it. And, because of the superwide-angle lenses, the cameras must be lowered down to the skin of the aircraft. (see Figure 1). It should be noted that aircraft with no cargo-hatch can be equipped with this feature at a relatively low cost. However, any modification of this nature must have the approval of the local transport authority.

Mode of Stabilization

The stabilization of a camera system in a cargo-hatch requires a base-plate and a camera-stand.

It is suggested that the base-plate be

transparent to provide visibility for exposure interval determination and for some visual control for the camera operator for navigation. The base-plate should also be rotatable to set the camera for either transverse or longitudinal position and to compensate for crab. And finally, it should have strength to carry the maximum potential load concentrated at the bearing points of the camera-stand levelling screws.

The camera-stand, supported by three levelling screws with shock absorbers, must be firm but light, and should leave the camera accessible for operational changes while mounted. The multi-camera stands must have some means of calibrating the alignment of cameras to ensure parallel optical axes. The stand with mounted camera(s) should be quickly detachable from the shock absorbers to facilitate protection of the camera during take-off and landing, and any necessary reloading or emergency repairs.

DESIGNS AND CONSTRUCTIONS

Basic factors influencing camera mount (base-plate and camera-stand) design include size of hatch in the aircraft and weight of camera assembly. Hatch size determines the maximum number of cameras possible; assembly weight determines gauge of base-plate.

Designs and dimensions discussed in this paper relate to the deHavilland Turbo-Beaver and Otter aircraft and the Nikon F 250 Motor Drive camera system. However, the principles of design relate equally to any aircraft and camera system which might be chosen. Whilst this paper does not provide blueprints, and data and specific dimensions are given only for the sake of clarity, actual blueprints may be made available on request.

Base-Plates

The base-plates for all single and multi-camera models are fundamentally the same, although they vary in *plate thickness*, number of ports and diameter, except for the four-camera mount which is also of larger diameter. Figure 2.

The maximum weight of a single F-250 camera unit is 2.75 kg (6 lb.). This weight and that of the camera stand are comfortably carried by an 8 mm (5/16-inch) thick plexi-glass plate at a hatch size of 46 cm (18 inches) in the Turbo-Beaver. Both the two and three-camera arrangements require a 10 mm (3/8-inch) thick plate.

Four F-250 cameras cannot be accommodated in a 46 cm (18-inch) hatch. The plate with the four holes illustrated in Figure 2 refers to the 68 cm (27-inch) hatch size of the Otter aircraft. This base-plate should be 13 mm (1/2-inch) thick. The smaller plates can also fit the Otter hatch with an adapter ring.

All models have a *metal ring* that fits exactly in the hatch and is held down by locking turn button devices. This ring accommodates a simple ball-bearing system on which the plexi-glass plate can be fully rotated and locked in any position for crab compensation and for choice of camera position (transverse or longitudinal). In the three-camera mount the shock absorbers are pushed to the periphery of the base-plate to extend the 46 cm (18-inch) hole to 50 cm (20 inches). This creates little difficulty, because there is adequate margin around the cargo-hatch to allow the plexi-glass to overhang. However, the high turn buttons in the hatch (standard in aircraft) have to be replaced by flat plate type fasteners whose bolts can be tightened through a hole drilled in the rotatable plexi-glass plate.

The three *shock absorbers* to minimize the transmitted vibration of the aircraft are located on the plexi-glass plate at the vertices of an equilateral triangle. They can be made up from soft rubber but manufactured absorbers are preferable. These are commonly used in mounting aircraft instrument panels and are available with different load ratings. The head is indented to provide firm seating for the cone-shaped levelling screw end, as illustrated in Figure 3.

Also, the base-plate is marked for the forward overlap controlling device which has been discussed by Zsilinszky (1969).

Camera-Stands

All models of camera-stands are constructed of metal (aluminum, brass or stainless steel). Cameras may be either *suspended* or *supported* within the stand.

Single-camera model - This design is based on the principle of a three point suspension. Two points are provided by the strap ring, while the third one is located at the tripod socket. See Figure 4a. Two L-shaped and threaded members are hollowed out to accommodate the strap rings protruding from the top of the camera (Figure 4b). These members hold the camera firmly, when it is pressed against them. Suitable pressure is provided at the tripod socket at the base of the camera by an adjustable and lockable clamp (Figure 4c). When the three points are firmly in position and secured by the clamp, the L-shaped members are tightened on the stand by locking collars. This operation of clamping the camera to the stand is done on the lap of the photographer and after take-off the assembly is placed on the base-plate.

Two-camera model - When two or more cameras are used for simultaneous exposures, it is essential that there is some means of calibrating the stand for parallel optical axes. It is most desirable that all film rolls run in the same direction, and that either camera may be picked up individually at any time during operation. Also, it may be required that one camera be raised or lowered while the other is stationary, if different focal length lenses are to be used. This allows any lens to be held flush with the skin of the aircraft at any time. A suitable stand, incorporating cradles for holding two individual cameras, is illustrated in Figure 5. The camera stand comprises two vertical plates, on to each of which a pair of cradles is fixed. The cameras are supported securely in the pairs of cradles. One of the double plates slides up or down on the other, governed by the two adjusting screws. The calibration of a multi-camera stand can be arranged by leaving one pair of cradles stationary with the other pair(s) adjustable. Adjusting is done by a slot device as required according to Figure 6. Rotatability around the vertical axis is not essential, because the parallel position of the relatively long cameras will inevitably minimize alignment discrepancies. The calibration is only a simple laboratory exercise. This camera-stand can be on the base-plate during take-off and landing, if the lens holes are covered for these critical periods. After take-off the cameras may be placed in the cradles, but will remain completely free for removal when necessary.

Three-camera model - This stand also provides supporting cradles for the cameras. However, it omits the option of the raising-lowering feature. Calibration device is included in two pairs of cradles (see Figure 7).

Four-camera model - The four-camera stand is designed for the 68 cm (27-inch) size hatch of the Otter aircraft. Structurally the stand is practically identical to the three-camera stand (compare Figures 7 and 8).

CONCLUSION

The designs discussed here are the outcome of experience - including both success and failure - in the operation of the 35 mm aerial photography. They are by no means unimprovable, but three are already fully operational. The single-, two- and four-camera mounts have been in use for some time. The three-camera version has just left the drawing board.

In practice, the single-camera mount is popular in routine photography (cutover or road mapping, delineation of fume, spray, fire or other damage and survey point identification) using the combination of a wide angle lens with black-white film. The multi-camera mounts are particularly useful in developmental work and in special operations, where the specific choice of films and filters can offer the advantage of multi-spectral imagery. Multi-spectral photography has been responsible for the success of numerous applied research projects, such as surveys of plantation success, *Fomes*

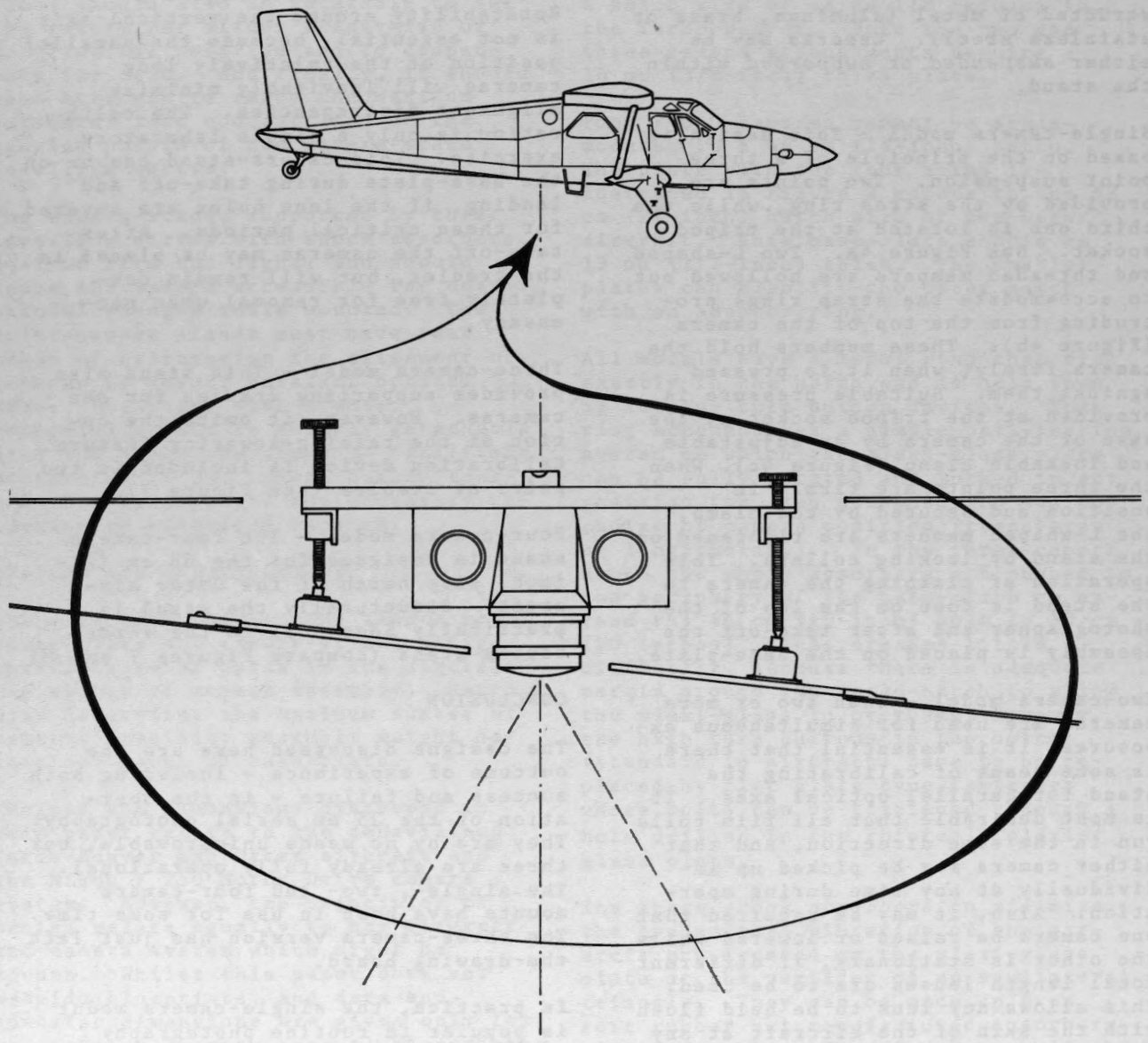


Figure 1. The position of the mounted camera in relation to the aircraft used.

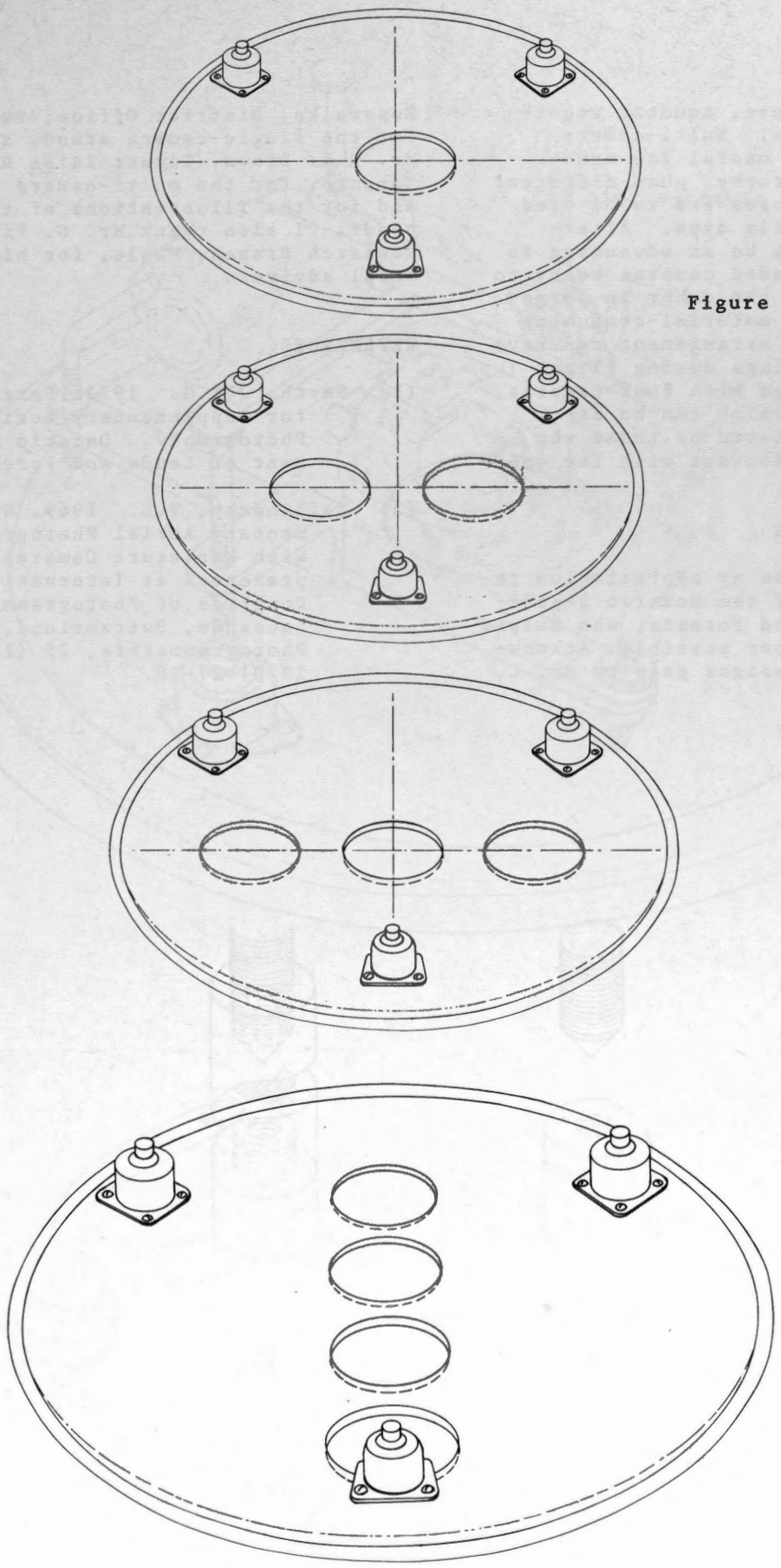


Figure 2. Base-plates.

annosus, snow geese, aquatic vegetation and others. Multi-camera mounts are also useful for monospectral photography, when different focal length lenses are to be used with the same film type. Alternatively, it can be an advantage to have several loaded cameras ready to shoot one after the other in large-scale and other material-consuming missions. This arrangement can save too many reloadings during flight (up to 1360 exposures with four cameras are possible), which can be especially appreciated by those who have often had contact with the emergency paper bag.

ACKNOWLEDGEMENTS

I wish to express my appreciation to those members of the Ontario Department of Lands and Forests, who helped to make this paper possible. Acknowledgement for designs goes to Mr. C. Rosewalka, District Office, Swastika, for the single-camera stand, and to Mr. R.R. Brown, Timber Sales Branch, Toronto, for the multi-camera stands and for the illustrations of this paper. I also thank Mr. G. Pierpoint, Research Branch, Maple, for his editorial advice.

REFERENCES

(1) Smyth, J.R.G. 1972. Parameters for Supplementary Aerial Photography. Ontario Department of Lands and Forests.

(2) Zsilinszky, V.G. 1969. Supplementary Aerial Photography with Miniature Cameras. Paper presented at International Congress of Photogrammetry, Lausanne, Switzerland, 1968. Photogrammetria, 25 (1969/1970) 27-38.

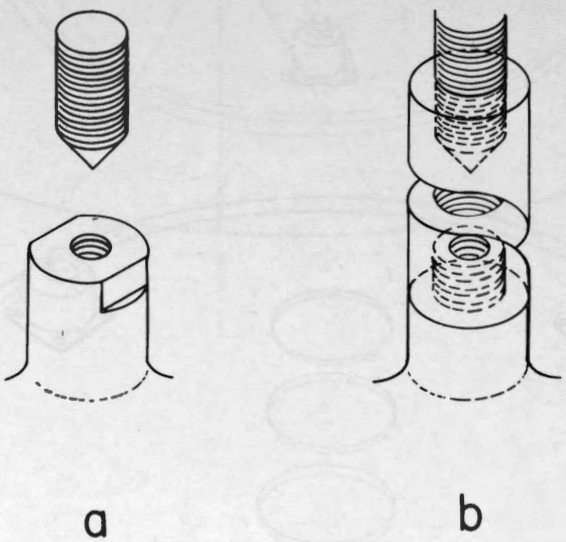
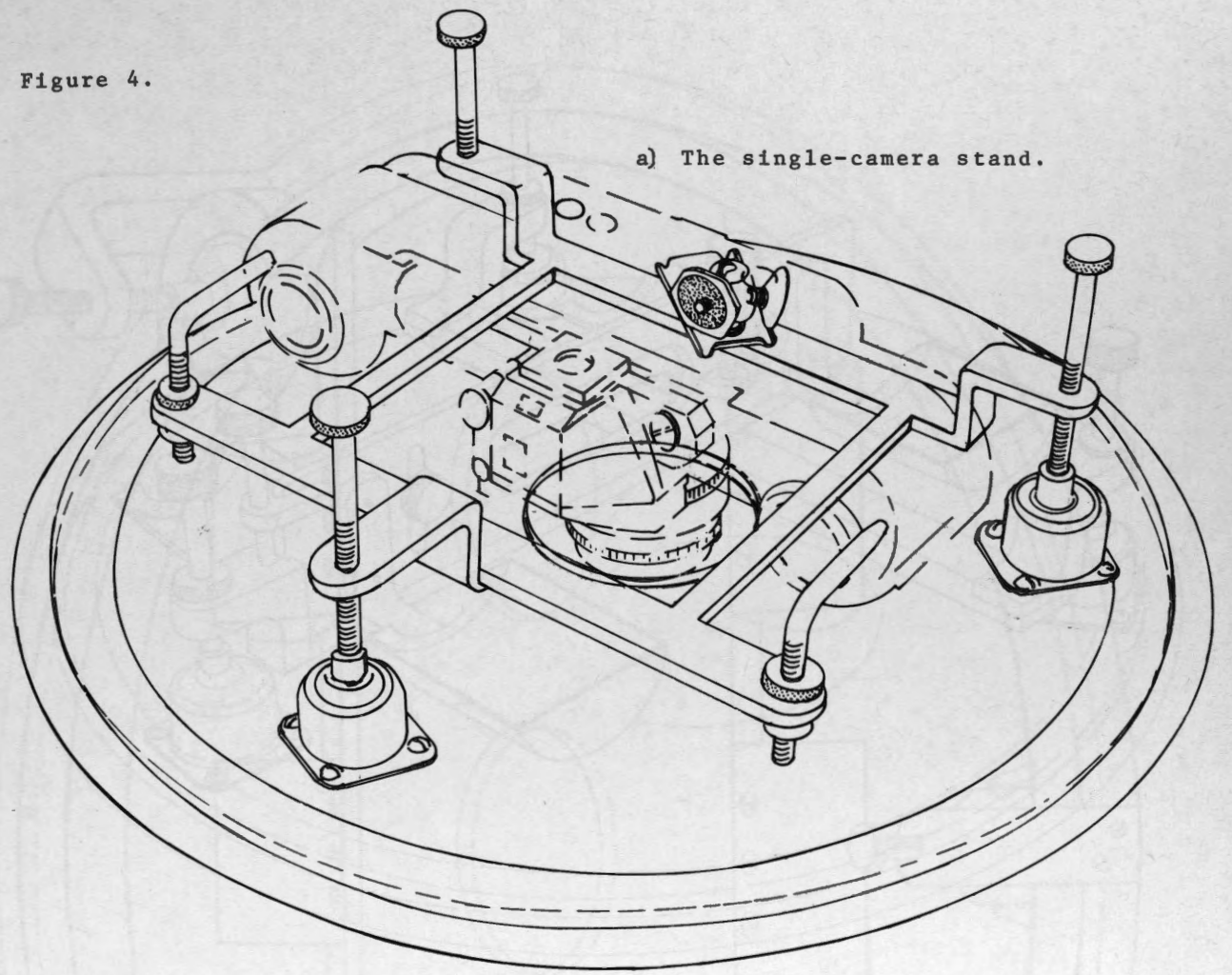
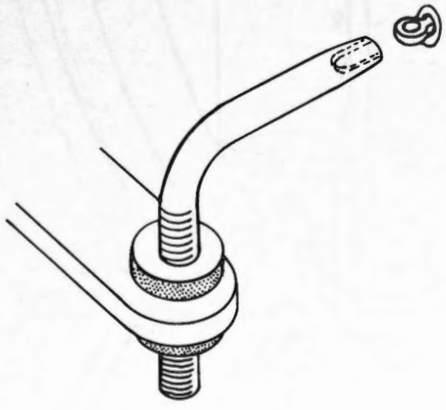


Figure 3. Two options for the connection between the baseplates and camera-stand:
a) free position.
b) locked position.

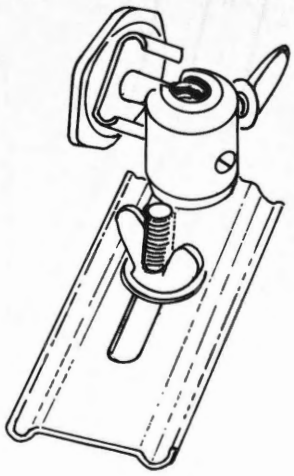
Figure 4.

a) The single-camera stand.





b) Close-up of the L-shaped suspending device.



c) Close-up of the camera clamp suspending device.

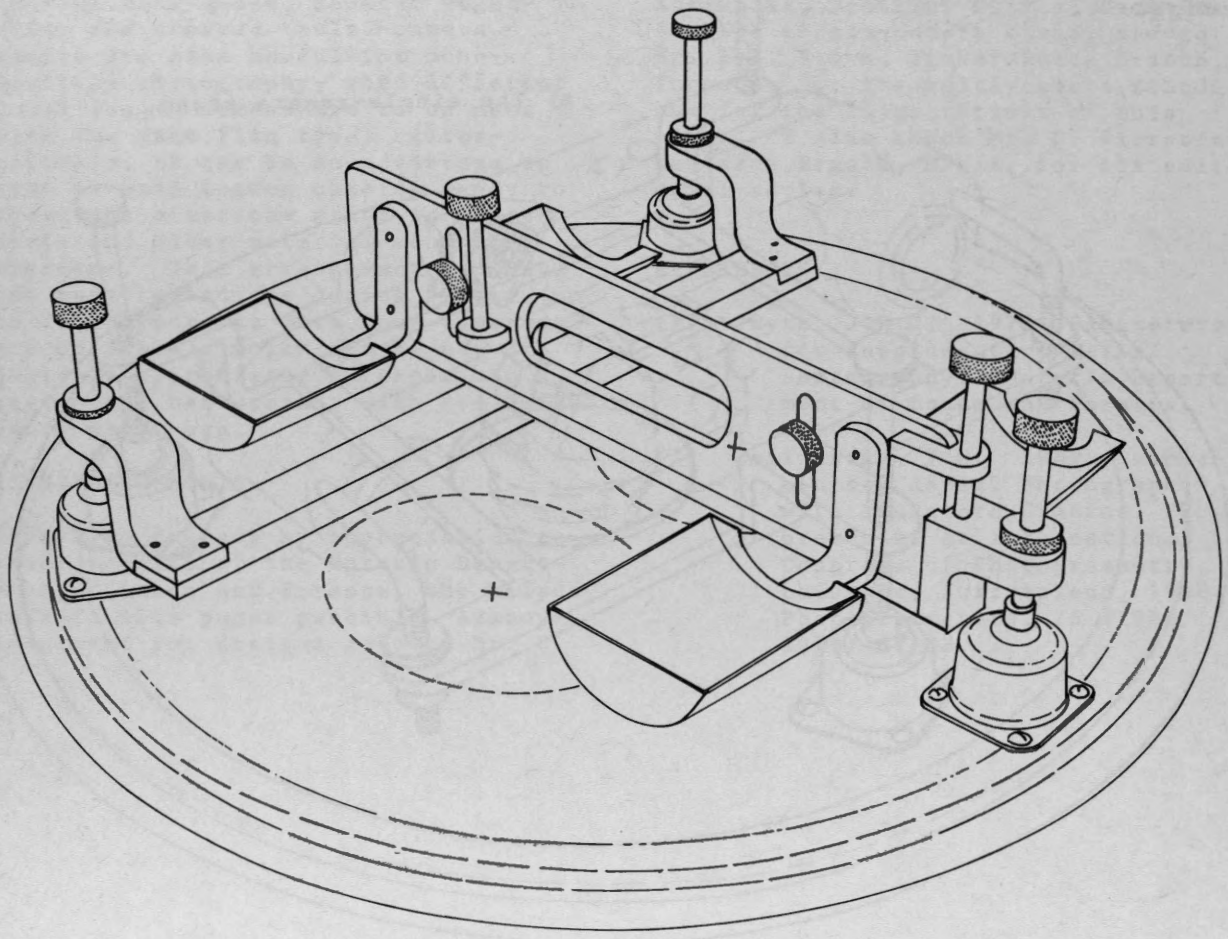


Figure 5. The two-camera stand.

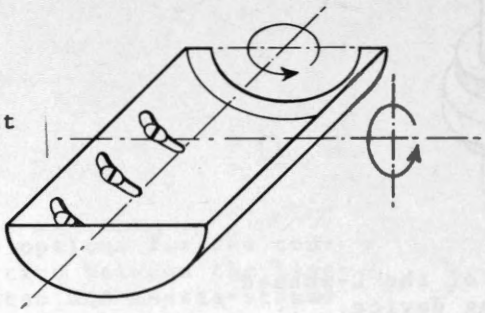


Figure 6. The movements for alignment of cradles.

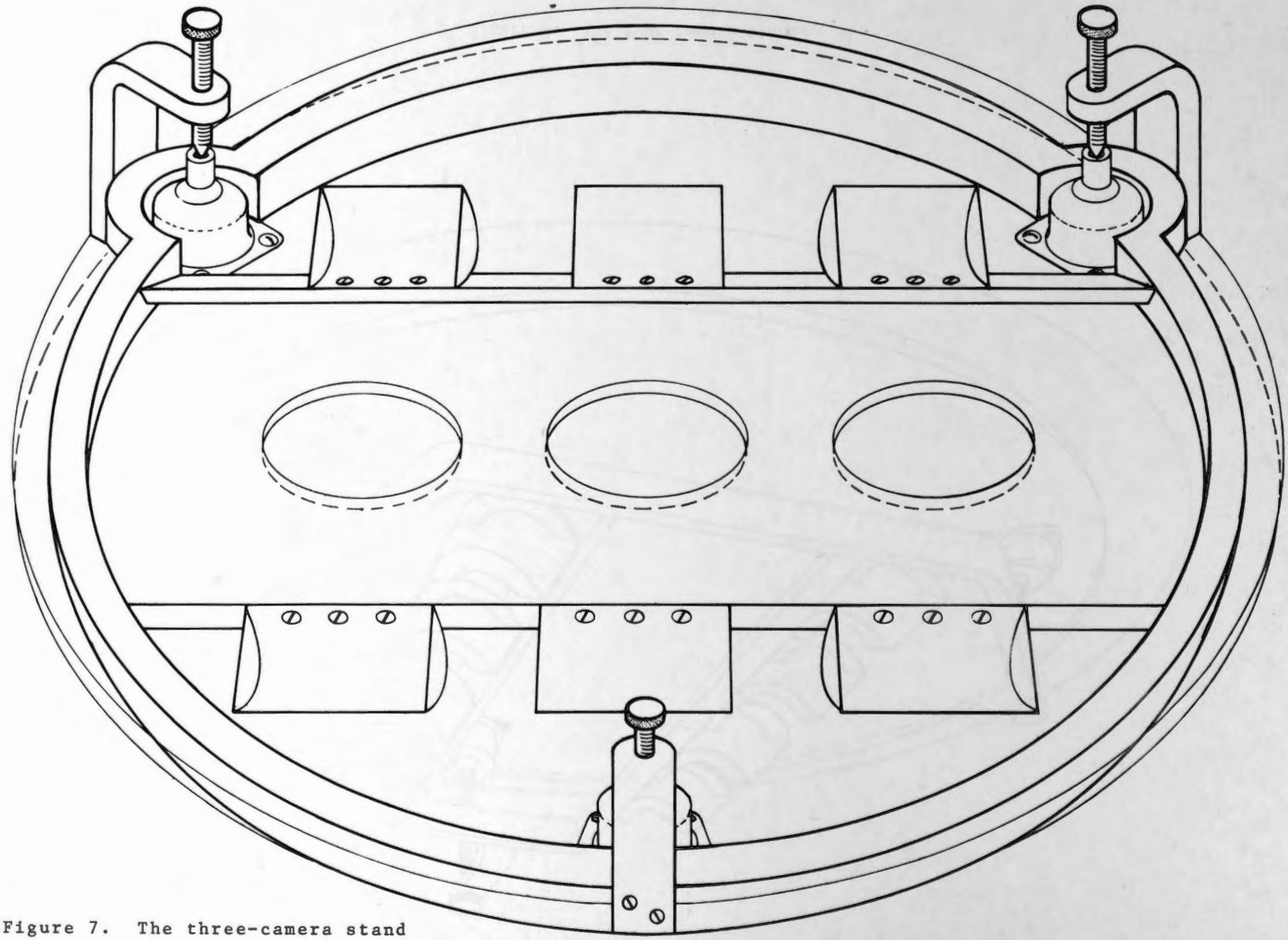


Figure 7. The three-camera stand showing eccentric positions of shock absorbers for camera clearance.

Figure 8. The four-camera stand.

TERRAIN ANALYSES (C)

ETALONNAGE AU SOL SUR PARCELLES EXPERIMENTALES DE RADIOMETRE BARNES PRT-5 DANS LA FENETRE 9.5 - 11.5 MICRONS.

F. Bonn, P. Clément, B. Bareilhe, C. Jalbert, Y. Leblanc, C. Trudel
Université de Sherbrooke
Sherbrooke, Québec

(English abstract at the end)

INTRODUCTION

La recherche en cours vise à la mise au point de clefs d'interprétation de documents infrarouge thermique pour la bande spectrale 9.5-11.5 microns, afin de dépasser le stade de l'interprétation qualitative. L'instrument utilisé est un radiomètre Barnes PRT 5 spécial.

METHODE:

Le choix de la fenêtre 9.5-11.5 microns: les radiomètres d'exécution courante utilisent la fenêtre 8-14 microns, laquelle inclut l'essentiel de l'émission terrestre. Mais une absorption par H_2O et CO_2 atmosphérique se manifeste, principalement entre 12 et 14 microns; elle varie suivant les conditions atmosphériques, la distance du détecteur à l'objet, etc. Nos intérêts s'appliquant surtout à la surface terrestre, visée généralement à basse altitude où l'absorption par l'ozone n'intervient pas, nous avons opté pour la bande spectrale 9.5-11.5 microns, correspondant au maximum d'émission terrestre entre -25° C et + 65° C; l'intervalle recouvre la majorité des conditions rencontrées sous nos climats, hormis quelques journées d'hiver. D'autre part, la courbe de transmittance du filtre utilisé sur le bolomètre découpe très abruptement la bande spectrale, sans minimum interne inférieur à 0.95.

L'utilisation de parcelles expérimentales: nos autres intérêts de recherche portant sur la connaissance et l'estimation de processus d'érosion sur versants, nous possédons des parcelles expérimentales équipées pour la mesure des exportations de matière et du bilan énergétique. La mesure de la radiation émise dans la fenêtre 9.5-11.5 microns est un bon moyen d'estimer, après conversion des données, le flux thermique émis par la terre à un moment donné.

L'interprétation d'un phénomène global, tel que la radiation émise, dépendant de plusieurs facteurs, nécessite un terrain connu et équipé en conséquence, afin de déterminer l'influence de chaque facteur et sa variation dans le temps. Les mesures ont donc été faites sur base comparative, entre une prairie semi-naturelle de 80 m^2 et une aire similaire artificiellement débarrassée de sa couverture végétale, l'ensemble étant sur un versant sud-ouest sur till limoneux.

Les variables mesurées ont été: la radiation solaire incidente; la température et l'humidité de l'air; la température du sol en surface, à 10, 20 et 30 cm de profondeur; l'humidité du sol à 10 et 25 cm; la radiation émise lue sur le radiomètre Barnes PRT-5. En plus des observations de routine, des mesures intensives (toutes les 2 ou 3 heures) ont été réalisées pendant plusieurs jours consécutifs: la première en mai, la seconde en juillet, la troisième en août et la quatrième en septembre-octobre 1971. Seule l'avant-dernière n'a pas fourni de données interprétables.

D'autres travaux ont été commencés: influence des barres rocheuses affleurantes et subaffleurantes sur la radiation émise; influence des variables hydrologiques sur la radiation émise par les eaux naturelles, ceci notamment dans le bassin de la rivière Eaton, Québec.

RESULTATS

Ceux discutés ici ne concernent que l'étude sur parcelles expérimentales et ne prétendent qu'à une présentation des problèmes.

D'un point de vue qualitatif, les facteurs influant sur l'émission par la surface de nos parcelles sont: la température de la surface du sol; le sens et la valeur du gradient thermique traversant le sol; la température de l'air; la nature de la surface du sol (état de la couverture végétale surtout); l'humidité du sol; la période de l'année (par l'intermédiaire de l'état de surface et de l'humidité du sol); la période de la journée (par l'intermédiaire du flux thermique traversant le sol, du taux d'activité végétale et de l'humidité superficielle du sol).

D'un point de vue quantitatif, certains paramè-

tres ont été mis en corrélation. La comparaison porte sur les deux parcelles et leur évolution respective dans le temps.

- Valeurs brutes de radiométrie infra-rouge (figures 1 à 3) de température de l'air (figures 3 à 6) et de température du sol (figures 7 à 9). Ces dernières indiquent le sens du flux thermique traversant le sol: flux ascendant lorsque la température de surface est inférieure à celle de 10 cm, flux descendant dans le cas opposé. La comparaison des températures en surface et à 10 cm permet de déterminer le changement de sens du flux avec une précision de ± 1 heure puisque les vitesses moyennes des flux thermiques sont de l'ordre de 5 cm/heure (C.W. Rose, 1966). En ce qui concerne l'humidité du sol, les variations horaires sont peu significatives; plus important est l'écart existant entre les deux parcelles au mois de juillet: le degré d'assèchement et l'épaisseur de la couche relativement assèchée sont plus grands pour la parcelle couverte.

- Mise en relation des phénomènes à périodicité semblable; c'est à dire liés directement à la radiation solaire, en particulier les phénomènes thermiques: températures de radiance de la surface des parcelles et température de l'air (figures 10 à 12); températures de radiance et températures de la surface des parcelles (figures 13 à 15). Le tableau suivant résume les relations existant entre ces paramètres. Toutes les corrélations globales sont hautement significatives. Il n'en est pas de même pour le détail par flux ascendant et descendant par suite du nombre insuffisant de données.

Relation	T^o radiance (y) - T^o air (x)	
	Equation de régression	Coefficient de corrélation
PC Mai	$y = -0.15 + 1.01\ x$	0.88
PD	$y = 0.77 + 0.85\ x$	0.89
PC Juillet	$y = -3.32 + 1.15\ x$	0.89
PD	$y = -5.99 + 1.51\ x$	0.82
PC Septembre	$y = -1.27 + 1.06\ x$	0.88
PD	$y = -0.40 + 1.00\ x$	0.89

Relation	T^o radiance (y) - T^o surface (x)	
	Equation de régression	Coefficient de corrélation
PC Mai	$y = -15.29 + 2.42\ x$	0.86
PD	$y = -4.51 + 1.27\ x$	0.92
PC Juillet	$y = 0.95 + 0.78\ x$	0.90
PD	$y = -7.96 + 1.27\ x$	0.93
PC Septembre	$y = -2.01 + 1.12\ x$	0.94
PD	$y = 0.40 + 0.92\ x$	0.91

CONCLUSIONS

La relation entre température de l'air et température de radiance de la surface de la parcelle couverte varie peu selon la période. Il n'en est pas ainsi de la parcelle désherbée pour laquelle la pente de la droite de régression est maximum en juillet, période de forte intensité radiative solaire sur sol nu et un minimum en mai (influence de la nébulosité).

Les variations des droites de régression exprimant la relation entre la température de radiance et la température des surfaces diffèrent plus nettement. Nous avons tenté de détailler ces relations selon le sens du flux thermique traversant le sol; la pente des droites partielles semble influencée par ce sens à certaines périodes: en juillet et septembre sur la parcelle couverte elle est plus forte par condition de flux ascendant, ainsi qu'en septembre sur la parcelle désherbée. Pour les autres séries, l'influence semble faible et de sens incertain.

L'émission d'énergie est "forcée" par l'existence de ce flux ascendant, phénomène essentiellement nocturne. Ce flux plus intense en saison chaude qu'au printemps par suite de l'accumulation d'énergie dans le sol. Ceci se traduit, toutes choses étant égales par ailleurs, par des températures de radiance plus élevée qu'en cas de flux descendant.

Cependant d'autres facteurs apparaissent dans l'interprétation des relations: la présence d'une couche relativement assèchée en juillet dans la parcelle couverte diminue l'émission infra-rouge. Au contraire, en mai, l'existence en surface d'un paillis humide, constitué par les tiges des herbes de l'année précédente,

l'augmente (cf. pente de la droite de régression). L'égalisation des conditions d'humidité et d'activité végétale en septembre-octobre se traduit par des droites très proches les unes des autres pour les deux parcelles, et ressemblant à celles de la relation température de radiance - température de l'air.

Le caractère encore trop espacé de ces observations ne permet pas encore de déterminer la part des diverses influences et les passages d'un type de répartition à un autre. L'enregistrement des données par automatisation de la station, but de la prochaine campagne, devrait l'autoriser.

Ainsi, on peut espérer déterminer les périodes et les heures où des relevés aéroportés par infra-rouge thermique auront le plus de chances de représenter une fraction constante des températures au sol, c'est à dire où les valeurs d'émissivités mesurées en laboratoire pourront être appliquées à l'interprétation des surfaces naturelles.

Ces recherches ont été partiellement subventionnées par le Conseil National de Recherches du Canada et le Conseil de Recherches pour la Défense, que nous remercions ici.

REFERENCES

(1) D. Auding, R. Kanth, Estimation for Sea Surface Temperature from Space, Remote Sensing of Environment. 1, pp. 217-220, 1970.

(2) F. Bonn, Applications de la radiométrie infra-rouge thermique à la détection de phénomènes géomorphologiques superficiels, Comm. 39e Congrès de l'A.C.F.A.S., Sherbrooke 1971.

(3) K.J.K. Buettner, C.D. Kern, The Determination of Infra-red Emissivities of Terrestrial Surfaces, Jal of Geophysical Research, vol. 70, #6, March 1965.

(4) A. Hadin, Elements of Modern Physics applied to the study of the Infra-red, Pergamon.

(5) J.O. Mattson, Thermal Patterns in the Landscape recorded with Infra-red Technique and Simulated in Model Experiments, Geografiska Annaler, 51 A, #4, 1969.

(6) R.B. Platt, J.F. Griffiths, Environmental Measurement and Interpretation, Reinhold, 235 p., 1965.

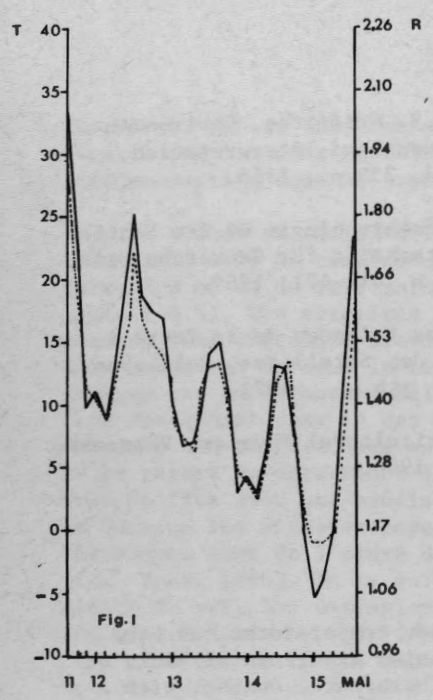
(7) J. Pouquet, Géomorphologie et Ere Spatiale, Zeitschrift für Geomorphologie, 13, #4, p. 414-471, 1969.

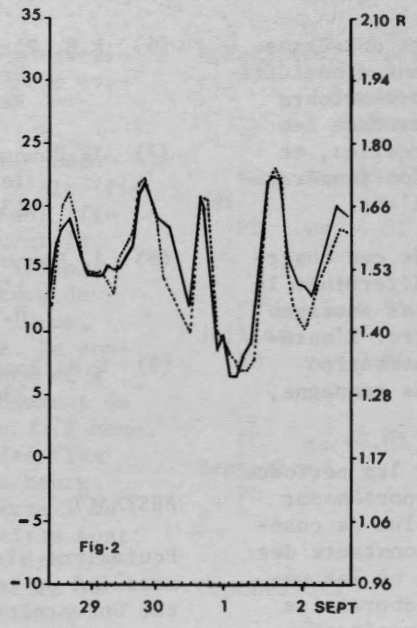
(8) J. Pouquet, Les Sciences de la Terre à l'heure des Satellites, Coll. Sup., P.U.F., 259 p. , 1971.

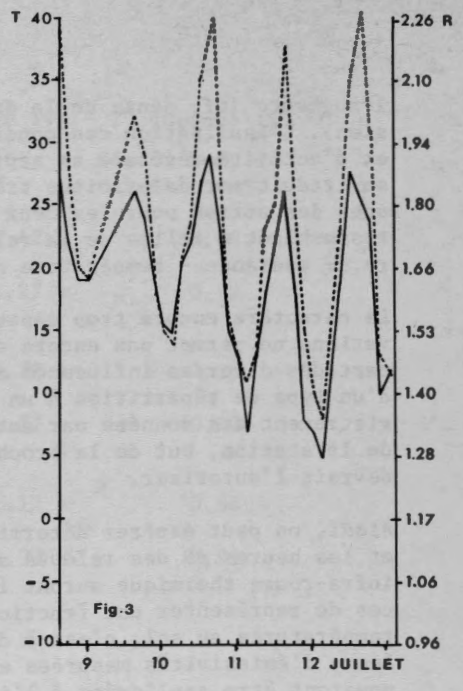
(9) C.W. Rose, Agricultural Physics, Pergamon, 230 p., 1966.

ABSTRACT

Equivalent black body temperatures has been measured on two erosion experimental plots at the University of Sherbrooke, Quebec, with a Barnes PRT-5 radiometer; the bandwidth is 9.5 - 11.5 microns in order to avoid partial atmospheric H_2O and CO_2 absorptions. The other measured variables are = solar incident radiation, air temperature and humidity, soil temperature and humidity at various depths. The results of 3 intensive observation periods (May, July,September) are here discussed. The most influent variables are: soil surface temperatures, direction and intensity of soil thermal gradient, soil surface conditions (mainly vegetation cover: one of the plots being bare and the other grass covered) and soil humiditv. Signifiant constant correlations between air temperatures and T^{BB} were obtained except for the bare plot where the linear regression coefficient is maximum in July and minimum in May. Correlations between soil surface temperatures and T^{BB} show several influences. All other factors being equal, soil I.R. emission may be increased by ascendant thermal flow through the soil, specially during Summer and Fall. Increased emission is due in Spring on vegetated ground to the wet straw cover whereas during Summer it is decreased by soil drying under active grass cover. Fall situation reflect equalizing conditions on both plots. Further accurate knowledge of these phenomenons will require automatic recording and permit determination of the most useful periods for remote sensing in this bandwidth.

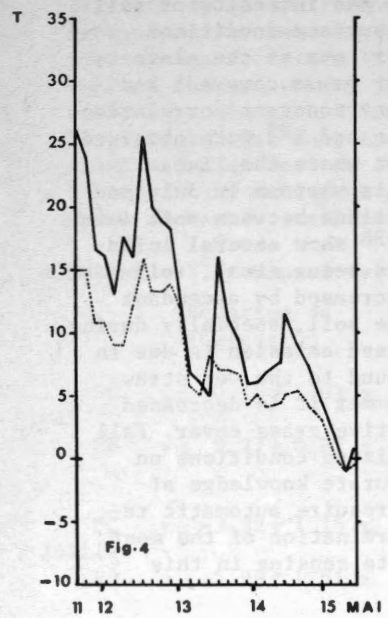


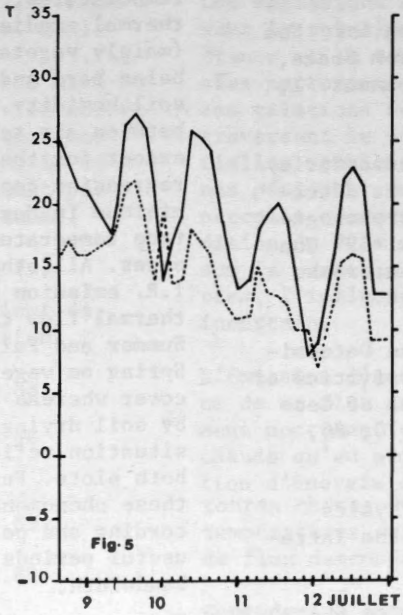


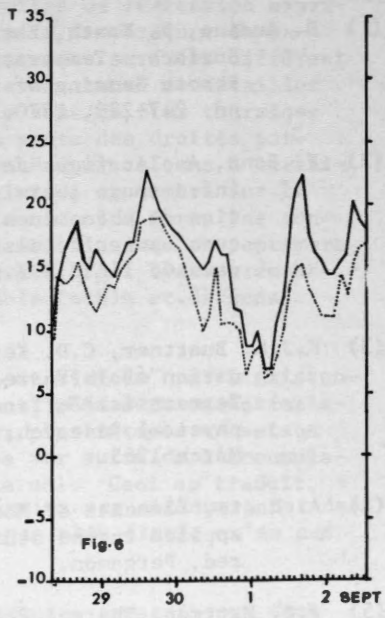


Figures 1 à 3: Températures de radiance.
..... Parcelle désherbée
Parcelle couverte

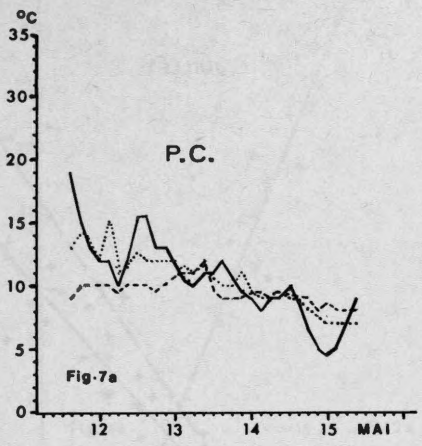
Echelle double = degrés Celsius et
m W - cm^{-2} - $ster.^{-1}$

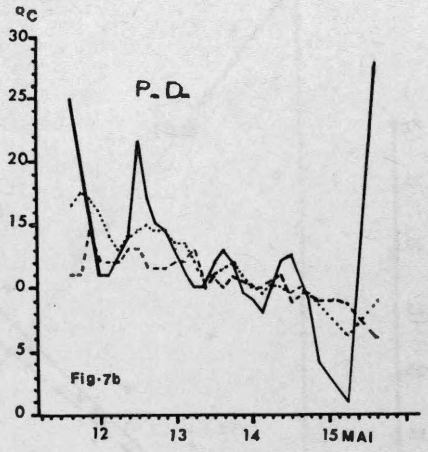


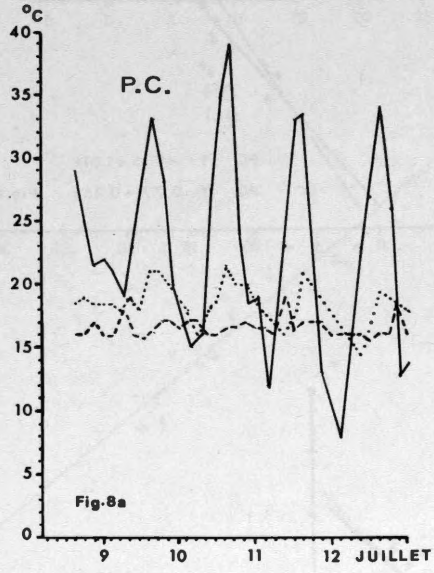


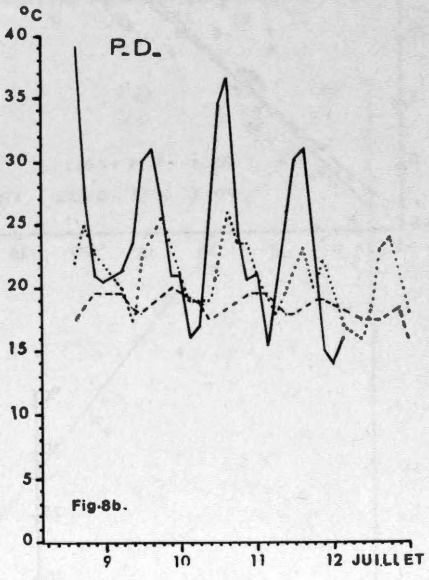


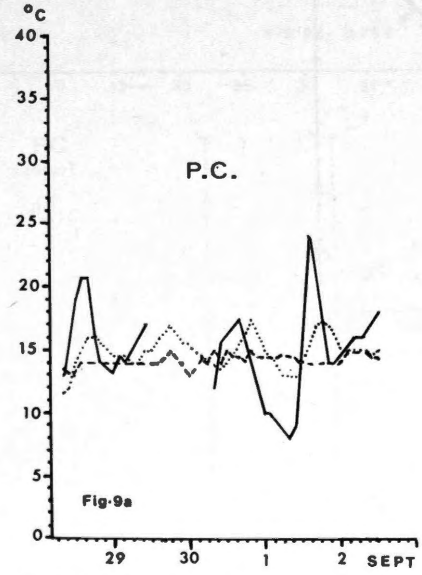
Figures 4 à 6: Températures de l'air
_____ thermomètre sec
---------- thermomètre humide

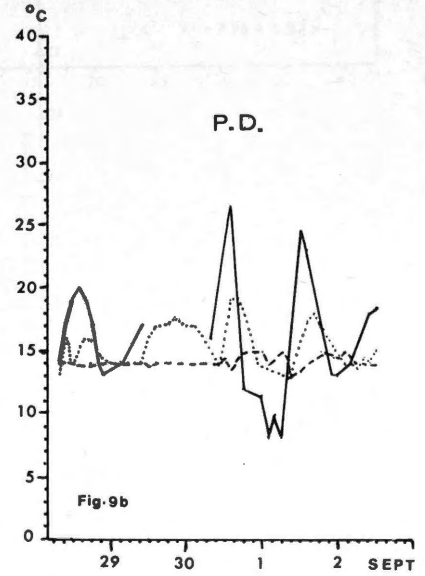




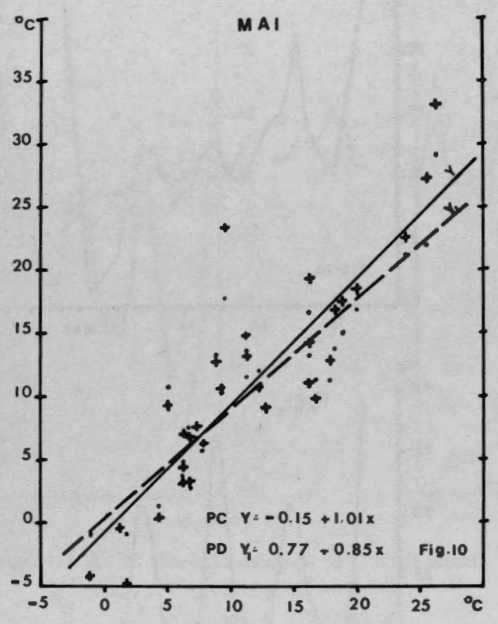


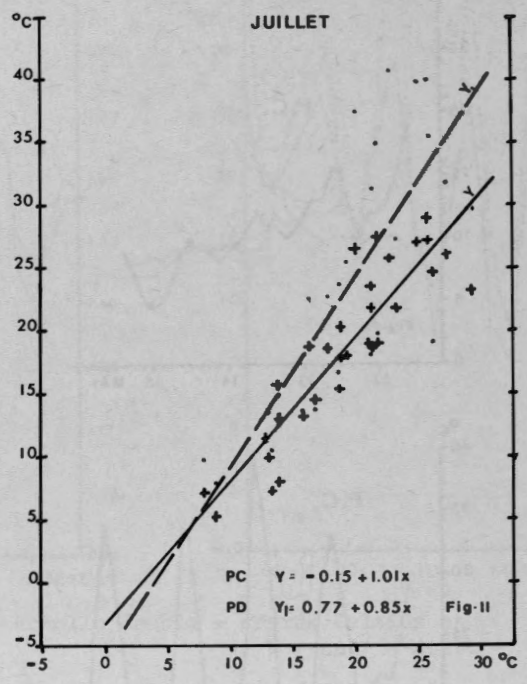


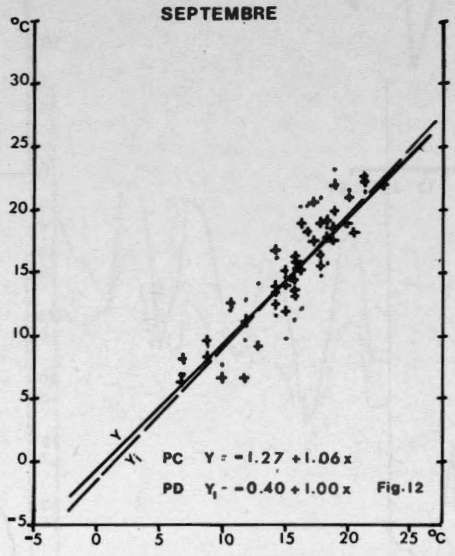




Figures 7 à 9: Températures du sol
PC: parcelle couverte _____surface
PD: parcelle désherbée..... 10 cm.
------- 30 cm.



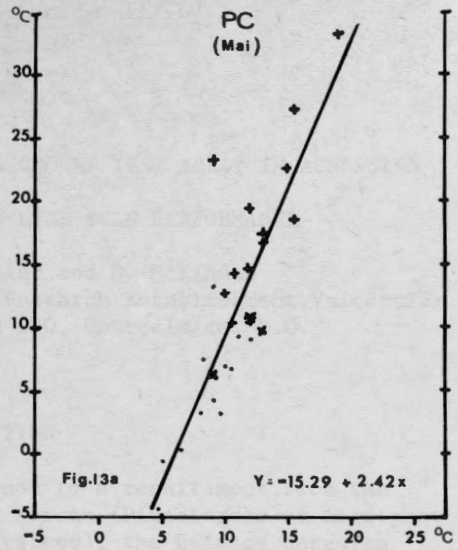


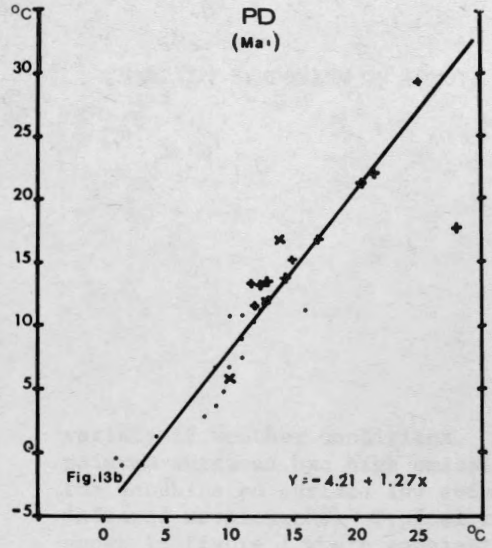


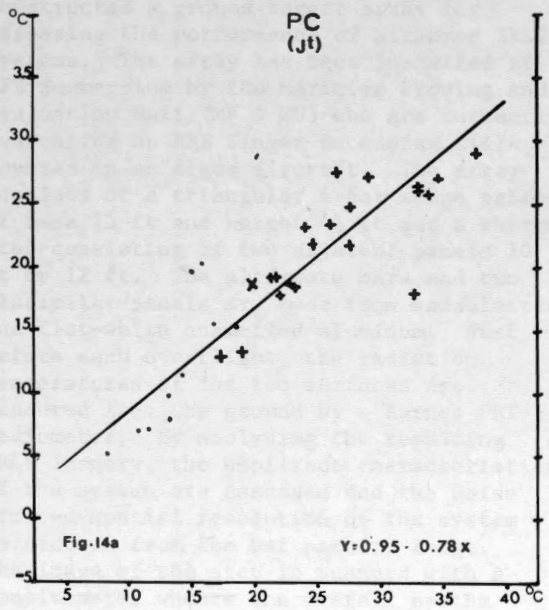
Figures 10 à 12: Relation températures de l'air (x) - températures de radiance (y).

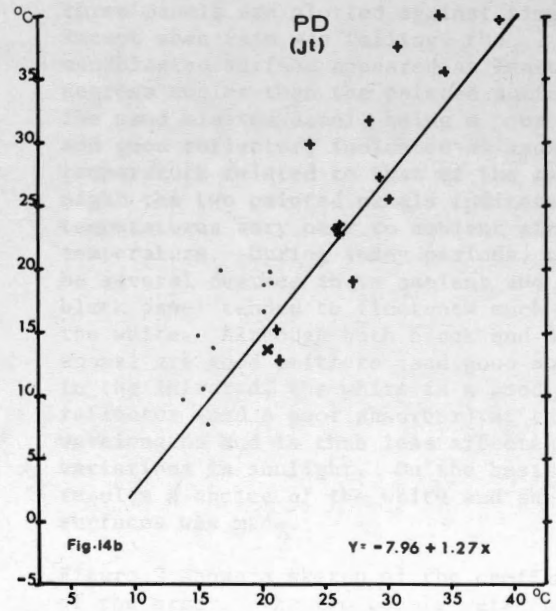
+ conditions de flux thermique descendant

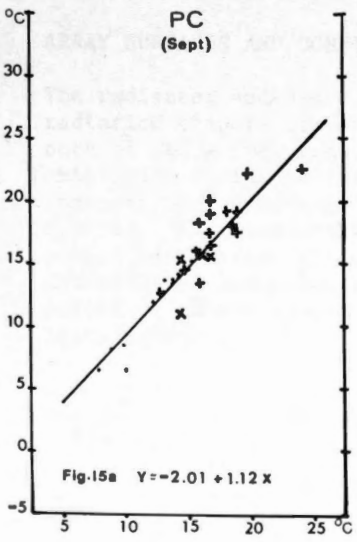
. conditions de flux thermique ascendant

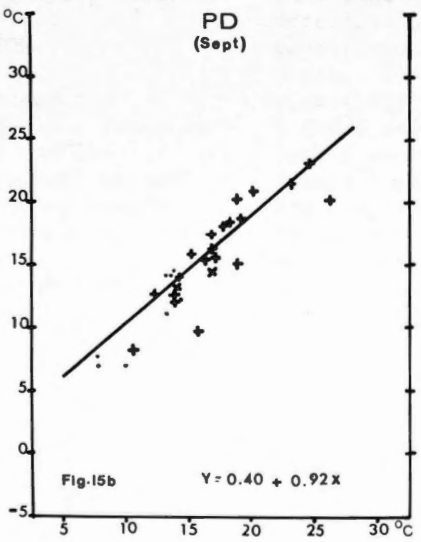












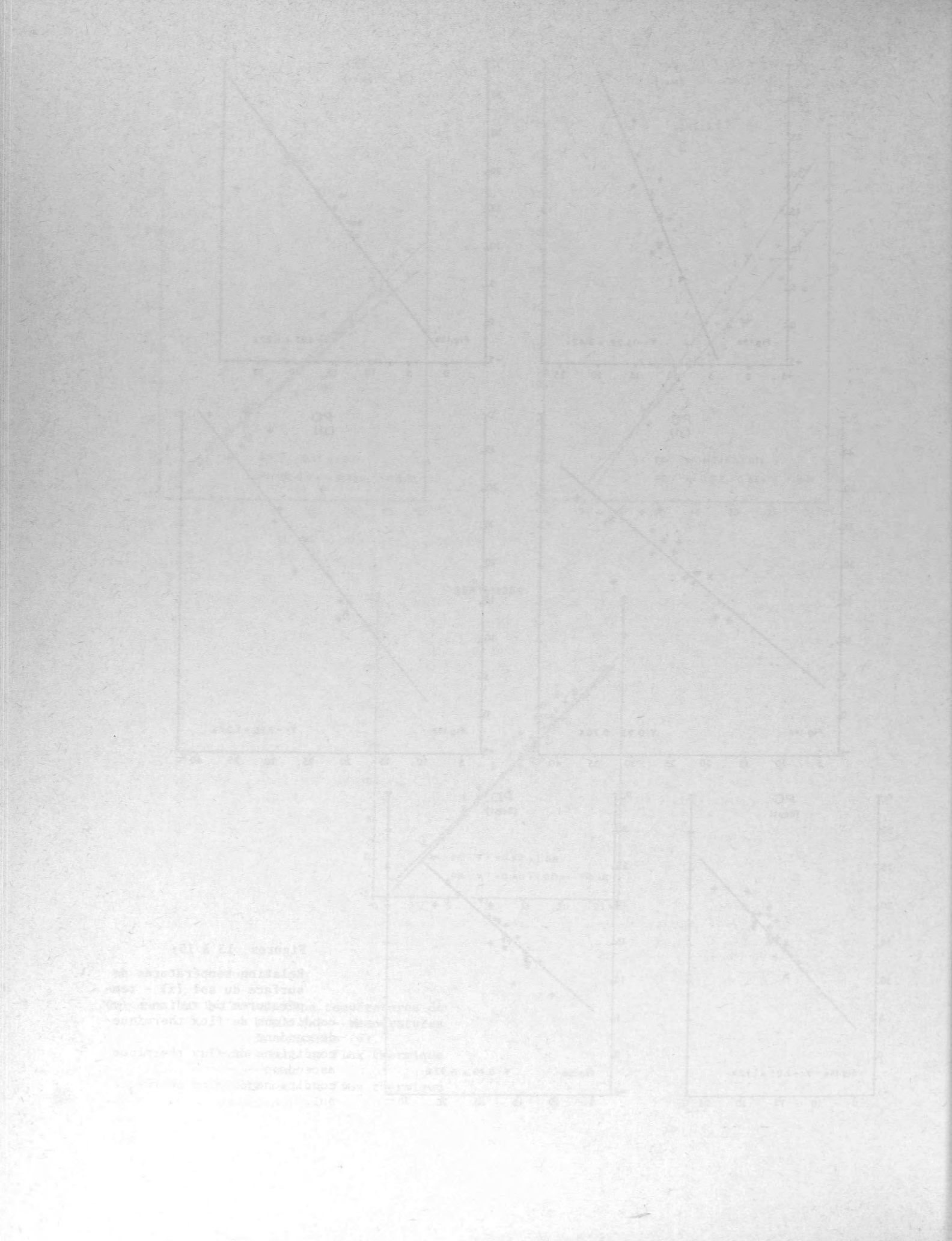
Figures 13 à 15:

Relation températures de surface du sol (x) - températures de radiance (y)

\+ conditions de flux thermique descendant

. conditions de flux thermique ascendant

x conditions de flux thermique nul.



USE OF A GROUND TEST ARRAY IN ASSESSING INFRARED LINE SCAN PERFORMANCE

G.A. Morley and D. McKinnon
Defence Research Establishment Valcartier
P.O. Box 880, Courcelette, P.Q.

INTRODUCTION

In response to a requirement from the Canadian Forces (Directorate of Aerospace Combat Systems), the Defence Research Establishment Valcartier (DREV) has constructed a ground target array for assessing the performance of airborne IRLS systems. The array has been installed at CFB Summerside by the Maritime Proving and Evaluation Unit (MP & EU) who are currently evaluating an HRB Singer Reconofax XIIIA mounted in an Argus aircraft. The array consists of a triangular 4-bar wedge pattern of base 15 ft and height 45 ft and a thermal step consisting of two adjacent panels 10 ft by 12 ft. The alternate bars and two dissimilar panels are made from sandblasted and flat-white enamelled aluminum. Just before each overflight, the radiation temperatures of the two surfaces are measured from the ground by a Barnes PRT-5 radiometer. By analysing the resulting IRLS imagery, the amplitude characteristics of the system are assessed and the noise limited spatial resolution of the system is deduced from the bar pattern image. The image of the step is scanned with a densitometer whence the overall system modulation transfer function is determined.

ARRAY SURFACES AND CONFIGURATION

The radiance, and hence the apparent radiation temperature of a body is a function both of its actual temperature and the emissivity of its surface. Apparent thermal contrast may be attained by varying either quantity. Various methods of producing actual temperature differences were investigated and rejected as either too expensive or too cumbersome for field installations.

A short series of measurements was made on panels of various emissivities laid out on the ground. Panels of black painted aluminum, white painted aluminum and sandblasted aluminum were prepared and their radiation temperatures were observed over a variety of weather conditions. The two painted surfaces had high emissivities and the sandblasted surface low emissivity at infrared wavelengths. Typical results are shown in Figure 1 where apparent temperatures, as measured by a Bofors IR camera, for the three panels are plotted against time. Except when rain was falling, the sandblasted surface appeared at least several degrees cooler than the painted surfaces. The sand blasted panel, being a poor emitter and good reflector, indicated an apparent temperature related to that of the sky. At night the two painted panels indicated temperatures very near to ambient air temperature. During sunny periods, they could be several degrees above ambient and the black panel tended to fluctuate much more than the white. Although both black and white enamel are good emitters (and good absorbers) in the infrared, the white is a good reflector (and a poor absorber) at the shorter wavelengths and is thus less affected by variations in sunlight. On the basis of these results a choice of the white and sandblasted surfaces was made.

Figure 2 shows a sketch of the configuration of the array. The two panels were chosen such that when scanned by the various detectors of the Reconofax XIIIA, steady-state conditions are reached while crossing each plate. The nominal instantaneous field of view (IFOV) of the detectors are between 1 and 4 mrad. Thus at 500 ft altitude the panels subtend 20 IFOV's for the 1 mrad detector and 5 IFOV's for the 4 mrad detector. The wedge shaped bar pattern was chosen as large as possible consistent with ease of field installation. With the aircraft at 500 ft altitude, the lowest spatial frequency of the bar pattern, at the base of the triangle, is 0.83 cycles/mrad which for most conditions should be lower than the resolution limit of the Reconofax XIIIA on all channels. Small hot sources in the form of hooded electric heat lamps were installed before and after the array along the line of flight and to each side. These are used to assist the flight crew in lining the aircraft

especially at night and to provide an accurate scale reference on the imagery. The array was constructed in sections of 1/8 in aluminum sheeting so that it may be dismantled and moved by two or three men in a day. Just before being overflown the radiation temperature of the array is measured with a Barnes PRT-5 hand-held radiation thermometer.

ANALYSIS

Each image of the array is analysed in the following manner: (a) the image of the panels is scanned with a microdensitometer yielding two points on the amplitude characteristic of the overall system and an edge scan for use in calculating the overall system modulation transfer function (MTF) and (b) the image is examined under a microscope and the limiting spatial frequency is found by determining the distance from the base of the wedge to where the bars blend together.

Figure 3 is typical imagery of array recorded on the four different channels of the Reconofax XIIIA late in the afternoon of 25 Nov 71. In all cases the sky was overcast and the temperature difference ΔT, between the two surfaces measured by the PRT-5 was close to $5^{\circ}C$. The limiting resolutions are .355, .445, .343 and .153 cycles/mrad for channels 1 to 4 respectively. The corresponding nominal IFOV's are 1, 2, 4 and 4 mrad. The detectors in Channels 1 to 3 are of Mercury Cadmium Telluride responsive in the 8 to 13 μm band and Channel 4 is of Indium Antimony responsive in the 3 to 5 μm band.

Figure 4 is a series of array images recorded using channel 3 of the Reconofax XIIIA. For each run a different gain was used. The Reconofax XIIIA has been fitted with a manual gain control as well as an AGC. During all these runs the temperature difference between the two surfaces as measured with the PRT-5 was greater than $15^{\circ}C$. There were broken clouds in the sky. It is noted that the density and hence the temperature of the white panel is very near to that of the snow in the surrounding field. The density of the sand blasted panel decreases with the increasing gain.

In figure 5, the density (as taken from the original film which is a negative in the sense that increasing density indicates increasing temperature) of the panels is plotted against system gain. These data were taken from the imagery of Figure 4.

The density of the sand blasted panels appears to reach a lower limit somewhat above the base + fog density of the film which was around density 0.35. This indicates amplitude limiting occurs in the system before the film, probably in the lamp or amplifier.

Limiting spatial resolution is a function of many factors, ΔT, atmospheric extinction, amplitude limiting, system noise etc. Figure 6 shows how limiting spatial resolution for high ΔT ($>15^{\circ}C$) varied with gain for Channels 2 and 3. The decrease in limiting resolution with increasing gain at the higher gain settings is probably due to the amplitude limiting indicated in Figure 5. The decrease in limiting resolution for the lower gains is probably due to the relatively increasing effect of film granularity as exposure contrast is reduced

Figure 7 is a microdensitometer scan of the two panels recorded on Channel 2 in Figure 4. These data were used to calculate the MTF of the overall IRLS system. A relatively large slit width of 40μm was used in order to smooth the effects of film granularity and other system noise. The data were entered in computer, where they were transformed into an edge in transmission space and differentiated to yield the uncorrected line spread function for which the MTF was calculated. This MTF was corrected by dividing the MTF of the slit alone into it.

The MTF calculated from the edge in Figure 7 is shown in Figure 8. A maximum is noted around a spatial frequency of 0.25 cycles/mrad corresponding to the ringing just to the left of the edge in Figure 7. The MTF cuts off at a spatial frequency just over 0.5 cycles/mrad. Beyond this frequency, false resolution and a phase reversal of the bars occurs. The actual limiting resolution as observed on the wedge bar pattern for this run was 0.45 cycles/mrad and is indicated on graph.

The computer programme is being modified to use a technique described by Jones (Ref. 1) and Jones and Yeadon (Ref. 2) developed for calculating the MTF of photographic systems from noisy edge scans in which the correction for densitometer slit width is made using a convolution integral. This technique incorporates digital smoothing and simplifies the calculation of the MTF in exposure space rather than in transmission space as described above. Plans are also being made for direct digital recording of the densitometer output to obviate the effort

required for manual reading of the densitometer scans.

CONCLUSIONS

The array has been used for limited number of runs, has performed as designed and has proven useful in determining the overall performance of an IRLS under operating conditions.

REFERENCES

(1) R.A. Jones. An Automated Technique for Deriving MTF's from Edge Traces.
PS&E vol 11, no. 2, Mar-Apr 67; pp 102-106.

(2) R.A. Jones and E.C. Yeadon. Spread Function Determination from Noisy Edge Scans.
PS&E vol 13, no. 4, Jul-Aug 69; pp 200-204.

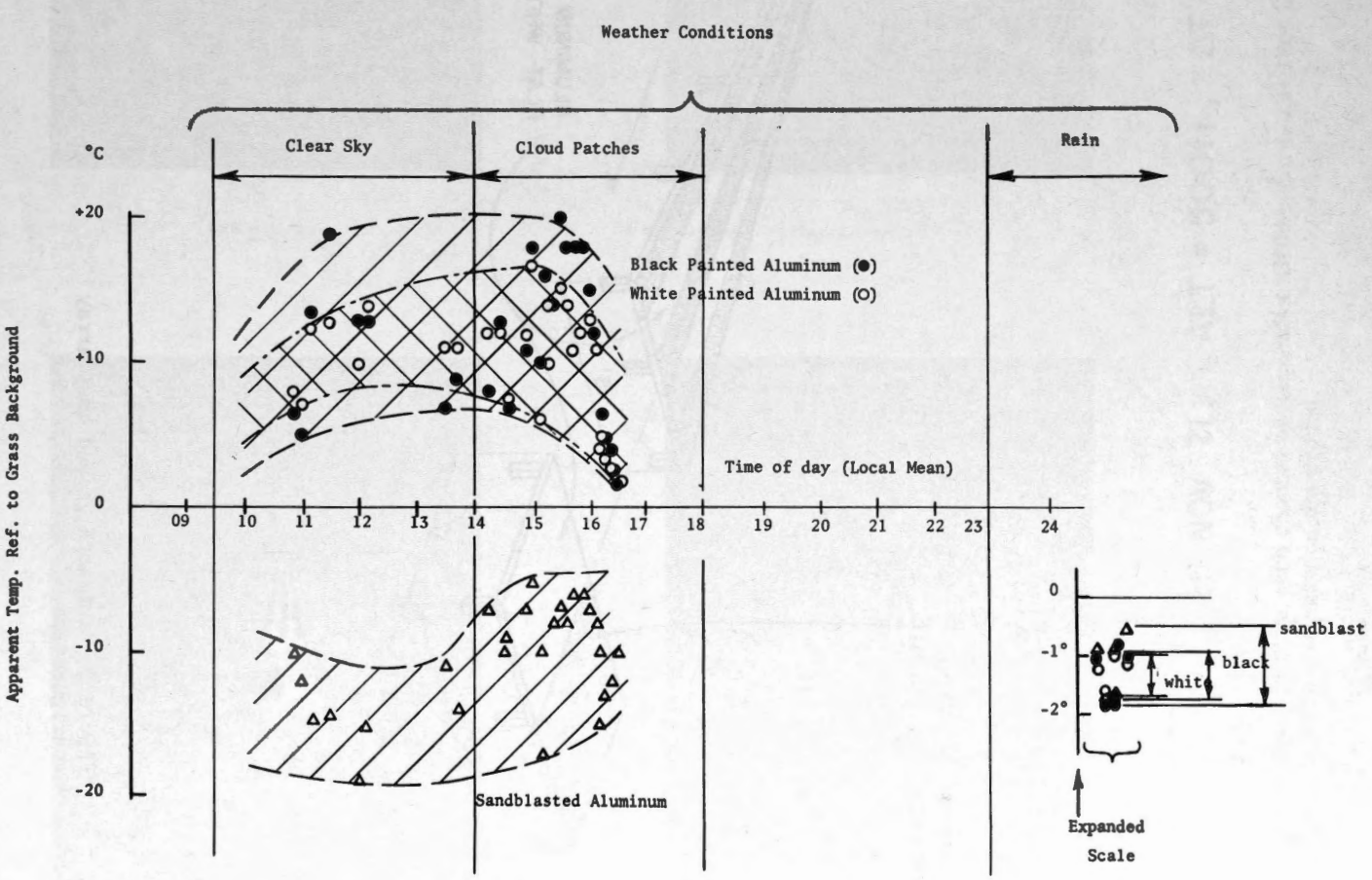


Figure 1 - Radiation temperature measurements of white painted, black painted and sandblasted aluminum panels using Bofors IR camera.

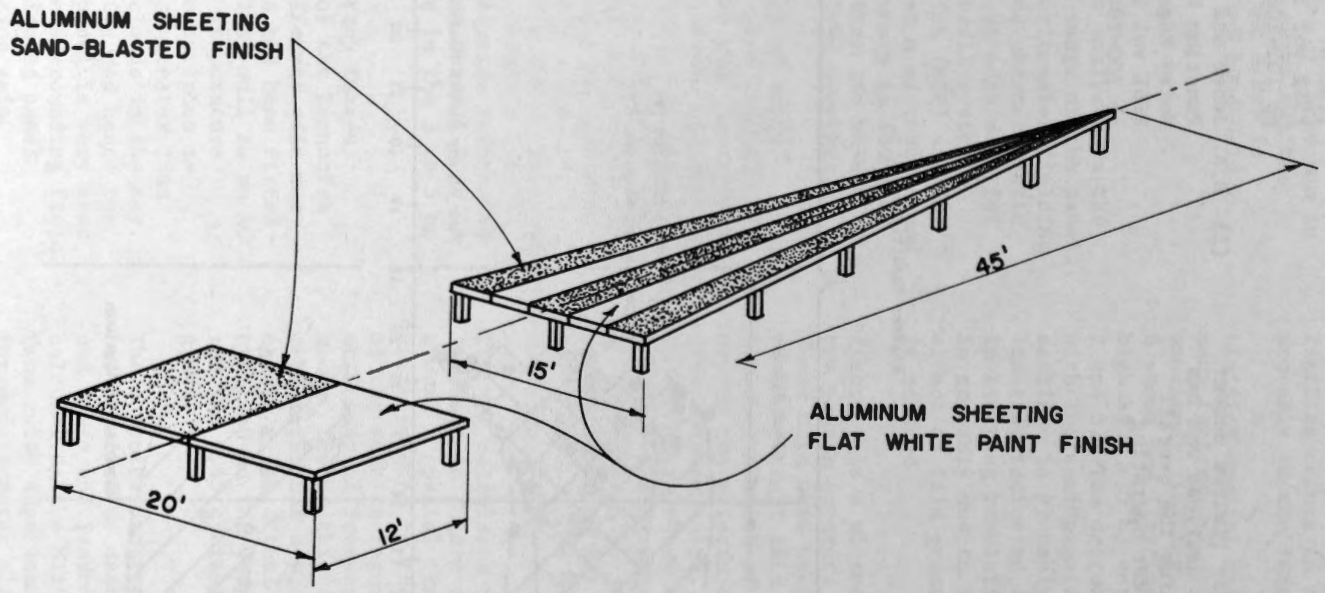
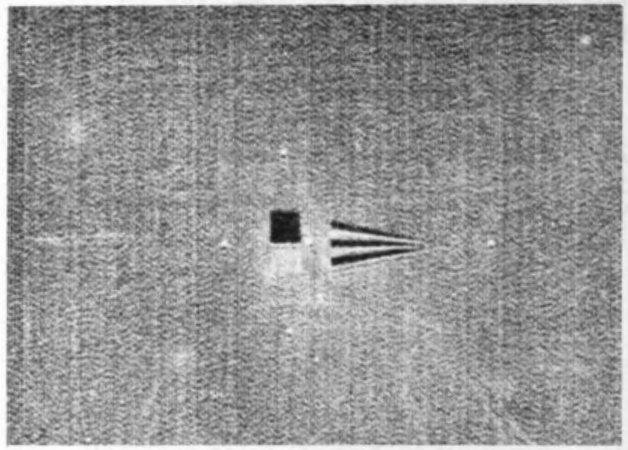
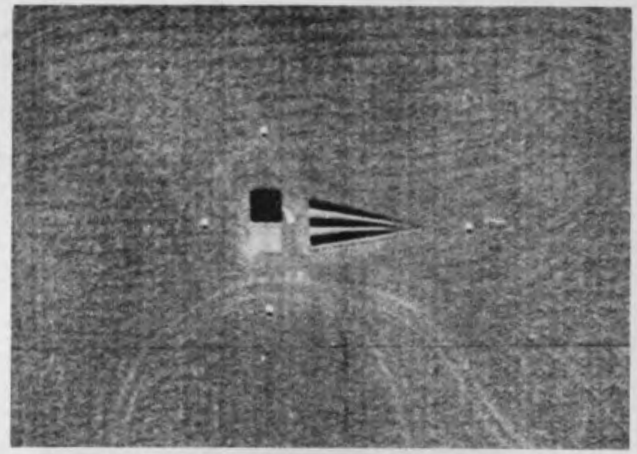


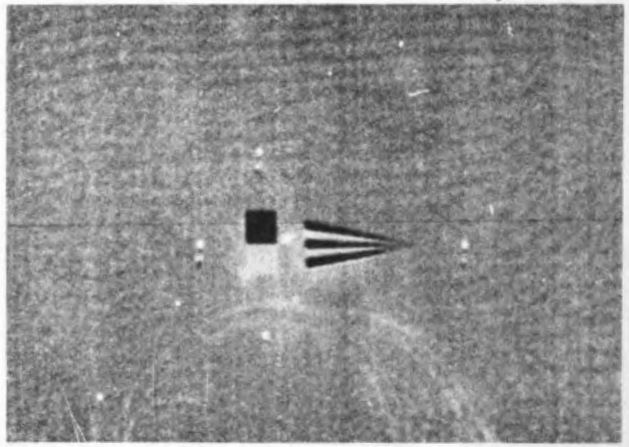
Figure 2 - Infrared ground test array configuration.



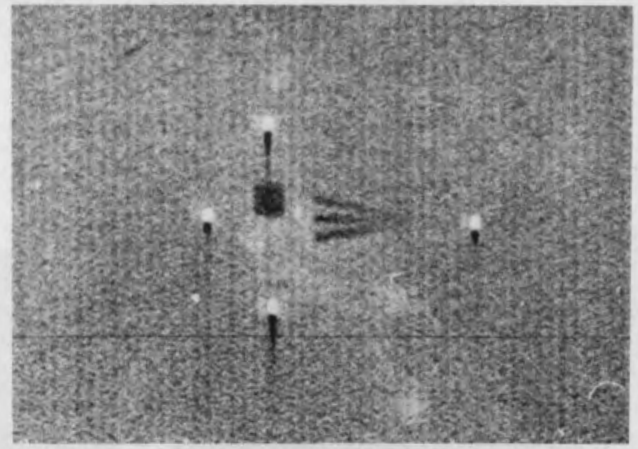
CH. 1



CH. 2



CH. 3



CH. 4

25 NOV 71, ALT = 500ft, $\Delta T \simeq 5^{\circ}C$

Figure 3 - Imagery from each channel of Reconofax XIIIA recorded at 1600 hrs, 25 Nov 71. The sky was overcast and there was snow on the ground.

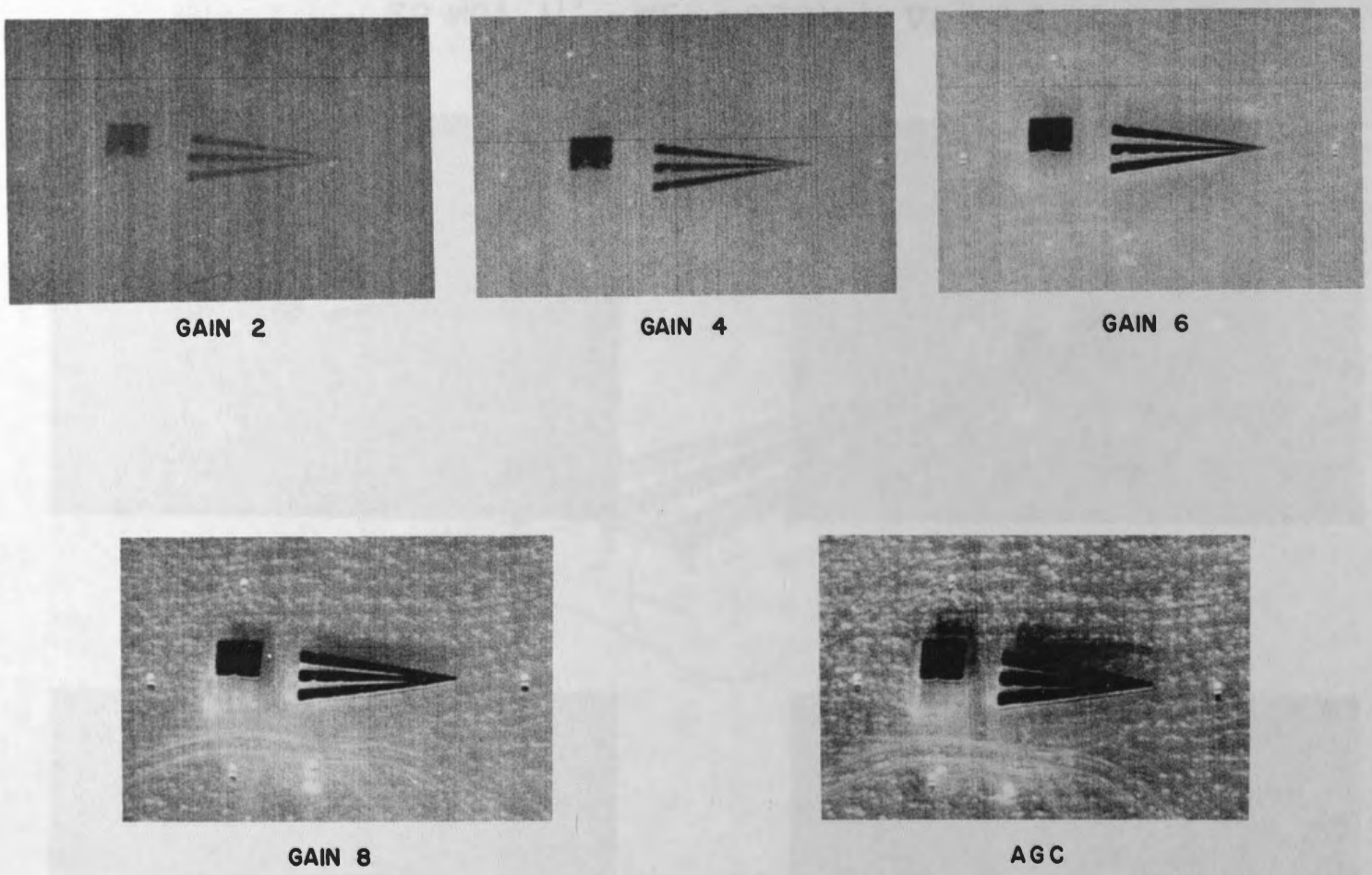


Figure 4 - Imagery from Channel 3 of Reconofax XIIIA using different gains, recorded at 1400 hrs, 24 Nov 71. The sky was broken and there was snow on ground.

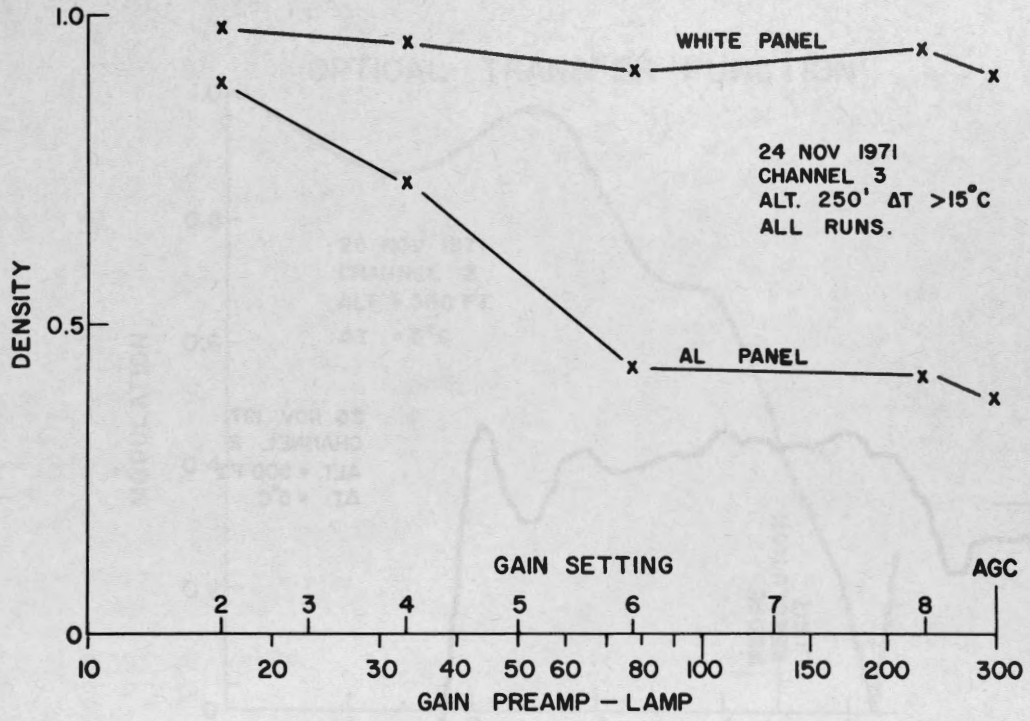


Figure 5 - Densities of white painted and sandblasted panels plotted against gain for imagery shown in Figure 4.

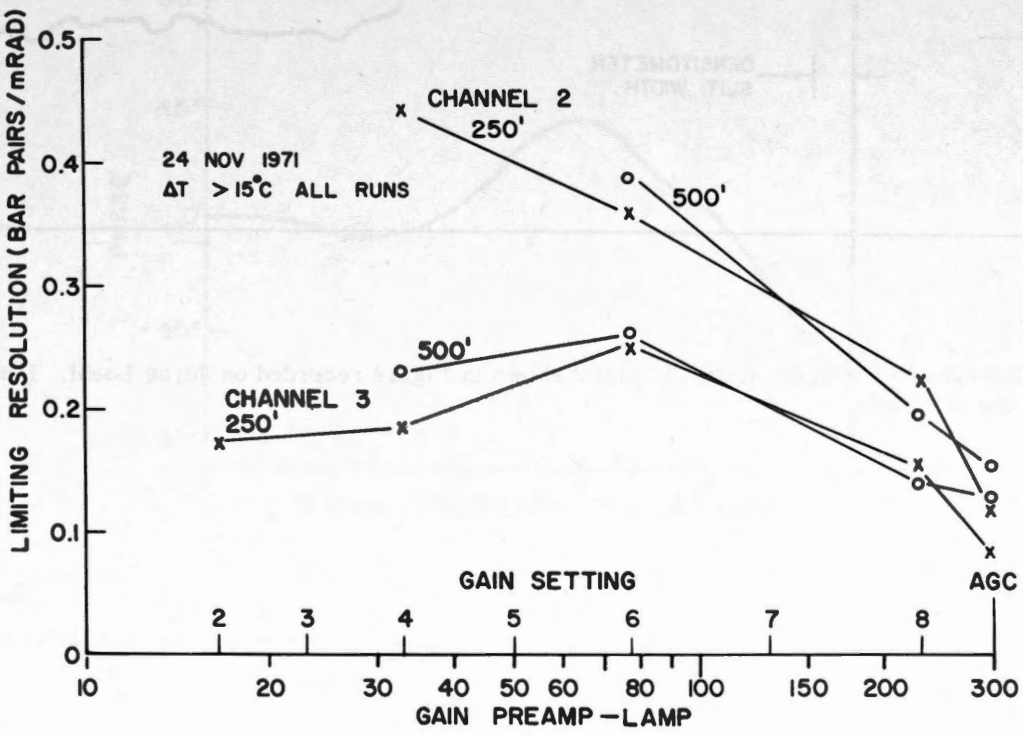


Figure 6 - Limiting resolution of Channels 2 and 3 of Reconofax XIIIA plotted against gain for imagery recorded 24 Nov 71. Data for bottom curve are from imagery in Figure 4.

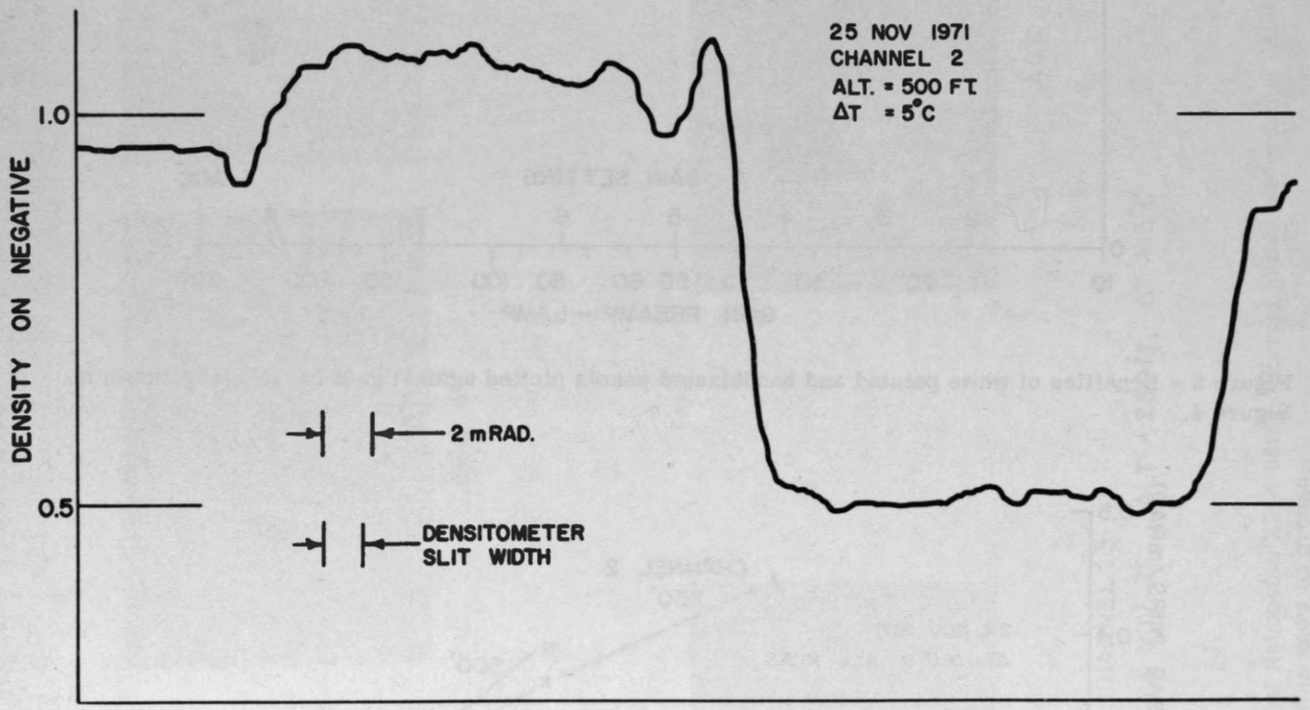


Figure 7 - Edge scan of imagery from Channel 2 shown in Figure recorded on Joyce Loebl. Densitometer using slit width of 40 um.

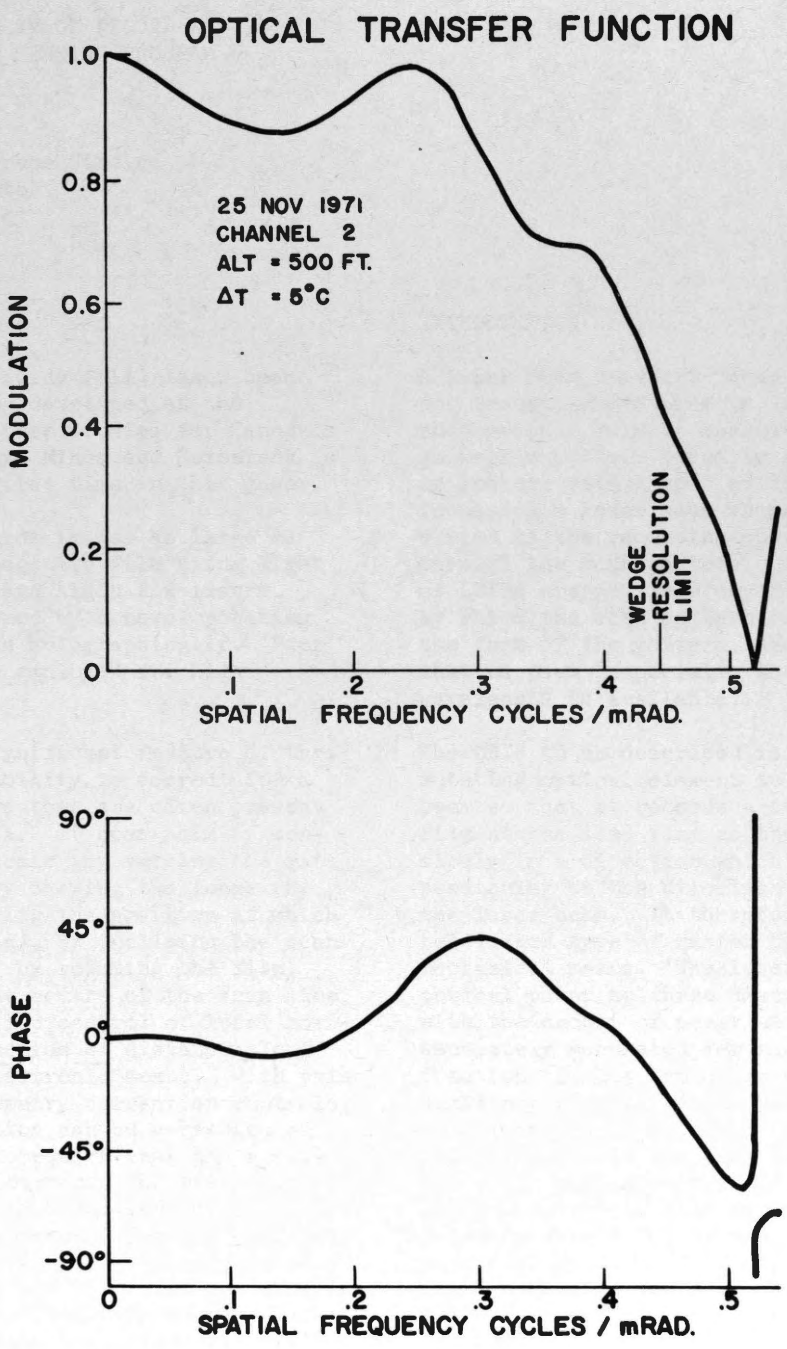
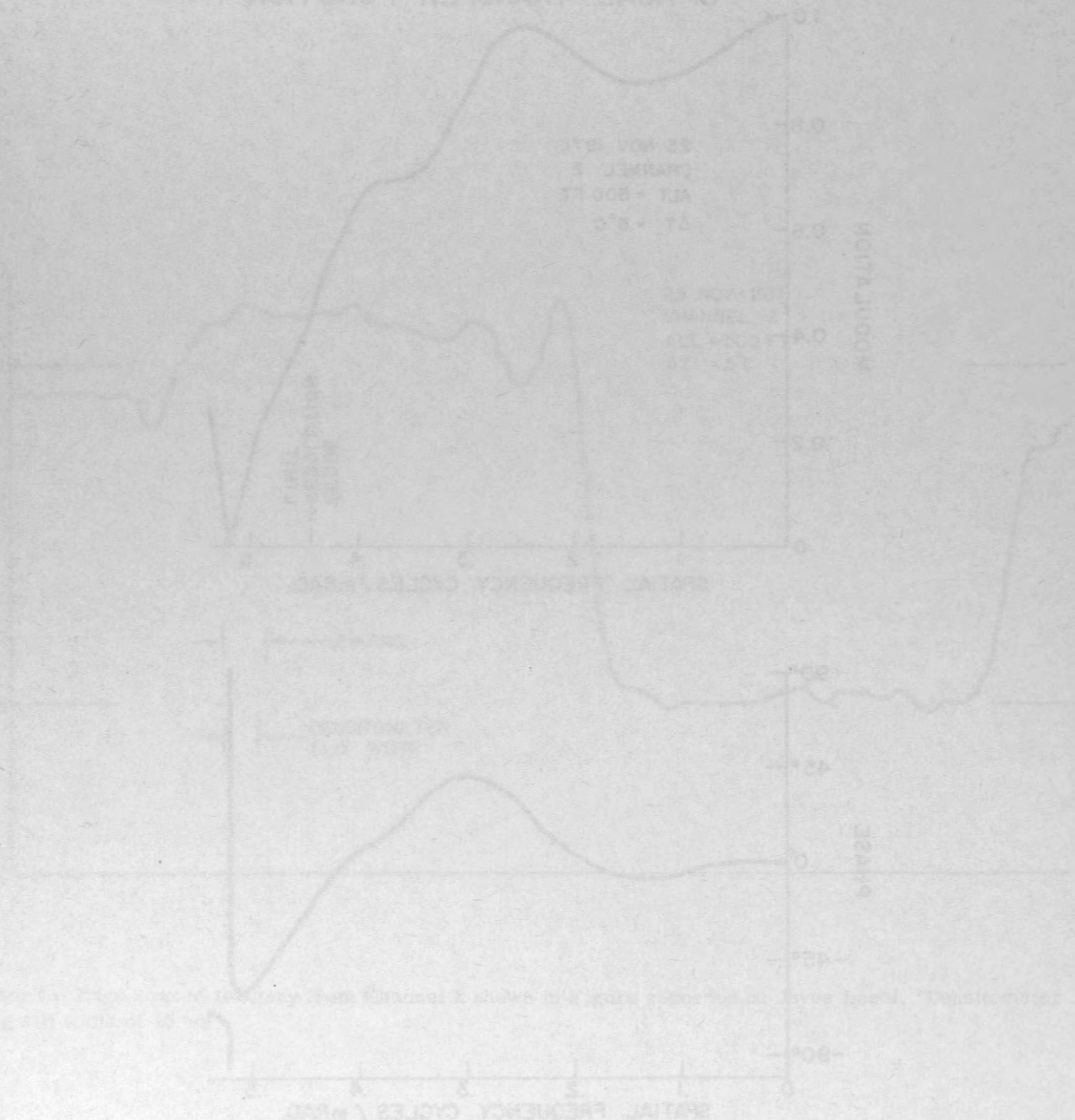


Figure 8 - MTF in transmission space of edge scan in Figure 7 after correction for densitometer MTF.



A COLOUR LASER BEAM IMAGE RECORDER FOR CANADA'S REMOTE SENSING PROGRAM

Dr. J. W. Locke
Institute for Aerospace Studies
University of Toronto
Toronto 181, Ontario

SUMMARY

A large format (9½" wide film) laser beam image recorder being developed at the Institute for Aerospace Studies for Canada's Department of Energy, Mines and Resources is described for the first time in this paper.

The instrument records images as large as 9 x 14" on colour negative film using light beams from Krypton and Argon ion lasers. These beams are formed by a novel rotating optical element made holographically. Four film transports are employed for high throughput.

Perhaps the most significant feature of this instrument is its ability to correct for a wide class of errors that are often present in line scanner data. It does this by controlling vertical scale (by varying the rate of film advance), by skewing the image (by electronically varying the position at which each line is started), by inclining the scan lines by up to ±15° by rotating the film transports about the centre of the scan line, and lastly by detailed control of local horizontal scale as function of distance along the scan line by electronic means. With this flexible set of geometry correction controls, any general distortion can be corrected so long as it does not require that the resultant scan lines be curved. The result is that data from the ERTS Multispectral Scanner can be completely corrected for the geometric distortions that originate with spacecraft angular rate errors and with scan nonlinearities. The instrument can cope equally well with line scanner data acquired by aircraft where its ability to correct for severe yaw angles and large amounts of S-bend horizontal distortion are important.

The image quality is compatible with recording 4000 pixels per 9"-long line. The geometric error introduced by the instrument itself is less than 1 part in 5000. Its writing rate is one 9 x 14" image per 25 seconds for each of the four film transports.

INTRODUCTION

A laser beam image recorder (LBIR) is a system for recording pictures as latent images on photographic film by scanning the film in a geometric pattern (usually a television type of raster) with a spot of light obtained by focussing a laser beam whose intensity is varied as the recording spot progresses through the scan pattern. Many possible types of LBIRs are possible depending on the means by which the scan pattern is brought about, the form of the pattern, and the type of laser that is used (especially whether more than one wavelength is available).

The LBIR to be described is one that uses a rotating optical element to deflect the laser beam so that it records a straight line on the film at the same time as the film is moved slowly in a direction which is nominally perpendicular to the direction of progression of the laser beam. It therefore generates a television type of raster by essentially mechanical means. The laser beam carries optical power at three discrete wavelengths with the amount of power at each wavelength separately modulated for the main recording function (it has two other wavelengths for auxiliary functions) and hence can record colour images on colour film in one pass. In brief, this LBIR has been designed to achieve extremely high geometric accuracy and resolution, extremely high radiometric and colourimetric fidelity and moderately high recording speed. It incorporates a surprisingly flexible set of controls by which, in response to commands from a control computer, deliberate geometric distortions of accurately-known form can be introduced into the image being recorded, generally to cancel known errors introduced by the original data source. Four film transports are provided in order that continuous-scan image data can be framed into discrete frames with nominally ten percent frame-to-frame overlap (two transports would be required for this) and in order that two different continuous-scan images can be simultaneously framed and

recorded by a multiplexing technique. Furthermore, each transport is equipped with a photodetector "behind" the film so that developed film already bearing images in transparency form can be scanned and digitized by a constant-power laser spot, converting the images from photographic to electrical form. The system is thus an image scanner and an image recorder.

The system should prove ideal for use by the Canada Centre for Remote Sensing for recording or scanning colour images from sensors on board spacecraft such as the ERTS A/B vehicles and aircraft. It combines, in a unique way, the extreme geometric fidelity that mechanical-scan LBIRs are potentially capable of without the usual rigid-raster inflexibility that has been characteristic of this device until now. The new dimension of geometry-perturbation control combined with the choice between direct one-step color recording or the image scanning and digitizing mode of operation makes this device extremely attractive to the Remote Sensing community. It is believed that this is the most complex LBIR ever constructed. In simpler form the system should be of commercial interest to computer users as a high quality graphics I/O device, to communicators as a graphics terminal (very attractive when used with communications satellites) and to users such as microcircuit manufacturers who wish to record precise patterns at high speeds. Table 1 summarizes the main characteristics of the LBIR. Fig. 1 shows the appearance of the main part of the instrument, the Optical/Mechanical Unit. (The electronic subsystems are concentrated in a separate console.)

GENERAL OPTICAL AND MECHANICAL DESIGN

Figure 2 shows the arrangement of the main elements of the scanning optics and the film transports. In the description below, the term "horizontal" will be used with respect to the scanning raster pattern to indicate the direction parallel to the scan lines and "vertical" for the orthogonal direction.

The heart of the system is a novel form of reflective hologram, designated the "holomirror", which is mounted on a high precision air bearing and driven by an electric motor. The "holomirror" has the property of simultaneously reflecting and focussing into several convergent beams the divergent beam of laser light that can be seen to illuminate it. It does this by diffracting the illuminating beam from a fixed, prerecorded pattern that has been generated on its nominally-flat surface during manufacturing by a holographic process. (That is, the pattern originates as the recording of the interference pattern of coherent light beams in a jig separate from the LBIR. The details of this process will be the subject of a future paper.) The holomirror pattern is such that there are twenty-four beams leaving its surface and lying close to a 120° cone. All of these beams come to a focus 19.100 inches from the rotational axis of the holomirror bearing. The beams are arranged in three closely-spaced (0.500 inches spacing) planes precisely perpendicular to the rotational axis in the peculiar set of azimuthal angles shown in the inset to Figure 2. Five wavelengths are involved and are present in the various beams as indicated in Figure 2.

The basic concept is simple. Only six of the beams are required to serve the main image-recording function. These are arranged to focus on the 19.100 radius circle in the "main-scene" plane at 60° intervals. The four films are held by vacuum against curved stainless-steel plattens spaced at 90° intervals on the same circle. The film is thus curved to conform to a 19.100 inch radius cylinder. At this radius the 9-inch useful width of the film subtends a nominal 30° at the rotational axis (actually 27° to allow for a synchronization interval). All of the beams rotate with the holomirror and hence the films are traversed sequentially in straight "horizontal" lines by the focussed radiation. If the film is moved slowly "vertically" the conventional scan raster results. Note that the six beams scan only one diametrically-opposite film transport pair at a given time and the scanning alternates between pairs with every thirty degree rotation of the holomirror. Thus two duplicate images are recorded on one diametrically-opposite transport pair and, in general, a different image can be recorded on the other pair also in duplicate.

The vertical drive for each of the film transports consists of a clamp on a precision carriage that grips the film at one point on each edge and pulls the film vertically as much as fourteen inches. When the clamp reaches the desired limiting position it opens, returns, reclamps and the cycle repeats. By using the diametrically-opposite transports in tandem (i.e. timing their advances suitably) the problem of framing continuous-scan data into overlapping frames can be solved. A mechanical shutter is

available at each transport to block the main-scene beams from the film when desired. During the overlap interval when it is required that the same scene material be recorded on both transports of a pair, both shutters are open. At other times only one or the other shutter is open.

The six main-scene beams contain the three wavelengths, 457.9, 530.9 and 647.1 nanometers that serve as the blue, green and red primaries for the colour recording process as discussed further below.

What are the other eighteen beams used for? The six "annotation" spots contain power only at 568.2 nanometers and can be used to record in monochrome alphanumeric characters above and below the framed main-scene images. These spots are in a plane offset by 0.500 inch from the main-scene spots so that a second mechanical shutter on each transport can control the recording time of the annotation information independent of the main-scene shuttering. The result is that different annotation data can be recorded on adjacent frames, a design goal. The twelve "grating" spots are not modulated and are used to illuminate the precision grating pattern, thereby providing a signal to a detector located behind the grating that is used as the horizontal position reference. This provides a function much like having a shaft position encoder of 5×10^4 counts/revolution. The grating beams contain power at 488.0 nanometers only. Since a grating signal is required continuously and space is not available for a longer grating, the grating length is such as to subtend 30° at the rotational axis and the grating spots are at the same 30° spacing, nominally, and must number twelve. For convenience in positioning the grating, the set of grating beams is rotated 15° from the position at which six of them would coincide in azimuth with the main-scene and annotation beams.

The film transports deserve some general comment. These are very long in the "vertical" direction because of the need to make a very gentle transition from the flat form the film must have on leaving the supply spool or entering the takeup spool to the 19.100 radius cylindrical form it has at the recording platten. (The transition lengths are such as to limit differential stretching of the film to one part in 10^4.) The vacuum plattens are operated at only -0.1 p.s.i. to provide a light hold-down force on the film and are polished to permit the film to slide over them. Two precision capstans provide the film tension only (the vertical rate is determined by the film-clamp carriage). As will be discussed further in the next section, one of the geometric distortion correction functions available is the ability to incline scan lines with respect to the horizontal by $\pm 15^\circ$. To implement this function, each entire film transport is mounted on an air bearing and rotates about the center of the main-scene scan line under the control of a servo-motor-driven lead screw. This is suggested in Fig.2 by the dotted transport rotated in "yaw" by the angle "θ".

GEOMETRIC CONTROL

Perhaps the most outstanding feature of this LBIR is the high standard of geometric accuracy to which it has been designed and the flexibility of the set of controls by which one can perturb the scan pattern from the normal raster.

What determines the absolute positional accuracy of this LBIR? The film transports have been designed to allow this to be determined solely by the absolute accuracy of the "vertical" and "horizontal" position references with a small degradation due to the finite position resolution of the yaw angular servosystem. The "vertical" position of the film is determined by the film-clamp which is on a carriage driven by a recirculating-ball-bearing lead screw. The carriage also carries the sensing head of an optical linear position encoder which provides the position feedback signal for the vertical servo. The control system used has zero steady-state position error at constant vertical velocity and thus the vertical accuracy is that of the feedback encoder, $\pm 1 \times 10^{-4}$ inch. The horizontal control system is referenced to the signal from the glass grating shown in Figure 2 by a control system which adds 1/32 of a grating period of position uncertainty to the maximum positional tolerance of $\pm 2 \times 10^{-4}$ of the grating crossings. The grating frequency is 455 lines per inch and the absolute positional accuracy of the horizontal control system is 2.7×10^{-4} inch. The finite resolution of the yaw angle encoder contributes 1.2×10^{-4} inches of vertical error at the extremities of scan lines degrading the vertical absolute accuracy to $\pm 2.2 \times 10^{-4}$ inches. Note that the nature of these errors is not such that they can be expected to change much with time.

Now let us consider the controls that are provided in this instrument that allow one to perturb the geometry of the raster from its

nominal form. The system is designed to process video data in digitized form with each scan line divided up into discrete picture elements, pixels, with the video data expressing the mean intensity to be recorded at each pixel location. We can consider each image as a nominally-orthogonal array of pixels which can be distorted from nominal form in the following ways:

(i) Vertical scale* is programmable in proportion to a control word over a 2:1 range and the scale may be changed every line if required. However, the servosystem that controls the vertical scale will not respond accurately to scale-variation spatial frequencies above ten cycles/inch.

(ii) Horizontal scale* is programmable in inverse proportion to a control word over a 4096:1 range. The scale may be changed arbitrarily 32 times per line without error.

(iii) The pixel array can be skewed arbitrarily by an offset in the horizontal position of the first pixel in a scan line programmable in direct proportion to a control word that can be changed arbitrarily with every line.

(iv) The angle of inclination of the scan lines with respect to their normal position perpendicular to the pixel columns (i.e. to the vertical), the yaw angle (θ), is programmable in increments of approximately six arcseconds up to $\pm 15^{\circ}$ in proportion to a control word.

These four controls are sufficient to provide for all anticipated distortions in data from line scanners on board aircraft and spacecraft (i.e. from the ERTS A/B Multispectral Spot Scanner) provided the distortions leave the scan lines straight. For example, in the common problem of correcting for the distortions created by vehicle attitude errors, control (i) is used to correct vertical scale for the effect of vehicle pitch rate, control (iii) corrects for the effect of roll rate and control (iv) for the presence of yaw angle.

The ability to correct the horizontal scale in detail along a scan line as noted in (ii) permits the correction of large amounts of S-bend, earth curvature, or instrument-introduced scan nonlinearities.

Correction (i) is implemented by a control system that controls the speed of the vertical leadscrew motor in such a way that the resultant vertical film velocity is always normalized by the instantaneous rate at which scan lines are being recorded. Thus the control word programs true scale not just vertical velocity, a significant point when a video data source such as a tape recorder may vary significantly in rate.

Correction (ii) is implemented by leaving the holomirror angular rate constant but varying the frequency at which the data is transferred from the source. This technique requires the storage of a significant portion of one line of data in a buffer memory so that the data for a particular pixel will be available at the time the LBIR recording spot reaches the required position for that pixel. The amount of buffer storage required is small compared with the storage required for other purposes at the Canada Centre for Remote Sensing such as the storage of twelve lines of ERTS A/B MSS scanner video data during line reordering. The control technique is purely electronic and employs a binary rate multiplier to multiply the grating signal frequency by the control word. The resultant correction to pixel locations is a true position correction independent of the rate at which scan lines are recorded. Furthermore, the resultant pixel positions are referenced to the grating and are as accurate as the grating.

Correction (iii) is also performed by processing the grating signal digitally and using it to determine the time at which pixel data is drawn from the buffer storage. The resultant offset is a true grating-referenced position-offset independent of line rate.

Correction (iv) is accomplished servomechanically. As noted earlier, each film transport is mounted on an air bearing and may be rotated about an axis perpendicular to the holomirror rotational axis, that passes through the centre of the main scene scan line on the film. The yaw angle, θ, is determined by a ball-bearing leadscrew actuator that can be seen in Figure 1. The actuator is driven by a servomotor in a control system that uses a shaft position encoder on the yaw leadscrew for feedback. The resultant angle is thus referenced to the lead progression of the screw. The control-word/yaw-angle transfer function is slightly

* Scale is defined as the pixel spacing, so that as scale increases a given object in the scene increases in size.

nonlinear as determined by the geometry of a linkage between the leadscrew nut and the film transport.

RADIOMETRIC AND COLOURIMETRIC DESIGN AND CONTROL

For remote sensing applications as well as others, the fidelity of the LBIR in transforming intensity data from electrical form to the exposure given the film in recording mode or in recovering density data in electrical form from image input in scanning mode, is just as important as geometric fidelity. By careful design, a LBIR can perform to a high standard in this respect as well. This LBIR is designed to accept intensity data quantized to eight bits and introduce no intensity error exceeding 1%. Before describing the control systems that achieve this, let us consider the lasers and wavelengths and film chosen for the system.

The films for which the LBIR is designed are Kodak Aerocolor Negative Film 2445 for colour operation and Kodak Plus-X Aerographic Film 2402 for black and white work. The critical choice is that of the Aerocolor 2445 material. It has been selected because it has adequate resolving power with a modulation transfer function above 90% at the highest anticipated data spatial frequency, 8.94 cycles/mm and from spectral sensitivity curves provided by Kodak, it is almost perfectly matched in spectral response to a set of three wavelengths available from ion lasers - 457.9 nm from Argon ion lasers and 530.9 nm and 647.1 nm from Krypton ion lasers that have been selected as the colour primaries. The choice of the black and white Kodak 2402 was based on its resolution and sensitivity being more than adequate for the purpose.

Three lasers are used in this LBIR to get high optical power reliably. These generate the above three colour primaries plus the annotation beam wavelength, 568.2 nm., and the grating horizontal position reference beam wavelength, 488.0 nm. The three lasers are: an Argon laser operating at 457.9 and 488.0 nm; a Krypton laser at 530.9 and 568.2 nm; and a Krypton laser at 647.1 nm.

The modulation of the intensities of the primary and annotation wavelengths is performed by first separating out the five desired wavelengths by prisms, directing the four beams to be modulated through Pockels-effect electro-optical modulators which act as electrically-operable attenuators. The modulators allow the video data to control the laser beam intensities. It is not sufficient to simply apply a control voltage to the modulator and use whatever intensity beam emerges from the modulator output, however, when one requires radiometric fidelity to better than one percent (and indeed to even keep errors below fifty percent) one must deal with the following properties of lasers and modulators:

a) While incredibly stable in wavelength, ion lasers like many other types, exhibit random output power drifts with time that are of the order of two to one long term (on a time scale of a month, say) with broadband intensity noise in the 10Hz to 2MHz frequency range of the order of one or two percent r.m.s. as a fraction of mean intensity. If no attempt is made to deal with the laser noise and drift, these effects will be observed at the output in direct proportion.

b) Electro-optical modulators of the transverse-field Pockels effect type have a nonlinear transfer function with the output intensity being a $\sin^2$ function of the control voltage.

c) The transverse-field modulators are highly temperature sensitive and in a normal laboratory environment show zero drifts of the same order as the full-scale electrical signal.

To combat these difficulties a rather complex feedback system is used wherein the output of the modulator is sampled by a photodetector and compared with an analog representation of the desired output. At low frequencies (below 80KHz) it is relatively easy to introduce sufficient loop gain to cancel errors due to laser noise and drift and to modulator nonlinearity. Furthermore, a simple control system can be used to generate a voltage that cancels out the temperature-dependent modulator offset. The tough problem is dealing with the nonlinearity and the laser noise and drift problems insofar as they affect the performance from 80KHz up to 10MHz. It has been found necessary to accurately cancel the nonlinearity of the modulator by a compensating non-linear preamplifier, a "predistorter", and introduce an electronic servo-attenuator that automatically normalizes the forward gain of the intensity control loop by the value of the mean laser intensity to ensure that the high frequency transfer function from video input to intensity output is independent of laser intensity drift over a two to one range.

The result is a system for converting from video input to optical output that is flat to 1% from d.c. to 1MHz and is down by 3db at 10MHz with an optical signal to noise ratio of 54db, 10Hz to 2MHz.

The outputs of the four modulators are recombined by a prism into the single, modulated beam that is focussed by the holomirror onto the films.

CONCLUSIONS

A complex but high performance laser beam image recorder will go into service in August 1972 at the Canada Centre for Remote Sensing that should prove ideal for the demanding image handling requirements generated by new sensor technologies. It will deal effectively and productively with multispectral data generated by most aircraft and spacecraft sensors. Variations on this form of LBIR may well find commercial use in graphical communications and computer I/O applications.

Table 1 LBIR Performance Summary

Format: 9x 14" on 9 1/2", 4 mil Estar base film

Radiometric Accuracy: ±1%, d.c. to 1MHz

Bandwidth: 1MHz ±1%, 10MHz -3db

S/N Ratio: 54db, 10Hz to 2MHz

	Horizontal	Vertical
Resolution: (limiting, not incl. film)	36 cycles/mm.	15 cycles/mm.
Absol.Geom.Accuracy:	$\pm 2.7 \times 10^{-4}$ inch	$\pm 2.2 \times 10^{-4}$ inch
Typical Raster:	2.20×10^{-3} inch/pixel 4086 pixels/9" width	3.12×10^{-3} inch/pixel 4490 pixels/14" ht.
Typ. Recording Rate: (on Kodak Aerocolor Negative Film 2445)	741×10^{3} pixels/sec.	81.6 lines/sec. 0.25 inch/sec.

LASER BEAM IMAGE RECORDER

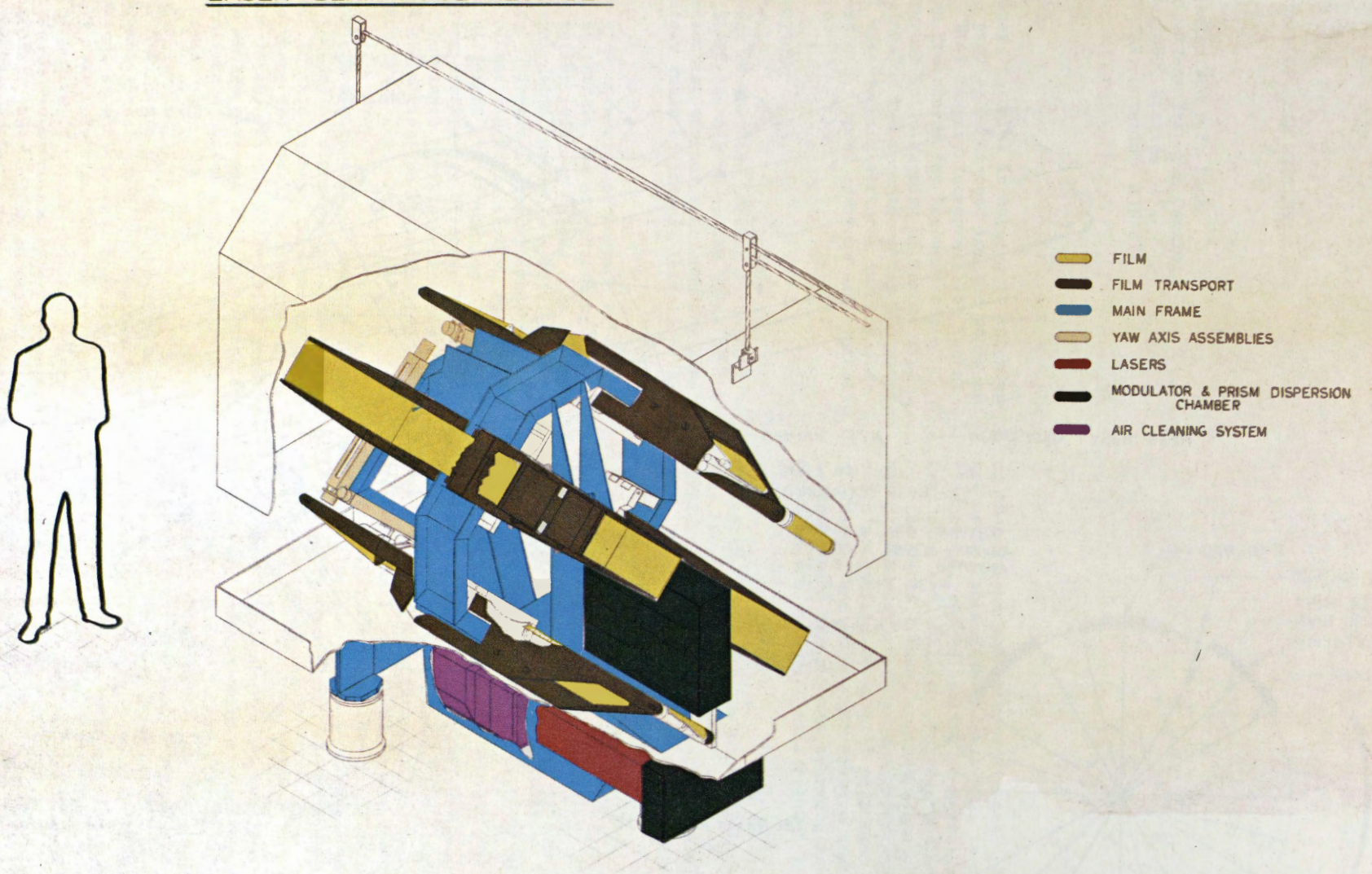


FIG. 1

GEOMETRY & LAYOUT OF LASER BEAM IMAGE RECORDER

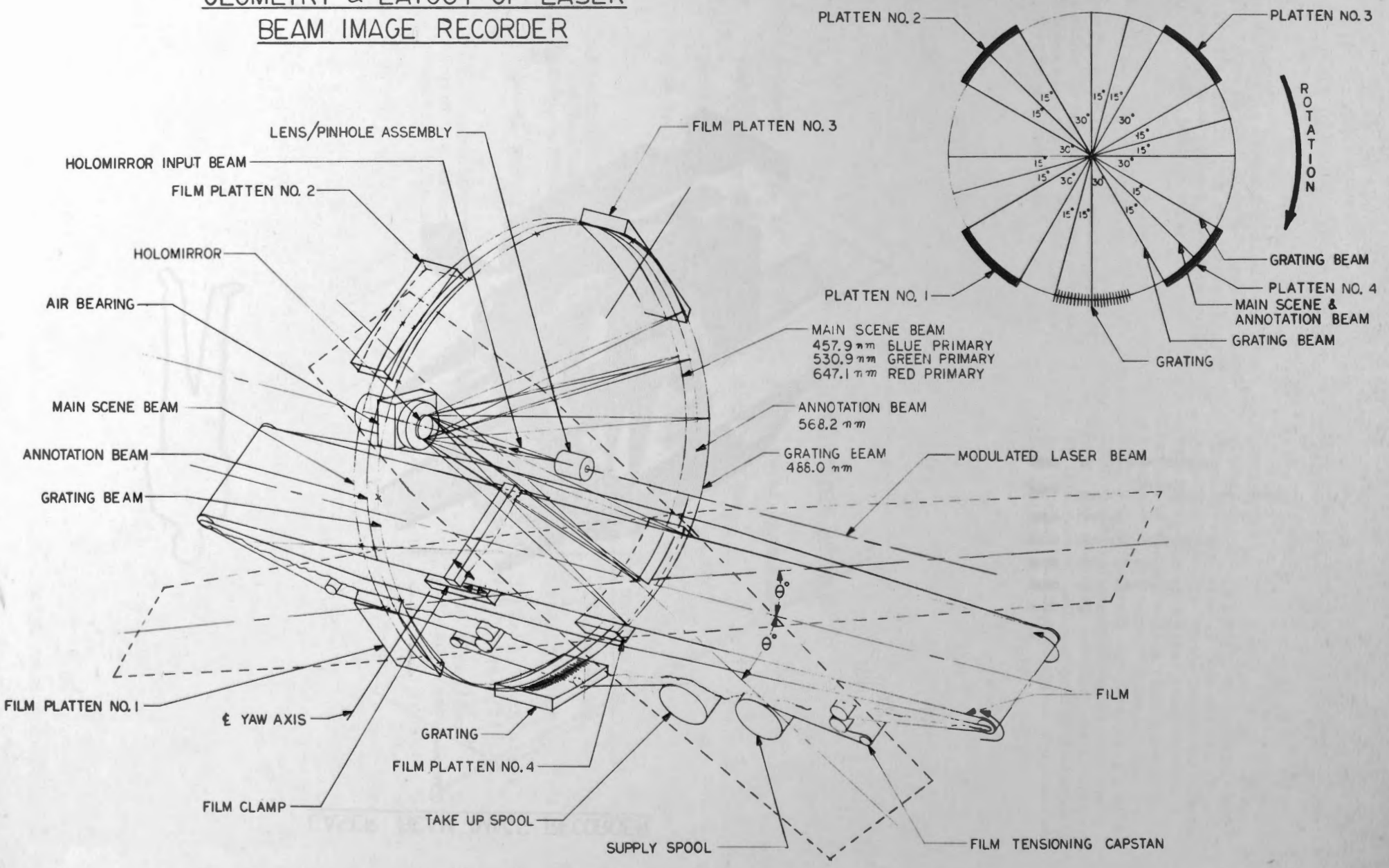


FIG. 2.

PRINCIPLES OF SLAR SYSTEMS

G.E. Haslam and D.F. Page
Communications Research Centre,
P.O. Box 490, Station 'A',
SHIRLEY BAY, Ontario. K1N 8T5

ABSTRACT

This paper, totorial in nature, is intended only to provide background for the paper entitled "SLAR Imagery Interpretation Techniques and Procedures for Reconnaissance Forestry and Soils Mapping of Remote Typical Regions". Two techniques for the generation of side-looking airborne imagery are explained here: real aperture and synthetic aperture.

INTRODUCTION

The advent of the Side Looking Airborne Radar (SLAR) has placed in the hands of the earth scientist another tool enabling him to remotely sense additional terrain parameters.

When we view an aerial photograph of say, a strip of terrain, we are observing the terrain reflectivity as mapped onto the photographic film through a specific transformation. Likewise, a radar image is a mapping of terrain reflectivity, but through quite a different transformation. It is the varying sensitivity of these transformations to intrinsic terrain parameters which provides the motivation for employing different sensors.

Radar systems are capable of imaging at any time of the day or night and in almost any weather. In this respect they are unique for uses where the timely collection of information is important. Modern side looking airborne radars have the ability to provide broad continuous aerial coverage on an economical basis.

To quote R.K. Moore of the University of Kansas Remote Sensing Lab[1], "Engineers and physical scientists tend to believe that they must understand the complete mechanism before it can be applied. Earth scientists, however, do not find this necessary. In fact, they seldom understand in detail the reasons for particular grey tones on photographs and infrared images; yet they are aerial photographs for a wide range of applications, and collecting of aerial photographs is a $2 billion per year business throughout the world. Thus, engineers and physical scientists must consider the importance to the earth scientists of spatial relations shown on images, and should not believe that failure to understand the physics behind a particular radar return situation will prevent its wide application".

Currently, there are two distinct types of Side Looking Airborne Radar systems: the REAL APERTURE system and the SYNTHETIC APERTURE system. The real aperture, being the simpler of the two, will be considered first.

REAL APERTURE SLAR

This radar is comprised basically of four elements: antenna, transmitter, receiver and recorder.

The antenna, mounted on an airborne platform, is directed perpendicular to the flight path. (See Fig. 1) Characteristics of the antenna are important; beam shape in the vertical plane is relatively broad so that it illuminates a swath of terrain to the side of the aircraft; in the horizontal plane, the beam shape is very narrow and fan-shaped.

The antenna is connected instantaneously to the transmitter and transmits a short burst of electromagnetic energy. (See Fig. 2) The energy propagates through space (at the speed of light) and strikes the terrain. Some of the energy is absorbed, the rest is reflected by the terrain. Of this energy, some is reflected in the direction of the antenna. The antenna by this time is connected to the receiver. The receiver converts this energy to an electrical signal which in turn modulates the intensity of the electron beam in a cathode ray tube. The beam of the cathode ray tube is swept

from left to right such that its position at any time corresponds to the position of the electromagnetic energy propagation across the strip of terrain. The cathode ray tube trace is recorded on photographic film which is moving past the face of the tube at a rate proportional to aircraft ground speed. It should be noted that the time when energy is returned to the antenna is related to the distance of the reflecting object from the aircraft. It is this fact that allows objects displaced in range to be separated on the film.

As the aircraft progresses along its flight path, the radar alternately transmits and receives. As a result, the reflectivity of the illuminated terrain is sequentially sampled and recorded on the photographic film.

A fundamental limitation of the REAL APERTURE SLAR system is the degradation of ALONG TRACK or AZIMUTHAL resolution with increasing range. The beamwidth of the antenna is fundamentally related to the size of the radiating aperture. (See Fig. 3) In addition, the beamwidth defines the along track resolution for real aperture systems. Therefore, the generation of high resolution imagery at long range requires the use of very long antennas and, of course, there is a limit to the length of antenna that can be transported by an airborne platform.

SYNTHETIC APERTURE RADAR

The synthetic aperture system circumvents the problem of having to fly very large antennas and has the ability to produce much higher along-track resolution. The technique is to fly a relatively short antenna and at discrete positions along the flight path, gather both phase and amplitude information about objects in the beam of the antenna. This information is stored and summed. The result is the same as if this information has been gathered by a very long antenna array; the resultant beamwidth corresponds to that of an antenna array of the same length as the distance over which the phase and amplitude information was collected. A limit to the achievable synthetic antenna aperture length is the accuracy with which the flight path can be flown. Apertures of several thousands of feet are not impossible, but at say, X-band, the flight path must be straight to about 1/2 an inch. There are schemes which electronically correct the flight path deficiencies; however, a high degree of flight stability and accuracy are still required.

This points out a rather fundamental difference in the way that real and synthetic aperture systems are deployed. Real aperture systems are generally flown at lower altitudes, say under 15,000 feet. Synthetic aperture systems, which do not suffer from azimuth degradation with range are generally flown at higher altitudes. Here they can get above the weather and fly the straight and stable flight paths required for high resolution.

REFERENCES

(1) R.K. Moore. Radar Imaging Applications. University of Kansas Remote Sensing Laboratory, Lawrence, Kansas.

(2) R.O. Harger. Synthetic Aperture Radar Systems, Theory and Design. Academic Press, New York, 1970.

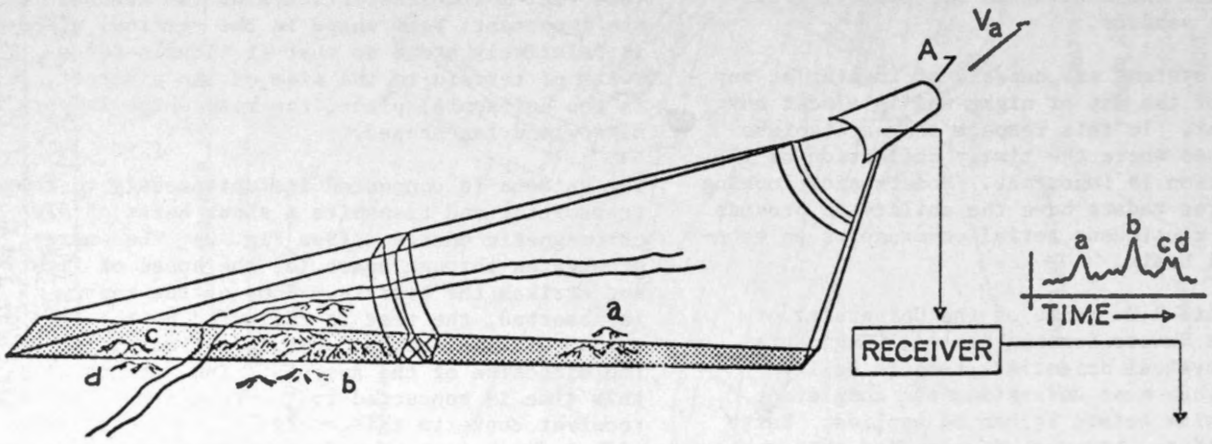


Fig. 1 SLAR antenna geometry. Antenna A is moving at velocity V_a.

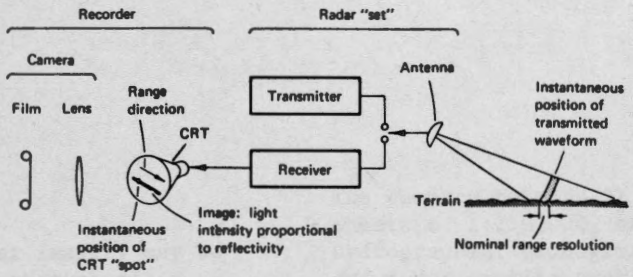


FIG. 2 An elementary range dimension imaging system.

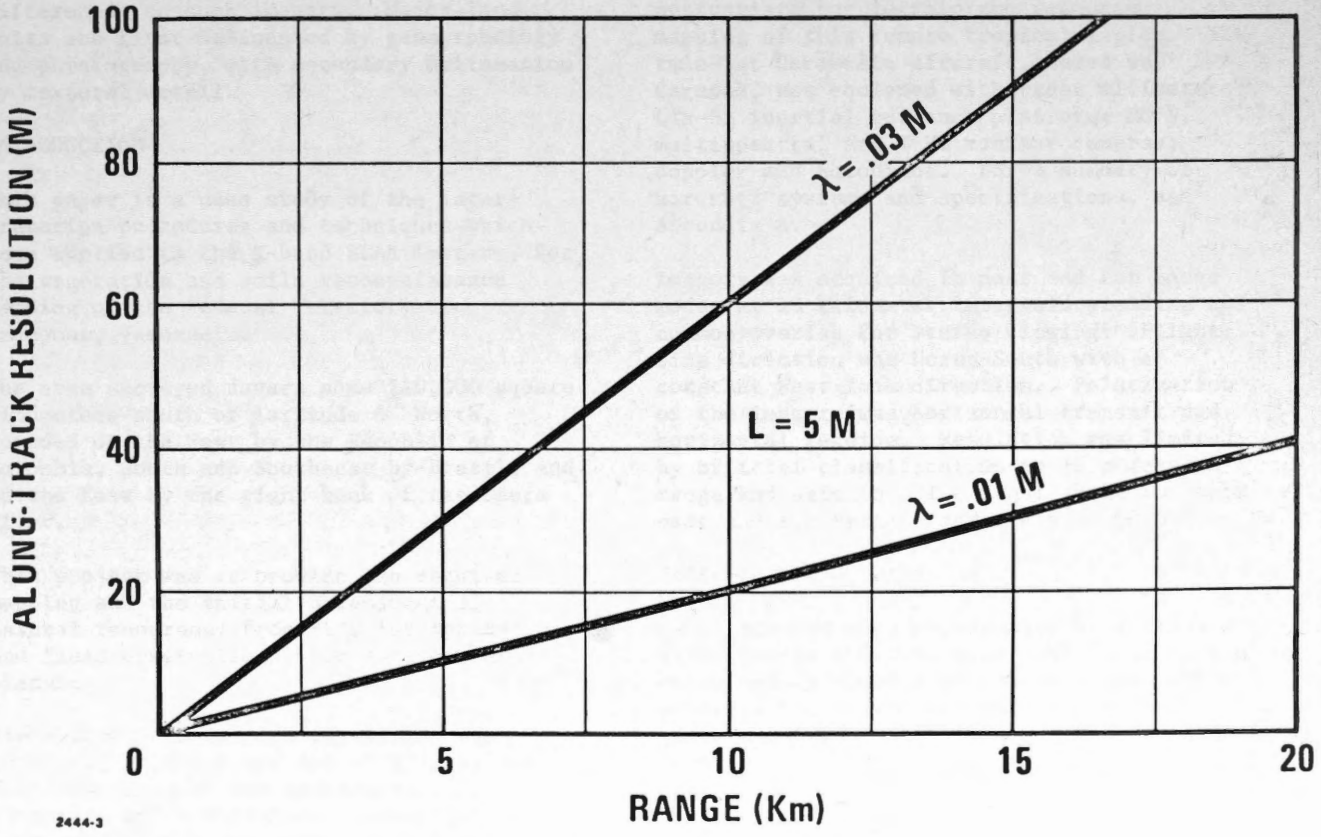
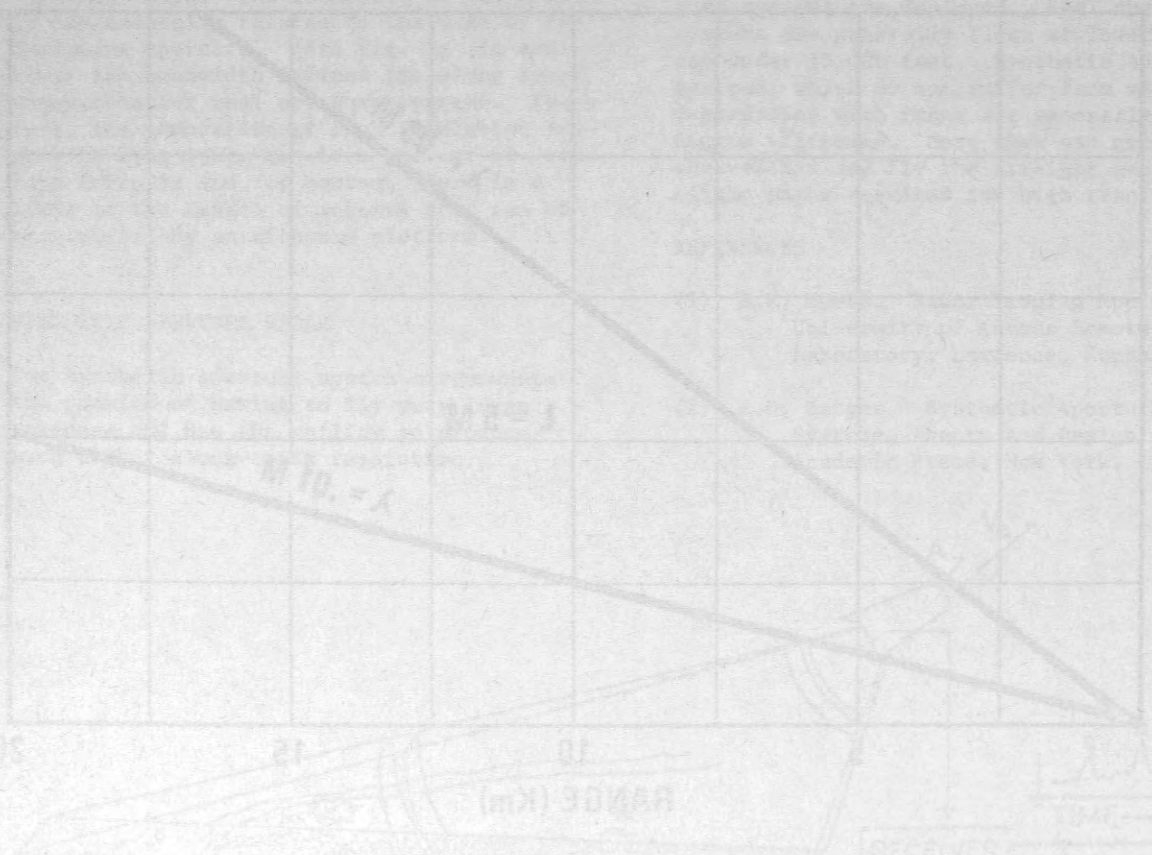


Fig. 3 Resolution of real aperture SLAR for antenna of length 5m operating at wavelength λ = 0.03m and 0.01m.



SLAR IMAGERY INTERPRETATION TECHNIQUES AND PROCEDURES FOR RECONNAISSANCE FORESTRY AND SOILS MAPPING

by

P.M. Addison[1]

ABSTRACT

Side-looking Airborne Radar imagery may be acquired day or night, or through cloud cover, when aerial photographic operations would be impossible. This imagery is particularly valuable for reconnaissance studies and the mapping of remote regions.

Special reference is made to soils and vegetation interpretation by stereo viewing and by the recognition of textural differences on such imagery. Major land units are first delineated by geomorphology and physiography, with secondary delineation by textural detail.

INTRODUCTION

This paper is a case study of the interpretation procedures and techniques which were applied to the X-band SLAR imagery, for the vegetation and soils reconnaissance mapping of the Federal Territory of Amazonas, Venezuela.

The area surveyed covers some 210,000 square kilometers south of latitude 6° North, bounded on the West by the Republic of Colombia, South and Southeast by Brazil, and on the East by the right bank of the Caura River.

This project was to provide the required mapping and the initial assessment of natural resources, from SLAR interpretation and field observations, for development planning.

Field work in soils and vegetation was conducted in April and May of 1971 by one team comprised of two pedologists, a forester, and a botanist. Investigations were conducted principally along the Rio Orinoco and some of its major tributaries.

The summary report, 21 radar mosaic map sheets at 1:250,000, and accompanying hydrographic, geological, soils, forestry and radar profile overlays, were delivered to the Government of Venezuela, according to contract, by International Aero Service Corporation on September 30, 1971.

ACQUISITION OF SLAR IMAGERY

The characteristics of Goodyear Aerospace synthetic aperature SLAR, were particularly appropriate for terrain and resource mapping of this remote tropical region. The twin-jet Caravelle aircraft, based in Caracas, was equipped with radar altimeters; LTN-51 inertial guidance platform; RC 9, multispectral and T.V. monitor cameras; doppler and autopilot. For a summary of aircraft systems and specifications, see Appendix A.

Imagery was acquired in near and far range modes at 18 kilometer intervals yielding 60% common overlap for stereo viewing. Flight line direction was North-South with a constant West look direction. Polarization of the imagery was horizontal transmit and horizontal receive. Resolution was limited by official classification to 16 meters on range and azimuth. The final scale for both near and far range imagery was 1:400,000.

PRESENTATION OF DATA

Radar mosaics were constructed at a scale of 1:200,000 to a U.T.M. grid with field control being mainly Venezuelan Boundary Commission geodetic and astronomic points. Only far range imagery was utilized for mosaic construction.

All interpretation of the SLAR imagery was compiled directly on the mosaics by the stereo viewing of the near-range strip and the appropriate far-range strip in the

---

[1] Consultant for remote sensing imagery interpretation and analysis, Ottawa.

mosaic at a like scale. Separate mosaics were used for hydrological, physical and structural geological, forestry and vegetation, soils, and radar profile interpretation procedures.

INTERPRETATION OF SOILS AND VEGETATION

The image interpretation may best be classed as of a "reconnaissance type", specifically designed to recognize productive and non-productive areas for agriculture and forestry development planning.

Data from ground observations were limited to the major navigable waterways. Some aerial photography provided supplementary coverage of the northern portion of the project area. A number of carefully catalogued overflights were used for control. Observations from recent and past expeditions into the Orinoco interior were carefully collated during interpretation.

A marked correlation between land types, vegetation and soils to the geomorphology and landforms became evident. Delineation relative to physiography and climate provided supplementary divisions. From these major divisions a logical "reconnaissance" classification evolved, and correlated closely with textural characteristics of the "forest canopy radar shadows".

Tone variations or radar reflection, bore greatest significance in the more open non-forested areas. This seems to be due to less return reflection on open areas and more scattering of radar energy on forested areas.

SOIL CLASSIFICATION

Most of the map units are broad soil associations regulated by local geomorphology and physiography. Stereo interpretation of these land units was correlated with some 64 field soil samples and profiles resulting in this classification.

Soils of seasonally or perennially wet plains

A-1 Alluvial soils, undifferentiated

S1 - Figure 1: This unit occurs along large rivers and is usually composed of ancient alluvial deposits which in this particular case is seasonally flooded by the Rio Orinoco.

A-2 Yellow and gray soils

S2 - Figure 1: This unit is largely ancient alluvial plains, slightly eroded, and periodically flooded in some regions. The recent evolvement of savannah soils is typical of this unit (possibly resulting from fire).

A-3 Sandy soils (poorly drained).

A-4 Marsh, swamp and highly organic soils.

Soils of eroded or sandy plains

B-1 Sands (well drained).

B-2 Red and yellow soils on highly dissected plains.

B-3 Eroded red and yellow soils, commonly with ironstone concretions, on nearly level to rolling plains, and associated rockland.

B-4 Eroded red and yellow soils on strongly rolling plains, and associated rockland.

Soils of uneroded and slightly eroded plains

C-1 Red and yellow soils on nearly level to rolling plains.

S3 - Figure 1: A typical uneroded plain elevated above the influence of the Rio Orinoco and having some traces (in the north and east of the unit) of the ancient peneplain.

C-2 Red and yellow soils on nearly level plains among steep hills.

S4 - Figure 1: A typical example of an eroded, nearly level plain, with some remnants of the peneplain comprising hills.

C-3 Red and yellow soils on low hills.

S5 - Figure 1: A typical soils unit associated with low, well eroded hills.

Soils of high plains and mountains

D-1 Red and yellow soils, moderately deep and associated rockland on steep hills and talus slopes.

S6 - Figure 1: A soils unit associated with moderate to steep hills and colluvium resulting from the erosion of talus. In this case soils are relatively deep.

D-2 Red and yellow soils, shallow and moderately deep, and associated with rockland and mountains.

S7 - Figure 1: An example of soils moderately deep in the valley to shallow on the hills, typically associated with the moderate erosion of the peneplain.

D-3 Rockland and shallow soils on rough mountains.

S8 - Figure 1: An example of shallow soils on high, rough mountain areas.

VEGETATION AND FORESTRY

The following classification of vegetation types and subtypes was developed by dividing the major recognizable features into lowlands, uplands, mountains, savannahs and marshes.

In this classification, single numbers represent a broad vegetation type, and a number accompanied by a letter indicates a subtype. Lowland forest is represented by the number "1", while a riparian or periodically inundated forest area is identified as "1a".

Classification of Vegetation Units

1 Lowland forest

1a Periodically flooded or riparian forest (Varzea Forest)

1b Permanently inundated or wet forest (Ipago Forest)

2 Upland forest

2a Highland riparian forest complex

2b Highland montane forest complex

3 Summit and Tepui vegetation

3a Montane savannah

4 Savannahs

4a Dry phase savannah

4b Wet phase savannah

5 Marsh and Aquatic vegetation

Complex types are identified by the use of primary and secondary numbers combined in accordance with their degree of relative importance.

Classification Description

1 . Lowland Forest. This vegetation class occurs on areas adjacent to major rivers and streams. In some cases it merges into relatively flat upland areas. Because of its situation it may be either periodically flooded during the wet season, or covered permanently by flood or stagnant waters of variable depth.

V1 - Figure 2: Typical "1" Lowland Forest with a smooth and uniform textured canopy. Tone appears mottled due to slightly undulating topography causing a variation of the SLAR's view angle.

V2 - Figure 2: This lowland Forest cover occurs on a slightly more elevated and fresher site than VI, resulting in an even, though less uniform, textured canopy. This texture is usually associated with larger individual crowns than VI.

1a. Periodically inundated lowlands or Varzea forests may develop into Gallery type of forest when accumulated silt carried by the rivers builds up into levees; vegetation then becomes characteristically different. Another type of gallery forest may also occur along streams in savannahs determined in this case by height of the water table. Such gallery forests are easily distinguished.

V3 - Figure 2: Lowland Forest "1a" or Varzea forests typically have a mottled and distinctly broken canopy texture. Small "holes" appearing in the broken canopy are usually small Indian settlements. Under stereo viewing the ceiba or kapok tree crowns, some achieving a diameter of 50 meters, stand out as individually resolved trees on the imagery.

1b. Permanently inundated or Ipago forests are almost perennially inundated. Appearance of the canopy is extremely uniform, due to limited number of species with almost similar heights and crown form. Such areas are frequently interwoven with channels. Separation of types 1, 1a and 1b is based essentially on observation of tonal quality indicating various reflectivity patterns derived from irregularity of the canopy and the presence or absence of tree storeys.

2 . Upland forests occur on well drained sites above the reach of flood waters. Terrain is usually flat to rolling. Forest characteristics on SLAR imagery indicate areas of broken continuity, the highest shadow corresponding to tallest growth of trees through the canopy level, and usually related to fresher and more fertile soil conditions.

V4 - Figure 2: The degree of terrain erosion and dissection is moderate to extreme in this Upland Forest type. Canopy texture varies with the broken terrain and the tree crown heights.

2a. Highland riparian complex forests occur along valley topography rising from a normal upland situation towards mountain conditions. This gives rise to a multiplicity of drainage channels associated with extremely high rainfall. Three main aspects can be recognized: V-type valleys, U-type valleys, and broad colluvial valleys.

In the V-type valleys, the slopes are covered by trees of fairly uniform height; in the U-type, vegetation is lush at the bottom of the valley and sparse on the rocky ledges and talus slopes bordering the almost vertical sides. The bottoms of these valleys occasionally show dry savannah areas possibly associated with exposure and lower rainfall.

The broad valleys have flat to slightly undulating topography. They display uniform tree vegetation of considerable size growing over moist soils, followed by a broken canopy resulting from the development of a two-storeyed forest over the colluvial material. Toward the head of such valleys, as they increase in altitude, the forest cover tends to merge with montane type forest vegetation.

V5 - Figure 2: This "2a" is a typical broad, U-type valley Upland Forest complex. Texture of the forest cover is rough or broken throughout the entire valley.

V6 - Figure 2: This "2a" is a V-type valley with moderately uniform forest canopy texture on the valley sides, and with a broken canopy texture in the valley bottom along the major drainage channel.

2b. Highland montane forest. This type develops at higher elevations, often on highly dissected and eroded peneplains, but more generally on less eroded mountains. Canopy texture appears rather uniform, in general, so that density of the crown pattern is somewhat related to that of the broad valley flats, except when affected by broken topography.

V7 - Figure 2: This "2b" Montane Forest type is on the steep talus and colluvial slopes of Cerro Duida, a major mountain and outstanding geological formation. Canopy texture is uniform except when broken by drainage channels.

V8 - Figure 2: A typical "2b" Montane Forest with a broken canopy occurring on a

highly dissected peneplain. This causes areas of high and low tonality relative to the aspect of the SLAR look angle.

3 . Summit and Tepui[2] vegetation develops at high altitude and therefore gives rise to contrasts in reflection which are largely due either to conditions of high humidity or to dryness of soils. Texture of imagery here proves more reliable than tonality.

Vegetation consisting of clumps of small trees in depressions, is easily distinguished from surrounding elements of bare rock with little or no vegetation. It can therefore be easily delineated since rocky surfaces give high reflection and details become obliterated.

V9 - Figure 2: A typical "3" Summit vegetation unit associated with little or no canopy reflection due to short, dispersed vegetation on a rocky background. These isolated "tepuis" plateaus occur 1500 to 2000 meters above surrounding terrain, almost continually under cloud cover.

V10 - Figure 2: This "3(2b)" is a combined Summit and Montane Forest vegetation group. The Montane Forest occurs on lower slopes and in the valleys where soils are more moist and deep.

3b. Highland savannahs are formed of low-growing shrubs that develop on either wet organic matter or dry coarse sands or pebbly rock material. In the first case, the vegetation will appear in darker tones than in the second instance where reflection is higher.

4 . Savannahs can be easily distinguished from the surrounding forest vegetation and their extension be determined with a high degree of accuracy. However, they possess variable tonal characteristics which help to determine whether they belong to a dry or a wet phase type.

V11 - Figure 2: A "4(1)" vegetation unit predominantly Savannah surrounded by Lowland Forest. The Savannah has a smooth, dark tone, indicative of clumps of short grasses and scattered trees associated with exposed, moist, surface material. This could be classified as "4b", Wet Phase Savannah, if it was not a complex type.

4a. Dry phase savannahs ordinarily develop on coarse rapidly drained sand flats, dry colluvium, rocky foothills or ironstone crusts. Vegetation is sparse. Reflection from the soils being very high, their extent can be determined with considerable accuracy.

V12 - Figure 2: This is predominantly a Lowland Forest type with Dry Upland Savannah indicated by the wedge. The narrow band of "gallery" forest along the drainage channels within the open savannah is of particular interest.

4b. Wet phase savannahs are found on finer soil material, mainly of alluvial origin, with flat topography or minor slopes. The vegetation is denser than on the dry phase type, due to better water-holding capacity of the soil. Lesser reflectivity causes them to appear darker on the imagery and therefore to be delineated with comparative ease.

5 . Marsh and Aquatic vegetation designates generally treeless areas permanently covered by moisture-loving plants and frequently interspersed with pools of stagnant water. In the imagery, they show up with a dark tonality which the

---

[2]Tepui, an Indian term used to describe high summit regions. ref. Maguire, Bassett, "Cerro de la Neblina, Amazonas, Venezuela.

vegetation slightly alters. Open water surfaces, on the other hand, appears black.

Summary of Vegetation

This is a general description only. Many subtle variations in texture and tone have not been included, as consistency could not be recognized throughout the project area. However, detailed study of these variations may hold the key to distinct vegetation communities.

The area distribution of vegetation was compiled by 9 hydrographic sub-regions.

In order to evaluate the relative importance of the various productive forest types and thus gain an appreciation of their significance in the nine Sub-regions of the Project, the following table indicates the area (km^2) distribution of productive forest based on the SLAR survey.

The total productive forest area is 93.6% of the surveyed area, and 84.8% of the Territory of Amazonas proper.

Related Costs of the SLAR Survey of Amazonas Territory

The Amazonas survey incorporated SLAR survey acquisition; construction of SLAR mosaics to Shoran, astronomical, and Boundary Commission control points; field work with associated logistic support; associated photo laboratory services; interpretation of hydrology, geology, soils, and vegetation, with final drafted overlays in each discipline. The cost of these services for the total 210,000 km^2, amounted to approximately 5 to 6 dollars per square kilometer.

CONCLUSIONS

SLAR imagery coverage of unexplored regions, in addition to providing remarkable physiographic and topographic detail, permits immediate identification and geographic localization of broad forest units.

SLAR coverage coupled with simultaneous field investigations allows determination of general features and ecological relationships of the forest vegetation, thus orienting future inventories of forest resources at considerable saving of time and expenditure.

I would like to thank Aero Service Corporation of Philadelphia and Goodyear Aerospace for survey data and imagery respectively, Spartan Aero Limited of Ottawa and the Forest Management Institute of Environment Canada for assistance, and in particular Dr. L.Z. Rousseau of Spartan Aero and Dr. A.J. Vessel of Aero Service Corporation for vegetation and soils data respectively.

REFERENCES

Summary Report by Aero Service Corporation to the Republic of Venezuela known as CODESUR f-5, SLAR Survey, Amazonas T.F. Venezuela. September 30, 1971. NOT PUBLISHED.

Foldats, E., 1962. La concentracion de oxygeno disuelta en las aguas negras. Acta. Biol. Venezuelica, 3:10:149-159.

Foldats, E. and Rutkis, E. (1969). Seulo y agua como factores en la seleccion de las especies de arboles que en forma aislada acompanar nuestros pastizales. Contribucion No. 46, Estacion Biologica de los Llanos.

Maguire, Bassett, "Cerro de la Neblina, Amazonas, Venezuela: A Newly Discovered Sandstone Mountain," Geog. Rev., Vol. XLV, No. 1 pp. 27-51, 1955.

Steyermark, J.A. (1966). Contribuciones a la flora de Venezuela, Pt 5, Acta Botanica Venezuelica, 1:3-4:9-168.

Williams, L., 1947. The Forests of the Upper Orinoco - Tropical Woods.

Area (km^2) Distribution of Productive Forest Types by Sub-Regions

Sub-Region	1	1a	1b	2	2a	2b	Total Productive Areas by Sub-Region
1	-	-	-	98	271	2,957	3,326
2	2,387	141	93	10,322	467	8,233	21,643
3	4,766	163	248	9,495	1,948	21,664	38,284
4	743	12	-	6,077	945	6,799	14,576
5	3,014	246	558	11,779	876	4,577	21,050
6	3,112	326	-	9,366	3,632	23,186	39,622
7	-	-	-	9,803	350	3,836	13,989
8	11,809	649	3,757	16,044	388	2,895	35,542
9	651	-	29	1,334	-	-	2,014
Total Area by Types	26,482	1,537	4,685	74,318	8,877	74,147	190,046 km^2
Total Productive Type Percentage	-	-	-	-	-	-	93.6%

APPENDIX A

Summary of Aircraft Systems and Specifications

OPERATIONS

Flight Direction: North-South orientation
Look Direction: West
SLAR Strip Annotation: The first four digits indicate degrees and minutes west of the Prime Meridian with the letter indicating flight direction.
Flight Line Spacing: 8 kilometers alternating near and far range.
Overlap: 60% minimum

AIRCRAFT

and Associated Systems

Aircraft (A/C): Caravelle Twin-jet
A/C Ceiling: 12,500 meters
A/C Range: 4,000 km +
A/C Payload: 21,000 kilograms
A/C Personnel: 6 crew
10 (max) Observers
Req'd. Runway: 2,300 meters, paved
A/C Velocity: Velocity is determined by the LTN-51 inertial guidance platform. The rate of SLAR film advance is also linked to the LTN-51 by a velocity - proportional analog voltage output.

A/C Radar Altimeters:

a) First return SLAR energy, directed via a vertical horn with a resolution of 20 meters and accuracy of 80 meters. Return is recorded in both digital and analog form.

b) The Stewart Warner Radar Altimeter has a resolution and accuracy roughly equivalent to the first return SLAR Altimeter. It is employed as a check and back-up piece of equipment.

A/C Altitude Control: Altitude is controlled by the manual setting of the altitude-hold mode of the Lear-Siegler L-102 auto-pilot. Reference altitude is adjusted to sea level over known sounding areas.

A/C Camera Systems:

a) Wild RC 9 camera
b) I^2S multispectral camera

c) Nadir observation closed circuit TV camera with Zoom lens.

GOODYEAR AEROSPACE
SLAR SYSTEM

Frequency: 9600 ± MHZ
Wavelength: 3.12 cm
Radiation: Coherent
Antenna Type: Both left and right side slotted phase array.
Resolution: Range 16 meters
Azimuth 16 meters

Note: Because of synthetic aperture, there is no loss of original resolution at maximum range.

Pulse Frequency: 1000 to 1600 pulses per second.

Pulse Length: Actual: 0.9 u sec.
Effective: 0.06 u sec.
Beam Width: Horizontal 1.4^{o}
Vertical 45^{o} with power distribution
Maximum Slant Range: 57.5 km
Look Direction: Direction optional, left and/or right, which may be switch in-flight.
Data Film: Type EK SO 3493, 5 inch width by 200 feet long.
Data Film Scale: range (or across-track) is 1:400,000 azimuth (or along-track) is 1:5555
Recording Capacity: approximately 1690 km.
Final Image Scale: is normally 1:400,000 but may be enlarged photographically after correlation of the data film.

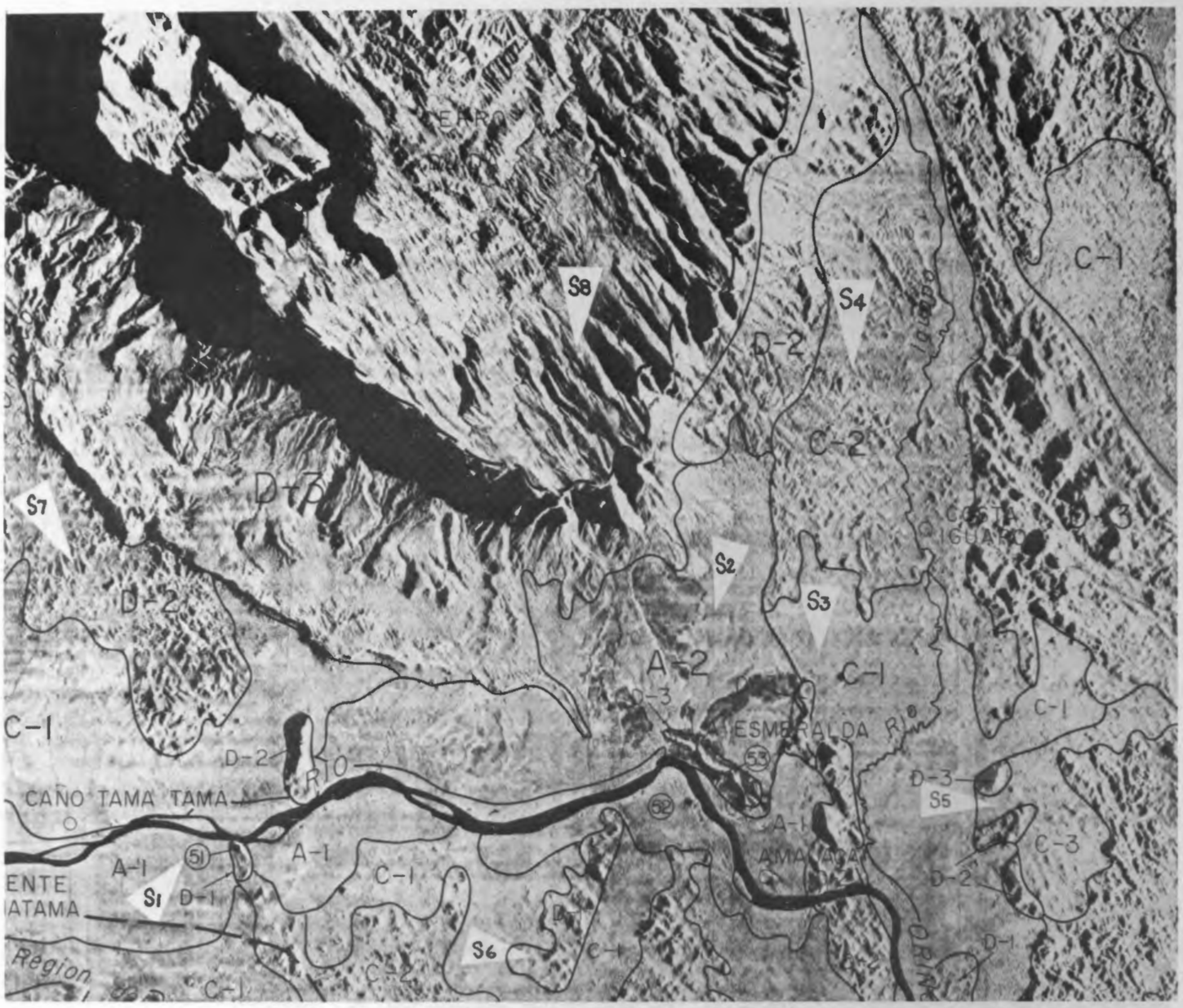

S8
S4
S7
S2
S3
S1
S5
S6
C-1
C-2
C-3
D-1
D-2
D-3
A-1
A-2
RIO
CAÑO TAMA TAMA
ESMERALDA RI
ORINO
Region

Figure 1

V8
V10
V9
V7
V4
V5
V6
V11
V2
V12
V3
V1
2a
2b
2b(3)
4(1)
1(4)
1
2
3
CAÑO TAMA TAMA
RIO
ESMERALDA
Region del Rio

Figure 2

SESSION II

PERCEPTUAL PRINCIPLES RELATED TO REMOTE SENSING *

Dr. M.M. Taylor
Defence and Civil Institute of Environmental Medicine
P.O. Box 2000, DOWNSVIEW, Ontario.

SUMMARY

Remote sensing systems may be regarded as extensions of man's natural senses. As such, they should be governed by the same principles that govern natural perception. Perception is described as a means whereby information useful for action is separated from the enormous mass of useless information, and encoded in a way suited to rapid evaluation of potential behaviour. This functional viewpoint leads to the idea that the "attention" of a "central processor" must be devoted at any one time to a small region of the environment and that a behaving organism should be provided with a large number of feature detectors which continually monitor the environment for items that deserve the attention of the central processor and send "alarms" when such features are detected. Attention should continually shift except when called by these alarms. The vigilance decrement is perhaps due to an inappropriate requirement that attention be deployed on a single display for continuous periods of time, and might be averted by transforming the display in such a way that targets appear in a manner to which natural "alarm" detectors are suited. Similarly, the provision of hardware feature detectors should be a major part of remote sensing systems used in searches for known target types.

INTRODUCTION

Remote sensing systems are used in many ways and in many environments. Wherever man cannot use his natural senses, whether because the thing to be observed is in the wrong place, or because he is not sensitive to the kind of energy with which it must be observed, there exists a situation where remote sensing is valuable. Remote sensing systems are used in airborne geophysical surveys, to monitor city pollution, to search for lost aircraft, to monitor conditions inside a nuclear reactor, to detect invading aircraft. We map, we seek, we monitor, and we react to events indicated by the instruments. We may record data for possible later use, or we may want to respond as quickly as possible to critical events. Whatever the operation, except in the simplest, algorithmically controllable feedback systems, a human sits at the centre of the system. Sometimes many humans are involved in a system whose overall functioning is controlled by one person at the centre; sometimes there is only one man involved. But in most cases, the remote sensing systems gather and display their information so that a human can act. He may read a map, the product of a remote sensing system, and decide to despatch an oil drilling party, or he may look at a radar screen and decide to send an interceptor to investigate an unidentified aircraft.

In this paper, "remote sensing" refers to any method whereby a human gains information about the world other than through the unaided use of his natural senses. Remote sensing systems are extensions of the natural human senses, and their effective design should integrate with the design of the natural perceptual system. Indeed, even if a computer were to be the controlling centre for a remote sensing system, the same principles would apply, since they are based on the utilization of information rather than being inferred from results of experiments on natural perception.

Although current uses of remote sensing systems bear mainly on mapping functions, in which data is brought back for later selection by the human, this paper considers primarily situations in which the data should be used immediately if they are to be most effective. To this end, a functional viewpoint of perception is described, a viewpoint which suggests some possible ways to make remote sensing systems in conjunction with the human more effective. Although the viewpoint appears reasonable, its assertions are frequently derived from everyday experience rather than from formal study. For this reason, and because this is not intended as a paper for psychologists, the viewpoint is asserted rather than argued, and references are held to a minimum.

The Problem of Perception

Perception has evolved as a method of coordinating the behaviour of the individual. It is not just a method of making an image to match the world so that a homunculus can look at the image instead of looking at the world. Some theorists go so far as to identify perception with the total set of behaviour for which the individual could be prepared, but it is not necessary to take such an extreme position in order to recognise that the evolutionary success of perception must be due to the survival and reproduction of the perceiving individuals. In any behaving system, whether it is an organism or an army, effective behaviour depends on getting valuable information to the right place at the right time, as well as on having proper methods of dealing with the information and reacting to it. According to this viewpoint, the content of perception is the information currently deemed valuable and relevant to possible action. The individual or organisation perceives in order to act, and perception is closely tied to action possibilities.

The fundamental problem facing any perceiving system is that there is too much information available in the world, and too little of it is relevant to the needs of the moment. We are bombarded from all sides with electromagnetic radiation of all wavelengths; the air vibrates in all directions over a wide range of frequencies; surfaces react differently to chemical and mechanical probing; we touch and are touched by innumerable objects. We can increase without limit the number of properties we choose to test, and we can increase without limit the resolution with which we measure. The relative positions and motions of objects give way to the relations among their constituent molecules, and thence to atoms and the uncertain relationships among their elementary particles. Some information is in the structural patterns of things as they are at any given moment, some in the patterns of events as they occur.

In principle, one could analyse all this available information, relate the different structures, and extrapolate correctly the probable consequences of all possible actions. In practice, matters are quite otherwise. The central decision processor cannot possibly perform in one lifetime the analysis of even the static structures. As for the changes introduced by ongoing events, the information they introduce must often be acted on in seconds if the individual is to survive. The great problem is how to select the useful from the relatively useless information, and how to minimize the search for structural relationships whose potential number is very large for even a small amount of input data.

Natural perceiving systems have evolved a number of techniques to deal with the information overload problem. So obvious as to be often ignored is the complete and arbitrary elimination of most of the electromagnetic spectrum and most of the mechanical vibration spectrum. Similar arbitrary elimination of information is accomplished by limitation in the resolution of the sensors. It is technically possible to design an eye with any desired resolving power, and an ear with any desired sensitivity. But resolution and sensitivity beyond certain limits becomes less and less useful and more and more costly. The elimination of data sources unlikely to prove useful or likely to be too expensive is a good design principle for a perceiving system as well as for a National Science Policy. A remote sensing system is the result of a policy decision to use a data source eliminated by nature, and must be regarded as increasing the data reduction problem facing the human observer.

Information is structure. A reduction in the types of structure sought implies a reduction in information processing requirements. Most structure in the world is short range in terms of the sizes of elements composing the structure. Neighbouring points tend to have the same brightness, and when they do not, the same difference probably can also be found between a nearby pair of neighbours. Distant points rarely correlate well. Short time spans yield higher correlations than do longer ones. Accordingly, it is usually more profitable to look for short range structure, because more will be found. This is especially true when one considers that the number of possible twofold relationships in a two-dimensional space varies with the fourth power of the range over which they are sought, whereas the number of actual relationships over the longer ranges is probably less in total than the number of shorter range relationships. The neural cells needed to implement one simple pair-wise relationship among all pairs of retinal receptors would occupy a volume of some 300 cubic metres, at a conservative estimate. This figure alone shows how few of the possible structural relationships can actually be detected in any real system operating on natural data.

The cost-effectiveness of structural searches

must decline very dramatically with increasing range. This fact is realized in the anatomy of natural perceiving systems, in that the elementary feature detectors typically have a high density of connection within a restricted neighbourhood and a very low connection density outside that neighbourhood. Even the so-called Fourier analyser elements may extend little more than a wavelength of the spatial frequency for which they are most sensitive. Larger-scale structure can readily be determined from the interrelations among the small-scale structures given by the primary feature detectors. In hearing, the probable relationships are among harmonically related features, and harmonic structure feature detectors should be common in the auditory system.

Wired feature detectors, at whatever level in the hierarchy, can respond only to the feature for which they are designed. The system designer has the option of incorporating wired detectors for all possible features or of eliminating unlikely features in favour of the ones that will most probably occur and be useful in practice. If the latter course is followed, as it must be in a device dealing with the complexity of nature, the system will be inflexible since the features discarded in the design may actually occur and be important on occasion. A system simplified by the elimination of possible fixed-feature detectors can be augmented by the addition of a variable, probably slow and non-algorithmic feature analyser which can look for new features of a kind not considered in the system design. This general analyser should then be capable in principle of discovering any kind of relationship that occurs in the environment. It should be used in most cases to determine complex or long range relationships, for relatively few of which would fixed feature detectors be economically feasible. The general analyser presumably forms part of the central processor which must decide on general courses of action on behalf of the entire individual, since it must take cognizance of relationships among effects from all data sources. It "thinks".

Information processing load reduction can thus be accomplished by ignoring possible data sources, and by ignoring possible relationships among data sources. The latter method of achieving reduction has a potentially severe difficulty in loss of adaptability, which can be mitigated by the incorporation of a general analyser capable of processing any relationships that might occur in the environment. Such a central processor pays for its capability with loss of speed. It also necessarily has to "pay attention" by considering at any moment only a small portion of the incoming data stream. This need to pay attention is the most significant aspect of the human central processor design from the viewpoint of the remote sensing system designer who wishes to provide the human with a useful instrument.

Interesting Things and Perceptual Modes

In order further to understand ways to accomplish information processing load reduction, we must consider how information is used in connection with behaviour. To this end, we can classify the things in the world that might be interesting in the sense that they might relate to current or subsequent behaviour.

Interesting things come in two major classes. These classes are not completely dichotomous, but shade into one another. Things in one class stay around long enough to be examined at leisure, while things in the other class are transient; they are here one moment and gone the next. What constitutes leisure depends rather on circumstances. A millionth of a second is a very long time in high-energy physics, but a month is short in the affairs of nations. Roughly speaking, if the thing being examined is there for a long time compared to the time taken to get all the essential data and to define the important relationships within the data, then it can be examined "at leisure". The processor can do many things before it has to pay attention to a stable thing. What constitutes leisure in the human time scale depends on circumstances, on what is being observed, and on what relations must be discovered in order to result in satisfactory behaviour. In ordinary perceptual experience, one would probably say that something which lasted only a few seconds was transitory, while something lasting a few minutes was stable.

Language usually classes things as "objects" or "states" on the one hand, and "events" on the other. Not all languages make the distinction in the same way. In English, there is a clear distinction between active events and stable states. The fact of change, however slow, is distinct from the state in any given time-frame. Objects, linguistically, have a permanent and individual existence, no matter how this permanency and individuality is belied by the real world.

In Hopi, according to Whorf (1940), the more prominent distinction is between conditions

lasting longer or less long than a cloud. Fast things, like a lightning flash, a "setting of a stone", or a spoken command, are classed one way, whereas things slower to change or vanish, such as a house, a season, or a migration would be in the other class. The distinction between object and activity is less clear than in English, while the distinction between short and long-lasting phenomena is sharper. Both ways of speaking reflect valid views of reality. English focusses more on the relationship between actor and the thing acted upon, on the differences between the fact of change and the momentary state, whereas the Hopi focusses more on whether the condition must be observed now or whether observance can be put off till later. The Hopi view is more consistent with the emphasis placed here on the functions of perception which depend on whether or not the phenomenon will wait until the observer can attend to it.

The World Map

One of the major functions of perception is the construction of a world picture, a "map" of the stable states and objects in the world. The construction of this map is often taken to be the only function of perception, and its accuracy to be a measure of the adequacy of perception. Evolutionarily, this cannot be so. The adequacy of perception must be measured by the survival of the species, and the survival value of such a trait must be determined by its ability to induce correct action quickly enough to meet the circumstances. The production of a map is obviously one function of perception, but not the sole function. Neither is maximum map accuracy a reasonable measure of the value of the map. It must be an action-centred map, organized and arranged so that possible courses of action can be tested with its aid. Irrelevant detail is detrimental for speed. The map is more than a picture, but less than reality.

The map constructed by perception shares many qualities with a paper map. Both fall short of a complete representation of the region they depict. Both emphasize relationships that are useful for the purposes at hand. just as a road map discriminates the surface qualities of the roads and may ignore the railways and the heights of land, while an agricultural map depicts fields of oats differently from fields of wheat, so the mental map must include coded representations of relationships that have been observed to exist in the world and are probably relevant to action. The labelling and coding by symbols and colours on paper maps are analogies of features which are often difficult to see in a photograph of the terrain. A skilled photo-interpreter must employ his skills before an aerial photograph can be turned into a map. Similarly, considerable coding and interpretation must be done before the raw sensory data stream can produce the relationships that go to make the mental map. The variability and complexity of these relationships suggest that they are largely discovered by the central processor. Indeed, since we have presumed that the central processor is responsible for the action decisions, it alone can know what features are currently relevant to action and should be incorporated in the map. This in turn implies the need for focussed attention in construction of the map. Map construction is relatively slow, partly because only a small part of the world can be scanned at one time, and partly so that the coding can be complete enough to permit fast map use when necessary for immediate action.

Neither the mental nor the paper map is used at the moment of its construction. Both serve as reduced and simplified sources of secondary information about portions of the world that may or may not be involved in later action. If action is required, however, the map makes it much easier to plan. Map data can be combined with current data, so that the terrain need not be resurveyed before each operation. The information is there and already correctly coded.

Attention and Alarm

Construction of the mental map requires focussed attention. Only a part of the map, and probably only a few features in that part, can be constructed at any one time. The world is too big and complex to be analysed all at once. Hence, the evaluative central processor must ignore most of the world most of the time. This does not mean to say that the whole individual ignores most of the world. The contrary is true. The wired feature detectors always operate on their portion of the data stream, regardless of where the attention of the central processor is deployed.

The focussing of attention implies a danger to the individual; events may happen and be missed if attention is not directed to them. This is the point of the distinction between things that can be examined at leisure and things that come and are gone. Attention can be deployed at will among the objects and stable states, but events must be examined

while they happen or they will be missed. Of course, a certain amount of storage permits ongoing events to be held briefly in memory until they can be evaluated. Such storage would have to act like a shift register holding a certain duration of the entire data stream. Some storage of this kind is necessary, but it would be very costly to provide a long duration of such a wide-band shift-register. It would also not solve the problem of events that require fast responses. The simpler answer is to make a design decision that certain features probably signal situations demanding the attention of the central processor, and whenever these features are detected, an alarm calls the attention of the central processor to the region containing the alarming feature.

Consider vision. At first glance, vision seems to be the perfect example of a pattern processing system. Most remote sensing systems use visual displays and rely on the observer's vision to sort out the useful from the useless information. Yet little of the visual system itself is devoted to pattern processing. The whole system is a good example of the focussing and alarm principles on which attention is based.

When the eyes look straight ahead, one can see almost a full hemisphere. However, very little pattern is visible in the outer ranges of the hemisphere, and only the central 1^{o} is really useful for pattern vision. Acuity declines dramatically with distance from the centre. Twenty degrees from central vision the resolution has declined by a factor of ten, which means that the information transmission capacity for pattern has declined by a factor of 100. This decline is not matched by a parallel decline in the receptor density. Twenty degrees out, the receptor density is at least as high as it is in the centre, and this density provides the technical capability for the visual resolution to be as good in the one region as in the other. But the fibres of the visual nerve are nowhere near as plentiful 20^{o} out as at the centre, as if the potentially available resolution had been deliberately discarded, after having been built in. It looks at first sight like a design by committee, an unusual thing to find in a system tried and tempered by evolution.

What are all the extra receptors good for? In terms of alarm systems, the answer is apparent. The visual periphery is exquisitely sensitive to movement. Movement can readily be resolved between positions which cannot be resolved on the basis of static acuity. Even 20^{o} out from the fovea, a movement can be seen between points separated by only 3 minutes of arc, and possibly less. In the further periphery, very little can be seen at all unless something moves, and when something does move it is very obvious. It draws the attention. It "alarms".

Motion is a reasonable feature to have as an alarm in the visual system. Most of the visible world is fixed at any one moment, and movements usually signal either an event that must be accorded some immediate reaction. A good proportion of movements deserve attention, and those portions of the world that do not move can readily be scanned when time is available. Those movements that do not signal interesting events can be dismissed after a quick attention shift determines their uninteresting nature, and even if they continue they may not thereafter raise an effective alarm. Their uninteresting nature has been coded and can be determined for or by the central processor without an overt attention shift.

The ability to dismiss an alarm is a necessary function of an attention-getting procedure based on fallible analysis. Since the alarm system is likely to respond to a situation that does not really require the immediate attention of the central processor, and to continue responding to situations to which the central processor has already reacted, the central processor must be able selectively to switch off or to ignore an alarm.

Alarms in vision can be effective even when the detected feature is as simple as motion. In hearing, the situation is different. Hearing is not primarily a mapping sense. Most things we hear signal events, but this very fact suggests that most auditory events are not behaviourally significant and should not draw the attention of the central processor. In hearing, significant events must be distinguished from non-significant events by the peripheral preprocessors before an alarm is given. This means that there must be alarm feature detectors much more sophisticated than the simple motion detectors of the visual system. Furthermore, since simple waveform equivalence is inadequate for determining an event in the world, fairly complex and probably hierarchical sets of feature detectors must be involved. The fact that one's own name can serve as an alarm (as is shown by the ability to hear it in a noisy cafeteria) indicates that alarm features can be not only complex, but also learned. This is probably true also of alarms in modalities other than hearing. It

suggests further that feature detectors themselves can be learned and do not have to be hard-wired genetically. For more complex features, most detectors are probably learned, probably in response to behavioural needs.

Implications for Remote Sensing Systems

The perceptual principles identified above carry several implications for the design of remote sensing systems whose data are to be used as they are gathered. They have little bearing on applications in which data are stored for later retrieval, as they are in mapping operations, for example.

The most immediate application is in the area of prolonged monitoring, or vigilance. The vigilance decrement is a term used to describe the inability of an observer to maintain his efficiency in detecting critical but infrequent events. The vigilance decrement seems to be a product of the technological age, and is usually associated with the need to pay continuous attention to a data source that does not provide much information and requires little reaction. According to the viewpoint on perception presented here, this is a most unnatural situation. Unless some event that needs immediate reaction or which requires a map update is going on, attention should not be devoted to a single data source. The primary task of the central processor is control of behaviour, which specifically requires continual restructuring of the world map. This may be accomplished by deploying attention among the various possible data sources, or by thinking about structures inherent in data already obtained. Mapping is prevented by paying attention to an uninformative data source. It seems that the conflict between the task requirements and the requirements of map updating is usually resolved by reducing the attention paid to the task. It is only fair to point out that this is a speculative hypothesis about the vigilance decrement, for which there is as yet no satisfactory published theory.

If this hypothesis concerning the cause of the vigilance decrement has merit, then remote sensing system designers, including the designers of radar and sonar systems, should beware of giving an operator an information-free and behaviour-free task. The operator should be allowed ample time to attend to matters other than the data display. Ideally, the possible presence of an interesting target should be signalled to the operator's alarm system. Motion perceived in the visual periphery would probably be adequate. A possible scheme to accomplish this in an early-warning radar involves time-compression. During each second of real time, the display would show a speeded history of the last period, say one minute, of the incoming data. Intruding aircraft would show as streaking lines, and noise clutter would show characteristically different motions. The streaks in motion associated with possible targets should be quickly seen, even by an operator occupied with a different task, provided that the screen was in his visual field.

Displays incorporating an alarm function, such as the time-compressed radar display, should have some facility for removing the alarm capability from an identified target. In the case of the radar display, for example, it is likely that the display would be digitally generated, and that a computer would be at the heart of the display. The operator should be able to identify to the computer the fact that he had detected the target, and the computer should react by computing the probable future track of that target. So long as it maintained a track near to that predicted from the history prior to the operator's response, the display should be kept dim, but if it deviated from expectation, the track should brighten until the operator again reacted to it. Permitting the operator to delete the alarm function from a detected target should enhance his ability to pick up new targets, while at the same time allowing him to keep track of the old. As was pointed out above, a necessary concomitant of an alarm system is the ability to turn off the alarm, and if the hardware does not permit this, then the operator's central processor will probably do it internally, so that new targets on the same display will be ignored.

A second difficulty with displays which rarely change their content of interesting information is that for behavioural purposes they become unchanging, with the consequence that it is more efficient for the central processor to work with the already coded map image of the display than with the incoming sensory data that needs coding. Reversion to memory mode may happen without the awareness of the individual, as witness the driver who goes through a new stop sign on a familiar route without ever seeing it. A driver to whom the route was new would always see that same stop sign. This phenomenon is well known to the police, who usually station an officer near every new stop sign. Inappropriate use of this memory image mode

of perception can be reduced by changing the information content of the display periodically, or by removing it from the observer's field of view for long enough periods that he cannot rely on his memory to reconstitute it. This technique is probably rarely effective in itself, but the possibility should be considered in a design when unchanging displays can happen.

A third consequence of the analysis of the perceptual system is the suggestion that there should be a major role in remote sensing systems for hardware feature detectors of any degree of sophistication, so long as they have a different probability of responding to targets and to non-targets. On the other hand, the analysis also suggests that attempts to provide fully automatic pattern recognition devices to discriminate any but the simplest targets in the simplest contexts are doomed to failure. The main function of hardware preprocessors should be to provide a fast, wideband and wide-field scan of the incoming data so that the preprocessor can identify portions of the data stream worthy of the human observer's attention. In a search task, for example, any reduction in data load is welcome. Preprocessors such as linear contour detectors might be useful in enhancing the likelihood that the observer would detect cultural artifacts. Man-made objects are more likely to be bounded by straight lines than are natural variations in intensity or colour. If a visual display were made to contain only the outputs of line detectors which covered the whole field, but which could be switched by the operator at will to a full-picture mode, the operator would be more likely to detect man-made objects in the wild than he would from a continuous full-picture display. Thought should also be given the possibility of using the good learning capabilities of the human auditory feature detectors.

The major principle involved in the use of feature detectors is that the observer should not have to attend to the incoming data unless there is a reasonable probability that it contains something interesting. But when such an occasion arises, the full data stream should be available to him. The feature detectors signal where to look; they do not presume to substitute for the pattern recognition skills of the man.

Feature detectors, or simple pattern recognition devices, can be seen to be very important in the design of remote sensing systems for real-time action. They are probably also useful in the slower scanning of imagery recorded during planned survey missions, provided that there is a definable distinction between interesting and uninteresting things in the imagery.

CONCLUSION

The principles of perception, seen from the viewpoint of information processing load, provide a useful guide to the design of remote sensing systems intended for human operators. It is not necessary or desirable to flood the operator with as detailed a picture of the world as the sensors permit. But when the world contains something the operator should see, the system should warn him, and permit him to see it in as much detail as possible. The remote system designer should neither overload the operator with information to be processed nor force him to pay attention to a single data source for long continuous periods of time.

REFERENCES

(1) B.L. Whorf. Science and Linguistics. Technol. Rev. 42, 1940. 229-231, 247-248. Reprinted in Whorf B. L. Language, Thought and Reality, 1956, MIT Press.

* DCIEM Research Paper No. 828

PHOTOGRAMMETRIC ASPECTS OF PRECISION PROCESSING OF ERTS IMAGERY

Dr. Vladimir Kratky,
Photogrammetric Research,
National Research Council of Canada.
Consultant to Canada Centre for
Remote Sensing, Ottawa, Ontario.

ABSTRACT

ERTS imagery as formed aboard the satellite by the Return Beam Vidicon (RBV) camera and Multispectral Scanning System (MSS), radio transmitted to the Earth and displayed by special reproducers, will be considerably distorted due to several factors. The character of the distortion depends mainly on the geometrical properties of the imaging system and on the dynamic flight conditions. The RBV geometry is analyzed with particular attention to the electronic distortion. Appropriate analytical formulation based on polynomial and projective transformation is shown to be capable of providing the scan line control essential to the precision processing system as adopted for the Canadian ERTS program. The geometry of MSS images is discussed in detail with respect to the distorting effects of the Earth rotation and the applied cartographic projection. The combined effect of these factors is assessed by means of expected errors and illustrated by several tables.

INTRODUCTION

The Earth Resources Technology Satellite (ERTS) project which is supposed to commence operations in the late spring this year, has been discussed in several already published studies and is also a subject of some other contributions to this Symposium.

Two basic systems will be used aboard the spacecraft for collecting image information: return beam vidicon (RBV) cameras forming consecutive central perspective images, and the multispectral scanner (MSS) continuously covering a 185 km wide ground swath by transverse scan lines. Both types of imagery will be transmitted to a ground receiving station, recorded on magnetic tape and transformed into photographs by special reproducing devices, either in the original raw geometric form or with introduced geometric corrections. An electron beam recorder (EBR) and a laser beam image reproducer (LBIR) have been designed for producing both the uncorrected and corrected version of the imagery. Either instrument can be properly controlled to introduce the corrections.

Generally, the image geometry can be improved only if one knows precisely the internal structure and function of the imaging, transmitting, recording and reproducing system. Furthermore, all parts of the system should be designed, calibrated and maintained in such a way that the metrical properties of the image do not vary with time. Only then the physical function of the system can successfully be fitted with a geometric or mathematical model relating the collected data with the object details.

In this regard, the existing photogrammetric theory and experience can be helpful in providing well tested schemes and procedures for geometric calibrations, for organizing measurement operations, for formulating and implementing analytical models.

From the photogrammetric standpoint, the RBV imagery can be treated in a similar way to high altitude aerial photographs, while the MSS imagery presents several additional serious problems in how to define its dynamic geometry and how to convert it into a conventional system. It should be pointed out that no stereocompilations can be performed because of specific conditions of the mission. The base-height ratio is extremely unfavourable and the terrain relief is too small with respect to the flying altitude. Thus, the photographs can be handled

only as individual photogrammetric images. Since the MSS system produces a continuous image along the track, its framing must be performed artificially, allowing for the 10% overlap desirable for optimum production of photo mosaics.

Several error sources of a physical character, such as terrain relief, Earth curvature and atmospheric refraction, affect the geometric fidelity of single photogrammetric images and may even set limits for the accuracy of ERTS photographs. More information on these aspects can be found in [1] or [2].

The corrected version of the ERTS image will represent a photomap and is considered to be a final product for the user-interpreter. At least at the present time, the ERTS photographs are not likely to be used otherwise. As an enhanced substitute of a map, the corrected ERTS photograph has to conform to certain rules which are usually applied in map compilation. This primarily means that all images are to be annotated and cartographically defined in a uniform system of rectangular coordinates, such as the Universal Transverse Mercator (UTM) system.

In comparison with aerial photographs the ground coverage of an individual ERTS image is extremely large. One can not simply disregard the metrical limitations of any cartographic projection defined over such an area. The internal distortions of the UTM projection should be assessed and taken into consideration especially for scenes located in border regions of a standard 6° UTM zone.

The basic principle of producing corrected images in both available image reproducers can be classified as digitally controlled scanning. Image input data stored on a magnetic tape sequentially modulate the intensity of the electron or laser beam, while the precalculated correction parameters are utilized to vary direction, offset and spacing of segmented scan lines, as well as the speed of feeding input data, as required by the analytically reconstructed model of distortions. This model is to be defined by analytical computations based on the knowledge of certain number of control points whose coordinates are compared with measurements on preliminary prints of uncorrected images. In general, the reproduction of a corrected image requires two passes of image data through the reproducer. The technology of this precision processing is outlined in the block diagram in Fig. 1 and more descriptive details can be found in [2].

It should be stated at this point that the following presentation does not contain any complete analysis of the whole ERTS photogrammetry and should rather be considered as a discussion of a few selected metrical problems.

ERTS ORBITAL GEOMETRY

Both the satellite and the Earth have their own motions whose effect is gradually and continuously summed up. While the satellite is progressing in its orbit, the Earth is rotating around its independent axis which is different from the axis of the satellite orbit. In addition, the spatial attitude of the satellite is subjected to inevitable variations.

The satellite position is determined by the close-to-circular orbit and is supposed to be continually tracked, but not directly controlled during the mission. The altitude variations of the orbit are predicted not to exceed the value of 0.5 km in the range of a single ERTS frame. The ephemeris data will be available all the time and gradual changes can be predicted. On the other hand, the attitude of the spacecraft has to be continually controlled to guarantee a proper system orientation towards the ground. Special subsystems provide such a control independently for all three basic attitude errors. The errors will then vary more or less continuously not exceeding the specified maximum 0.6° for yaw or 0.4° for pitch and roll.

For the purpose of the error assessment the analysis of the ERTS orbit can be performed anticipating the spherical shape of the Earth with the conclusions applicable locally to the Earth ellipsoid. In the first approach, the Earth rotation is neglected.

The orbit plane does not go directly through the poles; its polar inclination

ε amounts approximately to 9°. Fig. 2 illustrates the basic relations between the orbital parameters and geographical coordinates. The vertex V of the orbit represents the origin for the orbital travel distance ρ and, at the same time, determines the reference meridian for defining the change of the geographical longitude λ_s. The nominal position S_s of the satellite (not including the Earth rotation effect) is then defined by the orbital parameters ε, ρ or by the corresponding change of the geographical coordinates λ_s and $(90°-\Phi)$. The nominal heading of the satellite towards the local meridian is expressed by the angle H_s.

In accordance with the derivations in [2] the following relations define the nominal position and heading of the satellite

$$\sin \Phi = \cos \varepsilon \cos \rho \quad (1)$$

$$\cot \lambda_s = \sin \varepsilon \cot \rho \quad (2)$$

$$\tan H_s = \tan \varepsilon / \sin \rho \quad (3)$$

Some modifications of these basic formulas can be utilized for later analysis. The alternative formulation of the nominal heading H_s is

$$\sin H_s = \sin \varepsilon / \cos \Phi \quad (3a)$$

$$\cos H_s = \cos \varepsilon \sin \lambda_s \quad (3b)$$

which together with (3) and (2) determine the alternative expressions for λ_s

$$\cos \lambda_s = \tan \varepsilon \tan \Phi \quad (2a)$$

$$\sin \lambda_s = \sin \rho / \cos \Phi \quad (2b)$$

The Earth rotation affects the actual position of the subsatellite nadir point displacing it in the direction of the local geographic parallel. The latitude Φ stays unchanged, but the subsatellite track gradually deviates by H_e from the nominal direction H_s. With reference to derivations in [2] and in accordance with Fig. 2 the relevant changes are as follows

$$\lambda_e = \frac{\omega_e}{\omega_s} \rho \quad (4)$$

$$\tan H_e \doteq \frac{\omega_e}{\omega_s} \cos \varepsilon \sin \rho \quad (5)$$

where ω_e, ω_s denote the angular velocities of the Earth rotation and of the satellite, respectively. Definition of a subsatellite point and of the real ground heading of a subsatellite track at this point results from the combination of the above formulas, yielding

$$\lambda = \lambda_s + \lambda_e \quad (6)$$

$$H = H_s + H_e \quad (7)$$

Table 1 summarizes the values for the satellite or subsatellite position and heading in the nominal and real form.

RBV IMAGERY

General remarks

The RBV image is created in an instantaneous mode by a lens assembly as a central projection of the object. Despite the technical complexity of the image recording and transmitting, the collected data can be treated more or less similarly to standard photogrammetric frame pictures. This means, first of all, that the geometry of the image is considered as 'frozen' with respect to the short moment of exposure during which any movements of the system can be practically neglected.

Knowing the interior orientation of the imaging system and given the auxiliary housekeeping data on the position and attitude of the RBV cameras in the instant of exposure, one is able to perform the necessary image rectification of the image. The accuracy of this process depends on the quality of the given information. Generally much better results can be achieved when comparing the images with the available ground control. This implies that there must exist a certain number of well defined, identifiable control points whose coordinates are known in the UTM system, e.g. from existing 1:250,000 maps. From the comparison of the image and UTM coordinates the resection of the projective bundle of rays can be solved, which, in turn, defines the proper projection of the image into the reference UTM plane. In the particular case of RBV application, however, the spatial resection is simplified and substituted by suitable x,y-transformation of the image.

The interior orientation data of the RBV system are extremely important with regard to the electronic way of scanning, radio transmission and video recording of the original optical image. Since these operations are inevitably carried out with significant metrical distortions, a réseau technique is to be used for the calibration of the whole system. The regular array of 9×9 réseau crosses covering almost the complete viewing field of the camera makes it possible to define the electronic distortion for each réseau point. The lens distortion data can possibly be included into the same correcting process by modifying the réseau information.

Geometric considerations

The general flow of the image information in the process of acquiring and reproducing the RBV image is outlined in Fig. 3 and can be described in the following steps.

1. Terrain details are mapped in the UTM plane with coordinates X, Y.
2. When rotating the coordinate system by K one defines terrain coordinates $\bar{X}$, $\bar{Y}$ with the reference axes lined up with the nominal satellite heading in the UTM plane.
3. The optical image is created and briefly retained on the vidicon screen. Its coordinates x_V, y_V are defined by the central projection of the terrain scene and depend on the current satellite position and attitude, and on the lens distortion characteristic. Let us denote the respective changes of geometry by the symbols 'P, A, L'.
4. The vidicon image is now scanned, transmitted and recorded on video tape. The video information represents a stored image with coordinates x_T, y_T, significantly distorted by the deficiencies in each of the involved operations. This results in the change 'T' of the image geometry.
5. The video information is reproduced with the use of the EBR producing an image on 70 mm film. The coordinates x_R, y_R of the resulting image are additionally affected by the change 'R'. The common effect of both 'T' and 'R' is usually described as the electronic image distortion 'E'.
6. Since the EBR image is too small to be used directly for interpretations, the image is finally 3.7 x enlarged to the standard 9×9 in. size with the use of a special enlarger-printer. The inherent inaccuracy of this procedure will affect the final coordinates x,y by what we denote 'S' distortion.

All the described changes accumulate in the course of RBV image processing and would eventually degrade the geometric quality of the product, unless one tried to control the process by introducing proper corrections. The EBR controller (EBRC) is capable of modulating the scanning-printing process of the RBV image in such a way that the undesirable changes can be eliminated. For this purpose it is vital to have all the auxiliary information as quoted in the previous section. The measurements of ground control points performed on uncorrected or partially corrected images, are used for the refinement of the definition for 'P' and 'A' changes. At the same time the réseau measurement defines the remaining inherent changes 'L, E, S'. In the second pass of the video information through the EBR, the knowledge of 'P' and 'K' is used to construct the UTM and geographic annotation while 'A, L, E and S' are eliminated by introducing corrective changes 'C' which are controlled by the EBRC. In a symbolic way, one can write

$$C = -(A + L + E + S) \qquad (8)$$

and

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_V \\ y_V \end{pmatrix} + E + S + C = M\begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix} \qquad (9)$$

where M represents the required 1:1,000,000 scale of the final RBV image.

The correction changes 'C' have to be generated and properly synchronized with the feeding of the source information from video tape. This requires that the correction be formulated as a suitable x,y-transformation of the image. Its definition is to be broken down into two transformation phases T_R, T_G which deal with changes based on either réseau or ground control measurements. The first partial transformation can not be defined in a generally theoretical way since it is dependent upon the physical values of 'L, E, S'

distortion at réseau points and their distribution in the x,y-plane. The empiric polynomial transformation T_R of a suitable degree represents the best fitting form. The other partial transformation T_G, however, is determined quite definitely as the projective transformation. Its definition is possible only after the measured photo coordinates of ground control points are corrected with the use of the already found transformation T_R. The final control of 'C' corrections must be established as a composite transformation based on T_R and T_G. The effect of changes 'K, P' which are used directly for the UTM annotation of the image is to be taken out of the transformation T_G. If this partial transformation is denoted T_P one can write the following symbolic expression for the wanted resulting transformation T_C

$$T_C = T_R + T_G - T_P \,. \qquad (10)$$

Projective transformation

A rigorous solution for the spatial resection of an RBV photograph can only follow the removal of the electronic distortion and should be based on the well-known collinearity equations

$$\Delta X.\bar{z} - \Delta Z.\bar{x} = 0 \,, \quad \Delta Y.\bar{z} - \Delta Z.\bar{y} = 0 \qquad (11)$$

where ΔX, ΔY, ΔZ are UTM coordinate differences between the given ground control point and the projection centre, while $\bar{x}$, $\bar{y}$, $\bar{z}$ represent photo coordinates transformed with the use of a rotation matrix which is a function of the yaw κ, pitch ϕ, and roll ω. Since these elements should be very small in ERTS (0.6° for yaw, 0.4° for pitch and roll), the solution can be significantly simplified by linearizing equations (11) into the form

$$\bar{X}_o - \bar{X} = \frac{H}{f}\left(x + y\kappa + f\left(1+\frac{x^2}{f^2}\right)\phi + \frac{xy}{f}\omega\right)$$
$$\bar{Y} - \bar{Y}_o = \frac{H}{f}\left(y - x\kappa + \frac{xy}{f}\phi + f\left(1+\frac{y^2}{f^2}\right)\omega\right) \qquad (12)$$

where H - denotes the flying altitude and f - the calibrated principal distance of the RBV camera. The coordinates of the projection centre $\bar{X}_o$, $\bar{Y}_o$, the flying altitude H and the three attitude errors represent the total of six unknowns in the solution. Consequently, three points should be sufficient to obtain the theoretical solution but at least four ground control points stand for a reasonably efficient practical minimum. When neglecting terms smaller than the resolution element of the RBV system (estimated as approximately 70 to 100 m) one gets eventually

$$\bar{X} = \bar{X}_o - \frac{H}{f}x - H\phi - \frac{H}{f}y\kappa$$
$$\bar{Y} = \bar{Y}_o + \frac{H}{f}y + H\omega - \frac{H}{f}x\kappa \qquad (13)$$

Réseau transformation

The existing information on the electronic distortion of the RBV image is based mainly on investigations carried out by McEwen [3] and Wong [4]. One of their most important conclusions is that the RBV images are very stable in the geometry. Even though the magnitude of electronic errors is rather high (up to 1.3 mm in the final 1:1,000,000 scale), their vector distribution maintains certain invariant pattern which does not seem to change even over a reasonably long period of time. The vector distribution varies, of course, for different RBV cameras and EBR's. For a specific RBV-EBR combination an independent calibration is to be performed and regularly checked. For this reason, the following analysis fits just one particular case. It may render only general insight into the problem, but indicates at least in what order of magnitude the errors and rate of their change can generally be.

Since the réseau grid is formed by a 9×9 square array of crosses over the area of RBV picture, the degree of the polynomial formulation can be kept sufficiently high. A rectangular pattern of n×n selected réseau crosses can generally accommodate the polynomial transformation of the extended (n-1)th degree. The coefficients of transformation can be built up according to the gradually expanding following scheme

0° degree				1						parameters:	1
1°			x	xy	y						4
2°		x^2	x^2y	x^2y^2	xy^2	y^2					9
3°	x^3	x^3y	x^3y^2	x^3y^3	x^2y^3	xy^3	y^3				16
4°	x^4	x^4y	x^4y^2	x^4y^3	x^4y^4	x^3y^4	x^2y^4	xy^4	y^4		25
.....				⋮				.....			

The term 'extended' transformation means that the coefficients are included into the formula up to the desired order of a single coordinate x or y even if the order of the product $x^i y^j$ is higher than the specified degree. This arrangement makes it possible to fit the error distribution with the same degree of the function for any row- or column-wise selected group of n réseau points. The number of parameters amounts to n^2. The transformation is defined separately for x- or y-errors.

Theoretically, the degree of the selected transformation could be increased up to d=8, however, with the increase of parameters the redundancy of the solution is dropping. Furthermore, the definition of the fitting function is affected more by insignificant random changes and errors, while the only purpose of applying this transformation is to suppress the systematic irregularities of the image. Generally, all the 81 réseau points should be taken into account when solving the transformation parameters and the degree of the polynomial should be kept as high as necessary to bring the residual deviations down to the level of random noise. Several statistic means, such as Fischer testing can help in making the decision.

As a typical example the results for one of several comparative experiments will be presented here, which are based on réseau measurements performed at the University of Illinois and supplied by the courtesy of Dr. Wong. The réseau grid of the RBV camera X05197 was in this case connected directly to an RCA Laser Beam Image Reproducer with the output on 9 in. stable-base film. The original distortions after preliminary linear transformation are characterized by the r.m.s. error m_x = 424 µm and m_y = 478 µm in the final scale. The distribution of error vectors is shown in Fig. 4. The following table represents the decrease of residual errors when raising the degree of the polynomial transformation:

Polynomial degree	0°	1°	2°	3°	4°	5°	6°
m_x	424	399	311	80	34	34	30 µm
m_y	478	458	126	54	49	38	33 µm

All the transformations used the full number of 81 réseau points. For the purpose of completeness the table lists the results even for the lower degree transformations which are not seriously considered for any practical use. The same values are used for the graphical evaluation of the case in Fig. 5. The initial r.m.s. errors drop sharply with the use of 3rd and 4th degree transformation and stay practically without any significant improvement below the level of an RBV resolution element. The 4th degree polynomial transformation with 25 parameters can provide satisfactory definition of the electronic distortion leaving remaining residual errors at the level of m_x = 34 µm and m_y = 49 µm.

Fig. 6 represents the computer output for this case of polynomial fitting. It lists the points used for the solution, both groups of 25 parameters and displays the x- and y-deviations at the réseau points after the transformation. The internal accuracy of the solution is characterized by computed theoretical coordinate variances for an array of transformed fictitious points in the pattern of Fig. 7. The listed values cover the first quadrant (right upper corner) of the réseau grid with twice denser interval, extended partly beyond the réseau grid.

Scanning control

The EBRC is capable of providing the efficient control of scanning in accordance with the composite transformation (10). This can again be accomplished by means of polynomial transformation, but the original formulation is modified to represent the relevant changes in the direction of scan lines (these aim at the across-the-track direction parallel to y-axis). Considering the polynomial transformation of 4th degree and taking the x-coordinate as constant, one defines two transformation equations for every scan line

$$\bar{X} = A_o + A_1 y + A_2 y^2 + A_3 y^3 + A_4 y^4$$
$$\bar{Y} = B_o + B_1 y + B_2 y^2 + B_3 y^3 + B_4 y^4 \qquad (14)$$

where the newly defined coefficients A_i, B_i vary for individual scan lines again with 4th degree polynomial change of x

$$A_i = a_{i0}+a_{i1}x+a_{i2}x^2+a_{i3}x^3+a_{i4}x^4$$
$$B_i = b_{i0}+b_{i1}x+b_{i2}x^2+b_{i3}x^3+b_{i4}x^4. \quad (15)$$

Since the scanning can essentially be accomplished only in a straight direction, each scan line should be broken down into several straight segments which approximate the original non-linear functions (14) in a polygon mode. The shorter the segments are the better fit is obtained. The remaining systematic discrepancies in segmenting of scan lines should be lower than the basic resolution of the RBV system but, on the other hand, one can not gain on accuracy when these discrepancies are smaller than the random residuals of the transformation fit. Fig. 8 based on the data from the previous analysis gives an insight into the distribution of errors in the centres of segments for vertical scan lines with spacing equal to half the réseau interval. Each scan line is broken down into eight segments within the réseau grid. Only in the right upper corner the deviations for x slightly exceed the magnitude of the resolution element. If the number of segments is increased to 16 the deviations will drop down to approximately 25% of the listed errors.

It is obvious that it is not efficient to change the correction for each separate scan line. One can decide on a suitable interval of the stepwise alteration of the control depending upon the rate of changes in the x-direction (normal to scanning direction). For the analyzed example this change can be programmed for only each 10th to 20th line even in most critical peripheral zones.

Cartographic errors

The RBV image represents an instantaneous perspective view of the Earth scene with an internally rigid bundle of projecting rays and with practically negligible tilt of the camera. The corresponding area on the ground can, therefore, be outlined with regularly distributed four corner points whose connection lines on the Earth surface are slightly bent inwards the figure. With reference to the analysis of the Earth curvature error in [1] or [2] this bend (max. 50 m) can be practically neglected.

However, the RBV image is reproduced and transformed into the UTM plane. As long as the scene is located close to the central meridian of the UTM zone there is no further geometric degradation of the image. The distortion of the 185 km long side of the figure is smaller than 25 m. With the increase of the distance from the centre the distortion also increases. In accordance with the basic definition of the UTM projection the errors in scale and location of a mapped detail are as follows

$$dm = Y^2/2R^2$$
$$dX = dm.\Delta X = \Delta XY^2/2R^2 \quad (16)$$
$$dY = Y^3/6R^2$$

Considering the centre C of the RBV scene 300 km off the X-axis and setting the size of the frame $\Delta X=\Delta Y=185$ km, while the Earth radius is approximately $R=6378$ km, one derives the following table:

Y	dm	dX	dY
208 km	0.00053	98 m	37 m
300	0.00111	205	111
393	0.00190	351	249

These values determine the primary distortion dX, dY at nine standard points as shown in Fig. 9a. After introducing corrective shift and scaling, the distribution of errors improves as given in Fig. 9b. In this case the distance 300 km corresponds to the location at the border line of 6° UTM zone for the latitude $\Phi=24°$. For the southern part of Canada ($\Phi=42°$) the maximum absolute value of the UTM distortion is not larger than about 70 m which is practically negligible.

MSS IMAGERY

A comprehensive treatise of the precision processing of MSS images was presented in [2]. Only general outline of metrical problems is given here, followed by more detailed discussion in a narrower field.

Geometric considerations

The MSS system belongs to the category of panoramic cameras yielding dynamic geometry of the image. The image is formed by sequential scanning across the track of the flight and can be

classified as a hybrid projection. Individual scan lines render a perspective view at the terrain while the imaging of ground details aligned along the track is defined theoretically as an orthographic projection (see Fig. 10). The motion of the projection centre during individual scans can be neglected when selecting its average position for each scan line. Thus, a composite system of perspective plane bundles can be adopted as a representation of the proper geometry of imaging (Fig. 11).

There are several instrumental, physical and geometric factors that affect the simple relationship described above, causing geometric distortions of the image. Some of the factors are related to individual scan lines and can be defined apriori from a knowledge of the basic orbital data. These are system errors, such as panoramic distortion, non-linearity of the sweep etc.

Other distortions involve unpredictable factors and can not be coped with so easily. In this case, absolute and independent information has to be available to assess or control the deformation. This can be conveniently based upon the use of ground control data.

Scanning distortion due to Earth rotation

The Earth rotation is one of the main factors significantly affecting the fidelity of MSS imagery. While the satellite is progressing along its orbit, the Earth is rotating around its axis which is different from the axis of the orbit. As a result, the individual scan lines projected into the mapping plane are mutually shifted and rotated, thus causing a variable distortion of the image geometry. This is illustrated in Fig. 12. In the first approximation the relative change of the satellite position is defined by summing up the vectors $\overline{P_oP_S}$ and $\overline{P_SP_L}$ expressing the orbital and Earth rotation change, respectively. In accordance with Fig. 2 and with the derivations in [2] one can describe the linear change by the components

$$\Delta x = S_x R.d\rho \quad , \quad \Delta y = S_y R.d\rho$$

where

$$S_x = \frac{\omega_e}{\omega_s} \sin\varepsilon \tag{17}$$

$$S_y = -\frac{\omega_e}{\omega_s} \cos\varepsilon \sin\rho \tag{18}$$

An important conclusion can be drawn, stating that the rate of x-changes (the x-scale change) is constant for any part of the orbit so the spacing of scan lines is invariably increased, while the transversal y-shift of scan lines increases in proportion with the distance from the orbit vertex V.

In fact, the continuous increase of H_e (see increased Δ in Fig. 2) causes the subsatellite path to deviate gradually from its nominal direction when the subsatellite nadir point P_o is transferred into position P_E as shown in Fig. 12. At the same time the scan direction at P_E is kept perpendicular to the nominal satellite track and is at an angle H_s with the local geographic parallel. The additional displacement $d = \overline{P_LP_E}$ characterizes the curvature of the subsatellite track on the Earth sphere and can be derived by integrating the effect of sequential changes ΔH_e

$$d = \int_{\rho_o}^{\rho} \Delta H_e R.d\rho \ .$$

The natural curvature of the real subsatellite path is assessed by the derivative of H_s with respect to ρ as follows

$$\frac{dH_s}{d\rho} = -\sin\varepsilon \frac{\sin\Phi}{\cos^2\Phi} \ . \tag{19}$$

According to Table 1 the values H_e are always small ($\leq 4^\circ$) so that tan H_e can be substituted by H_e in radians which yields

$$\Delta H_e = \frac{\omega_e}{\omega_s} \cos\varepsilon (\sin\rho - \sin\rho_o) . \tag{20}$$

When solving the integral for d with the use of (20) one gets

$$d = R \frac{\omega_e}{\omega_s} \cos\varepsilon (\cos\rho_o - \cos\rho - \Delta\rho \sin\rho_o \tag{21}$$

Scanning pattern in the UTM plane

The above analyzed dynamic characteristics of the scanning geometry are subject to change when projecting a part of the MSS continuous strip into the UTM plane. In accordance with the theory of the UTM projection the local convergence of the geographical system $(\Phi,\Delta\lambda)$ towards the X, Y coordinate system is defined by

$$C = \Delta\lambda \cos\Phi + r \qquad (22)$$

where the difference $\Delta\lambda$ is related to the central meridian of the zone and r represents remaining higher order terms which can be neglected in this analysis.

According to Fig. 13 the heading related to the new reference axis X is defined by the value K

$$K = H - C \qquad (23)$$

Similarly, the nominal heading in the UTM system is redefined as

$$N = H_s - C \qquad (24)$$

and for the direction of scanning S, which always stays normal to the nominal heading, one gets

$$S = N - 90^\circ = H_s - 90^\circ - C \qquad (24a)$$

These three basic parameters of MSS scanning could be analyzed by appropriate substitutions for H, H_s and C into the previous equations. However, it is simpler to restrict the analysis only to changes dK, dN, dS which occur with respect to the basic change $d\rho$ along the track.

It follows from equation (1)

$$\frac{d\Phi}{d\rho} = -\frac{\cos\varepsilon \sin\rho}{\cos\Phi}, \qquad (25)$$

from (2) and (2b) we get

$$\frac{d\lambda_s}{d\rho} = \frac{\sin\varepsilon}{\cos^2\Phi} \qquad (26a)$$

and from (4)

$$\frac{d\lambda_e}{d\rho} = \frac{\omega_e}{\omega_s} \qquad (26b)$$

so that

$$\frac{d\lambda}{d\rho} = \frac{\sin\varepsilon}{\cos^2\Phi} + \frac{\omega_e}{\omega_s} . \qquad (27)$$

The rate of change for H_s is already given in equation (19). Similarly, from (5) and (1) it results when setting $\cos^2 H_e \doteq 1$

$$\frac{dH_e}{d\rho} = \frac{\omega_e}{\omega_s} \sin\Phi \qquad (28)$$

so that

$$\frac{dH}{d\rho} = -\sin\varepsilon \frac{\sin\Phi}{\cos^2\Phi} + \frac{\omega_e}{\omega_s} \sin\Phi . \qquad (29)$$

Finally, one derives the following relation from (21) with the use of (25) and (27)

$$\frac{dC}{d\rho} = -\Delta\lambda\cos\varepsilon \sin\rho - \sin\varepsilon \frac{\sin\Phi}{\cos^2\Phi} - \frac{\omega_e}{\omega_s} \sin\Phi \qquad (30)$$

The first two terms represent change $dC_s/d\rho$ due to satellite orbiting, the remaining term expresses the change $dC_e/d\rho$ due to the Earth rotation. It is interesting to note that the second term in (30) is identical with (19), which can be interpreted that meridional convergence of the UTM system cancels the curvature of the nominal track in the geographical system when mapping into the UTM plane. On contrary, the last term in (30) is opposite in sign to (28) which means that the natural curvature of the real subsatellite track on the Earth sphere will double when mapping into the UTM plane. Point P_E in Fig. 12 is shifted into the final position P by the amount $d = \overline{P_L P_E}$. The final curve $\overline{P_o P}$ represents the real subsatellite track in the UTM plane.

Nominal satellite track - Nominal track as defined by ρ (Φ,λ_s) and H_s represents an orthodroma on the sphere and can, therefore, be unrolled into a suitable plane as a straight line. In the UTM plane the following change of (24)

$$\frac{dN}{d\rho} = \frac{dH_s}{d\rho} - \frac{dC_s}{d\rho} \qquad (31)$$

yields with the use of (19) and (30)

$$\boxed{\frac{dN}{d\rho} = \Delta\lambda \cos\varepsilon \sin\rho} \qquad (32)$$

This rate of change depends on the distance $\Delta\lambda$ from the central meridian of the zone, and on the traveled distance ρ along the track. Table 2 lists the values $dN/d\rho$ for different

parameters Φ and $\Delta\lambda$. The maximum rate for the zone border region at the equator reaches

$$\max \frac{dN}{d\rho} = 0.052 \doteq 1/20$$

which can bend the nominal track over the interval covering approximately one frame (185×185 km) by the value of $\Delta N = 5'$.

The rigorous double integration of (32) with respect to ρ would determine the lateral deviation Δ at the end or Δ_c in the centre of the frame as shown in Fig. 14. However, since this integration is not quite easy, a simple estimate can be used instead by considering the rate $dN/d\rho$ constant over the range of a frame. In that case

$$\Delta N = \int_{\rho_o}^{\rho} dN = c\Delta\rho \qquad \text{and}$$

$$\Delta = R \int_{\rho_o}^{\rho} \Delta N d\rho = \tfrac{1}{2}\, cR\Delta\rho^2 \ .$$

In fact, the rate $dN/d\rho$ is always decreasing in the direction of the track. Consequently, the simplified estimate Δ represents a stricter criterium than the real value Δ based on the rigorous integration of the variable rate.

This means that the maximum lateral deviation of the nominal track deformed in the UTM plane will always be smaller than $\Delta_c = \frac{1}{4}\Delta$, and in accordance with

$$\Delta_c = \tfrac{1}{8}\, cR\Delta\rho^2 \qquad (33)$$

we get

$$\Delta_c < 34 \text{ m} \ .$$

This is far below the resolution element of MSS.

Real subsatellite track - When differentiating (23) we get

$$\frac{dK}{d\rho} = \frac{dH}{d\rho} - \frac{dC}{d\rho} = \frac{dH_e}{d\rho} - \frac{dC_e}{d\rho} + \frac{dN}{d\rho} \qquad (34)$$

which results in

$$\boxed{\frac{dK}{d\rho} = 2\,\frac{\omega_e}{\omega_s} \sin\Phi + \Delta\lambda \cos\varepsilon \sin\rho} \qquad (35)$$

The first term in (35) is always positive for northern hemisphere and represents the combined effect of the natural bend of the track on the Earth sphere and of the meridional convergence being dependent only on the geographical latitude. The other term also varies with the latitude and, in addition, with the longitudinal distance $\Delta\lambda$ from the central meridian of the zone. The maximum composite rate will occur at the east of the zone where both terms of (35) are positive. After substituting for $\sin\Phi$ from (1) into (35) the location of the maximum curvature of the track can be determined as $\Delta\lambda = 3^\circ$ and $\rho = 20^\circ$ ($\Phi = 68^\circ$). This maximum curvature amounts to

$$\max \left|\frac{dK}{d\rho}\right| = 0.15 \doteq 1/6.6$$

which in accordance with (33) determines

$$\max \Delta_c < 100 \text{ m} \ .$$

Conclusion can be drawn that the maximum angular bend of the track within the range of a frame can in the UTM plane come close to $\Delta K = 15'$ while the maximum lateral deviation of the track from a straight line is always smaller than 100 m. Table 3 lists the rates $dK/d\rho$ for different locations in a UTM zone. With the exception of the left lower corner all the values are positive. Accordingly, over the Canadian territory the track is always bent in the same way (with concave shape on the oriented map).

Scanning direction - The scanning process maintains the direction normal to the nominal track, even though the scan line is displaced by the effect of Earth rotation. In accordance with Fig. 12 the scan line at the point P goes perpendicular to the virtual track passing through P_E. As a matter of fact, it is deviated by the convergence angle γ towards the initial scan direction at the point P_o. In analytical terms and according to (24a)

$$\frac{dS}{d\rho} = \frac{dH_s}{d\rho} - \frac{dC}{d\rho} = -\frac{dC_e}{d\rho} - \frac{dN}{d\rho}$$

and after substituting from (30) and (32) we get

$$\boxed{\frac{dS}{d\rho} = \frac{\omega_e}{\omega_s} \sin\Phi + \Delta\lambda \cos\varepsilon \sin\rho} \qquad (36)$$

which is an expression similar to (35) from the previous section. The first term is only half the magnitude of the relevant part in (35). This again can be attributed to the fact that the Earth rotation affects the position of the satellite but not the direction of scanning. The rate of change for S lies always between the rates for K and N.

The maximum absolute rate $dS/d\rho$ is found for $\Delta\lambda = 3°$ and $\rho = 36°$ ($\Phi=53°$), namely

$$\max \frac{dS}{d\rho} = 0.09 \doteq 1/11$$

which means that maximum convergence of scan lines within a frame is always smaller than 9':

$$\max \Delta S = \max \gamma < 9' .$$

Table 4 completes the insight into the analysis.

Total distortion of an MSS frame in the UTM plane

The previous sections show how each scan line as an individual entity is displaced and rotated in the UTM plane by the combined effect of the dynamic factors involved and of the geometric limitations of the cartographic projection. The analysis should still be completed by considering the exterior geometry of an individual scan line projected on the ground and mapped into the UTM plane.

Since the subsatellite nadir point travels 215 m during one scan, all the scan lines projected on the ground become primarily slanted. The slant angle which is constant and relatively small (4') represents an error caused by the orbital dynamics. The mapping process, however, may still introduce another inaccuracy by bending individual scan lines in an analogous way to the above discussed bend of the nominal subsatellite track. An analysis could prove that this error is far below the accuracy expectations and can, therefore, be omitted in the final assessment of the geometric distortion of the MSS imagery.

Although the changes by the attitude errors and their variations represent a significant distortion factor, they will not be treated in this presentation and a mere reference to [2] is made for additional information.

Fig. 15a summarizes, for the area of one MSS frame, what was analyzed in previous sections. The frame is stretched out in the direction of the track by the value Δx which represents a constant affine change of approximately 1% for a scene located in any latitude. The skew of the frame caused by Δy gradually increases with the progress of the satellite due south, and reaches its maximum of approximately 7% relative change at the equator. Both these errors were defined by the factors S_x, S_y in equations (17), (18). The curvature of the real subsatellite track causes the longitudinal bend of the frame as characterized by the value Δ at the point E in Fig. 15a or by Δ_c at the point C in Fig. 15b. Both these values can be derived from the curvature $dK/d\rho$ of equation (35) as listed in corresponding Table 3. Finally, the convergence of scan lines deforms the frame by the value of γ which depends on the integration of the rate of change $dS/d\rho$ as expressed by (36) and tabulated in Table 4.

All the mentioned changes of the frame are tabulated in Table 5 with respect to the geographic latitude Φ. Fig. 15b then indicates how the original discrepancies of Fig. 15a are distributed when orienting and scaling the frame in along-the-track direction. The remaining errors, however, have to be eliminated only by the controlled process of the MSS reproduction in scanner-printers.

Precision processing of MSS images

While RBV images have to be reproduced with the use of the EBR which is capable of line segmenting, MSS imagery is produced with full line control in either EBR or LBIR. In order to produce MSS corrected images one has to formulate a mathematical model of its geometry. It is convenient to use two basic steps in formulating the solution. First, a preliminary correction is defined using all the internal, apriori definable distortions, such as sweep non-linearity and panoramic distortion. Only then the internally corrected nominal image can be transformed and oriented externally when compared with ground control and other absolute information. The

analytical terms of the solution adopted for producing the corrected image should also cover the effect of the above discussed Earth rotation and cartographic distortions. In other words, one has to solve the absolute orientation of a complete imaging model which inherently defines all relevant dynamic factors. Such a mathematical model for the MSS precision processing was described in detail in [2]. The model includes the functional formulation of the pseudo-periodical variations by means of either a harmonic analysis or a polynomial transformation.

REFERENCES

[1] A.P. Colvocoresses. ERTS-A Satellite Imagery. Photogrammetric Engineering 1970, No.6, pp.555-560.

[2] V. Kratky. Precision Processing of ERTS Imagery, Technical Papers from the 1971 ASP Fall Convention, San Francisco 1971, pp.481-514.

[3] R.B. McEwen. Geometric Calibration of the RBV System. Paper to 7th International Symposium on Remote Sensing of Environment, The University of Michigan, Ann Arbor 1971.

[4] K.W. Wong, E.V. Gamble, R.E. Riggins. Geometric Analysis of the RBV Television System. Civil Engineering Studies, Photogrammetric Series No. 33, University of Illinois, Urbana, Ill., July 1971.

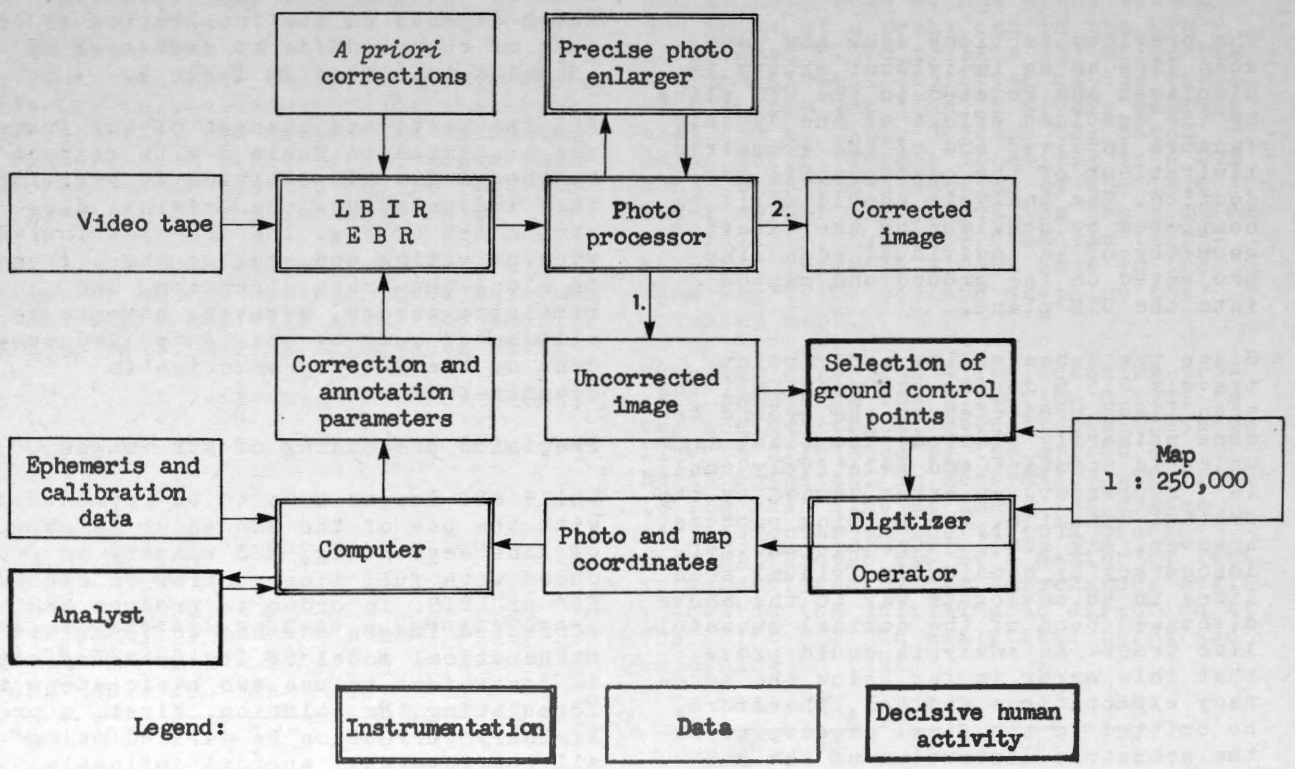


Fig. 1 ERTS precision processing

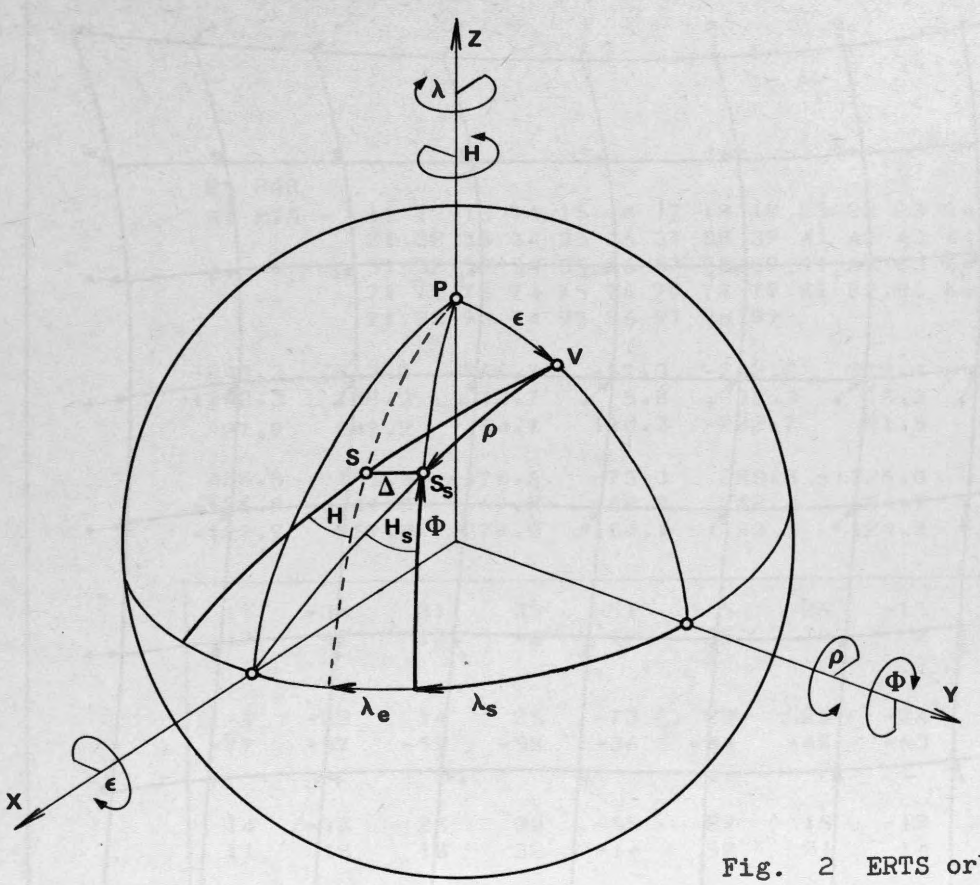


Fig. 2 ERTS orbital imagery

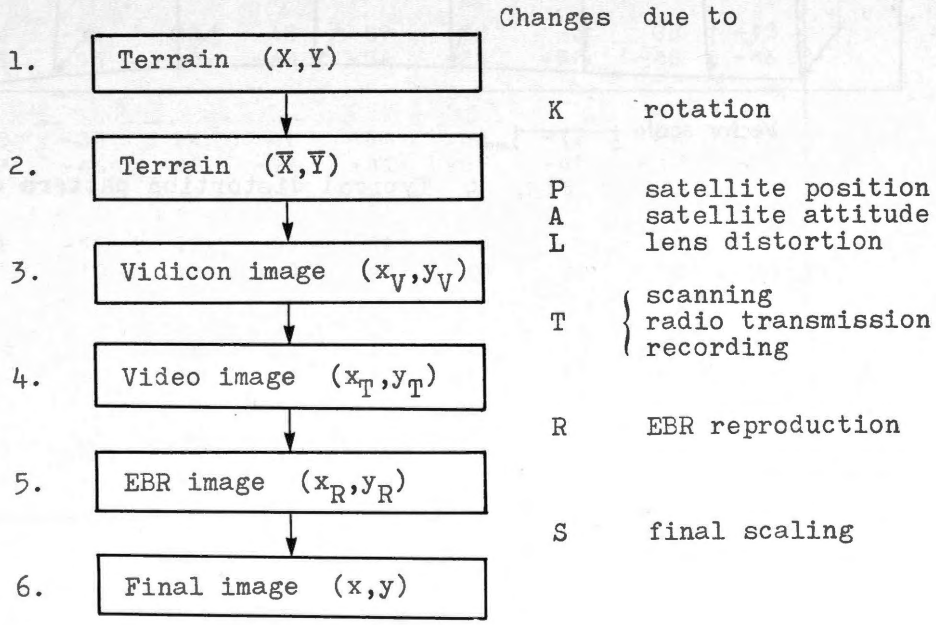


Fig. 3 Flow of information in RBV image processing

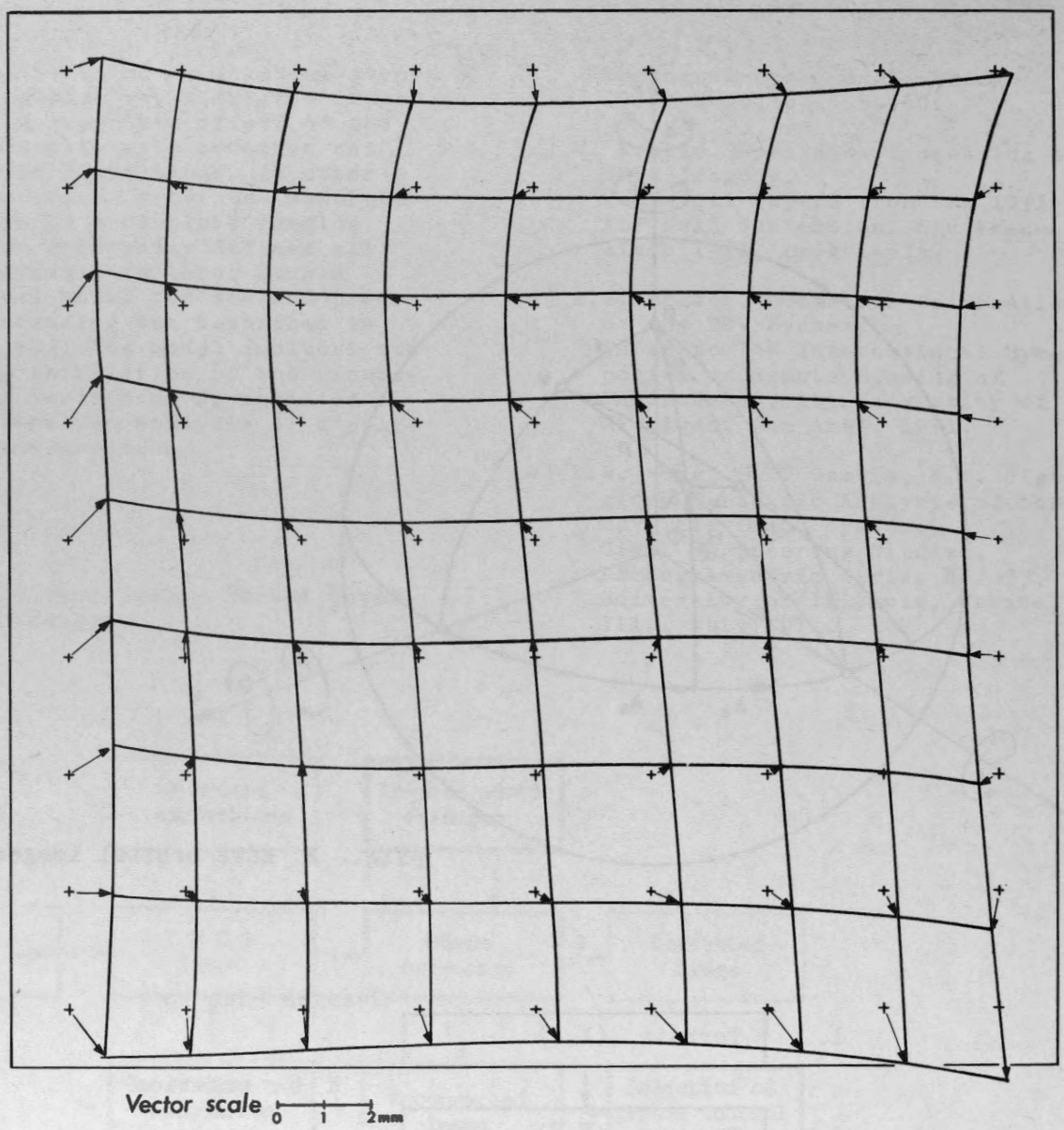


Fig. 4 Typical distortion pattern of the RBV image

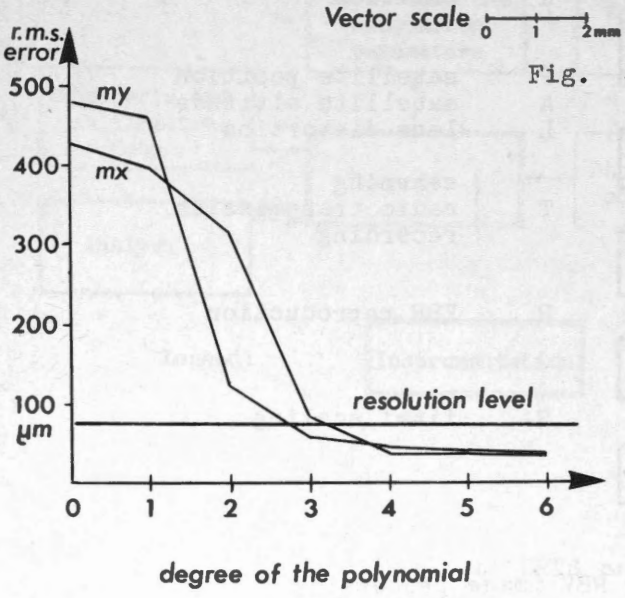


Fig. 5 Residual RBV errors

25 PAR
81 PTS - 11 12 13 14 15 16 17 18 19 21 22 23 24 25 26 27 28 29
31 32 33 34 35 36 37 38 39 41 42 43 44 45 46 47 48 49
51 52 53 54 55 56 57 58 59 61 62 63 64 65 66 67 68 69
71 72 73 74 75 76 77 78 79 81 82 83 84 85 86 87 88 89
91 92 93 94 95 96 97 98 99

-243.3	813.7	-966.7	-57.3	-242.2	279.9	394.6	56.9	-229.9
-1342.3	788.3	12.7	5.8	233.3	16.3	72.5	518.7	291.7
27.9	247.2	54.7	110.3	-202.7	21.5	-27.1		
408.8	155.5	-170.6	73.3	283.8	-1126.0	44.9	-504.2	-367.2
-526.9	-41.3	-69.8	-60.2	262.4	654.7	70.9	-159.6	25.9
-122.9	259.8	279.0	184.1	80.5	124.8	-143.5		

11	-36	31	33	-51	5	26	-15	9
19	49	10	-4	19	35	19	2	24
1	-22	14	25	-73	28	26	-24	-15
-77	-57	-59	-58	-36	-44	-48	-63	-48
14	-18	25	30	-55	29	18	-19	22
11	18	10	32	-14	52	21	14	24
-4	-28	14	36	-60	29	2	-20	10
112	78	79	89	-6	72	62	62	63
20	-35	20	45	-59	20	-3	28	-13
-55	-55	-30	-40	-36	-28	-50	-48	-46
12	-35	14	47	-54	-2	-12	20	-19
-49	-46	-42	-26	-31	-25	-41	-42	-37
12	-52	19	49	-37	3	-7	14	-4
5	13	15	-3	42	14	36	54	36
14	-30	26	34	-51	18	-13	22	0
45	40	24	3	43	-12	-13	10	-4
7	-27	12	37	-54	28	-14	10	-7
-22	-7	-27	-7	3	-10	-10	1	-6

Residual deviations

dx dy in um

ST. DEV. 34.4 49.2

Fig. 6 Residual errors after the 4th degree transformation

25 PAR
81 PTS - 11 12 13 14 15 16 17 18 19 21 22 23 24 25 26 27 28 29
31 32 33 34 35 36 37 38 39 41 42 43 44 45 46 47 48 49
51 52 53 54 55 56 57 58 59 61 62 63 64 65 66 67 68 69
71 72 73 74 75 76 77 78 79 81 82 83 84 85 86 87 88 89
91 92 93 94 95 96 97 98 99

V$_o$-MATRIX

1.69	1.60	1.44	1.44	1.75	2.16	2.22	1.98	3.88	16.43
0.40	0.38	0.34	0.34	0.41	0.51	0.52	0.47	0.91	3.88
0.20	0.19	0.17	0.17	0.21	0.26	0.27	0.24	0.47	1.98
0.23	0.22	0.19	0.20	0.24	0.29	0.30	0.27	0.52	2.22
0.22	0.21	0.19	0.19	0.23	0.28	0.29	0.26	0.51	2.16
0.18	0.17	0.15	0.15	0.19	0.23	0.24	0.21	0.41	1.75
0.15	0.14	0.13	0.13	0.15	0.19	0.20	0.17	0.34	1.44
0.15	0.14	0.13	0.13	0.15	0.19	0.19	0.17	0.34	1.44
0.16	0.16	0.14	0.14	0.17	0.21	0.22	0.19	0.38	1.60
0.17	0.16	0.15	0.15	0.18	0.22	0.23	0.20	0.40	1.69

0.248

Fig. 7 Theoretical coordinate variances for transformed points

25 PAR
81 PTS - 11 12 13 14 15 16 17 18 19 21 22 23 24 25 26 27 28 29
31 32 33 34 35 36 37 38 39 41 42 43 44 45 46 47 48 49
51 52 53 54 55 56 57 58 59 61 62 63 64 65 66 67 68 69
71 72 73 74 75 76 77 78 79 81 82 83 84 85 86 87 88 89
91 92 93 94 95 96 97 98 99

SEGMENT. ERRORS IN μM

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	34	40	44	48	51	53	55	58	60	63	66	69	72	75	78	80
-10	-11	-12	-14	-15	-17	-18	-18	-17	-16	-13	-9	-3	5	14	25	39
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	22	26	30	32	34	36	37	38	40	41	42	44	46	47	49	50
-7	-9	-11	-12	-14	-16	-17	-18	-18	-18	-16	-14	-11	-7	-2	6	15
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	12	15	17	19	20	21	22	22	23	23	24	24	25	26	26	27
-7	-9	-11	-12	-14	-15	-16	-17	-18	-19	-19	-18	-18	-16	-13	-9	-4
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	3	5	6	7	8	9	9	9	10	10	10	10	10	10	11	11
-12	-13	-13	-14	-15	-15	-16	-17	-18	-19	-20	-21	-21	-21	-21	-20	-18
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-6	-5	-3	-2	-2	-1	0	0	0	0	0	0	1	1	1	2	3
-20	-19	-18	-17	-16	-16	-16	-17	-17	-18	-20	-21	-22	-24	-25	-26	-26
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-11	-10	-10	-9	-8	-8	-7	-6	-6	-5	-5	-4	-4	-3	-1	0	3
-31	-27	-24	-21	-19	-18	-17	-16	-17	-17	-18	-20	-21	-23	-25	-27	-29
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-14	-15	-14	-14	-13	-12	-10	-9	-8	-7	-5	-4	-2	0	2	6	10
-47	-39	-32	-27	-23	-20	-18	-16	-15	-15	-16	-16	-18	-19	-21	-24	-27
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-17	-18	-17	-16	-15	-13	-11	-9	-7	-5	-2	1	4	8	12	18	25
-66	-53	-42	-34	-27	-22	-19	-16	-14	-13	-12	-11	-12	-12	-14	-16	-19
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

1st 2nd --- segment

18.5 13.9

Fig. 8 Residual errors in segmented scan lines

49	102	176	dX	in m	-53	0	74
37	111	249	dY		28	0	36
0	0	0	X		0	0	0
37	111	249			28	0	36
-49	-102	-176	Y		53	0	-74
37	111	249			28	0	36

Fig. 9 UTM distortion at the zone border

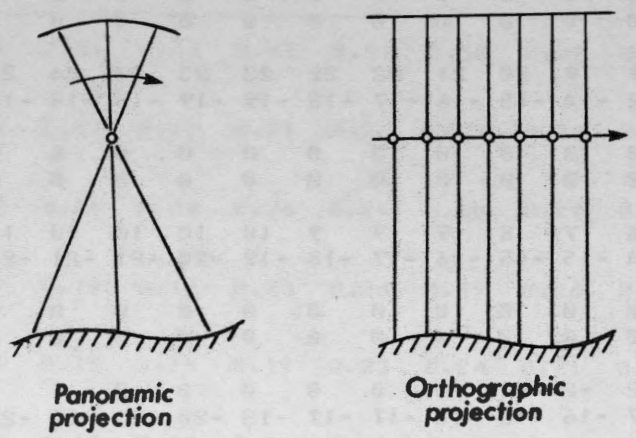


Fig. 10 Hybrid geometry of MSS imagery

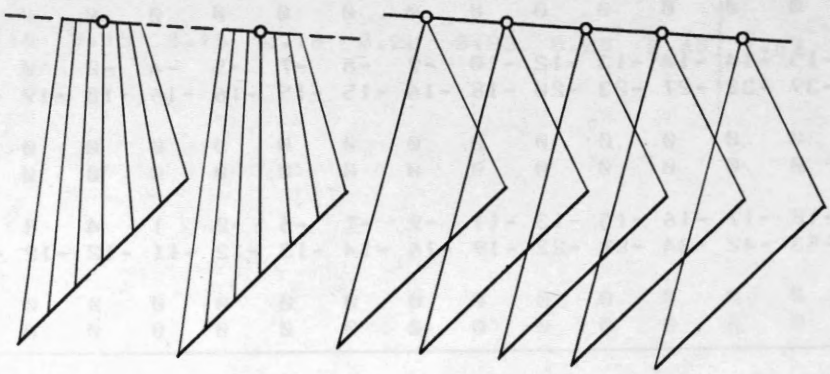


Fig. 11 Dynamic geometry of scanning

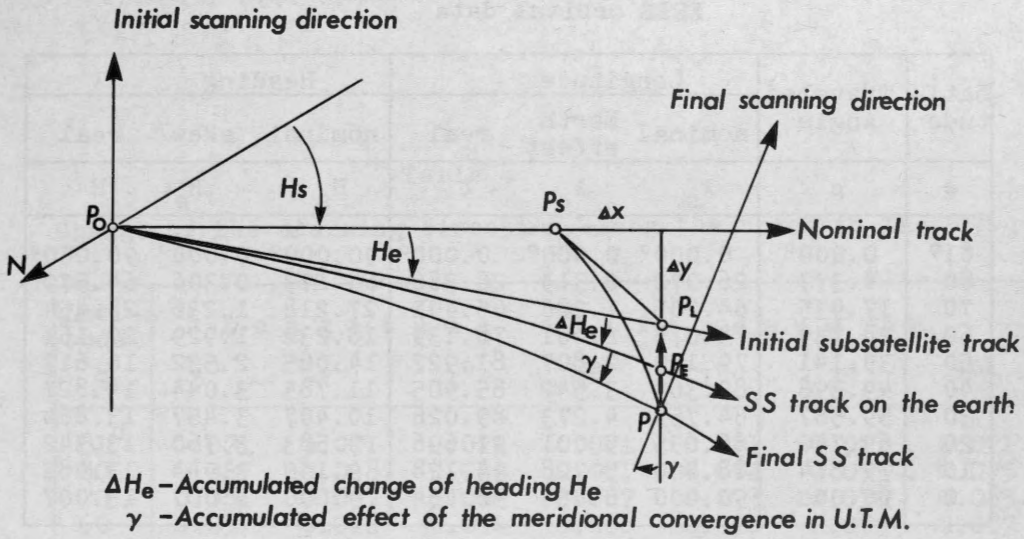


Fig. 12 Changes of the subsatellite track

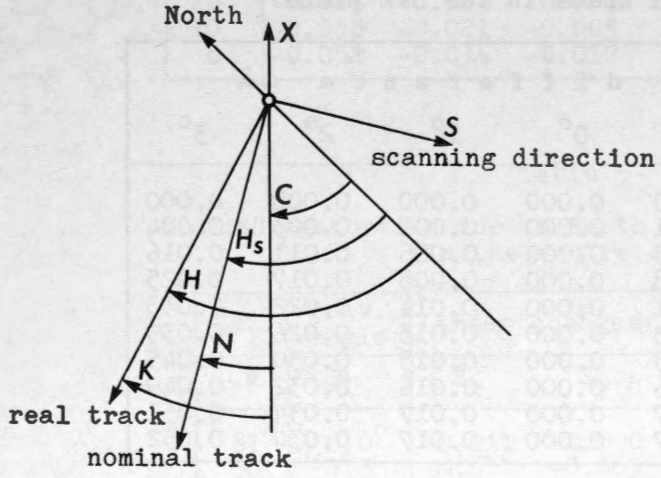


Fig. 13 Scanning parameters in the UTM plane

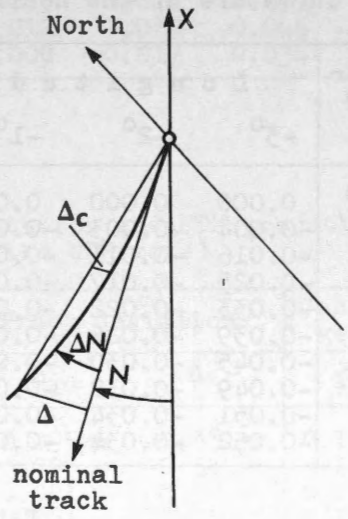


Fig. 14 Curvature of the nominal track

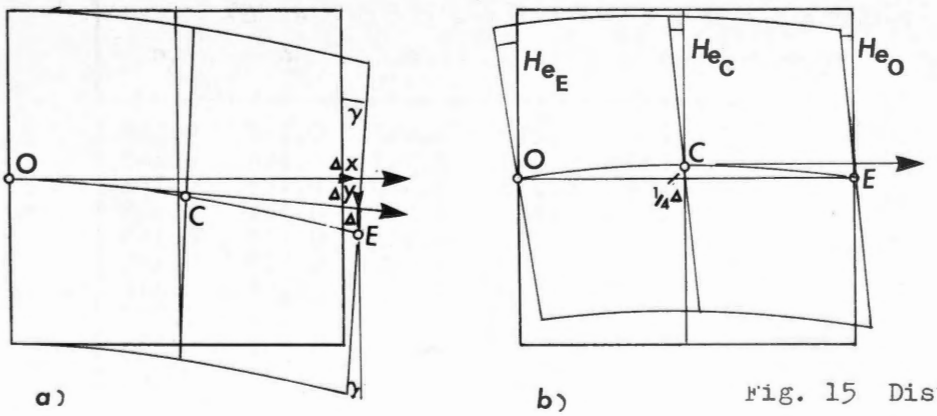


Fig. 15 Distortion of a MSS frame

Table 1

ERTS orbital data

Latitude	Traveled Angle	Longitude			Heading		
		nominal	Earth effect	real	nominal	skew	real
ϕ	ρ	λ_s	λ_e	λ	H_s	H_e	H
81°	0.000°	0.000°	0.000°	0.000°	90.000°	0.000°	90.000°
80	4.377	26.072	0.314	26.386	64.273	0.306	64.579
70	17.935	64.205	1.286	65.491	27.218	1.236	28.454
60	28.739	74.078	2.061	76.139	18.232	1.929	20.161
50	39.141	79.120	2.807	81.927	14.085	2.532	16.617
40	49.398	82.363	3.542	85.905	11.783	3.044	14.827
30	59.587	84.753	4.273	89.026	10.407	3.457	13.864
20	69.740	86.695	5.001	91.696	9.583	3.760	13.342
10	79.874	88.400	5.728	94.128	9.140	3.944	13.085
0	90.000	90.000	6.454	96.454	9.000	4.007	13.007

Table 2

Curvature of the nominal track in the UTM plane

Latitude	Longitude difference $\Delta\lambda$						
Φ	-3°	-2°	-1°	0°	1°	2°	3°
81°	0.000	0.000	0.000	0.000	0.000	0.000	0.000
80	-0.004	-0.003	-0.001	0.000	0.001	0.003	0.004
70	-0.016	-0.011	-0.005	0.000	0.005	0.011	0.016
60	-0.025	-0.017	-0.008	0.000	0.008	0.017	0.025
50	-0.033	-0.022	-0.011	0.000	0.011	0.022	0.033
40	-0.039	-0.026	-0.013	0.000	0.013	0.026	0.039
30	-0.045	-0.030	-0.015	0.000	0.015	0.030	0.045
20	-0.049	-0.032	-0.016	0.000	0.016	0.032	0.049
10	-0.051	-0.034	-0.017	0.000	0.017	0.034	0.051
0	-0.052	-0.034	-0.017	0.000	0.017	0.034	0.052

Table 3

Curvature of the real subsatellite track in the UTM plane

Latitude	Longitude difference $\Delta\lambda$						
Φ	-3°	-2°	-1°	0°	1°	2°	3°
81°	0.142	0.142	0.142	0.142	0.142	0.142	0.142
80	0.137	0.139	0.140	0.141	0.143	0.144	0.145
70	0.119	0.124	0.129	0.135	0.140	0.145	0.151
60	0.099	0.108	0.116	0.124	0.132	0.141	0.149
50	0.077	0.088	0.099	0.110	0.121	0.132	0.143
40	0.053	0.066	0.079	0.092	0.105	0.118	0.131
30	0.027	0.042	0.057	0.072	0.087	0.101	0.116
20	0.001	0.017	0.033	0.049	0.065	0.081	0.098
10	-0.026	-0.009	0.008	0.025	0.042	0.059	0.076
0	-0.052	-0.034	-0.017	0.000	0.017	0.034	0.052

Table 4

Change of the scanning direction along the track in the UTM plane

Latitude Φ	Longitude difference $\Delta\lambda$						
	-3°	-2°	-1°	0°	1°	2°	3°
81°	0.071	0.071	0.071	0.071	0.071	0.071	0.071
80	0.067	0.068	0.069	0.071	0.072	0.073	0.075
70	0.051	0.057	0.062	0.067	0.073	0.078	0.083
60	0.037	0.046	0.054	0.062	0.070	0.079	0.087
50	0.022	0.033	0.044	0.055	0.066	0.077	0.088
40	0.007	0.020	0.033	0.046	0.059	0.072	0.085
30	-0.009	0.006	0.021	0.036	0.051	0.066	0.080
20	-0.024	-0.008	0.008	0.025	0.041	0.057	0.073
10	-0.038	-0.021	-0.005	0.012	0.029	0.046	0.063
0	-0.052	-0.034	-0.017	0.000	0.017	0.034	0.052

Table 5

MSS distortion due to Earth rotation and UTM projection
(in the centre of the UTM zone)

Lat.	Trav. angle	Linear changes		Track deviation		Convergence
Φ	φ	S_x	S_y	Δ_c	Δ	γ
81°	0.0°	0.011	0.000	-0.095km	-0.381km	0.118°
80	4.3	0.011	-0.005	-0.095	-0.380	0.118
70	17.9	0.011	-0.021	-0.091	-0.382	0.112
60	28.7	0.011	-0.034	-0.083	-0.333	0.103
50	39.1	0.011	-0.045	-0.074	-0.293	0.090
40	49.4	0.011	-0.054	-0.062	-0.245	0.076
30	59.6	0.011	-0.061	-0.048	-0.190	0.058
20	69.7	0.011	-0.067	-0.033	-0.129	0.039
10	79.9	0.011	-0.070	-0.016	-0.063	0.019
0	90.0	0.011	-0.071	0.000	0.000	0.000

A QUICK-LOOK READOUT FOR ERTS MULTISPECTRAL SCANNER IMAGERY

R. Barrington and R. Hutchinson,
Communications Research Center, Ottawa.

J.S. MacDonald and D.S. Sloan,
MacDonald, Dettwiler and Associates Ltd.,
Vancouver.

INTRODUCTION

The ERTS Satellite carries two separate camera systems, one employing return-beam vidicons (RBV) as the sensing devices, and the second employing a multispectral scanner (MSS) which utilizes an oscillating mirror as the scanning mechanism. Both of these systems can be regarded, from a user point of view, as very high altitude television camera systems which produce images in three or more spectral bands of a 100 nautical mile wide strip of terrain along the satellite's path.

Receiving stations, such as the Canadian ERTS station being built at Prince Albert, Saskatchewan (PASS), receive information directly from the satellite when it is within sight of the station. Figure 1. shows a block diagram of the PASS system as it is presently planned. Basically, video data is received from the satellite and recorded on magnetic tape for shipment to a central processing facility in Ottawa where it is converted into multispectral imagery.

Since the elapsed time between the recording of the information and its subsequent conversion into pictorial form can be as long as several days, a need exists for some means of monitoring system performance at the receiving site. As shown in Figure 1, such a facility is provided in the RBV system by a so-called Quick-Look Device (QLD). A block diagram of the QLD appears in Figure 2. It consists simply of a very high resolution CRT system with an automatic camera for photographing the face of the CRT. The RBV system is well suited to this type of monitoring because of the following factors:

1. It uses analog transmission.
2. Its spectral bands are sent sequentially.
3. Conventional electronic raster scanning is used.

This makes it possible to use a simple television-type reproduction facility such as that depicted in Figure 2 to capture all of the RBV data.

In contrast to the RBV system, the MSS system uses simultaneous digital transmission of all its spectral bands, and does not use a conventional raster scanning technique. Thus the problem of designing a QLD for the MSS system is somewhat more difficult.

MULTISPECTRAL SCANNER SYSTEM

As shown in Figure 3, the MSS camera consists of a rocking mirror and telescope system which focuses the image simultaneously onto 26 photosensors. The data from these sensors is transmitted to the ground, and comes out of the Demultiplexer as 25 digital data streams. As shown in Figure 3, the first 24 sensors are arranged in 4 columns of 6 sensors. Each column corresponds to a different spectral band. Thus as the mirror moves the image past the sensors, 6 scan lines are produced simultaneously, one from each of the 6 sensors (2 for the thermal IR) in a given column. Since the data from 6 scan lines are received at the same time, the simple scanning technique employed in the RBV Quick-Look system will not suffice.

The most natural scanning method for this type of scan would seem to be the zig-zag scan similar to that shown in Figure 4, however several other properties of the MSS system make this approach impractical. As mentioned previously, the MSS system uses a rocking mirror to achieve the horizontal scan. This scanning is not stable; there may be anywhere from 3000 to 3400 picture elements in a given sweep. Thus if the QLD were to reproduce each set of lines as it is received, there would be severe line centering error due to this line-length variation. Also the MSS system transmits all of its spectral bands simultaneously. A simple QLD can produce imagery from only one of these bands at a time, so that if the remaining spectral bands are to be reproduced,

the system must be capable of accepting input from the digital tape recorder. Since the output from each sensor is recorded on a separate track of the recorder, tape skew must be allowed for.

MSS QUICK-LOOK DEVICE

All of the above factors lead to the conslusion that an MSS Quick-Look system must be capable of buffering a single scan of information. Such a system is shown in block diagram form in Figure 5. It uses essentially the same display and automatic camera system as the present RBV-QLD, augmented by auxillary electronics designed to handle the requirements of the MSS system.

The output data format from the demultiplexer is shown in Figure 6. Video data is output during approximately one half of the scan interval, the other half being occupied by preamble and calibration data. The video data is stored in the buffer memory during the half of the scan when it is available, and is read out to the CRT system during the other half of the scan. The selection of the 6 channels corresponding to the desired spectral band is accomplished by the INPUT SELECT circuit. The thermal IR data is handled by demultiplexing channel 25 and distributing the data such that each IR element is the same size as 9 elements from the other bands. The CONTROL AND TIMING circuit synchronizes the system to compensate for the skew present in the signals from the tape recorder. Data from the 6 channels of the selected spectral band is thus read simultaneously into 6 4K sections of memory. The line length information present in words 3773, 3774 and 3775 is stored by the LINE LENGTH circuits.

When the data from a single scan has been stored, the 6 segments of memory are read out sequentially to the CRT system. The CONTROL AND TIMING circuits generate the horizontal and vertical sweep signals in digital form which are fed to the QLD through D/A converters. In addition, the LINE LENGTH information is used to determine on which horizontal sweep-word the output of data commences. This corrects for variations in the sweep time of the rocking mirror in the satellite. The horizontal sweep contains a non-linear transfer function to compensate for the sinusoidal sweep of the rocking mirror. Using read-only memories in the SCAN LINE SELECT circuits it is possible to compensate for variations in sensitivity between individual sensors in the satellite.

With the proposed scheme, which has no capability to store more than a single scan of data, a continuous-motion film transport is required if all the data is to be captured easily. The presently proposed RVB-QLD uses a step-change type of transport which requires a few tenths of a second to change film. Several scans of information would be missed during the film-changing operation using this type of transport. For system monitoring functions, this problem is not a serious drawback, but if the Quick-Look Device is to be used as a source of rapidly available medium-quality imagery for users, then a continuous-motion transport is necessary.

CONCLUSIONS

The proposed MSS Quick-Look Device has the capability of producing imagery with the major causes of distortion and error removed. This imagery will certainly be adequate for monitoring system operation. The suitability of the image quality to potential user requirements remains to be tested, however it is our feeling that there are many applications where the imagery produced by the MSS-QLD will be adequate, and where its rapid availability will be an advantage.

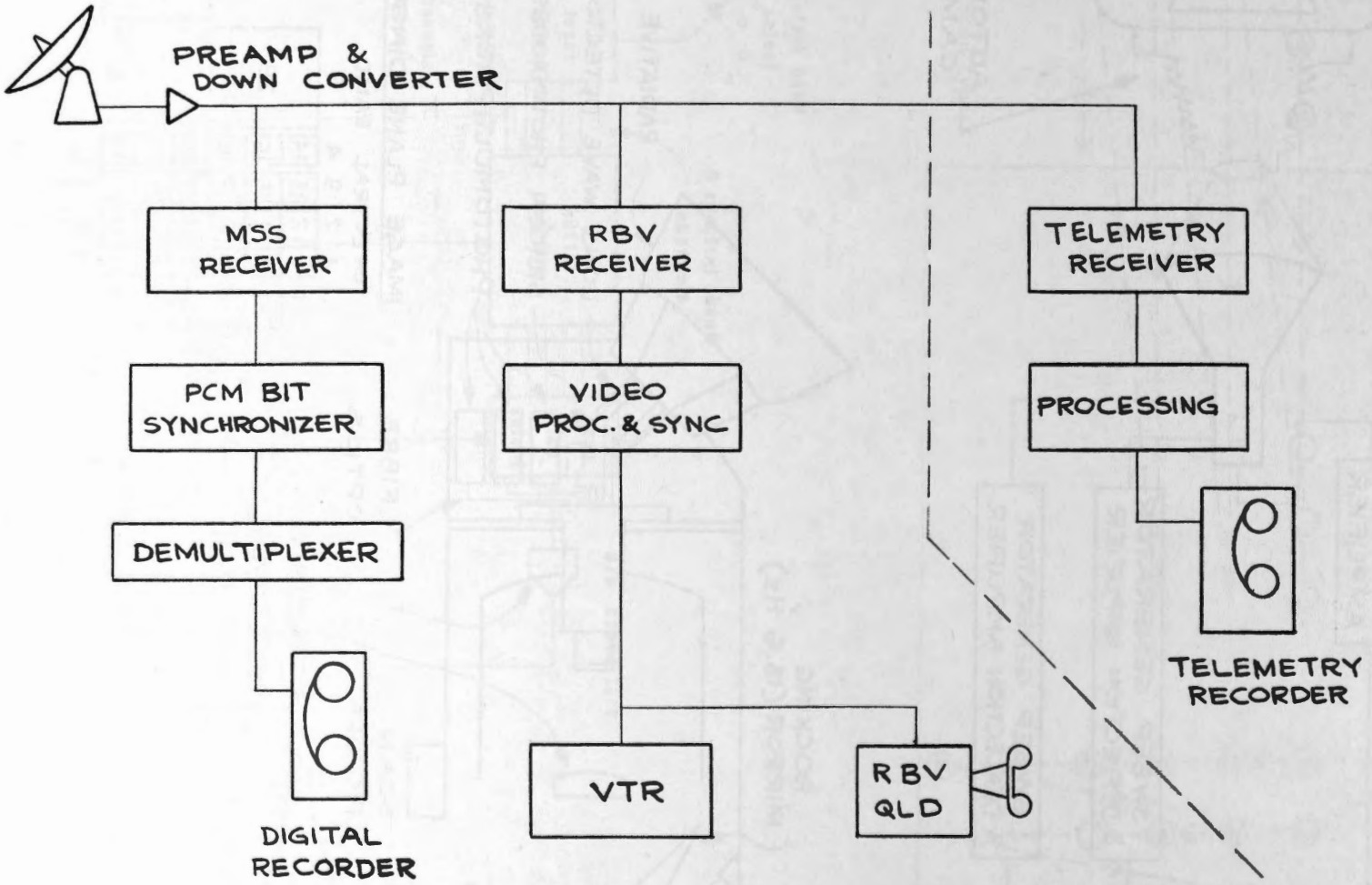


Figure 1. Pass block diagram

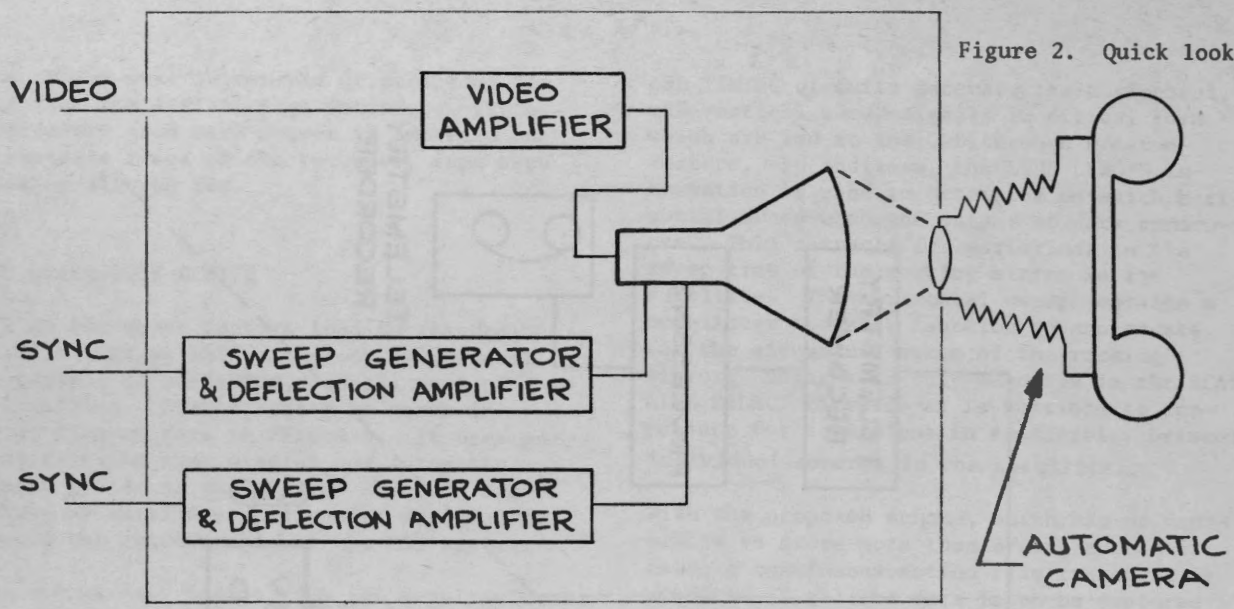


Figure 2. Quick look device

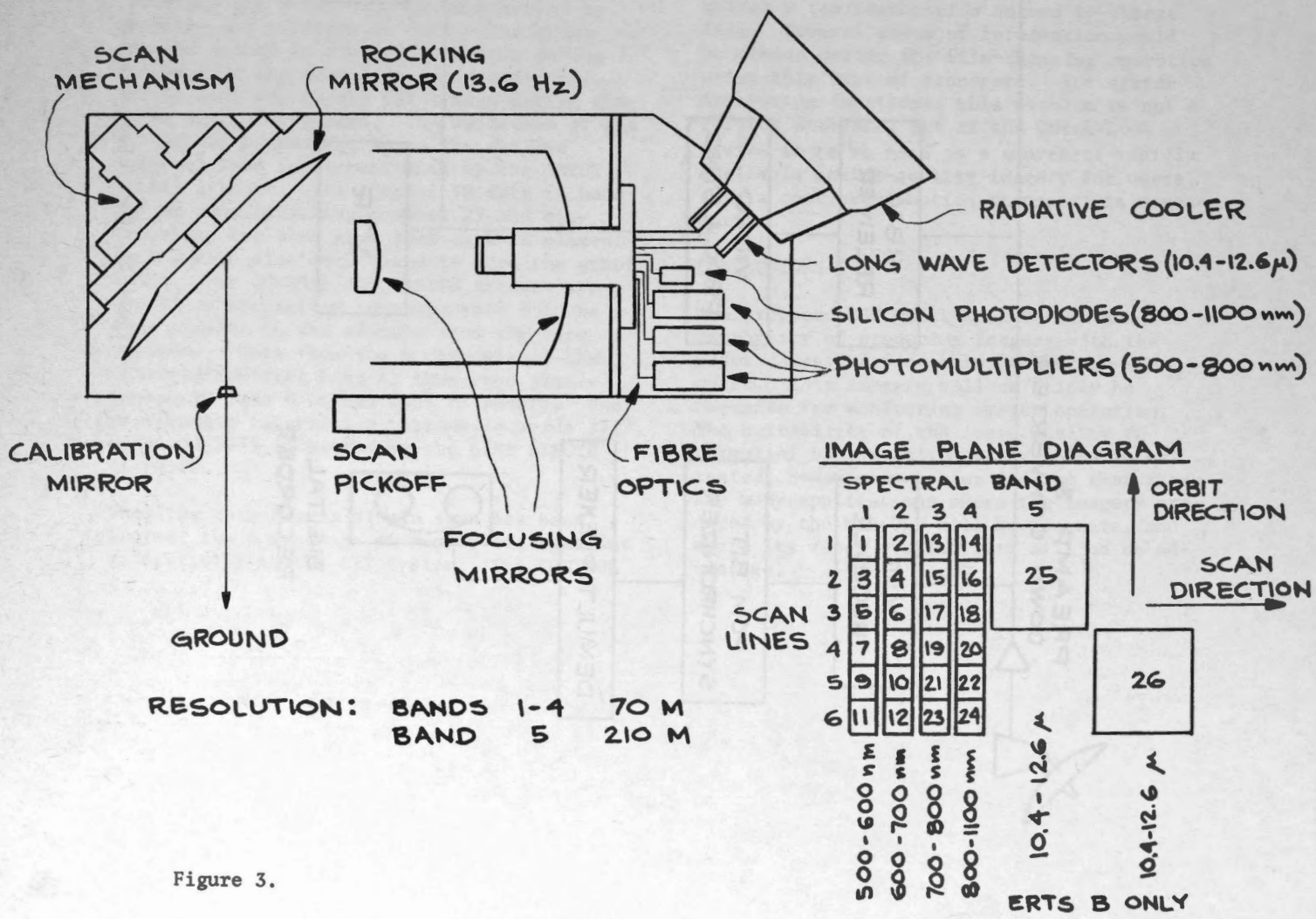


Figure 3.

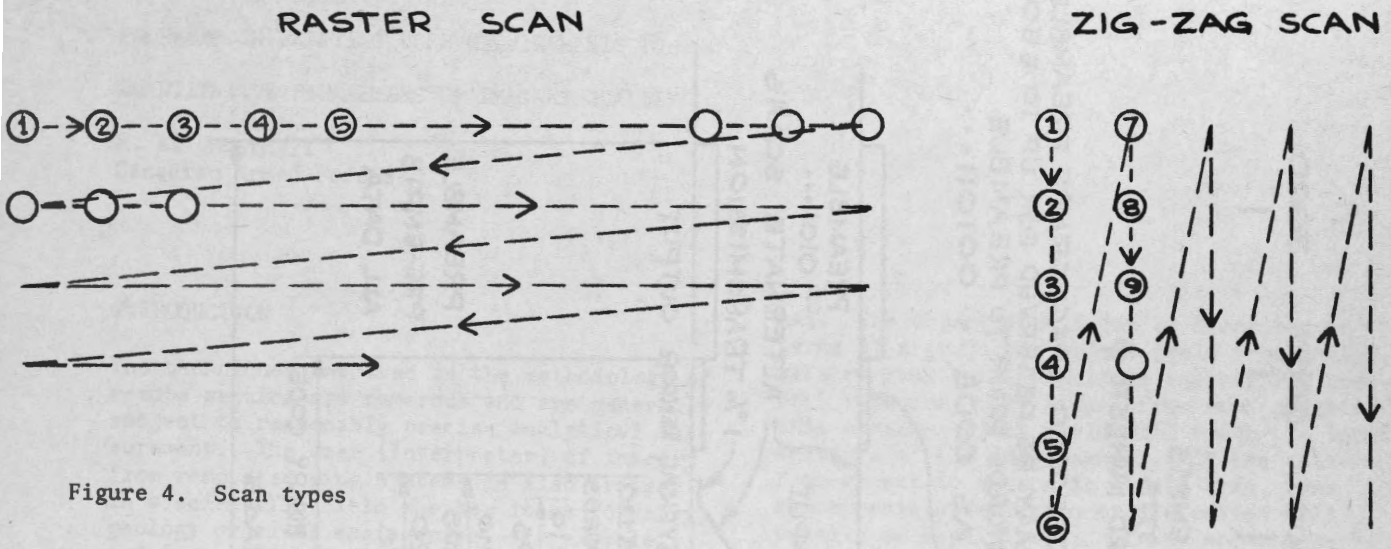


Figure 4. Scan types

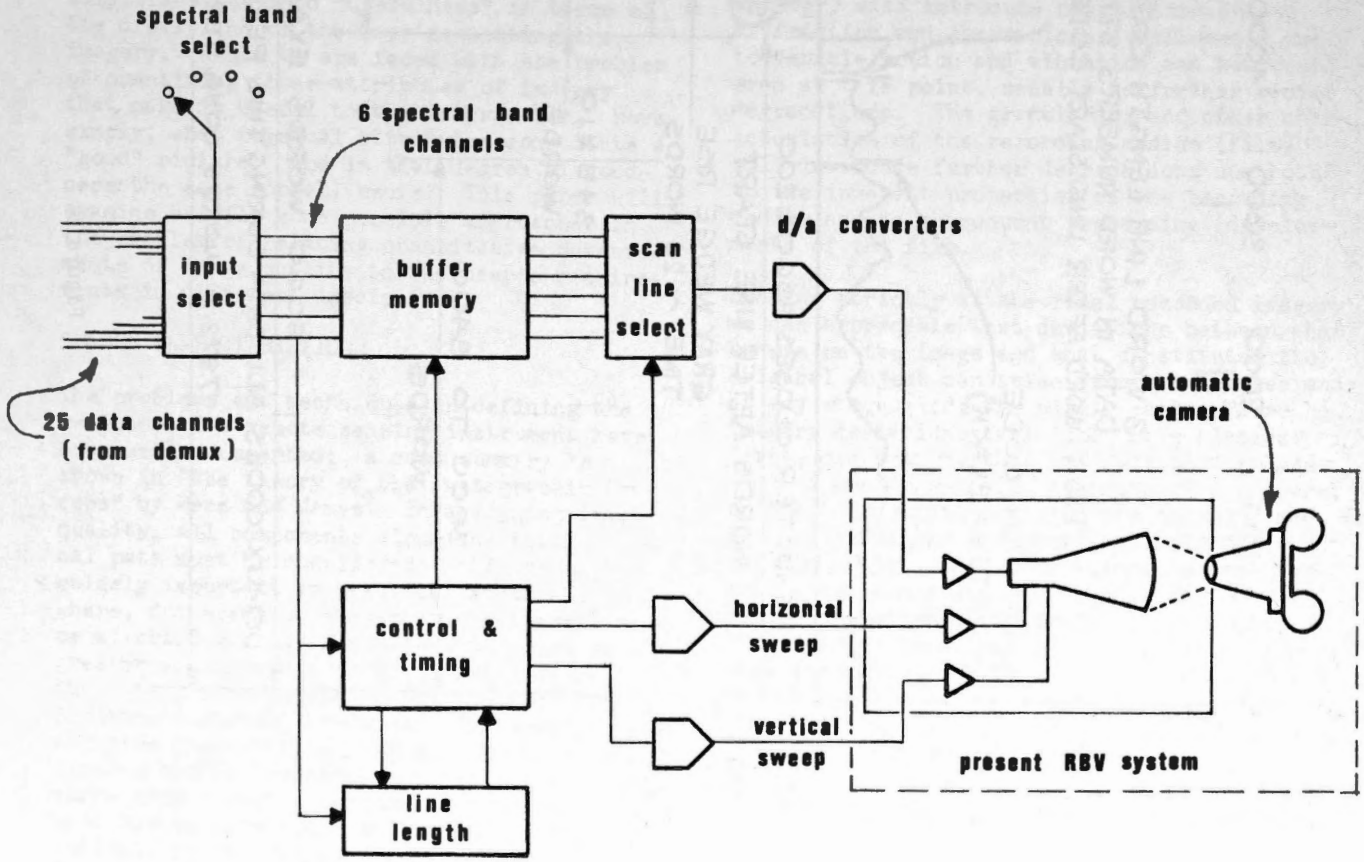


Figure 5. MSS quick look processor

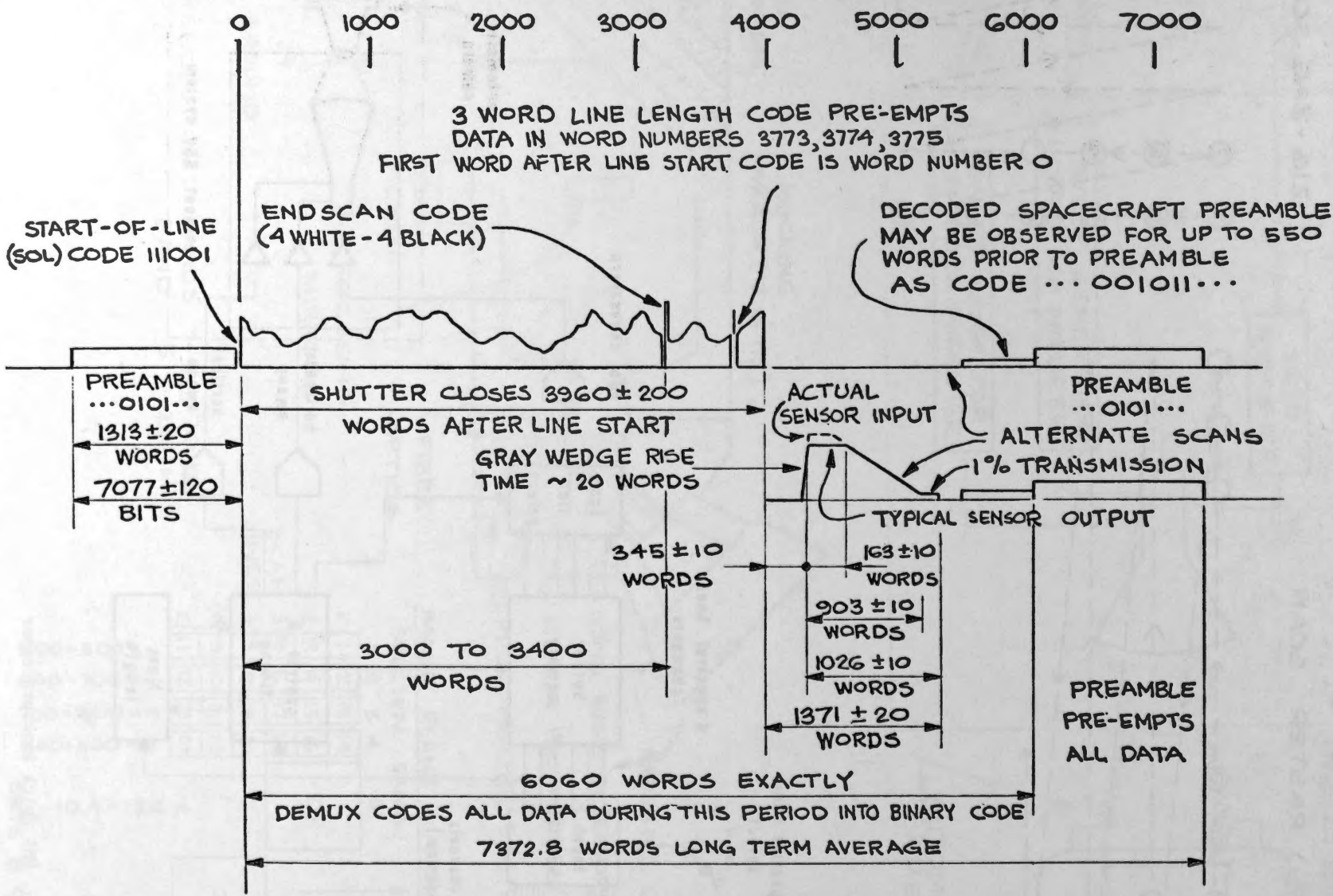

0
1000
2000
3000
4000
5000
6000
7000
3 WORD LINE LENGTH CODE PRE-EMPTS
DATA IN WORD NUMBERS 3773, 3774, 3775
FIRST WORD AFTER LINE START CODE IS WORD NUMBER 0
START-OF-LINE
(SOL) CODE 111001
ENDSCAN CODE
(4 WHITE - 4 BLACK)
DECODED SPACECRAFT PREAMBLE
MAY BE OBSERVED FOR UP TO 550
WORDS PRIOR TO PREAMBLE
AS CODE ··· 001011 ···
PREAMBLE
···0101···
1313±20
WORDS
7077±120
BITS
SHUTTER CLOSES 3960±200
WORDS AFTER LINE START
ACTUAL
SENSOR INPUT
PREAMBLE
···0101···
ALTERNATE SCANS
1% TRANSMISSION
GRAY WEDGE RISE
TIME ~ 20 WORDS
TYPICAL SENSOR OUTPUT
345±10
WORDS
163±10
WORDS
903±10
WORDS
3000 TO 3400
WORDS
1026±10
WORDS
1371±20
WORDS
PREAMBLE
PRE-EMPTS
ALL DATA
6060 WORDS EXACTLY
DEMUX CODES ALL DATA DURING THIS PERIOD INTO BINARY CODE
7372.8 WORDS LONG TERM AVERAGE

Figure 6.

PROBLEMS IN RELATING USER REQUIREMENTS TO QUANTITATIVE PARAMETERS OF IMAGERY QUALITY

M. A. Underhill
Canadian Armed Forces

INTRODUCTION

The techniques involved in the methodology of remote sensing are numerous and are generally subject to reasonably precise analytical measurement. The user (interpreter) of imagery from remote sensing systems is also trained in a scientific field whether it be forestry, geology or civil engineering, for example, and is used to dealing with precise analytical measurements in his own discipline. Unfortunately the quantitative terms which define image quality do not easily or formally translate into "usefulness" in terms of the discipline of the user requesting the imagery. Hence we are faced with the problem of quantizing those attributes of imagery that make it useful to the interpreter. More simply, what physical attributes constitute a "good" picture? And is that degree of goodness the same for all users? This paper will examine briefly some possible approaches in the problem of relating quantitative measurements of image quality to the user's requirements in different disciplines.

SENSOR CHARACTERISTICS

The problems and techniques in defining the response of a remote sensing instrument have been well documented; a good summary is shown in "The Theory of the Photographic Process" by Mees and James[1]. In analysing image quality, all components along the total optical path must be considered; this is particularly important in practical applications where, for example, atmospheric attenuation or aircraft vibration may have as great or greater an influence on the final imagery than the design parameters of the sensing instrument itself. Consider, for example, an airborne cartographic camera. Other imaging sensors can be treated by an analogous technique even though a portion of the signal path may be electrical rather than purely optical, as in the case of the camera. Proceeding along this hypothetical signal path for a camera, we can list the parameters of the original object and those components which can lead to significant deviations in the signal. The object itself can be described in terms of a projected object field (spatially) with varying tonal amplitude and varying spectral response. The signal from each point of this complex object field will reach the input aperture of the image sensor. If the path from object to sensor is at all long, then atmospheric attenuation or distortion will usually be significant; if the acceptance angle of the sensor is very wide then the attenuation or distortion may well vary across the angular field. Within the sensor itself each element (whether lens, aperture stop, or shutter) will introduce degradation due to diffraction and aberrations. Influences due to vehicle motion and vibration can be considered at this point, usually as further vector degradations. The granularity and other characteristics of the recording medium (film) will introduce further degradations due both to the inherent properties of the recording medium and to subsequent processing (development) of the film.

Looking strictly at the final recorded imagery we can appreciate that deviations between what we see on the image and what constituted the original object can arise from many causes and at any stage along the signal path. These imagery deviations fall into three classes: first, distortion, that is, spatial displacement of one image point to another as compared to the true relationship in the object plane. These distortions are usually attributable to vehicle motion, lens aberrations, sensor geometry, and non-stability (shrinkage) of the recording medium. The second class of deviations can be described as attenuations of spatial frequencies. Common causes are diffraction limitations, vibration, granularity of the recording medium, and optical aberrations. The last class of deviation concern spectral response. In the case of black and white film, for example, spectral differences in the object are integrated (in this instance, the integrating function being the spectral response curve of the film-filter combination). In "colour" systems, deviations arise because the recording medium for each colour has a finite band-width that is non-linear, the

colours do not add additively, and under visual viewing conditions, the viewing conditions themselves seriously affect the apparent image colour.

Several image quality parameters are well defined and can be briefly mentioned. The System Transfer Function (of which Modulation Transfer Function (MTF) is perhaps the most useful) is analytically extremely useful in that each element in the signal path can be described by a characteristic transfer function and, if the system is linear, the transfer functions can be cascaded to provide a measure of the overall response in terms of spacial frequencies. Another measure of image quality is acutance which is a measure of edge response. A further quantitative parameter is granularity which, if considered as a measure of noise, is equally applicable to electronically processed images as well as film or paper images. Further subjective qualitative descriptions are resolution, graininess (which is the visual analogue to granularity) and sharpness, the visual analogue to acutance. Last, we must mention image contrast, density, and density range. The foregoing attributes commonly used to describe image quality are derived from the theory of the optical process, and as such are valuable to the sensor designer. Each of these attributes can be specified in the design stage and, within reasonable limits, the manufactured instrument will meet that design criteria. Further, it can be generally stated that the higher the value for each of these parameters, the better the quality of imagery. Unfortunately, both in theory and economic practice, it becomes impossible to maximize one attribute without lowering the quality of another[2]. The decision as to which attribute to maximize and which can be suppressed leads the image sensor technician or designer to ask the question - what is the imaging system going to be used for?

USER REQUIREMENTS

User requirements can be looked at in two ways: in terms of the task in hand, i.e. for a forester, the number and type of trees - a measurement which obviously does not relate very usefully with the commonly used imagery quality parameters; or we can look at what the interpreter feels are meaningful quantitative measurements of image quality. R.N. Colwell[3] shows three lists of interpreter's requirements of image quality, each having been published in Photogrammetric Engineering: list one - angular field, definition, distortion, character of emulsion, altitude, ground speed, vibration, character of illumination; list two - graininess, sharpness, resolving power, tone reproduction; list three - the tone or colour contrast between an object and its background, image sharpness characteristics, and sterioscopic parallax characteristics. It is interesting to note that few of the listed terms are ones immediately useful to an image sensor designer, or for that matter, relate to any of the commonly measured parameters of imagery.

A large portion of the problem becomes evident at once if we attempt to visualize such a complex object as an apple tree in terms of spatial and spectral frequencies, geometric relationships, and edge sharpness - obviously an enormously difficult task. Consider some of the attributes in the process of interpreting an apple tree. Size is most essential in relation both to its surrounding objects and the geometry of the imaging system; thus, we would not expect an object the size of the Empire State Building or the size of a daisy to be an apple tree. We can eliminate many objects within the image based on this criterion alone. The shape of an apple tree (from a vertical view) can be classed as that of all trees, that is, circular; however, we would not expect the geometric precision characteristic of some conifers. The texture of the image will assist in distinguishing trees of differing leaf size, and the gross texture will help identify the general distribution and size of twig and branch structures. Cast shadow will aid in determining tree "squatness" and trunk characteristics. Tone and colour relative to the background will further help in identifying the object as an apple tree. Object relationship describes the congruency of objects in the image, thus, the apple tree is located relative to an orchard and not too distant from a water source. This attribute is one of the most important for interpretation and yet one easily overlooked in determining sensor requirements. Figures 1, 2 and 3 illustrate the importance of this attribute.

Further functions independent of image quality but essential to interpretation are time, location, and meteorological conditions under which the image was obtained. These factors alone would rule out the possibility of an apple tree, for example, occurring in imagery taken over Baffin Island. Each of the above interpretative parameters suggests various image requirements. In the example given, distortion would probably be of little importance, but good edge sharpness and high spatial resolution would be required. Fidelity of tone reproduction would be of the utmost importance. Completely different parameters are required,

for example, if the interpretative task is production of a map of an urban area.

RELATIONS BETWEEN IMAGE PARAMETERS AND INTERPRETABILITY

A number of studies have been made showing the effects of varying one or more image parameters on the subjective image quality. Few of these studies have carried the task further to consider the end user's requirements from the imagery. One of the first was Selwyn and Tearle[4] relating variation in resolution of aerial lenses with image quality. Macdonald[5] treats several aspects of image quality in relation to general interpretability. Higgins and Wolfe[6] presented a paper relating definition (clarity of detail) to resolving power and acuity. Several other papers treating specific aspects of image quality to visual image appearance are by: Tupper and Nelson[7], Simmons[8], Barrows[9], Scott[10], Macdonald and Watson[11], and Stultz and Zwig[12]. Brock[13,14] in several of his publications and papers presents an excellent review of qualitative relationships between image measurements and image interpretability. Some of this work is further updated by Noffsinger[15]. The referenced papers all treat the problem of relating image parameters and interpretability from the point of view of the imagery sensor engineer, and thus are not immediately useful to the interpreter unless he is also reasonably knowledgeable of quantitative imagery analysis techniques. Let us look at a few other practical techniques in bridging the gap between interpreter's requirements and available means of imagery analysis.

In some applications of remote sensing the nature of the interpretative task does permit analysis of the object in spatial and spectral terms which can, at least emperically, be related mathematically to the characteristics of the imagering system. If such a relationship is possible then it becomes feasible to consider automated data handling systems based on the concept of mathematically determined pattern and tonal recognition. Considerable work in this field has been done as it pertains to classification of agricultural crops[16]. The type of interpretative task that naturally leads to this technique is that in which the objects to be categorized are of a large homogeneous area and of which texture, tone, and spectral response are their most important and characteristic attributes. Another task permitting automated interpretative techniques is of feature plotting where the geometric position of the feature is, a priori, unknown, but the characteristic of the feature itself is well defined, such as railroad tracks, buildings (as a general class), and blacktop roads.

More generally, the nature of the interpretative task is too complex, psychologically and physiologically, to permit a reduction of its nature into mathematical terms. Even here, an empirical formulation can be obtained by correlation of statistically determined pictorial "goodness" on one hand, and a function whose variables are the common imagery parameters on the other hand. An example of this technique is given by Stultz and Zwig[12] where the interpretative task in this instance was pictorial quality in a salon photographic display. The obvious problems with this approach are that a new formulation is required for each different interpretative task, and the statistical problems in correlating the required data from a large number of observers are considerable. Another paper by Bennett, Winterstein and Kent[17] relates resolution, scale, granularity, and contrast to the interpretability of various types of military targets. In this instance, the training of the interpreters and the degree of a priori knowledge of the targets is also related to the ability to interpret the imagery. Although the paper does not present an empirical formulation relating these factors, such a formulation is implied as feasible from the data presented.

From a very practical point of view, two simple solutions remain to be considered for solving the problem of relating the user's requirements to the imagery characteristics. The first solution is, for each image sensor, to prepare a catalogue of imagery over typical terrain of interest to each possible user of that image sensor. This technique is simple and requires on the part of the interpreter no knowledge of image quality measurements, nor does it require much knowledge on the part of the image sensor technician of the details of the interpretative task. The obvious pitfalls are threefold: first, that with any number of sensors and any number of users (such as is the situation with a national remote sensing programme serviced by a single operational agency) the size of the imagery catalogue quickly becomes very large if it is to include representative images of all interesting target and terrain types; secondly, there is the real danger that sensor settings or modifications could be omitted from the catalogue which might lead to significant increases in interpretability for a certain task; third, such a catalogue provides only general qualitative information to point to required improvements when new sensors must be designed or purchased. Notwithstanding the listed pitfalls, such a catalogue of imagery can be a

useful tool in assisting the selection of sensors for a specific user's task.

The second solution is to derive a target that can be used for evaluating a sensor or improvement to a sensor system and which relies on the same psychological and physiological actions in its reading as does the contemplated interpretative task. In many respects the common tri-bar or annulus resolution targets fall into this category.

RESOLUTION TARGETS AS PREDICTORS ON IMAGE INTERPRETABILITY

It is worthwhile to consider in some detail the relationship between resolution and interpretability for a specific task, not because a single valued resolution reading is superior to other image quality parameters, but rather because modifications to the nature of the resolution target show a possible approach to bridging the gap between quantitative image quality and user's "image goodness" requirement. In the experiment to be described, three types of targets are used, and the resolutions from these targets are correlated with the specific interpretative task of identification and recognition of vehicles (vertical view). The following is an extract from a paper by the author[18].

The most frequently encountered types of resolution target are the U.S. Air Force tri-bar target (called "USAF target" for the remainder of this paper) and the National Research Council annulus target as devised by L.E. Howlett[19]. Both of these targets consist of simple geometric forms presented in a geometrically decreasing order of size. Both targets, as normally presented, are of periodic form and hence exhibit some of the optical phenomena associated with periodic targets. Because the nature of each target element is known a priori, determination of resolution involves a scan of the decreasing sizes of target elements until the detection of the expected pattern is no longer possible. Donaldson and Gough[20] have devised a set of alphnumeric block letter characters (figure 4) of virtually equal visual identification. If such block letters were presented in a random array, both in terms of character type and size, then use of such an array as a means of determining resolution would involve identification of individual characters with no a priori knowledge of the character type, little or no periodicity to the characters, and no judgement of the resolution value until after the complete array were scored. The purpose of the experiment was to determine if the resolution from the block letter target related better than resolution from the annulus or USAF target to the task of identifying and recognizing vehicles at threshold conditions. For the purpose of this experiment, recognition is defined as classification of vehicle type, and identification is defined as naming the vehicle from a list of known vehicles.

EXPERIMENTAL PROCEDURE

For the laboratory experiments, a Speed-Graphic camera fitted with a 70 mm roll-film back and a 152 mm Kodak Ektar lens was used, rigidly held in the "ASTM Standard Camera Alignment"[21]. A second focal length lens was available by removing the front element of the Ektar lens, giving a degraded lens system with a calculated focal length of 96.07 mm. The target array (as shown in figure 5) was placed in "ASTM Standard Orientation No. 1 for a Plane Surface"[21] with respect to the camera and was so designed that with the distances used between camera and targets, no part of the array was more than 10 degrees off the optic axis. This target array was illuminated to give essentially uniform exposure over the entire array.

For the reasons shown by McGee[22] and Carmen[23], it was decided to restrict this investigation to low-contrast targets. The paints used produced an average reflection density difference of 0.26 ± 0.02; with the background density being 0.92 ± 0.02. These gray paints were essentially neutral as measured by a MacBeth Quantalog RD-100 Densitometer. The same paints were used for the construction and painting of all resolution targets and the models.

The USAF resolution targets were of the tri-bar type with spaces between bars equal to the bar width. Each bar was 5 times its width. The tri-bar elements were arranged in serial orders of 10 mm, 9, 8, 7, 6, 5, 4, 3.5, 3, 2.5 and 2 mm bar widths. Target parameter "d" was taken as twice a bar width.

The annulus resolution target annuli are presented in pairs, each annulus to have an inner diameter equal to 1/3 the outer diameter, and each annulus to be separated from all other annuli by at least 1/3 of the larger annulus diameter. The annuli pairs were arranged in serial order of 49 mm, 42, 36, 30.5, 24, 21, 18, 15, 12 and 9 mm outer diameters. Target parameter "d" was taken as 2/3 of the outer diameter.

The seven alphanumeric characters chosen were made up in each of sizes 20 mm, 25, 30 and 40

mm to a side. For each negative made, 25 of the 35 possible size-character combinations were randomly selected and displayed in a 5 x 5 matrix. Target parameter "d" was based on the length of character side.

The models chosen as recognition and identification targets are shown in Table 2, together with their reference number and pertinent dimensions. Of the 25 models, 10 were selected and presented in random order in two rows of five within the target array. Each model would appear only once within any given negative. The models selected were painted and presented under the same contrast conditions as used for the resolution targets. These models were considered representative of typical civilian and military vehicles that a photo-interpreter might be required to recognize or identify and, in each case, the model was considered sufficiently accurate in scale and detail so as to be representative of an actual vehicle. Note that the models were presented with a minimum of clues as to background, cast shadow (i.e. under conditions of flat lighting) and were not necessarily in scale with respect to one another.

Inasmuch as this experiment was designed to relate to the very practical problems encountered in military photo-reconnaissance, the factors chosen as variable to achieve differing resolution results are from those that might be encountered under actual conditions. Note that the change in photo-optical conditions are, in themselves, relatively unimportant except to provide a varied means of changing the resolution and that their effects are uniform across the whole film surface of interest. One hundred and forty negatives were exposed, varying "f" stop, film type, object distance, flare exposure (produced by a secondary exposure with opal glass held over the lens), lens degradation (with accompanying change of focal length) and focus; from these negatives, 30 were chosen (examples shown in figures 5 and 6) by the criteria of a model recognition range greater than 0% and less than 100%. The three films chosen (Improved Tri-X type 5063, Tri-X type 8403 and Panatomic-X type SO-136) were representative of medium to high speed films currently in use for photo reconnaissance. A typical negative is shown in figure 6.

The chosen negatives were presented to two groups of readers, trained and untrained, in a random order, together with the scoring sheet, close-up photographs of the models used and scoring instructions. All untrained readers used a twenty power binocular microscope on a light table of variable intensity. Each observer was allowed to vary the light intensity as desirec, use whatever time was required and, in each case, was not required to give a forced response.

An initial analysis of the data was made on an IBM 1620 computer to check for gross errors and detect areas of missing data. The final analysis was conducted on an IBM 320.

STATISTICAL ANALYSIS

Each response variable was calculated in the form of a resolution value (or an analogous value for model recognition and identification) with a dimension of mm^{-1}.

$$R = \frac{O \times 304.8}{f \times d}$$

R = resolution (in mm^{-1})
O = object distance (ft)
f = focal length (mm)
d = target parameter (threshold)(mm)

The annulus and tri-bar "d" values were determined from the mean of all observers' values for each negative.

In determining a "d" value for the block letter targets it was realized that although the alphanumeric characters chosen were essentially equal in terms of visual recognition, they did not necessarily have the same characteristic under a photographic operation. To eliminate any error from inherent differences in recognition, a weighting function was determined for each letter type as follows:

$$W_i = \frac{C_i \times N}{C_t \times N_i}$$

W_i = weighting factor, letter type i
$C_t = \sum$ all correct letter responses (all types and sizes)
$C_i = \sum$ all correct letter readings (type i)
N = total of letters presented
N_i = total of all type i letters presented

By analogy with the annulus type of target, the block letters should have "d" values of 8, 10, 12, 14 or 16. To determine the actual "d" value, a modification of a scoring technique used for finite multi-choice quizzes was devised. Because there are a possible 7 responses for any one position on the matrix, "wrong" responses are penalized by a factor of 1/(7-1). This scoring technique gives a range of "1" for all correct and "0" for guessing (a ratio of 1 correct to 6 wrong). The final "d"

value is determined as follows for each negative:

$$SB_j = \frac{(DA - \frac{SO}{6})}{DB}$$

SB = score for letter size j

DA = $\sum W_i$ for letters correct of size j

SO = $\sum W_i$ for letters wrong of size j

DB = $\sum W_i$ for all letters presented of size j

$$"d" = 18 - \sum_{j=1}^{5} (2.4 \times SB_j)$$

This technique thus gives a "d" value of 18 for a situation of no response or guessing in all of the block letter matrix positions, and a "d" value of 6 for a matrix of all correct responses. Note that, for this experiment, negatives that showed either perfect or "18" scores were not used.

In a manner similar to that used for the block letter targets, the following scoring technique was used to determine model recognition and identification:

$$R = \frac{Z1 - Z3/24}{Z1 + Z2 + Z3}$$

R = score for identification
Z1 = number of models correct
Z2 = number of models missed
Z3 = number of models wrong

$$S = \frac{Y1 - Y3/5}{Y1 + Y2 + Y3}$$

S = score for recognition
Y1 = number of models correct
Y2 = number of models missed
Y3 = number of models wrong

$$T = \frac{P1 + P2 + P3 + P4}{4 + (4 \times R)}$$

P = model size parameter x weighting factor (ident)

$$U = \frac{Q1 + Q2 + Q3 + Q4}{4 + (4 \times S)}$$

Q = model size parameter x weighting factor (recog.)

P1, P2, P3, P4 = smallest four values of P

Q1, Q2, Q3, Q4 = smallest four values of Q

$$\text{MODEL IDENT} = \frac{\text{Cbj Dist (ft)} \times 304.8 \times R}{\text{Foc Length (mm)} \times U}$$

The size parameter on which model recognition or identification is based was either model length, mean of length plus width, square root of projected area or square root of length times width.

From each of the negatives, a mean resolution and model identification or recognition for each parameter choice and, for each negative, a mean square variation represented by:

$$(X_i - \bar{X})^2$$

X = resolution response
Y = model identification or recognition

was calculated. A standard regression analysis technique was employed in determining both linear and quadratic lines of best fit. To avoid treating the data as a bi-variate case, the X parameter (resolution) was coded into equi-spaced cells by taking the integer part of the average resolution plus 0.5. Besides the regression co-efficients for each case, the following statistics were also calculated: a calculated "F" for lack of fit; a S_{yx} term based on the square root of the residual following from the analysis of variance; an error term and degrees of freedom represented by the error term and the average value of the mean square variance previously calculated for each negative. Finally, the correlation co-efficient R was calculated for each set of X, Y data by the formula:

$$R = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\left[\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2\right]^{\frac{1}{2}}}$$

RESULTS

The complete results show that, in fact, no significant difference exists between the various means of determining target size parameters nor was the line of best fit quadratic but rather was linear. The results shown in Table 1 are averages of the values computed for each of the 4 different ways of expressing model size parameters.

TABLE 1

Avg. 'R' and $S_{yx}/\bar{Y}$ values for RECOGNITION

Target		Trained $\bar{X}$	Trained Std. Dev.	Untrained $\bar{X}$	Untrained Std. Dev.
Block - Letter	R	.7636	.014	.7046	.010
	S	.3524	.010	.4008	.010
Tri-Bar	R	.6785	.014	.6845	.010
	S	.3989	.010	.4151	.010
Annulus	R	.6556	.010	.6131	.010
	S	.4166	.010	.4415	.010

Avg. 'R' and $S_{yx}/\bar{Y}$ values for IDENTIFICATION

Block - Letter	R	.5614	.026	.6500	.010
	S	.4813	.028	.5454	.033
Tri-Bar	R	.3856	.034	.6317	.010
	S	.5597	.024	.5775	.024
Annulus	R	.3950	.026	.5871	.010
	S	.5352	.028	.6004	.024

Note: R = Correlation co-efficient
S = S_{yx}/Avg. Y

Table 1 shows the main results of interest to this experiment. With 30 data points, the degrees of freedom for correlation are 28 and hence, at 90% confidence (alpha equal to 0.10) all correlations above the value of 0.3070 are significant, in other words, we have found a significant correlation for each X, Y test shown in the table. Because the range of X values does not vary significantly, we can consider the confidence interval on Y at the point $X_o = \bar{X}$ in each case. Under this condition, the expression for the confidence interval on Y reduces to:

$$\hat{Y} \text{ plus, minus } t_{n-2,\, a/2} S_{yx} (1/n)^{\frac{1}{2}}$$

For each of the 48 cases, n equals 30, hence the whole expression reduces to:

$$\hat{Y} \text{ plus, minus } K \times S_{yx} \quad \text{where K is a constant.}$$

Thus, it is reasonable to consider intercomparing values of S_{yx}. Further, because we are assessing X as a predictor of Y and different ranges and average Y values are encountered in the different cases, we can divide the term S_{yx} by avg Y for each case to allow for a comparison between cases of different Y values.

Tables 3 and 4 show the weighting factors for each Block Letter and Vehicle Model Type. Note that the weighting factors for Vehicle Models were determined by a technique analogous to that described for Block Letters. Table 5 shows the photographic parameters of the thirty negatives selected for the experiment.

DISCUSSION OF RESULTS

Table One shows several significant results: first, there is a significantly better correlation for recognition than for identification. This would tend to show that the reading of all three types of resolution target is basically one of recognition. Secondly, there is a significantly better correlation for the block letter resolution target than for either the annulus or for the USAF targets. Thirdly, the difference in response between trained and untrained observors is not significant in every case; however, it would appear that in each case the trained observors show better correlation for recognition, but poorer correlation for identification. It is quite obvious from the results that a simple target can be devised which correlates considerably better to the task of identification and recognition of vehicles. It has not been proved whether this significant change in correlation would be true under actual field conditions (data is presented in the original paper which was not amiable to quantitative statistical analysis, but which would indicate that such increase of correlation does carry over into field conditions), nor has this reported improvement in correlation been proved for interpretative tasks other than vehicle interpretation; however, one would expect intuitively a similar result for related fields which likewise rely on spatial rather than tonal characteristics.

CONCLUSION

The paper, as a survey, has pointed out some of the problems in relating user requirements to image parameters and has indicated several practical and theoretical approaches that could warrant further development. In preparation of this paper and in particular in the original preparation of the experiment outlined, it became obvious that many of the problems relate to the non-common terminology between the various disciplines involved and to the lack of full understanding of the role the psychology, physiology and physics of sight plays in the interpretative task.

ACKNOWLEDGEMENTS

The author would like to express his thanks to the Rochester Institute of Technology for permission to publish extracts from his thesis material, and to the staff of the Defence Photographic Intelligence Centre for their participation in experiment on which the thesis was based.

REFERENCES

1. Mees, C.E.K. and James, T.H.: The theory of the Photographic Process, 3rd ed., The Macmillan Co., New York, 1966.

2. Manual of Photographic Interpretation, American Soc. of Photogrammetry, Washington, 1960 p. 58.

3. ibid, p. 52

4. Selwyn, E.W.H. and Tearle, J.L.: The Performance of Aircraft Camera Lenses, the Proc. of the Phys. Cos., 58:493, Sept 1946.

5. Macdonald, D.E.: Resolution as a Measure of Interpretability, Photogram. Eng., 24:, Jan 1958

6. Higgins, G.C. and Wolfe, R.N.: The Relation of Definition to Sharpness and Resolving Power in a Photographic System, J. Opt. Soc. Am., 45:121, Feb 1955

7. Tupper, J.L. and Nelson, C.N.: The Effect of Atmospheric Haze in Aerial Photography Treated as a Problem in Tone Reproduction, Photogram. Eng., Vol 6, No 2, 1955

8. Simmonds, J.L.: A Quantitative Study of the Influence of Tone-Reproduction Factors on Picture Quality, J. of Photo. Sc. & Eng., Vol 5, No 5, Sept 1961

9. Barrows, R.S.: Factors Affecting the Recognition of Small, Low-Contrast Photographic Images, J. of Photo. Sc. & Eng., Vol 1, No 1, Jul 1957

10. Scott,R.M.: The Present State of the Art With Regard to Detection Recognition, Applied Optics, 3:13, Jan 1964

11. Macdonald, D.E. and Watson, J.T.: Detection and Recognition of Photographic Detail, J. Opt. Soc. Am., 46:715, Sept 1959

12. Stultz, K.F. and Zweig, H.J.: Roles of Sharpness and Graininess in Photographic Quality and Definition, J. Opt. Soc. Am. 52:45, Jan 1962

13. Brock, G.C.: Image Evaluation for Reconnaissance, Applied Optics, 3:11, Jan 1964

14. Brock: G.C.: Physical Aspects of Air Photography, Longmans, Green & Co., Edinburgh, 1952

15. Noffsinger, E.B.: Image Evaluation - Criteria & Applications, J. Soc. Photo-optical Inst. Eng., Vol 9, No 1: 32, Nov 1970

16. Holmes, R.A.: Data Requirements and Data Processing Earth Resources Surveys, J. Soc. Photo-optical Inst. Eng., Vol 9, No 2: 52, Jan 1971

17. Bennett, C.A., Winterstein, S.H. and Kent, R.E.: Image Quality and Target Recognition, Human Factors Society, 10th Annual Meeting, Paper #1

18. Underhill, M.A.: Evaluation of an Alphanumeric Target as a Means of Determining the "Resolution" of a Photographic Reconnaissance System, MSc Thesis, Rochester Institute of Technology, 1968

19. Howlett, L.E.: Photographic Resolving Power, Can. Jour. of Research, 24:A: 15, July 1946

20. Donaldson, K.C. and Gough, H.O.: The Determination of a Set of Alphanumeric Characters of Equal Recognizability, BSc Thesis, Rochester Inst. of Technology, 1967

21. McCamy, C.S.: Proposed Recommended Practice for Description and Selection of Conditions for Photographic Specimens, Phot. Sci. Eng., 10:185, Jul 1966

22. McGee, H.A.W.: The Effect of Target Contrast of the Focus and Performance of the Metrogon Lens, Photogrammetric Engineering 18:848

23. Carmen, P.D. and Carruthers, R.A.F.: Brightness of Fine Detail in Air Photography, J. Opt. Soc. Am., 41:305, May 1951

TABLE 2

Table of models selected as representative targets

(Dimensions in mm)

Model Code No	Class Code No	Description	Length	Width	Mean of L + W	Sq. Rt. Proj. Area	Sq. Rt. L x W
0	0	Do not know					
1	3	Stake truck (6 wheel)	74	25	49.5	43.0	43.0
2	3	Landrover Stn Wagon	67	27	47.0	42.5	42.5
3	3	Hopper truck	66	26	46.0	41.5	41.2
4	6	Mustang Sedan	70	27	48.5	43.5	43.5
5	3	Dumper truck	51	27	39.0	37.0	37.1
6	3	Jeep pick-up	65	25	45.0	40.0	40.2
7	6	Jeep 4 x 4	55	28	41.5	40.5	39.3
8	5	'Saladin' Arm. Car	61	29	45.0	42.0	42.0
9	2	Snow-trac	55	33	44.0	41.5	42.7
10	1	Small tractor (farm)	49	29	39.0	30.0	37.7
11	1	Bulldozer Cat D8	58	32	45.0	38.5	43.1
12	2	Gr Bt 3-ton LKW 4 x 4	49	17	33.0	29.0	28.9
13	2	U.S. M3 half track	50	22	36.0	33.0	33.2
14	5	Gr Bt Recon Sct car	45	24	34.5	34.0	32.8
15	4	USSR Stalin tank	53	25	39.0	36.5	36.3
16	3	U.S. 2½ ton 6 x 6	49	19	34.5	30.5	30.5
17	5	U.S. M40 S.P. Cannon	54	26	40.0	37.5	37.5
18	4	U.S. M59 lt tank	47	26	36.5	35.0	34.9
19	3	Gr Bt Ambulance 3-ton 4 x 4	50	18	34.0	30.0	30.0
20	4	USSR T-34 lt tank	47	24	35.5	34.0	32.1
21	2	U.S. M3A1 10 ton half track	48	18	33.0	28.0	29.4
22	4	Centurian tank Mk III	52	26	39.0	37.0	36.7
23	4	USSR T-54 heavy tank	48	26	37.0	35.5	35.3
24	4	Ger HS20 APC	44	20	32.0	30.0	29.7
25	6	U.S. Jeep 4 x 4	33	15	29.0	22.5	22.2

Class Code No	Description
0	Do not know
1	Tractor
2	Personnel carrier
3	Truck
4	Tank
5	Armoured car
6	Small cars

TABLE 3

Untrained Observers

Letter	Total	Cor.	Wrong	Missed	Pcnt Cor	Pcnt Wr	Pcnt Miss	Weight
1	760	406	272	82	53.42	35.79	10.79	0.8449
2	786	526	163	97	66.92	20.74	12.34	1.0584
3	808	574	109	125	71.04	13.49	15.47	1.1235
4	783	535	136	112	68.33	17.37	14.30	1.0806
5	840	439	302	99	52.26	35.95	11.79	0.8265
6	807	476	240	91	58.98	29.74	11.28	0.9329
7	666	490	122	54	73.57	18.32	8.11	1.1636
TOTALS:	5450	3446	1344	660	63.23	24.66	12.11	1.0043
MODEL IDENTIFICATION								
1	132	72	45	15	54.55	34.09	11.36	1.3559
2	95	38	42	15	40.00	44.21	15.79	0.9943
3	50	16	28	6	32.00	56.00	12.00	0.7954
4	87	35	39	13	40.23	44.83	14.94	1.0000
5	67	37	19	11	55.22	28.36	16.42	1.3727
6	117	38	53	26	32.48	45.30	22.22	0.8073
7	36	18	11	7	50.00	30.56	19.44	1.2429
8	131	76	29	26	58.02	22.14	19.85	1.4421
9	87	45	36	6	51.72	41.38	6.90	1.2857
10	131	124	2	5	94.66	1.53	3.82	2.3529
11	109	51	31	27	46.79	28.44	24.77	1.1631
12	66	37	14	15	56.06	21.21	22.73	1.3935
13	110	25	38	47	22.73	34.55	42.73	0.5649
14	64	42	11	11	65.63	17.19	17.19	1.6313
15	81	13	44	24	16.05	54.32	29.63	0.3989
16	94	10	58	26	10.64	61.70	27.66	0.2644
17	130	44	52	34	33.85	40.00	26.15	0.8413
18	65	22	21	22	33.85	32.31	33.85	0.8413
19	51	27	18	6	52.94	35.29	11.76	1.3160
20	65	2	29	34	3.08	44.62	52.31	0.0765
21	80	20	22	38	25.00	27.50	47.50	0.6214
22	117	19	58	40	16.24	49.57	34.19	0.4037
23	65	17	29	19	26.15	44.62	29.23	0.6501
24	71	3	37	31	4.23	52.11	43.66	0.1050
25	79	46	10	23	58.23	12.66	29.11	1.4474
TOTALS:	2180	877	776	527	40.23	35.60	24.17	0.9747
MODEL RECOGNITION								
1	240	185	38	17	77.08	15.83	7.08	1.2606
2	343	179	128	36	52.19	37.32	10.50	0.8535
3	606	419	176	11	69.14	29.04	1.82	1.1308
4	594	315	213	66	53.03	35.86	11.11	0.8673
5	195	125	60	10	64.10	30.77	5.13	1.0483
6	202	110	82	10	54.46	40.59	4.95	0.8906
TOTALS:	2180	1333	697	150	61.15	31.97	6.88	1.0085

TABLE 4

Trained Observers

Letter	Total	Cor	Wrong	Missed	Pcnt Cor	Pcnt Wr	Pcnt Miss	Weight
1	520	302	110	108	58.08	21.15	20.77	0.9667
2	540	358	67	115	66.30	12.41	21.30	1.1035
3	555	350	60	145	63.06	10.81	26.13	1.0497
4	540	360	57	123	66.67	10.56	22.78	1.1096
5	580	269	183	128	46.38	31.55	22.07	0.7720
6	555	318	110	127	57.30	19.82	22.88	0.9537
7	460	296	97	67	64.35	21.09	14.57	1.0710
TOTALS:	3750	2253	684	813	60.08	18.24	21.68	1.0037
MODEL IDENTIFICATION								
1	90	28	25	37	31.11	27.78	41.11	1.1164
2	65	9	17	39	13.85	26.15	60.00	0.4969
3	35	2	10	23	5.71	28.57	65.71	0.2051
4	60	18	8	34	30.00	13.33	56.67	1.0766
5	45	18	4	23	40.00	8.89	51.11	1.4354
6	80	10	19	51	12.50	23.75	63.75	0.4486
7	25	10	0	15	40.00	0.0	60.00	1.4354
8	90	60	7	23	66.67	7.78	25.56	2.3923
9	60	26	14	20	43.33	23.33	33.33	1.5550
10	90	84	1	5	93.33	1.11	5.56	3.3493
11	75	50	6	19	66.67	8.00	25.33	2.3923
12	45	13	4	28	28.89	8.89	62.22	1.0367
13	75	9	10	56	12.00	13.33	74.67	0.4306
14	45	24	3	18	53.33	6.67	40.00	1.9139
15	55	5	4	46	9.09	7.27	83.64	0.3262
16	65	2	17	46	3.08	26.15	70.77	0.1104
17	90	25	12	53	27.78	13.33	58.89	0.9968
18	45	8	4	33	17.78	8.89	73.33	0.6380
19	35	4	2	29	11.43	5.71	82.86	0.4101
20	45	1	2	42	2.22	4.44	93.33	0.0797
21	55	2	9	44	3.64	16.36	80.00	0.1305
22	80	2	12	66	2.50	15.00	82.50	0.0897
23	45	1	5	39	2.22	11.11	86.67	0.0797
24	50	0	0	50	0.0	0.0	100.0	0.0
25	55	7	8	40	12.73	14.55	72.73	0.4567
TOTALS:	1500	41	203	879	27.87	13.53	58.60	0.9041
MODEL RECOGNITION								
11	165	139	4	22	84.24	2.42	13.33	1.6180
2	235	76	61	98	32.34	25.96	41.70	0.6211
3	415	224	108	83	53.98	26.02	20.00	1.0367
4	410	204	35	171	49.76	8.54	41.71	0.9556
5	135	90	18	27	66.67	13.33	20.00	1.2804
6	140	48	44	48	34.29	31.43	34.29	0.6585
TOTALS:	1500	781	270	449	52.07	18.00	29.93	1.0284

TABLE 5

Table of Selected Negatives Used for Readings

Neg Code	Object dist (ft)	Exposure 'f' stop	Speed (sec)	Flare 'f' stop	Exposure speed (sec)	Variation from best visible focus (in mm)	Focal length (mm)	Film type	Date	Transmission Min	Densities Max
AA	125	4.5	1/200	nil	nil	0	152	5063	30-3-67	0.76	1.10
AB	125	8	1/100	nil	nil	0	152	5063	30-3-67	1.04	1.25
AC	125	32	1/5	nil	nil	0	152	5063	30-3-67	0.56	0.83
AY	70	4.5	1/200	nil	nil	1+	152	5063	30-3-67	0.44	0.62
AZ	70	8	1/100	nil	nil	1+	152	5063	30-3-67	0.44	0.66
BO	50	32	1/5	nil	nil	0	96.07	5063	4-4-67	0.88	1.14
BU	50	32	1	nil	nil	0	96.07	5063	4-4-67	0.98	1.21
BX	30	32	1/5	nil	nil	0	96.07	5063	4-4-67	0.42	0.64
CA	30	4.5	1/25	nil	nil	0	96.07	S0136	4-4-67	1.06	1.51
CC	30	8	1/10	nil	nil	0	96.07	S0136	4-4-67	0.70	1.17
CJ	50	8	1/10	nil	nil	0	96.07	S0136	4-4-67	0.55	0.98
CM	50	11	1/5	nil	nil	0	96.07	S0136	4-4-67	0.68	1.20
CN	50	11	1	nil	nil	0	96.07	S0136	4-4-67	2.10	2.50
DA	90	4.5	1/25	nil	nil	2+	152	S0136	4-4-67	0.62	1.06
DB	90	8	1/10	nil	nil	2+	152	S0136	4-4-67	0.27	0.58
DD	90	4.5	1/25	4.5	1/50	2+	152	S0136	4-4-67	0.73	1.15
DE	90	8	1/10	4.5	1/50	2+	152	S0136	4-4-67	0.40	0.66
DS	120	8	1/10	4.5	1/25	0	152	S0136	4-4-67	0.74	1.02
DV	120	8	1/10	4.5	1/10	0	152	S0136	4-4-67	0.70	0.89
DX	120	4.5	1/100	nil	nil	0	152	8403	6-4-67	0.45	0.78
EK	90	8	1/50	8	1/25	0	152	8403	6-4-67	0.60	0.67
EL	90	16	1/10	nil	nil	0	152	8403	6-4-67	0.52	0.68
EC	90	8	1/50	nil	nil	0	152	8403	6-4-67	0.50	0.68
EP	90	32	1	nil	nil	0	152	8403	6-4-67	0.79	0.97
ER	90	8	1/50	nil	nil	2+	152	8403	6-4-67	0.44	0.63
ES	90	32	1	nil	nil	2+	152	8403	6-4-67	0.73	0.95
FA	70	8	1/50	nil	nil	1-	152	8403	6-4-67	0.52	0.68
FC	70	32	1	nil	nil	2+	152	8403	6-4-67	0.78	0.93
FK	30	32	1	nil	nil	2-	96.07	8403	6-4-67	0.91	1.05
FL	30	32	1	nil	nil	1-	96.07	8403	6-4-67	0.90	1.06

Note: Plus and minus indicates focus racked forward or back of point of best visual focus

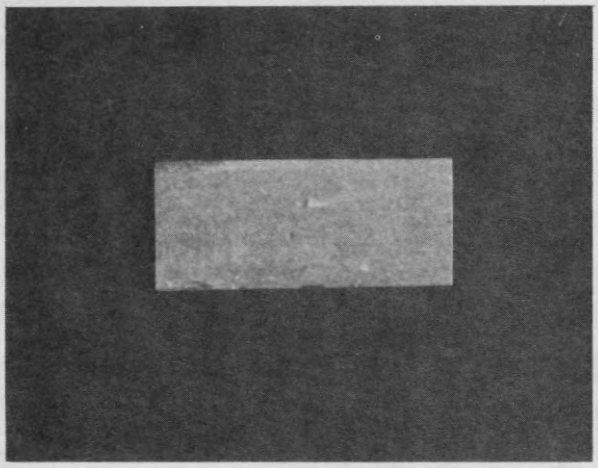


Figure 1. Object at threshold of identification , no surround detail

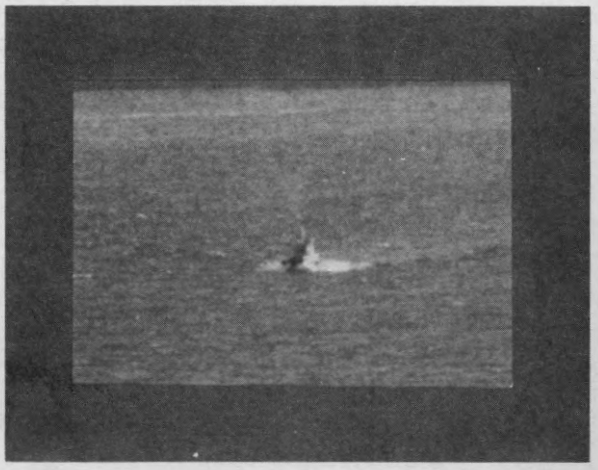


Figure 2. Object with detail, no surround detail

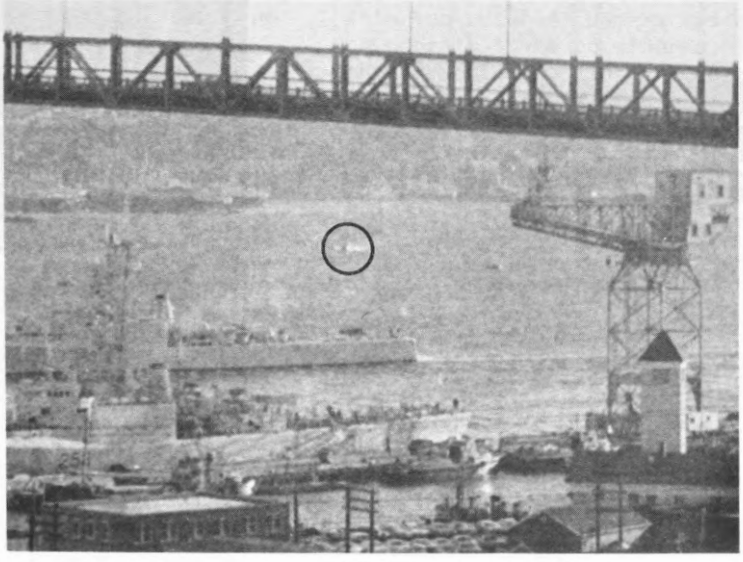


Figure 3. Object at threshold of identification, with surround detail

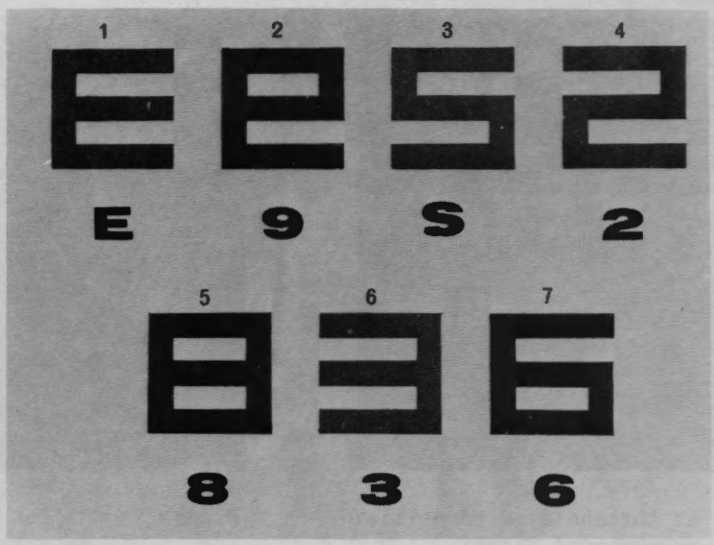


Figure 4. Block-letter design fir resolution target

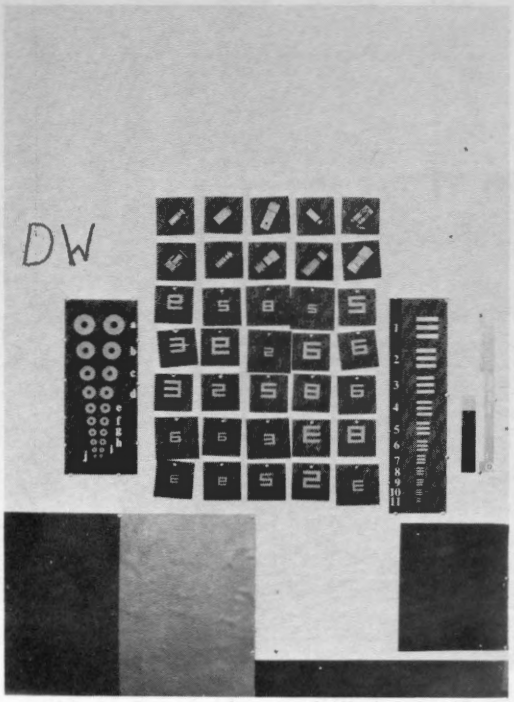


Figure 5. Detail of typical target presentation, resolution trials

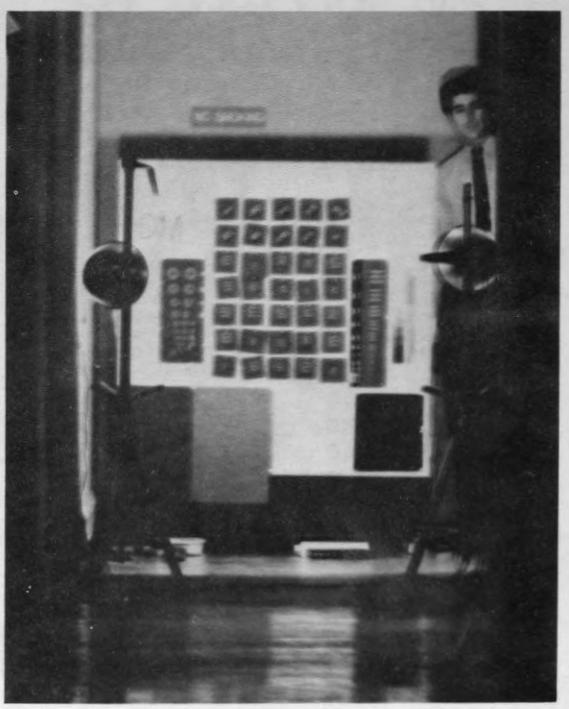


Figure 6. Typical degraded image (8X magnification) for threshold resolution determination

GEOMETRIC CONSIDERATIONS IN REMOTE SENSING

Dr. Eugene E. Derenyi,
Associate Professor,
Department of Surveying Engineering,
University of New Brunswick, Fredericton, N.B.

ABSTRACT

In order to judge the potentials of image-forming remote sensors properly, both the spectral characteristics and the geometric aspects must be considered. With respect to the latter one, two factors are of primary importance:

1. Geometric or spacial resolution, which in case of an optical mechanical scanner is a function of the instantaneous angular field of view and for a radar system depends on the accuracy of the time measurement, on the slant range and on the beam width.

2. Geometric fidelity, which is influenced by the inherent distortions of the sensor, distortion characteristics of the photographic material, terrain and environmental conditions, and by fluctuating in the attitude, altitude and velocity of the airborne vehicle carrying the sensor.

Both factors are discussed in detail. From a geometric point of view, the performance of unorthodox image forming sensors is inferior to that of modern frame camera systems. Results of a theoretical investigation and of a test conducted on infrared scanner imagery are presented as a proof. The use of stabilized platforms and analytical or analogue image restitution is suggested to improve the geometric fidelity of dynamic imagery.

INTRODUCTION

Sensors which operate outside the visible light region such as infrared scanners and side looking radar, are gaining an increased acceptance in the field of photo interpretation. They are capable to operate day and night, under adverse weather conditions and are able to reveal information which otherwise would remain undetected.

Spectral characteristics of these sensors are reasonably well known and intensive research is being performed at many agencies on this subject. However, little unclassified information is available on their geometric characteristics and investigation along these lines began only recently. Nevertheless, one must be familiar with the geometric properties of a sensor in order to judge its potentials properly.

Two aspects have to be considered in this connection:

1. Geometric or spacial resolution.
2. Geometric fidelity.

RESOLUTION

Resolution is a rather complex phenomenon which involves not only the basic characteristics of an imaging system but also other variable factors such as the shape, size and contrast of the target.

In a conventional camera system there are two components to be considered, the lens and the photographic emulsion. Electronic parts perform an important function in unconventional imaging devices and the signal to noise ratio is an important criterion. One should also note that, in most cases, the resolution of the latter types is different in longitudinal and lateral direction with respect to the flight path.

For optical-mechanical scanners the theoretical limit of resolution is defined by the instantaneous field of view, usually specified in angular units. The best ground resolution is obtained when the scanner points vertically downward. Then it decreases as a function of $\sec \theta$ in longitudinal direction and of $\sec^2\theta$ in lateral direction, where θ is the scan angle [Holter, et al., 1962].

For side looking radar systems the ground resolution in lateral direction is limited by the accuracy of time measurement and as such it is proportional with the time duration of the pulse. In longitudinal direction it is proportional to the slant range and to the

beam width [A.S.P., 1966].

Table 1 presents a numerical comparison between the ground resolution of all three sensor types discussed. Values for the frame camera are based on the performance of a typical superwide angle mapping camera. The instantaneous field of view of the scanner is taken as 2 mrad. by 2 mrad., the pulse duration of the radar is assumed to be 0.1 microsec. and the beamwidth is 0.1°, all of which are typical values for an unclassified sensor. Furthermore all values are given in meters and are valid for a flying height of 1000 m.

TABLE 1

Sensor	Direction	Angle of view 0°	30°	45°	60°
Camera	Longitud.	0.14	0.20	0.40	1.10
	Lateral	0.14	0.20	0.40	1.10
Scanner	Longitud.	2.0	2.4	2.8	4.0
	Lateral	2.0	2.8	4.0	8.0
Radar	Longitud.		2.0	2.5	3.5
	Lateral		30.0	21.2	17.3

It is apparent from these results that the resolution of an unclassified IR or radar sensor is far inferior to the resolution of a modern aerial camera.

GEOMETRIC FIDELITY

The geometric fidelity of an image is influenced by the following factors:

1. Inherent distortions of the whole instrument package which produces the image (interior orientation). This factor can only be ascertained by a thorough calibration of the system. Thus far none or very little research has been done towards the calibration of unconventional imaging systems.

2. Distortion characteristics of the photographic material. This problem is common to any sensor which produces a photographic image including the frame camera and is well documented in the literature.

3. Distortions due to natural causes such as atmospheric refraction and earth curvature. Image refinement procedures developed for conventional aerial photography are readily available to deal with this problem.

4. Sensors which operate outside the visible light region employ scanning techniques to collect the reflected or emitted energy from objects with extended dimensions. In the direction of flight, the scanning is usually induced by the forward motion of the carrying vehicle, whereby the visual reconstruction of the sensed energy becomes a continuous strip image. Therefore, angular oscillations of the vehicle and variations in velocity and altitude displace the image points from their correct geometric position. The magnitude and direction of the displacement varies from image point to image point. It is a function of the changes in the position and orientation of the sensor with respect to a fixed coordinate system(exterior orientation) and of the location of the image point in question within the strip. Displacements are more pronounced along the edges than at the centre.

As an example, Table II lists the errors introduced into the ground position of points, situated at various scan angles, due to a displacement of the corresponding image points [Derenyi and Konecny, 1964]. A change of one degree of arc is assumed in the pitch (ϕ), roll (ω) and yaw (κ), the three components of the angular orientation, and a change of ten meters in the altitude. The flying height is assumed to be 1000 m above terrain. The effect of the individual components and the total position error are listed separately. All units are in meters and also represent the displacements in mill ($^{\circ}/_{\circ\circ}$) of the flying height.

TABLE II

Scan Angle	0°	30°	45°	60°
Pitch	18	18	18	18
Roll	18	24	35	70
Yaw	0	10	17	30
Height	2	6	10	17
Total	26	32	44	80
Roll Comp.	18	21	27	39

Table II indicates that:

1. The roll is by far the largest contributor to the displacement. Therefore roll compensation is highly desirable and is in fact employed in certain models of IR scanners. Table II also lists the total position error in case of roll compensation.

2. The pitch has a constant effect throughout the strip and thus introduces a scale error.

3. The yaw, which does not distort a frame photograph, contributes considerably to the position error in strip imageries.

4. The magnitude of the displacements even

that with roll compensation is far beyond the resolution limit of the sensors.

Although a one degree tilt and a 10 m altitude error can occur in frame photography as well, however, these values are valid for the entire frame. The resulting displacements do not impair interpretation and can easily be corrected by established photogrammetric procedures. In case of a dynamic imaging device the magnitude of the attitude and altitude error can fluctuate rather rapidly especially at a low altitude and at a slow air speed. The velocity of the angular oscillation of a light aircraft under such conditions can reach several degrees per second. Consequently displacements can vary from a minimum to a maximum value within a small segment on image and the image may become distorted to such an extent that identification becomes difficult. The problem is even more acute when attempts are being made for the metric evaluation of dynamic imageries.

Three alternatives are open to deal with this problem:

1. Accept the inherent limitations of dynamic imageries without making any efforts towards improvements. Needless to say that this is an incorrect approach. It reduces the significance of remote sensors which otherwise, due to special detection capabilities, could become a valuable asset in the detecting, measuring and representing the geometric and physical features of our environment.

2. Place the sensor on a stabilized platform to reduce the effect of aircraft oscillations. This is certainly a valid approach. However, it could produce acceptable results at a reasonable cost only when the imagery is utilized solely for interpretation. For metric applications the distortions must be reduced to a value which is below the limit of resolution, which can only be achieved with a high cost.

3. Develop suitable analytical and/or analogue procedures for the restitution of dynamic imagery. Although in this case it is impossible to achieve a strict solution, nevertheless the geometric fidelity can be greatly improved through such means.

Serious efforts are being made in this direction at the Department of Surveying Engineering, University of New Brunswick [Derenyi, 1971].

The geometric fidelity of remote sensor imageries presently available is demonstrated by the following test: Identical points were selected on aerial photographs and two strips of infrared scanner imagery covering the same terrain. The coordinates of all image points were measured in a comparator and ground coordinates were deduced analytically. It should be noted that the effect of the panoramic distortion was also corrected. Coordinates computed from the two IR images were then compared with those originating from the aerial photographs and with each other. Table III lists the maximum, average and the root-mean-square (RMS) point errors obtained. All values are given first in meters and then in mill (°/₀₀) of the flying height (1500 m). The scanner in question was roll compensated.

TABLE III

Point Error	Photo-IR1 m	Photo-IR1 o/oo h	Photo-IR2 m	Photo-IR2 o/oo h	IR1-IR2 m	IR1-IR2 o/oo h
Max.	+66.1	+43.1	+41.7	+27.7	-33.2	-22.1
Ave.	+11.5	- 7.7	+ 8.5	+ 5.5	- 1.5	- 1.0
RMS	±17.7	±11.8	±13.4	± 8.9	± 7.9	± 5.3

A comparison of these values with those presented in Table II indicate that discrepancies can mainly be attributed to the uncorrected effect of the changing exterior orientation of the sensor. A perfectly straight and level flight, without velocity and height variations was assumed in the absence of any pertinent information. It is also apparent that the point errors are significantly larger than the resolution limit of the sensor.

CONCLUSIONS

The results of theoretical investigations and of tests conducted on real imagery indicate that, from a geometric point of view, the performance of unconventional image forming sensors is inferior to that of present day camera systems. Furthermore, the sources of distortions are rather complex and not yet thoroughly investigated. Unconventional aerial imageries must therefore be employed with these limitations kept in mind. At the same time one must strive for improved instrumentation and image restitution techniques in order that the unique potentials of these sensors may be fully utilized.

ACKNOWLEDGEMENT

This investigation has been supported by a National Research Council of Canada and by a Defence Research Board of Canada grant in aid of research.

REFERENCES

(1) American Society of Photogrammetry, 1966 "Manual of Photogrammetry", Third Edition.

(2) Derenyi, E. and G. Konecny, "Geometry of Infrared Imagery", The Canadian Surveyor, Vol. XVIII, No. 4, p. 279-290.

(3) Derenyi, E., 1971 "An Exploratory Investigation Concerning the Relative Orientation of Continuous Strip Imagery", Department of Surveying Engineering, University of New Brunswick, Fredericton, Canada, Research Report No. 8.

(4) Holter, M.R., S. Nudelman, G.H. Suits, W.L. Wolfe and G.J. Zissis, 1962, "Fundamentals of Infrared Technology", The MacMillan Co., New York.

THE INTERPRETATION OF ARCTIC IMAGERY

I.H.S. Henderson and R.J. Brown
Defence Research Establishment Ottawa
and
W.A. Young
Litton Systems (Canada) Ltd.

INTRODUCTION

Two aspects of problems encountered in the interpretation of remotely sensed Arctic imagery have been examined. These are the environmental factors leading to degradation of imagery and a review of machine techniques which could aid in the processing of imagery.

Literature surveys in both areas have shown gaps in our knowledge and have resulted in specific suggestions for further work. These are discussed in the paper.

The vastness of the Canadian Arctic, the absence of a large and permanent population and, above all, the absence of sunlight for a very large proportion of the year in the high Arctic, render very difficult searches for parties in distress and the gathering of information for maintenance of control. Remote sensing from aircraft or satellite affords one means for carrying out search, surveillance and reconnaissance in the Arctic. However, widespread and effective use of remote sensing awaits more complete understanding of the potentials and limitations of sensors and sensing.

We have recently completed two literature surveys designed to provide partial answers to the questions, "What sensors can be used and for what proportion of the time?", and secondly, "What techniques using machines would speed up or facilitate the interpretation of sensed data?". The first question is important because the Canadian Forces require an ability to obtain information and carry out searches in virtually all weather and light conditions or in the presence of cloud cover. This implies, in general, a necessity for sensing or imaging in the non-visible portions of the electromagnetic spectrum.

The variety of sensor systems can provide complementary views of the terrain by virtue of diverse spectral sensitivities and imaging geometries. However, the disturbing characteristics of the atmosphere, and the radiometric properties of the terrain differ significantly in visible, infrared, and microwave spectral regions. Each sensor system requires specialized interpretive procedures, and knowledge of its unique characteristics.

ATMOSPHERIC ATTENUATION AND INTERPRETATION

The principal modes of atmospheric attenuation are absorption and scattering. Absorption influences electromagnetic wave propagation very little in the visible and near-visible region (0.3 to 1.0μm) or in the microwave region (beyond 1.35cm). However, in the intervening range, absorption plays a major role. From about 1 to 15μm there are several strong absorption bands, the main "windows" for remote sensing purposes being from 3 to 5μm and 8 to 14μm. There is very little transmission from 15 to 1000μm, and some transmission from 1000μm to 1.35cm.

The effective width of a transmission window is very sensitive to the concentrations of the absorbing species whose presence defines the extremities of the window. This has important implications in the determination of temperature by remote use of infrared radiometry or line scanner techniques. Temperature measurements can be seriously in error when there is a variable amount of water vapour or carbon dioxide between sensor and target, these two absorbers being those which define the limits of the 8 to 14μm window.

The attenuation by atmospheric scattering may be summarized by the statement that microwave radiation penetrates clouds, infrared penetrates haze but not clouds whereas visible radiation penetrates neither. As a general rule of thumb, atmospheric scattering is maximum when the radii of the scatterers are of the same order of magnitude as the wavelength of the propagating radiation. Ideally, one would like to determine the atmospheric conditions during the sensing mission and compute the effects on the sensor. However, because of the magnitude of the computational task and the lack of detailed

knowledge of absorption line shape it is in many cases not practical to calculate exactly the effects of the atmosphere on propagating radiation. Such a computation is further complicated by the constantly changing atmospheric temperature and water vapour profiles. Thus, for the indefinite future, it will be necessary to establish the correction parameters experimentally from a comparison of remotely sensed data with those determined locally by ground truthing, together with corrections for the atmospheric conditions prevailing at the time of measurement.

A number of definite gaps in available information may be cited. In some Arctic locations there is very little information on such common factors as temperature, precipitation, wind, solar and long wave radiation (inward and outward radiation), type and proportion of cloud cover, fog, visibility, concentrations of ozone and carbon dioxide, how the air temperature varies from year to year with changes in the incoming radiation, the relative contributions to the atmospheric moisture content by evaporation from open leads and by sublimation of snow and on the formation and persistence of temperature inversions over the ice cap. Insufficient experimental work has been carried out on transmission, emission, and backscatter of infrared radiation under cold weather conditions, including the effects of winter fog, snow and ice crystals. In fact, even in the temperate regions, very little has been done to measure the backscattering coefficient in the infrared. Further study of rapid variations in the refractive index of water droplets with wavelength in the 8 to 14μm region is required to provide an understanding of the transmission of infrared, and particularly laser radiation through aerosols, and for the development of a model for atmospheric scattering and absorption under Arctic conditions.

The reflection characteristics of snow and ice under a variety of weathering conditions have not been sufficiently investigated nor have there been adequate determinations of the thermal contrast one may expect between a target and its background after radiation has traversed a given path through the atmosphere under given meteorological conditions. Most of the transmission studies have been on the attenuation of a beam of infrared radiation under these conditions rather than on target contrast. In relation to this, little quantitative information exists on the reduction of surface contrast with surface wind speed and the rate of dissipation of a given temperature gradient at a given ambient temperature under various wind conditions and the pronounced variations in spectral signatures which result from both diurnal and seasonal changes. Finally, data are lacking on the amount of attenuation one would expect for an electromagnetic beam, at microwave and infrared frequencies, propagating through falling and blowing snow.

PATTERN RECOGNITION

The other factors of importance, in Arctic reconnaissance, namely the enormous area to be sensed and the sparsity of search support from ground personnel, must result in the accumulation of far more sensed data and imagery than can possibly be adequately or usefully examined by a limited number of trained interpreters. Thus, there is a requirement for methods of discarding most of the redundant information so that observers can derive information while it is still useful, for example, in time to make it worthwhile to send in rescue parties following location, on imagery, of a downed aircraft.

The human photointerpreter is the vital key to all aerial reconnaissance tasks at the present time and will be for the foreseeable future. Although many advances are evident in machine data processing, the human interpreter has not been superseded. The interpreter employs trial and error, intuition, logical deduction, and a wealth of background experience and knowledge in the achievement of his task. Interpretation is based, in part, on features not resident in the picture, and also upon the viewer's purpose. Many pattern recognition studies of relevance to aerial reconnaissance problems have been reported in the literature, and a conclusion typical of most of them is that "the method shows great promise". However, an operationally workable system remains to be demonstrated. Satisfactory procedures have been developed only for simplified restricted problems, and suffer from many limitations. It is evident that the significant features have not been clearly defined, the important relationships are poorly understood, and the available remote sensing data base is inadequate.

One of the major limitations of automatic spatial pattern recognition is sensor capability. This is particularly true in the Canadian Arctic where winter darkness, low sun angle and cloud cover preclude good photography during much of the year. Thus it is necessary to rely on the infrared line scanner, sideways looking airborne radar and low light level T.V., all of which provide

less spatial resolution than photographs. Furthermore, since some of these sensors image in normally non-visible portions of the electromagnetic spectrum, intensive efforts to provide acceptable ground truth are necessary.

Ideally, a man/machine system concept should be adopted in which the pattern recognition and data processing functions are performed by the system element which is best suited to the task. The human becomes the pattern recognizer. The machine is the data processor, which performs all data handling and image processing tasks except direct pattern recognition. The machine should provide preprocessing of all available support data which the interpreter demands, and should display it in the most easily assimilated and interpreted form. Preprocessing attempts to modify the transfer function of the data retrieval system to compensate for the effects of system noise, distortion, blurring, etc., and for environmental effects of atmosphere, illumination, terrain structures, etc. Non-restorative preprocessing is also employed to remove irrelevant and redundant information or to enhance edges, contours, and gradients by a "non-realistic" transformation. Preprocessing is not generally selective against uninteresting objects and often generates artifacts which require perceptive interpretation. Conversely, all data derived from the imagery by the human should be input to the computer through the on-line hardware which he operates in the performance of his interpretive task. Other machine assists could be the storage and retrieval of information, semi-automatic entry of mission flight data, mensuration and computation (of object size, height, ground position, distance between objects, etc.), rectification, reorientation, plotting, and compilation and print-out of formatted reports.

The central and most difficult problem in pattern recognition is that of feature extraction from a background of irrelevant detail. This process must precede recognition. An ideal feature extraction would be independent of noise and limited variations in object size, orientation, contrast and topological deformation. In the past, feature extraction systems have had most success in the field of character recognition, limited success in target acquisition systems and less success in other areas. There is currently no unique feature selection technique applicable to all pattern recognition problems, and no general theory to guide the choice of relevant features. These features are generally tested for effectiveness in terms of their contribution to the probability of correct recognition or in terms of a specified selection criterion, the choice of which may be empirical.

Gestalt approaches to pattern recognition treat patterns as unified organized entities, and assume that analysis into characteristic elemental features destroys the pattern. Mask matching techniques and its variants, which include optical matched filtering, fall into this class, and are highly sensitive to object size, orientation, and shape, and also to translation of the correlation filter. Even though techniques are available to circumvent scale and orientation sensitivity, mask matching offers little prospect of direct utility to pattern recognition problems in Arctic reconnaissance because the interesting patterns lack the required uniformity.

Time represents a very powerful and valuable dimension in which to discriminate and identify patterns. Change detection is very difficult to automate, however, because time-displaced imagery usually cannot be placed in registration, and point-by-point comparison cannot be employed. Sophisticated pattern recognition techniques are required to effect a spatial correlation of equivalent image elements in the two images before detailed changed detection can be implemented.

There is no simple solution to any of the Arctic reconnaissance and surveillance problems using pattern recognition, particularly in areas such as the detection of men in distress, small-scale camps and human activities, etc. However, deliberate attempts to be conspicuous to a known pattern recognition system by parties in distress greatly enhance their chances of detection.

We recommend that pattern recognition studies be concentrated, at least for the present, in the areas of preprocessing or image enhancement, spectral discrimination of targets, discrete target detection and identification, and change detection.

In conclusion, we make a plea for an improvement in one aspect of the interpretation problem that does not form part of our review. A very important component of remote sensing data is accurate positional information, effectively annotated by computer or otherwise to the data. Data or imagery often lose much of their usefulness through lack of an accurate navigational fix at the time of data acquisition. It must be

recognized that navigational equipment employed on reconnaissance missions is often inadequate to provide the required accuracy.

REFERENCES

(1) R.J. Brown, DREO report in preparation.

(2) W.A. Young, "A Pattern Recognition Survey with Implications for Arctic Reconnaissance and Surveillance", Final Report, DSS Contract Serial No.2GRO-99, March 31, 1971.

CAN WE TEACH COMPUTERS

TO SEE?

T. Kasvand,
Control Systems Laboratory,
National Research Council of Canada,
Ottawa, Ontario.

ABSTRACT:

The number of pictures obtained by satellites, from aerial photography, from bubble chambers, microscopes and so on is nearly unlimited. Ideally we would like this picture data to be analysed by computers.

On specific but limited problems fair success has been achieved. Thus, bubble chamber pictures are processed in large numbers, several types of micro-biological objects are recognized by computers, attempts have been made to classify fingerprints, to recognize objects on aerial photographs, etc.

By pooling the experience gained from successes as well as failures, the author tries to outline the requirements for a fairly general pattern recognition system which can be taught to recognize given objects in the pictures.

INTRODUCTION - WHAT IS THE PROBLEM

The question raised, i.e. can we teach computers to see, raises many additional questions. Seeing to us is a perfectly natural everyday activity that we take for granted. However, the ability to see is not in the eyes, nor is the old proverb correct which states that "beauty is in the eyes of the beholder". Seeing is a very complicated information processing task which we have acquired during hundreds of millions of years of competitive existence. Attempts to solve similar information processing problems by using computer equipment is called pattern recognition. The computers are used for their logic capabilities and consequently serious pattern recognition problems could not be investigated until computers became freely available. This trend seems to be reversing now, huge amounts of computer equipment is idling, but every microsecond utilized has to be paid for. Since the solution to a pattern recognition problem cannot be found in a handbook, nor is there a set of ready-made formulas to be exploited, large amounts of computer time may be required just to determine the nature of the problem. Computers are by no means ideal for pattern recognition tasks, but they are the best we have at present.

At this point it might be suitable to include a few anatomical facts about our "hardware" for seeing[1]. In our eyes we have about 10^8 light sensitive rods and cones (i.e. photocells). The output from these is "collected" (in the sense of logical operations) and transmitted to the brain via about 10^6 transmission paths (along the two optic nerves). In the brain there are some 10^{10} active elements or neurons, which communicate with 10^2 to 10^3 others. The transmission speed along the "wires" (called axons and dendrites) which interconnect the active elements (neurons) varies between 1 m/sec to 300 m/sec. The entire system operates in parallel with a "cycle time" in the order of 0.3 sec. We know practically nothing about the details of how the information is processed in the brain. The knowledge we have is on the "speculative systems philosophy" level and will be mentioned further on for comparative purposes.

Pattern recognition is a very general name for methods of extracting useful information from pictorial material. The word picture may stand for a printed page, a roll of paper from a cash register, a medical x-ray, a picture taken through a microscope, an aerial or satellite photograph etc. The number of possible pictures to which pattern recognition techniques may be applied is very large. It is not certain that the useful information can be extracted from the picture using our present knowhow. However, without trying and without adequate equipment one cannot make the necessary experiments since the present pattern recognition "theory" is inadequate to deal with practical problems.

The fundamental difference between the usual data processing or information retrieval and extraction of information from a picture (i.e. pattern recognition) is the nearly total absence of control over the "data base". It is meaningless to think about elaborate coding schemes or input data files, since they simply do not exist in a picture. The input information from which the processing has to start is the picture itself and any attempt to force a data structure onto the picture is an admission of failure at the start.

In certain very restricted cases it is feasible to force a data structure onto the picture. The best examples are commercial mark sense, magnetic ink and optical character readers. They work well only when the specifications regarding character size, shape, thickness, tilt, position and contrast relative to the paper are according to specification. However, on a microscope slide we cannot ask the bacteria to line themselves up like soldiers and to present their "best face" forwards so that some elementary technique can count or classify them. Nor can we expect all farmers to make all their fields of the same size and shape and lay them out relative to some well-defined reference marks (and perhaps even put a code in the top left corner of each field identifying the crop type) so that a machine extracting the data from aerial photographs can use simple logic.

To continue the example of the aerial photographs and the farmers, there is of course a certain amount of control over the pictures and much additional information available. Thus, the position and elevation where the picture was taken is known, the spectral range, the sun angle, the atmospheric conditions, the time of day, the general layout of the area (map), etc. are all known. In the mechanical analysis of the picture, some or all of this additional information may have to be included to allow unique decisions to be made.

It is well known that when one is asked a definite question for which one has to find the answer in a picture, the person's search stragedy varies according to what information is wanted. The search procedures are based on experiences and context as to what to look for to answer the question (2). It may be difficult to believe that our own excellence in pattern recognition is largely based on context and experience but, when during experiments the data is presented out of context, our own ability to recognize approaches or even sink below that of a simple algorithm. Furthermore, pictures do not necessarily contain the explicit information we claim we can extract from them.

Even though attempts are made, we are by no means sure how to program context and experience in a useful manner. However, an easy way to introduce context is to involve the human operator in the automatic pattern recognition system and let him decide what to do when the computer gets into difficulties. The amount of information in the picture is very large from the information theory point of view. The actual information which we want is very much less and will vary depending on what is wanted.

PRESENT APPROACH TO PATTERN RECOGNITION PROBLEMS

The present approaches to pattern recognition problems might be best characterized as cautious. The initial hopes of solving such problems did not materialize. The problems are more difficult than were initially expected, or perhaps more correctly stated, the initial very simple algorithms could not be extended and generalized. The number of different approaches or algorithms is about as large as the number of persons who have tried to solve these problems.

The "classical" pattern recognition theory defines the problem as consisting of three or four separate sub-problems. These are called:
a) Picture scanning to convert the pictures into a computer-readible form.
b) Preprocessing, which does something to the picture to make the next step easier.
c) Feature extraction or extraction of certain measures on the objects in the picture.
d) Classification or recognition of the objects, based on the measures obtained in step c.
This breakdown is, of course, rather general, but only for step d, i.e. the classification, does there exist a theory. The classification step is important but it is a relatively minor subproblem. Furthermore, the pattern recognition problem is far more subtle than the above breakdown indicates, since it is a closed-loop process involving both learning and hypothesis testing.

To produce a successful application of pattern recognition at present requires a compromise by either dividing up the problem between an operator and a machine or by simplifying the problem to a level suitable for the machine alone.

The simplification is carried to an extreme in the magnetic ink and optical character recognition machines. Similar, but not so restrictive simplifications are possible in many other cases where control exists over the contents and quality of the picture (3). With careful sample preparation, photographing and picture selection many biological and other types of problems can be solved (4,5).

However, picture quality cannot always be maintained in production runs or the data from the less than perfect pictures is sooner or later wanted. The relatively elementary processing procedures cannot handle these cases. Additional logic may be added to the programs, but this only moves the problem one step further back where attempts will be made to relax the quality specification even further. This procedure, it would appear, could be iterated a sufficient number of times to produce the desired solution. In practice such a "cut-and-try" procedure does not converge beyond a few rounds of modifications. In fact it starts to diverge producing worse results than the original simple logic did since it is not possible to comprehend all the combinations of logic and input picture data.

A rather sophisticated set of logic statements will have to be arranged in some kind of hierarchy, preferably starting with the simple (and fast) logic, and when this fails, the machine falls back on to more and more complicated (and slower) logic rather than giving up. Such a procedure has been found to give good results in practice, but it depends on the ability of the logic to detect that it has failed(3). This is not always the case.

At present many (but not all) pattern recognition problems can be solved by using a man-machine system (17). The automatic part of the system should be able to extract about 85 to 95% of the wanted information from pictures, if they match the quality specifications upon which the programs are based. If higher accuracy is wanted, the system will have to display its results to the operator and ask advice and demand correction or verification of the data it has extracted. The rules for man-machine interaction will vary from problem to problem. In general they may be characterized by the following phrases:

"Hello operator, I got troubles here, help me out!"

"Hello operator, this is what I found, are you happy?"

For further details see ref. 6.

The 85 to 95% recognition accuracy should not be interpreted as a must. Very useful picture processing systems can be constructed where the operator does most or all of the "recognizing" and the machine carries out the measurements, the bookkeeping and sometimes collects data for "learning" programs (7,8,9).

A GENERAL SOLUTION (10)

A time invariant picture may be defined as a two-dimensional function (i) of its two spatial coordinates x and y, i.e.

$$i = f(x,y) \qquad (1)$$

where i is the illumination or transparency and x and y are arbitrary coordinates assigned to the picture. The function i may be assumed to represent gray levels only, since each spectral range may be analysed separately. Colours actually help rather than hinder. Normally all the variables are quantized into finite ranges, giving the so-called picture matrix

$$i^{\Delta} = f(x^{\Delta}, y^{\Delta}) \qquad (2)$$

which is stored in the computer. It is best to assume that no additional data besides i or i^{Δ} is available about the picture to prevent the temptation of constructing specific algorithms which cannot be generalized later.

Of course, now the question arises: What can be computed from equations (1) or (2)? In principle, any kind of mathematical operation may be applied, but one has to ask what purpose such a procedure is to serve. If the spatial gray level frequencies are of interest, these can be obtained from a two-dimensional Fourier transform or from holographic techniques. If gray level statistics are needed, a histogram will do, cross- and auto- correlations may be computed etc. Such techniques, however, treat the picture as a unit and tend to lose the individual characteristics of the objects in the picture. It is the individual objects in the picture that have to be located and identified. Consequently any procedure which tends to ignore the individual objects or object pieces and their interrelationships is of limited use in pattern recognition.

Standard pattern recognition procedures require an object to be isolated before recognition is attempted. Isolation is accomplished by contour following or by

segmentation algorithms developed for specific pattern recognition problems. After the object or a part of it has been isolated it is normalized in size, and possibly put into a standard position and orientation, before being presented to a classification algorithm.

A picture normally contains a number of objects of varying gray levels, different sizes and shapes, in arbitrary positions and orientation. If, in addition, the objects touch or overlap, the number of possible combinations becomes too large to handle with specific separation algorithms. It appears as if one has to know which object one is dealing with before it can be separated from the others. Of course, if the object is known, its recognition is no longer needed. The problem appears to become completely intractable if the objects to be recognized are not known at the time of programming of these procedures.

Before attempting to define what should be computed from the picture, one has to know what transformations an object in the picture may undergo without necessarily becoming another object. Clearly an object may be of varying size and differ in its gray level, it may be anywhere in the picture (translated) it may be rotated and it may be distorted to a certain extent. Besides these variations, the object may be partially obscured by other objects, it may appear as a mirror image of itself, it may have noise, obscuring lines, shadows and texture on its surface, it may even be broken into pieces, etc. It is meaningless to hope to store all the possible versions of an object in the computer. Only one or a few descriptions of the "ideal" version of the object could be stored. It is more meaningful or logical to store descriptions of pieces of objects only. Symbolically, we may express the transformations as follows:

$$O^i = DRSGO^n + T \qquad (3)$$

where G represents the scaling of gray levels, S is related to the variation of the size of the object, R represents rotation or orientation and D stands for distortion. O^n is the object as seen in the picture, while O^i is its "ideal" version. The variations in the object caused by noise, texture, etc. are not included in equation (3). More about this later.

In general, it can also be assumed that an object in the picture consists of a certain number of pieces or atoms which are in certain positional relationships with respect to each other. At least some of the atoms are "visible" at any given instant. Obviously, if no part of an object is visible in the picture, it may as well be assumed that the object is not present. Symbolically expressed,

$$O = a_1, a_2 \cdots a_k \cdots a_n \qquad (4)$$

where O = an object
a_k = k-th atom of the object.

The atoms $a_1, a_2, \cdots a_n$ have certain positional, size and orientation relationships among each other. Consequently, equation (4) should not simply be interpreted as a subset of atoms $a_1, a_2 \cdots a_m$, where $m > n$.

We are accustomed to talking about "ideal" atoms, but these are never observed. What actually is seen in the pictures is a "non-ideal" version of the atom. The relationship between the ideal atom (a^i) and its non-ideal version (a^n) may be expressed as

$$a^i = D\,R\,S\,G\,a^n + T \qquad (5)$$

where, as above:
G = scaling for gray levels
S = size transformation
R = rotation transformations (orientation)
D = change due to distortion
T = translation from an unknown position to a known position.

On the picture we can only hope to locate a^n and from it compute a version of a^i, before any comparison or recognition is attempted. However, the operations D, R, S, G are unknown, and we do not know where the atom is located, i.e. T is unknown. Equation (5) may, in principle, be inverted and any "seen" version of the atom computed, but attempts to try all possible combinations of D^{-1}, R^{-1}, S^{-1}, G^{-1} and T for every area found in the picture fails due to excessive computational requirements.

The problem takes on a different character if the computer representation of the seen atom (a^n) is constructed in terms of variables for which we know the effects of D, R, S, G, without having to know which ideal atom a^i it resembles most. In other words, with proper selection of atom representation, it should be possible to normalize the atom before its recognition. Thus, the area in the picture which is to represent an atom is to be found before attempting recognition, and the atom has to be described in terms in which the effects

of D, R, S, G, T are either known or can be ignored. The variables in terms of which an atom is described will be called "point features". The atom will be located by using a so-called τ function. For simplicity, the discussion will be limited to gray level pictures containing essentially two-dimensional objects.

The "Point Features"

According to requirements defined above, the point features will have to be computed from the picture without knowing anything about picture contents and they will have to possess known relationships to the D R S G and T transformations.

The conditions placed on the entities to be computed from the picture (the point features) do not define them uniquely. In previous papers (10,11) it has been argued that the contour of an object and the curvature of the contour are the primary point features. The contours of objects in the picture are obtained by differentiating the picture function, smoothing the results and by finding the local maxima of a function of the various spatial derivatives. For discussions of this problem see (12). The curvature is computed as the change of direction of the contour over unit length of contour.

Difficulties in computing the point features may easily be encountered in noisy pictures and the computations tend to be very time consuming without special hardware, but they can be done. The resultant point features have known relationships to the transformations and the first objective or step in recognizing the objects has been accomplished.

It is interesting to note that one set of logical operators in the eye detects short line sections of given slope (13). There are also gradient operators. The author has not encountered any information relating to general curvature detectors. However, a contour may be approximated by straight line sections and the curvature may be computed (if needed) from sequential sections.

The Γ Function

Even though the point features are necessary, they are not sufficient. This becomes clear from the following considerations.

The objects in the picture can be in any position and orientation, of any size, etc. How is one to "pin down" an object or a part of an object under these conditions? Nothing is known about the location, orientation etc. of the objects. However, some form of object description is needed before any kind of comparison becomes possible with stored information in the computer.

This is a very real problem which is normally circumvented by various methods. In optical character recognition, for example, the size, position, orientation etc. of the letters is predefined within very narrow tolerances which must not be violated if successful results are even hoped for. None of these machines can read the characters as they appear in a bowl of alphabet soup.

However, the more realistic problems are of the alphabet soup variety, i.e. the objects are in random locations and orientations, etc. A much milder, but still quite restrictive condition is now introduced. The objects in the picture are assumed or required to be separated from each other and also preferably well contrasting against the background. Under these conditions various contour following or "boxing in" operations become feasible. The objects are now "pinned down" allowing various rotation, normalization etc. techniques to be used to force them into some "standard" form. After this operation the "classical" pattern recognition techniques, like feature extraction, decision space formation etc. become applicable.

Thus, the second problem is to find regions in the picture where objects or object fragments (atoms) might be located. This is accomplished by defining a function (the τ function) whose local extrema can be computed without needing to know what the picture contains.

The purpose of the τ function is to locate atoms in the picture, since the objects will have to be located and if need be fragmented into atoms before any recognition should be attempted. The τ function is used to define an area η (over the object) which is to be called an atom and it has to locate a point $e(x,y)$ which is to serve as the origin for a co-ordinate system in which a description of this area is formed.

Symbolically expressed

$$e(x_o,y_o) = \text{local extremum of } [\tau(p(x,y))] \text{ over } x,y \quad (6)$$

where $e(x_o, y_o)$ = the centre or focal point of the atom

$p(x,y)$ = point features

The area η is bounded by a curve $\omega(x,y)$ such that the point features inside η, or functions computed on the point features, exhibit an extremum property or a limit. Obviously, the translation T has now become known.

To illustrate the idea, assume that

$$\tau = \text{distance to closest contour} \quad (7)$$

This definition of τ fragments objects with clearly defined boundaries, without surface texture and noise. Thus, by computing the distance to the closest contours ($\tau(x,y)$) over the entire picture, we immediately see that

$$e(x_o,y_o) = \text{local max}_{x,y}\ \tau(x,y) \quad (8)$$

gives points (x_o,y_o) at the local peaks. The boundary $\omega(x,y)$ for the neighbourhood η is given by the bounding contours (distance = 0) and the location of min ($\tau(x,y)$) if no contour exists. Several other types of τ functions are described in Reference 10.

The most critical aspect of the τ function is not its mathematical form, but the requirement that it has to fragment the objects in the same way, irrespective of how the objects are orientated and irrespective of their size and position. It should not be too sensitive to minor distortions and noise and it should be independent of the number of objects in the picture. It does not matter if many different additional atoms are formed by the inter-object spaces, provided that the atoms of an object remain relatively invariant. The absolute minimum requirement is that at least one atom of each object can be located, normalized and later identified irrespective of "what has happened" to the object.

There is a possibility that such a mechanism exists in the visual system. The so-called fixation experiments (14) have shown that objects looked at "fall into pieces" where the pieces are logical building blocks out of which the object in question may be constructed. It is not a stationary process but fluctuates apparently randomly, some pieces vanish, others become visible, new types of pieces and combinations are formed, etc. The phenomenon is not visible without experimental aids. It may be that the same mechanism is in action when one stares for example at a checker board, where various squares tend to form into different patterns.

It appears as if the visual system is forming various tentative descriptions of areas in the picture where a more or less "coherent" description can be obtained. These "coherent descriptions" we presumably have seen before or remembered in one way or another. It is an open question whether a fragment is noticed during these experiments which has never been seen before. The next stage in the automatic processing is to form "coherent" descriptions of neighbourhoods in the picture.

Atom Description

The τ function has defined an area η over the picture, bounded it by $\omega(x,y)$, and has given a point $e(x,y)$. It now remains to construct an easily manipulable representation of the atom. Many different representations are, of course, possible, see Reference 10. The following one has been extensively studied, it is used in the programs and has been found to give acceptable results.

At $e(x_o,y_o)$ place the origin of a polar co-ordinate system (r, θ). In each θ direction θ_j we find at least one contour at distance r_j unless r_j is limited to a maximum value, or the edge of the picture is encountered. At the contour all the point features are available. Along the radius r, from zero to r_j we have a set of gray levels $i(x,y)$.

A simplified version of the atom, used in the present programs, consists of the following table:

x_o, y_o = location of atom given by (9) $e(x,y)$

$\theta, r, |g|, \varphi, c$ = a 5 tuplet of numbers giving the angual directions θ and the radii r (to the contour in the polar co-ordinate system), the magnitudes of the illumination gradients $|g|$, their angles φ and the curvatures of the contours c. $\theta = 0^o, 15^o, \ldots 345^o$ at present.

The size of the atom is now defined as the average ($\bar{r}$) of the radial directions r. Obviously all the radii can now be scaled by $\bar{r}$,

and the curvatures c by $1/\bar{r}$. The gradient may likewise be scaled by the average gradient $|\bar{g}|$. Thus, the gray level transform G and the size transform S have been compensated for. Partial compensation for distortion (D) is obtained from the relative insensitivity of the location of the extremum (x_o, y_o). However, given enough distortion, one atom can be changed to another atom. Thus, the effect of the distortion should be considered during atom recognition where the seen atom is compared to a list of "ideal" atoms. The effect of rotation R is directly available in θ and φ. At present the atom is not normalized for rotation.

Object Learning and Recognition

To summarize, the point features can be computed without any knowledge of picture contents. Based on the point features, it is possible to define functions (τ) which can fragment complicated objects into atoms without the need to know what the picture contains. For each of the unknown atoms a description can be formed and normalized after which recognition of the atom is attempted. It may be preferable to try to formulate atoms which are recognizable by human as being certain parts of objects, but this is not a necessary condition.

Now a very interesting set of both programming and philosophical problems has arisen. However, before dwelling more deeply into these problems a simple object learning procedure is outlined which has been programmed and found to give rather interesting results.

The necessary and sufficient conditions for the outlined procedure to work are best illustrated by a highly simplified example of a system which would be capable of learning to recognize objects.

(1) "Show" to the computer an object (O) on neutral background and isolated from other objects. This will ensure that the atoms formed all belong to the object O, simplifying programming for automatic learning. The background is not important if the operator teaches the machine online.

(2) Let the object O be fragmented by the τ algorithm and form normalized atom descriptions (a_k) for each, i.e.,

$$O = a_1, a_2, \ldots, a_k, \ldots, a_n$$

(3) Each of the atoms a_k are compared with the already known "ideal" atoms in machine's memory. Since the atoms were normalized, the comparison may simply consist of a distance computation (in the θ, r, $|g|$, φ, c space) between the memory atoms and the seen atoms. Thus, if the memory atoms are $A_1, A_2, \ldots, A_j, \ldots A_m$, then we find for example that

$a_1 = A_1$ with a reliability or probability p_{11}

$a_1 = A_2$ with a reliability or probability p_{12} etc., for all a_k and A_j

The best matches between the a-s and the A-s are selected. Now two possibilities can occur:

(i) The match is not "good enough" (i.e. a_k is too far from A_j or p_{kj} < some limit) in which case the seen atom is a "new" one and is stored as a new memory atom A_{m+1}. At start, when the machine "knows nothing", all the sufficiently different seen atoms are stored as memory atoms.

(ii) The match is acceptable. In this case the seen atom is "recognized" as a version of a memory atom.

(4) Even if the machine found some unknown atoms in the object O, the object may still be "known" to the machine, since atom interrelationships have to be taken into account, and all atoms are not of equal importance. Thus the atoms should be weighted according to their importance in separating objects in a given recognition problem, based on what the machine already "knows", and the best matching atom combination should be chosen.

For the machine to learn to recognize the object O, it has to be programmed to form a description of the atoms and their interrelationships. In the present programs the teaching stage is controlled by operator. There appears, however, to exist no unsurmountable obstacles to programming the machine to form "its own" atom interrelationships.

(5) For the machine to recognize the object O, it is sufficient that one of the atoms a_1 to a_n is identified correctly, and that there exist instructions for starting a search from the known atoms to the other as yet unlocated and unidentified atoms. This is essentially an hypotheses testing procedure where, starting from the identified atoms, an hypothesis is formed as to what objects the machine may be dealing with. Based on the guide-lines of how to proceed

from atom to atom, the machine can be programmed to verify, reject and alter its hypotheses.

The above simplified system is confused by objects with surface texture, pictures containing much noise, or where intentional "camoflaging" effects have been used, since the formation of the atom is not under the guidance of "what the machine expects to see", nor are alternate descriptions of the same area formed. Thus an additional "feedback loop" has to be introduced between the atom formation process and the atoms that the machine has already seen. These experiments have to wait until access to a larger (than 8K core) computer becomes possible.

An hypothesis testing procedure is observed or at least assumed to be the plausible explanation of human behaviour when they are faced with recognition problems. Under normal circumstances there is only one meaningful solution to a recognition task, i.e. there is only one way of "seeing" the object. However, there are pictures which are ambiguous, i.e. there are two or more ways of interpreting the information in the picture. These effects are usually called optical illusions (15). A person looking at such a picture sees either one or the other of the interpretations, they may alternate, but both interpretations are never seen at the same time.

In case of brain damage, a large variety of interpretive defects are observed. The phenomena range from very subtle to rather gross. In the simpler cases it appears as if various "subroutines" have been ruined. To mention a few as examples (16), a person may not be able to distinguish between M and W. When asked to draw a man, the arms, legs, eyes, nose etc. are all jumbled up in the picture. Their relative positions are not correct. An object which has been "crossed out" by drawing a couple of lines across it is totally unrecognizable, etc.

The philosophical aspects of object learning and recognition brings out all the problems of artificial intelligence. Such a discussion would become both lengthy and speculative. Very few actual results are available to serve as guide posts or mile stones towards the solution. These are problems which have to be solved before we can elevate the computing machines from their present "number crunching" role to more useful and intelligent devices.

CONCLUSIONS

An attempt has been made to put the past and present efforts in pattern recognition into a perspective.

The present method of solving pattern recognition problems may be characterized as specific solutions applied to specific problems. If the necessary control exists over the contents of the picture to be processed, an 85 to 95% accuracy in recognition is feasible. Beyond this accuracy, or when the picture is too complicated for automatic analysis, the pattern recognition system has to include a human operator. The operator's task is to help the computer out when it gets into difficulties or to verify the results. Such interactive programs are quite simple to write and require very little mental effort on the part of the operator.

Pattern recognition systems which are not constructed as specific solutions to specific problems are being investigated. (18) Much more emphasis has to be placed on these systems since the specific solutions cannot be refined indefinitely and the pattern recognition problems are increasing both in number and complexity.

The "general system" as outlined, requires a considerable amount of programming, but there are no major conceptual difficulties to be overcome. In practice, a myraid of detailed problems are encountered. However, without even a modest attempt at constructing such a system, we do not know how close we are to solving the really interesting problems in pattern recognition.

APPENDIX: SOME EXPERIMENTS WITH THE "GENERALIZED SYSTEM".

In order to test out the proposed "general" method, a simplified version of it was programmed. The experiments have been carried out on a small (8K) process control type computer connected to a controllable flying spot scanner. The picture can be resolved into about 4000 x 4000 resolution elements and it is treated as a read -only memory.

The programs consist of three major parts, called analysis, teaching and recognition. Besides these, there are display programs which allow the operator to interact with

the computer and which are photographed for later study.

Analysis

This section contains programs for computation of the point features and the τ function. The atom is formed and compared with atoms in memory.

The picture is scanned and displayed on a memory scope for operator purposes. The operator may advise the machine as to which part of an object he thinks is an atom. The approximate co-ordinates are passed to the τ algorithm which decides whether the area can be called an atom. The operator cannot override the decision of the τ algorithm. If an atom is found, its description is formed, normalized and compared against a list of memory atoms. The best match is computed and this memory atom is displayed for operator comments. If the atom was not recognized, (which happens at the start when the memory is zeroed) or the operator informs the machine that the recognition is not good enough, this atom is stored as a new memory atom, and is now "perfectly" recognized. If the match is acceptable, the atom is entered in a "list of seen atoms" which contains data about the location of the atom, its size and orientation with respect to the corresponding memory atom, its name, etc. This list is used in the subsequent teaching process. The above procedure is repeated until the machine has seen a sufficient number of atoms in terms of which a complicated object can be described. With the help of a "memory map" of the seen areas, and a limit on recognition accuracy, the analysis stage can be run automatically.

Teaching

The teaching of object inter-relationships is at present under operator control. In the first version of this program there are only three position operators (a) point to; (b) equal; (c) logical "and".

(a) With the pointing operation the machine is taught expressions of the type: if you see atom a_1 at (x_o,y_o) go search for atom a_2 in location (x,y). The distance and angular direction from a_1 to a_2 is stored in terms of the size and orientation of a_1. Consequently the representation is size and rotation invariant. The pointing operator is chiefly used to guide the machine from "less important" to "more important" atoms of an object.

(b) The equal operator allows us to specify that two objects (O_1 and O_2) are the same ($O_1 = O_2$) or to specify which atoms compose an object.

(c) The logical "and" operator defines the logical and positional relationships of the atoms out of which a complicated object is composed. Thus, we may say that object $O = a_1 \cdot a_2 \cdot a_3$ which means that object O consists of atoms a_1, a_2 and a_3 which are in a certain positional arrangement with respect to each other. The machine uses the positional relationships to guide itself from the already seen to the yet to be discovered atoms to verify the existence or non-existence of a given object.

The logical "or" is achieved by repeating "and" statements. "Negation" is possible but has not been included in all programs. Compound statements of the form $O_1 = a_1 \cdot a_2$, $O = O_1 \cdot a_3$ are allowed, as well as the definition of the hierarchy levels for objects. The atoms are at the lowest hierarchy. The weights for atoms and the "majority and" are not included at present due to memory limitations.

When a sufficient number of statements about the object has been given the teaching is terminated. However, whenever the machine gets into difficulties during the recognition stage, the teaching may be continued. In principle, it is possible to program the machine to allow it to "teach itself". This aspect can be programmed since sufficient experience has been obtained, but the present computer system is inadequate.

Recognition

In the recognition part of the program the machine uses the above positional relation statements to guide itself from the already found atoms to atoms which the machine expects to find in certain positional relationships with respect to the found atoms, in order to identify all (or the majority of) atoms belonging to one of the desired objects.

In the beginning of a recognition experiment, the hierarchy of the desired objects is specified. The machine may be guided to a starting point in the picture or started from a random number generator. The starting

point is given to the τ function, which tries to locate an atom. If an atom cannot be found, a new random (or given) starting point is used, until an atom has been located. The atom description is formed, normalized, and the atom is compared with memory atoms. The best matching memory atom is selected. Now the machine looks in its list of language statements to decide what to do next. If a statement exists involving the found atom it will "look for" the connected atom. This process continues until one of the desired objects is recognized or the machine has exhausted all its language statements.

Comment

The system is working and is being taught various practical pattern recognition problems. In one of the first experiments, the machine was taught to recognize nerve-fiber cross-sections, see Reference 4. This it performed adequately. When watching the machine during a recognition experiment, one gets the feeling as if the machine "understood" the problem. However, it also became immediately obvious that if many references of where to go from one atom have been given, the machine also has to be provided with the requirement of finding the shortest path to the goal. Otherwise the machine verifies the existence or non-existence of too many atoms before the atoms belonging to the desired object have been "seen". This can be accomplished with proper "teaching" but a goal seeking algorithm is needed. Many other weaknesses in the program exist, but only now can they be clearly defined.

REFERENCES

1. Polyak, S. The Vertebrate Visual System. Univ. of Chicago Press, 1955.
2. Yarbus, A.L. Eye Movements and Vision. Plenum Press, 1967.
3. Kasvand, T. Pattern recognition applied to measurement of human limb positions during movement. NRC, LTR-CS-41, Nov. 1970.
4. Kasvand, T. Pattern Recognition Applied to the Counting of Nerve Fiber Cross-Sections and Water Droplets. Methodologies of Pattern Recognition, S. Watanabe, p. 317.
5. Hamill, P. Experiments in pattern recognition - Current work on automatic chromosome analysis. NRC, LTR-CS-29, April 1970.
6. Kasvand, T. Man-machine interaction in pattern recognition. Second Man-Computer Communication Seminar on Interactive Graphics. Radio and E.E. Div., NRC. May 31-June 1, 1971.
7. Lipkin, B.S. Rosenfeld, A. Picture Processing and Psychopictories. Acad. Press, 1970.
8. West, E.C. Polly, A system for the analysis of scientific photographs. Canadian Research and Development, Nov.-Dec. 1971.
9. Kasvand, T. Computer processing of films of a conical object in a wind tunnel. NRC, LTR-CS-68.
10. Kasvand, T. Experiments with an On-Line Picture Language. International Conference on Frontiers of Pattern Recognition, Honolulu, Jan. 12-18, 1971.
11. Kasvand, T. Some thoughts and experiments on pattern recognition. Chapter 18, Automatic Interpretation and Classification of Images, A. Grasselli.
12. Ratliff, F. Mach Bands. Holden-Day Inc. 1965.
13. Hubel, D.H. Wiesel, T.N. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. J. Physiol., Vol. 160, 1962.
14. Evans, C.R. Some studies of pattern perception using a stabilized retinal image. Brit. J. Psychol., 1965, pp. 121-133, Vol. 56, No. 2.
15. Attneave, F. Multistability in Perception. Scientific American, Dec. 1971.
16. Luria, A. Higher Cortical Functions in Man. Basic Books Inc. 1966.

17. Kasvand, T. Hardware and software requirements for pattern recognition applications. Submitted to Canadian Journal of Operational Research and Information Processing (INFOR)

18. Kanef, S. Picture Language Machines. Acad. Press. 1970.

THE BLOCK ADJUSTMENT OF COLOUR IN HIGH-ALTITUDE PHOTOGRAPHY

S.H. Collins,
Associate Professor,
Air Photo Laboratory,
School of Engineering,
University of Guelph.

ABSTRACT

This paper describes a complete program of density calibration for colour and multi-spectral high-altitude photography. The central technique is a 'block adjustment' of colour, considered as a multi-dimensional variable, over a complete block of frame photography; in a manner analogous to the block adjustment of point location in analytical photogrammetry. The images of selected ground areas are located in the overlaps between frames and between flight-lines. The colours of these 'pass points' are used to determine correction functions for colour variations that occur within and between frames throughout the roll. The part of the sun-angle effect which is due to the nature of the terrain is suggested as a powerful discriminant for automatic photointerpretation.

A method is also described for the absolute calibration of the whole block for ground radiance. The method is photographic, and it ties the radiance values to a great variety of known terrain and cover types on a regional basis. The value of this work in providing comprehensive ground truth for satellite imagery over a large region is discussed.

INTRODUCTION

The interpretation of photographic imagery can be made quite satisfactorily, for many purposes, on the basis of tones and colours that are only qualitatively related to the reflections of light from the terrain. The usual methods of photointerpretation are examples of this. While most interpreters could do better work if the photos they used were produced under better control, they do not need and could not effectively use imagery in which the densities were quantitatively related to ground radiance.

When any attempt is made to quantify tones and colours of the image, however, it is imperative that the relationships between the ground radiance and the final response at the image should be fully understood and placed on a completely quantitative footing. Many attempts have been made over the last four or five years to relate crop types, soil moisture and similar variables to photographic density or colour values. Even the attempts that were not simply ludicrous were doomed to very low levels of correlation and confidence; and the variations of tone and colour that are described below, when uncontrolled, have forced most investigators to realize the futility of the whole proceeding.

It has also been proven however that spectral radiance can in fact be used successfully for discrimination between crop types, at least. The successful method, developed at the University of Michigan, involves the ratios of the energies received in various spectral regions by a scanner, without absolute calibration. This scanner technique is still severely limited in the scope and accuracy of its determinations, and is not at present capable of giving metrically accurate maps. It is being improved however and may soon make an important contribution to the automatic thematic mapping at which it is aimed.

There has been a full realization in photogrammetry and photointerpretation of the complimentary nature of airborne surveys and ground surveys or 'ground truth'. The technical development of remote sensors has not been accompanied by a similar provision of ground truth, and is foundering on that obstacle right now. The reasons for this are obvious; remote sensors were spawned in war and bred in space, and have only recently been turned to look at the earth with peaceful intent. The men who have carried out the development have not had close contact with the earth sciences concerned with ground truth, in general, and they have had even less contact with photogrammetry and cartography. They are only now beginning to realize that their instruments cannot be put to effective use until a complete ground truth system is worked out.

The work described below is far from complete, and it may be only partially successful. However, it presents at least a valid hope of establishing accurate relationships between terrain features, terrain radiance and the photographic densities that they provide. The important point is that a single roll of high-altitude wide-angle frame photography, if suitably controlled and calibrated, can provide an extension of ground truth over about 10,000 square miles that may include several distinct physiographic regions. The purposes of the techniques described below are as follows:

1. To permit the comparison of terrain colour between widely separated points.

2. To provide values of spectral radiance, within the photographic region of the spectrum, over the whole block of photographs.

3. To permit inference of ground truth over the whole block by extrapolation from the images of test sites.

4. To provide measures of confidence for the colour comparisons, the radiance measurements and the ground-truth inferences.

5. To provide all the information on a metrical base that will assist in relating satellite imagery to an accurate map.

ERTS and Skylab will carry internally-calibrated sensors that will record the spectral radiance of large resolution elements as seen from orbital height. In order to test the capability of the satellite sensors for identification of terrain elements, it will be necessary to have a great variety of well-organized regional information about the terrain. The block of high-altitude airborne photography, when calibrated and related to test sites, will be an ideal medium for organizing and presenting this information, and for permitting its transfer to a good map base.

Generalized Block Adjustment

The fundamental task of classical photogrammetry is to provide from two or more perspective photos a model of the terrain in a constant known scale, correctly oriented with respect to a given datum plane. The analogous problem for energy transfer and recording (as opposed to point location) is to provide a radiance model of the terrain, in which the absolute radiance of a terrain element is known for every wavelength and for every angle of view. The size of the terrain element must be as large as the resolution element of the sensor. As in photogrammetry, the density of the information across the field of view may be low or high, but it must be obtainable with high accuracy.

The analogy with classical photogrammetry can be carried a little further. Relative orientation of two or more metrical airphotos permits the creation of a true space model of the terrain in three dimensions. Absolute orientation then follows to give the space model a definite scale and to level it to a datum. The analogue to relative orientation in our case is the process that will permit the detection of differences in the colour of the terrain between one field point and another, freed of all effects that result from the method of sensing or the position of the sensor. The analogue to absolute orientation is the actual calibration of photo colours or densities in terms of the radiance of the terrain.

In photogrammetry, each photo point is specified by two spatial coordinates. In monochromatic photography, the density of a resolution element can be considered as a third coordinate. In tri-pack colour photography three densities can be measured independently of one another. In the case of multi-spectral photography with four identical cameras working in four spectral regions, a resolution element will be specified by a six-dimensional vector with two spatial coordinates and four density coordinates.

The terrain point itself is specified for the purposes of classical photogrammetry by its three space coordinates. Its specification in terms of radiance is far more complicated; for as well as involving three or more spectral coordinates, the radiance usually differs in different directions. This last property has been considered as an insurmountable difficulty by some workers, but may prove to be useful in the automatic discrimination of different types of terrain.

It is clear that the complete relative and absolute block adjustment is a formidable undertaking, involving the solution of a number of difficult problems. The whole technique has been laid out in this way in order to clarify research goals, and there is reason to hope that a well-equipped and well-staffed organization may be able to accomplish most of these goals.

The Relative Adjustment Procedure

The relative adjustment procedure would ideally consist of the following steps:-

1. Sensitometric exposures should be applied at the start and end of each roll, at least. It would be better to apply them also at many points along the roll.

2. A set of terrain types representative of the region is chosen. 'Pass points' are located within examples of each type on every frame on which the type occurs. Many pass points are located on every frame, particularly where there is a multiple overlap between frames and between flight-lines. These points are coded in such a way that the images of a particular point on the terrain can be identified in every frame in which they occur.

3. Density measurements are made at every pass point. At the same time, density measurements are also made at a large number of other points according to the requirements for the extension of ground truth over the region. For example, if regional trends in soil colour are to be established, many measurements will be made on this type of terrain element at locations determined by the physiography of the region.

4. The density values are coded and linked to the positions of the points within frames and to the frame number. Then all data are submitted to a multivariate analysis in which the variability of the data are accounted for. The functions to be fitted in the analysis and the other means employed to account for variability will require a great deal of work. The start that has been made on this work is described below.

5. Ideally, a stereocomparator should be used for locating the pass points and measuring their plane coordinates. A scanning microdensitometer should then be used to measure the densities, and the output of the microdensitometer should include a means of identifying the plane coordinates of any desired point. This technique would be particularly desirable for multi-film work, to save the labour of transferring so many points from one film to another. In any case, it would be most desirable to be able to carry out all of the usual procedures of analytical photogrammetry along with the 'block adjustment' of colour.

Variables of the Adjustment

The variables that must be considered in the block adjustment may be classified in two ways: according to their physical origin or according to the type of correction that must be applied. The first classification is as follows:-

1. Temporal variation of illumination. It may take about two hours to expose one roll of high-altitude photography, and the intensity, colour and angle of the sun's illumination change in this time. The changes in clear weather will be progressive from the beginning to the end of the roll.

2. Temporal changes in the terrain. Changes in soil and plant moisture can occur in a two-hour period. These changes will usually be small if the photos are taken in the middle of the day.

3. Angles of View relative to the azimuth of the sun and relative to the vertical. These are the most difficult effects to analyse, but they are also the most interesting. The mean radiance of the terrain is greatest in the direction of the sun, because no shadows are seen from that direction. This creates a 'hot-spot' in one quadrant of each frame. The hot-spot is more pronounced in high-altitude than in low-altitude photos. Part of this may be due to the lower resolution of the high-altitude photos for shadow detail at ground scale. That it is not primarily due to the atmosphere is proven by photos taken over water. There is no hot-spot in such photos, although of course there is a strong reflection of the sun in the opposite quadrant of the photo.

This hot spot effect is one of the sources of photographic density variation that is being analysed by Jane Law of the Air Photo Laboratory at the University of Guelph. The photography used for the analysis was taken from 60,000 feet, using a 6" lens and 2443 I.R. colour film. A graphic representation of the hot spot effect together with the radial lens/filter effect is shown in Figure 1. Figure 2 illustrates the hot spot effect alone. The models were constructed from visual density readings.

The variations in effective terrain radiance with the elevation of the sun and the

azimuth of the camera relative to it may prove to be most useful in the identification of terrain features. As an example of this, consider a green pasture and a green deciduous woodlot beside it in early summer photography. In the region of the hot-spot their colours might be nearly identical. In other regions the woodlot will expose varying proportions of shaded area to view, while the dense pasture growth does not. The amount of shadow detail may serve as a discriminant of cover types particularly when the flying height is so great that the height of the cover cannot be distinguished stereoscopically.

4. Effects of the atmosphere. Scattering by the atmosphere is not homogeneous, and back-scattering counter to the direction of the sun's rays is stronger than scattering at right angles to the sun's rays. The effect appears to be small in clear weather but may not always be negligible. It might be difficult to separate the effect from variations of terrain reflectance unless part of the flight is over water or over calibrated ground targets at different angles of view. There is also a purely radial atmospheric effect because of the greater air path, and consequently more scattered light, at the wider angles of view.

5. Lens and filter effects. These are constant in time and should create a purely radial variation about the principal point. They are quite large, however, and may be quite complicated functions of the radius when all colours are considered. Figure 3 illustrates the radial density variation induced by a 6" lens, a minus blue filter and an anti-vigretting filter when used to expose film at 60,000 ft.

6. Film and processing. It is difficult to achieve constant colour balance in colour photography, and sensitometric exposures along the roll would be helpful in correcting the changes that occur even along the roll. The whole block adjustment process could be carried out with greater precision and with greater dynamic range in multi-spectral monochrome photography, where process control is far easier and where the sensitometry of the individual films can be chosen almost at will.

7. Sensitometric effects in small detail. When working on image areas that are not much larger than the smallest possible resolution element for the system, adjacency effects in development and scattering of light in the film emulsion must be considered. No correction seems possible for these effects, but it may be wise to consider them in attempting to visualize how small details should be rendered.

The second classification is the one that must be used in actual experiment. It relates the variable to the effect it produces at different points within the frame and along the roll:-

1. Variations along the roll. These may be due to variations of film and processing, to temporal variations of illumination and to temporal changes of the terrain. Sun angles also change during the flight, but can be largely ignored if all azimuths are measured relative to the sun. The frame number is the coding for roll position.

2. Radial variations within frames. The lens and filter effects are the predominant causes of purely radial variation of colour, although atmospheric haze causes radial change as well.

3. Tangential variations within frames. There are no purely tangential effects, because this type of effect is associated with the varying reflections and shadow effects of the terrain. Adjustment functions are required in which the tangential coordinate and radial coordinate appear together.

Photo point locations can be measured in either a polar or a rectangular coordinate system. In either case it is convenient to take an origin at the principal point and the x-axis (or zero-angle axis) in the azimuth of the sun. This choice removes both the effects of crab of the camera in flight and of the slow change of the sun's azimuth throughout the roll.

Absolute Radiance Calibration of the Block

The output of the relative adjustment procedure is a set of adjusted colour (density)

values for each point of the terrain that has had an entry in the program. At this stage, the relationships between the film colour values and the actual spectral radiances of the terrain are not known. An internal sensitometric calibration of the film-processing combination is probably essential for the complete success of the whole procedure, but it is not sufficient to establish accurate values of ground radiance. There are too many ways in which the light is altered in quantity and quality on its passage from the ground to the film.

Ideally, one would like to have large grey test-panels of six or eight different reflectances set up within the region. The spectral radiance of these panels could be established, at the time when they are photographed, for all vertical and azimuth angles (it would be convenient of course if the targets could be considered as perfectly diffuse reflectors within the angles accepted by the camera). If the sensitometry of the film and processing has been established, only one set of targets would be required. Otherwise, they would have to be placed at different locations throughout the region.

Ground targets of any size are expensive to set up and very difficult to maintain in good condition. In this case the size is very large, as will be seen from the following calculation. Consider photography from an altitude of 15,000 meters with a 6" mapping camera. Only the center of the field of view need be considered, because the relative adjustment procedure should compensate quite accurately for radial effects. The resolution element of the lens-film combination, even for a single-layer emulsion, will be at least 0.020 mm in size. For good densitometric measurements the ground targets should give images at least five times this size, preferably larger. Each ground panel would then have to be at least 10 meters square.

The work being carried out by Dr. Dieter Steiner of the University of Waterloo, with the help of the Air Photo Laboratory of the University of Guelph, offers a really useful and comprehensive method of calibrating high-altitude imagery. In his basic research into automatic terrain recognition, he has set up suitable test-panels and has mounted four 70mm cameras as a multi-spectral group in an aircraft of the Ontario Department of Lands and Forests that Victor Zilinsky is sometimes able to place at his disposal. The films are processed under sensitometric control. In each flight, whatever its particular purpose, exposures are made over the test targets and over the experimental farms of the University of Guelph, where there is a great variety of recorded ground truth on soils, crops, crop diseases and yields. This type of photography, carried out at relatively low altitudes, can be used for the absolute calibration of high-altitude photography in terms of ground radiance. It is even more important, however, that it will provide the only feasible and economical method of establishing the detailed characteristics of a large number of areas within the high-altitude block. This will calibrate the block in a quite different sense, one that will be important to all remote sensing techniques that are carried out over the region.

Progress in the Research

The details of the techniques and the scientific results will not be included in this paper, which is only intended to show the overall plan of the research. The complete process of relative and absolute block adjustment will take several years to develop, will not reach its full potential until high-altitude and low-altitude airborne missions are carried out in coordination with the over-flight of a Skylab research team or something similar.

A partial relative adjustment has been carried out for a block of photographs taken over the eastern Lake Ontario region in connection with the International Field Year for the Great Lakes. The camera was a Wild RC8, the film was Kodak 2448 false-colour IR, and the flying height was 20,000 meters. The date of the photography was early July, 1970. The radial and tangential variations of three-dimensional colour seem to have been characterized quite well, but the computer programs have not been developed yet to the point where full power of the block adjustment process can be realized.

The absolute calibration procedure has been established as far as the photography is concerned, but has not yet been carried out because no simultaneous low- and high-altitude photography is available. The only problems that remain arise from the need of instrumental means of rapid density measurement coordinated with location measurement.

Summary

There is now no question that satellite imagery is important in certain applications, the most noteworthy being the prediction of weather phenomena. There is also no question

that ERTS and Skylab will uncover other valuable applications. The value of the imagery that these satellites will provide will be greatly enhanced if, on a regional basis, a wealth of ground truth is presented for comparison. The resolution element of much of the imagery will be so large that the ground truth will have to be correlated to ground radiance values integrated over quite large areas.

Aerial photography at very high altitude seems the ideal medium for collecting and displaying the required ground truth. It can easily be related to a metrical base, and each frame is large enough at ground scale to facilitate the integration that will be necessary. At the same time, it is sufficiently detailed that the extension of ground truth over the region can be partly carried out by photointerpretation, and that the photography itself has great value for a variety of mapping purposes.

This paper presents the general plan of a method for turning blocks of high-altitude photography into accurately calibrated records of ground radiance from which the integration to large resolution elements of satellite imagery can be carried out. Along with the radiance information, the method will permit the simultaneous and well-coordinated display of specific items of ground truth, chosen to meet the needs of the potential users of the high-altitude and of the satellite imagery within the region. The method will assist in answering the question that has arisen in every method of photointerpretation and remote sensing: "Within the area of the ground covered by this small portion of an image I'm looking at, what are the actual details on the ground and how do they make their contribution to the energy that is being received?"

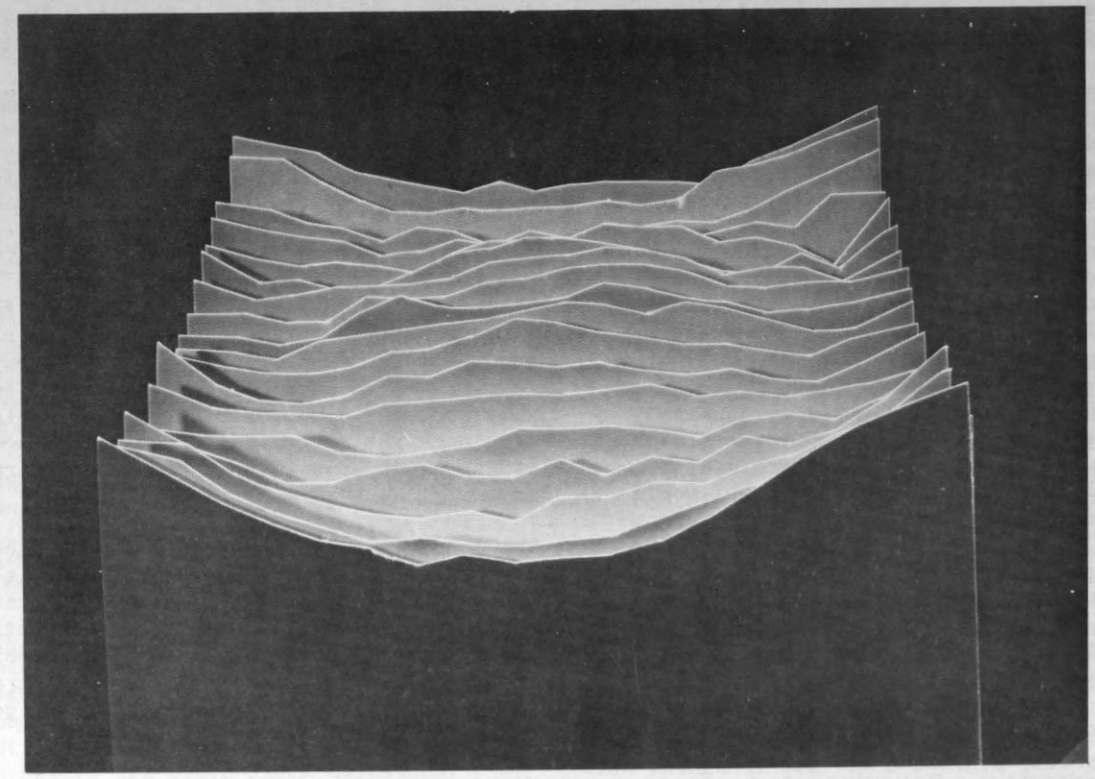


Figure 1. A model representing variations in visual density caused by the hot spot and radial lens/filter effect, on a single frame of photography. ASL 60,000 feet.

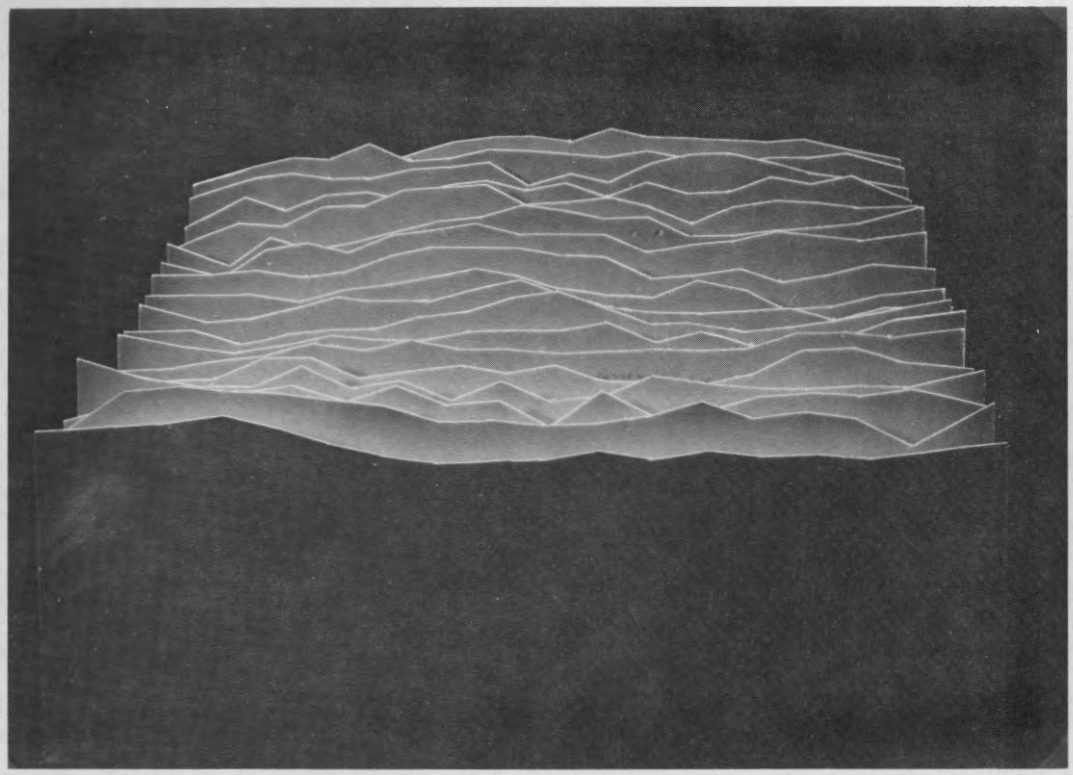


Figure 2. A three dimensional model of the hot spot effect alone.

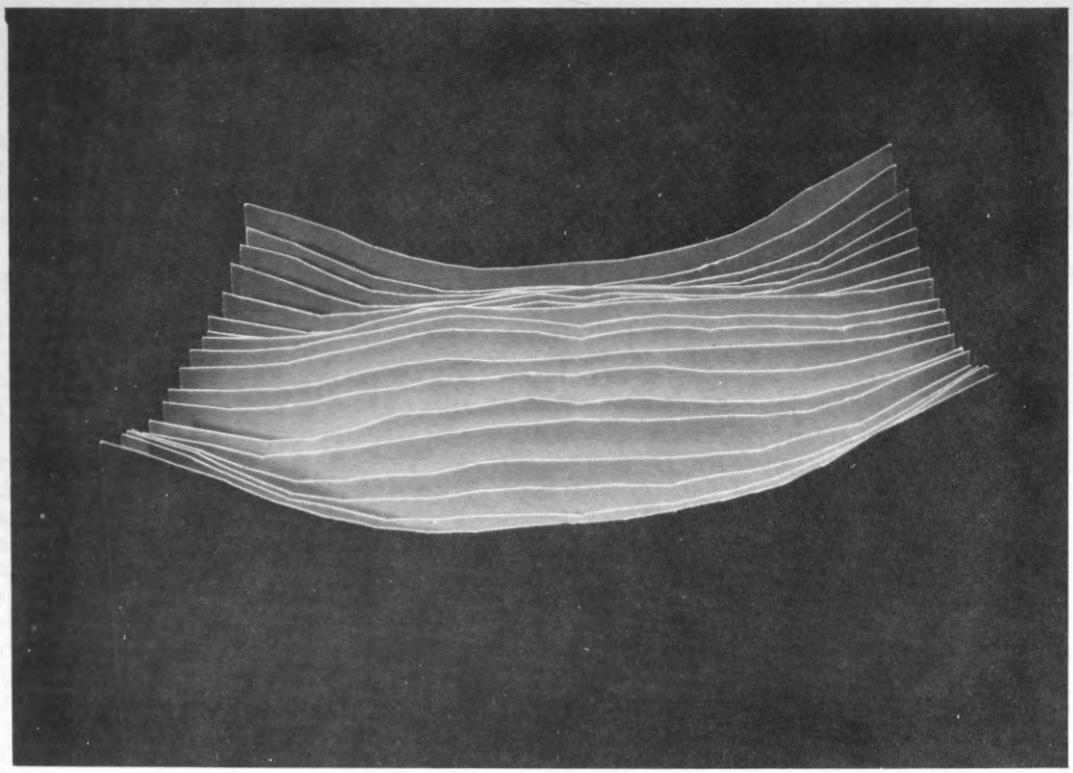


Figure 3. A model of radial variation in visual density induced by the lens filter effect.

MINIMUM DISTANCE CLASSIFICATION IN REMOTE SENSING

A.G. Wacker
Associate Professor
Electrical Engineering Department
University of Saskatchewan
Saskatoon, Saskatchewan

D.A. Landgrebe
Laboratory for Applications of Remote Sensing
Department of Electrical Engineering
Purdue University
West Lafayette, Indiana

SUMMARY

The utilization of minimum distance classification methods in remote sensing problems, such as crop species identification, is considered. Minimum distance classifiers belong to a family of classifiers referred to as sample classifiers. In such classifiers the items that are classified are groups of measurement vectors (e.g. all measurement vectors from an agricultural field), rather than individual vectors as in more conventional vector classifiers.

Specifically in minimum distance classification a sample (i.e. group of vectors) is classified into the class whose known or estimated distribution most closely resembles the estimated distribution of the sample to be classified. The measure of resemblence is a distance measure in the space of distribution functions.

The literature concerning both minimum distance classification problems and distance measures is reviewed. Minimum distance classification problems are then categorized on the basis of the assumption made regarding the underlying class distribution.

Experimental results are presented for several examples. The objective of these examples is to: (a) compare the sample classification accuracy (% samples correct) of a minimum distance classifier, with the vector classification accuracy (% vector correct) of a maximum likelihood classifier; (b) compare the sample classification accuracy of a parametric with a nonparametric minimum distance classifier. For (a), the minimum distance classifier performance is typically 5% to 10% better than the performance of the maximum likelihood classifier. For (b), the performance of the nonparametric classifier is only slightly better than the parametric version. The improvement is so slight that the additional complexity and slower speed make the nonparametric classifier unattractive in comparison with the parametric version. In fact disparity between training and test results suggest that training methods are of much greater importance than whether the implementation is parametric or nonparametric.

INTRODUCTION

A fairly common objective of remote sensing in connection with earth resources is to attempt to establish the type of ground cover on the basis of the observed spectral radiance. The examination of systems capable of achieving this objective shows that a certain duality of system types exists. Landgrebe[1] refers to the two types as image-oriented systems and numerically-oriented systems. The duality exists promarily for historical reasons as a consequence of the independent development of photographically oriented and computer oriented technology. The primary distinction between the two system types is that in image oriented systems a visual image is an essential part of the analysis scheme while in numerically oriented systems the visual image plays a secondary role. In Fig. 1 the location of the "Form Image" block in relation to the "Analysis" block characterizes the two system types.

In numerically oriented remote sensing systems it is frequently possible to design the data collection system in such a manner that classification becomes a problem in pattern recognition. This situation prevails if one attempts to study earth resources through the utilization of multispectral data-images. The term multispectral image (i.e. without the modifier data) is used to refer to one or more spectrally different superimposed pictorial images of a scene. The modifier data is added to indicate that images are stored as numerical arrays as opposed to visual images.

To obtain a multispectral data-image of a scene, the scene in question is partitioned on a rectangular grid into small cells (pixels) and the radiance from each pixel for each wavelength band of interest is measured and stored. The set of measurements for a pixel constitutes the measurement vector for that pixel. A multispectral data-image for a scene is simply the complete collection of all measurement vectors

for the image. The spatial coordinates (i.e. row and column number) of each pixel are of course also recorded to uniquely identify each measurement vector. Fig. 2 depicts the situation.

The methods used to generate multispectral data images can conveniently be divided into two categories. In the first category, film is used to record the image. The film is subsequently scanned and digitized to produce a data-image. The multispectral property is obtained either by scanning several images photographed through different spectral windows, and overlaying the data; or by utilizing color film and separating the spectral components during the scanning procedure. In the second category the image is generated electrically and stored in an electrically compatibile form, usually on magnetic tape as either an analog or digital signal. The electrical signal to be stored can be generated by a number of different systems, the multispectral scanner and return beam vidicon probably qualify as the two most common examples. For the scanner the multispectral property is obtained by filtering of the spectral signal collected through a single aperture prior to recording, or by the superposition of several unispectral images collected through different apertures.

As already stated, pattern recognition techniques can serve as the basis for affecting classification of multispectral data-images. Much of pattern recognition theory is formulated in terms of multidimensional spaces with the dimensionality of the space equal to the dimensionality of the vectors to be classified. This vector dimensionality is, of course, determined by the number of attributes or properties of each pixel to be considered in the classification (e.g. number of spectral bands). Classifying a multispectral data-image by classifying the observation vectors from such an image on a pixel by pixel basis falls naturally into this common pattern recognition framework. In contrast to this vector by vector approach there are classification schemes which collectively will be referred to as "sample classification schemes". In such schemes all vectors to be classified are first segregated into groups (i.e. samples) such that all the vectors in a group belong to the same class. The whole group of vectors is then classified simultaneously. The minimum distance method considered is one such classification scheme.

In utilizing sample classification schemes two distinct problems can be identified. The first is concerned with partitioning the measurement vectors into homogeneous groups, while the second is concerned with the classification of these groups. Except for the comments in the next paragraph consideration is restricted to the second problem.

It frequently occurs for multispectral data-images that many of the adjacent measurement cells belong to the same class. For example in an agricultural scene each physical field typically contains many pixels. In fact it is precisely this condition that prompts the investigation of sample classification schemes. In such situations the physical field boundaries serve to define suitable samples for problems like crop species identification, and it is in this context that sample classifiers might also be referred to as per-field classifiers. It is apparent that for the situation just described one method of automatically defining samples is to devise a scheme that automatically locates physical field boundaries in the multispectral data-imagery[2,3]. For the minimum distance classification results presented later, physical field boundaries will actually be used to define the samples, but the field boundaries are located manually rather than automatically. A second and perhaps more promising approach to the problem of defining samples is via observation space clustering. In this approach vectors from an arbitrary area are clustered in the observation space, and all the vectors assigned to the same cluster constitute a sample irrespective of their location in the arbitrary choosen area. In this case the term "fields" no longer seems appropriate and consequently the term sample classifier is preferred over the term per-field classifier.

It is apparent that sample classification schemes cannot be used in all situations where a vector by vector approach is possible. A basic requirement is that the data to be classified can either be segregated into homogeneous samples, or occurs naturally in this form. Where the minimum distance scheme can be applied it intuitively has several potential advantages over a vector by vector classifier; in particular it is potentially faster and more accurate.

It seems logical that provided the time required to automatically define the samples is not too great, then sample classifiers should be faster than a vector by vector classifier. This is of considerable importance in utilizing a numerically-oriented remote sensing system to survey earth resources because a characteristic of such surveys is the tremendous volume of data involved. One would also anticipate that the vector classification accuracy (% vectors correctly classified) for vector by

vector classifiers would be lower than the sample classification accuracy (% samples correctly classified) for sample classifiers. The reason for this is that in sample classifiers all the information conveyed by a group of vectors is used to establish the classification of each vector, whereas in vector by vector classifiers each vector is treated separately without reference to any other vector. In a sense sample classifiers utilize spatial information because vectors a are classified as groups, which naturally have some spatial extent. No spatial information is used in vector by vector classifiers, consequently, sample classifiers should perform better since spatial information is certainly of some value.

MINIMUM DISTANCE CLASSIFICATION

Problem Formulation

In a certain sense minimum distance classification resembles what is probably the oldest and simplest approach to pattern recognition, namely "template matching". In template matching a template is stored for each class or pattern to be recognized (e.g. letters of the alphabet) and an unknown pattern (e.g. an unknown letter) is then classified into the pattern class whose template best fits the unknown pattern on the basis of some previously defined similarity measure. In minimum distance classification the templates and unknown patterns are distribution functions and the measure of similarity used is a distance measure between distribution functions. Thus an unknown distribution is classified into the class whose distribution function is nearest to the unknown distribution in terms of some predetermined distance measure. In practice the distribution functions involved are usually not known, nor can they be observed directly. Rather a set of random measurement vectors from each distribution of interest is observed and classification is based on estimated rather than actual distributions.

It is necessary to define more precisely what constitutes a suitable distance for minimum distance classification. Mathematically the terms "distance" and metric are used interchangeably. For our purpose it is convenient to distinguish between the two terms. In essence all that is required for a well defined minimum distance classification rule is a measure of similarity between distribution functions which need not necessarily possess all the properties of a metric. The term distance refers to any suitable similarity measure while the term metric is used in the normal mathematical sense. More specifically metric on a set S is a real valued function $\delta(.,.)$ defined on S X S (X indicates cartesian product) such that for arbitrary F,G,H in S

(a)	$\delta(F,G) \geq 0$	1
(b)(1)	$\delta(F,F) = 0$	2
(2)	If $\delta(F,G) = 0$ then $F = G$	3
(c)	$\delta(F,G) = \delta(G,F)$	4
(d)	$\delta(F,G) + \delta(G,H) \geq \delta(F,H)$	5

A distance, as used herein, is defined to be a real valued function $d(.,.)$ on S X S such that for arbitrary F,G,H in S at least metric properties a,b(1) and usually b(2) and (c) hold. For theoretical proofs it is in fact often desireable to require that d be a true metric while in practical application such a restriction is usually not necessary.

Not only are distances between individual distribution functions of interest but since each class could conceivably be represented by a set of distribution functions the distance between sets of distributions is also of interest. Definition 1 defines the distance between sets of distributions.

Definition 1 - Let the distance d(F,G) be defined for all F,G, in A, where A is an arbitrary set of cdf's of interest. If A_1 and A_2 are non-empty subsets of A then the distance $d(A_1, A_2)$ between the sets A_1 and A_2 is defined as

$$d(A_1, A_2) = \inf_{\substack{F \in A_1 \\ G \in A_2}} d(F,G) \qquad 6$$

Note that definition 1 applies to finite and infinite sets of distribution functions. Of course, if the sets are finite then taking the infimum is equivalent to taking the minimum.

Furthermore, if each set consists only of a single distribution function then the distance between the sets is precisely the distance between the distribution functions. The distance between a distribution function and a set of distribution functions is also included as a special case. It is necessary to make some comments about the usage of the notation d(F,G). Some of the distance measures considered are expressed in terms of probability density functions (pdf's) rather than cumulative distribution functions (cdf's). The convention adopted is that the notation d(F,G) is still used and referred to as the distance between cdf's, even though the distance is expressed in terms of the densities

of F and G (i.e. in terms of f and g).

The minimum distance classification scheme can now be formally defined. It is convenient to use a decision theoretic framework for this purpose. In general to specify a problem in this framework it is necessary to specify:

(a) Z - the sample space of the observed random variable.

(b) Ω - the set of states of nature; that is, the set of possible cdf's of the random variable. If the functional form of the cdf is known, then Ω can be identified with the parameter space.

(c) $\overset{*}{A}$ - the action space; that is the set of actions or decisions available to the statistician.

(d) $L(a,F)$ - loss function defined on $\overset{*}{A}X\Omega$ which measures the loss incurred if $F\epsilon\Omega$ is the true state of nature and action $a\epsilon\overset{*}{A}$ is the action taken.

The general formulation of the minimum distance problem in this framework follows:

(a) $Z = E^q$ (q-dimensional Euclidean space)

(b) $\Omega = [\Omega^{(1)}, \Omega^{(2)}, \ldots, \Omega^{(k)}]$ where $\Omega^{(i)}$ is the set of possible distribution functions for the ith class, i = 1, 2, ..., k.

(c) $\overset{*}{A} = [a_1, a_2, \ldots, a_k]$ where a_i is the decision to decide the random sample to be classified belongs to the ith class, i = 1, 2, ..., k.

(d) $L(a,F) = 0$ if $F\epsilon\Omega^{(i)}$ and action a_i was taken

$L(a,F) = 1$ otherwise.

A decision rule is a function defined on Z and taking values in $\overset{*}{A}$. The minimum distance decision rule is given by definition 2.

Definition 2 - Let $\underline{Y}$ be the vector of all sample observations. The minimum distance decision rule $D_{MD}: Z \rightarrow \overset{*}{A}$ is $D_{MD}(\underline{Y}) = a_i$ (i.e., decide the random sample to be classified belongs to class i) in case

$$d(\tilde{F}_N, \Lambda^{(i)}) = \min_{j = 1,\ldots,k} d(\tilde{F}_N, \Lambda^{(j)}) \qquad 7$$

Where $\Lambda^{(i)}$ is the set of cdf's selected to represent the ith class and $\tilde{F}_N$ is a sample-based estimate of the cdf of the random sample to be classified.

Several items in definition 2 require clarification. The vector $\underline{Y}$ includes not only the random sample to be classified, but also any other observations used in the classification procedure. For example, if training samples are used for each class, these are included in $\underline{Y}$. The sets $\Lambda^{(i)}$ also require comment. $\Lambda^{(i)}$ may be the set of all possible distributions for class i (i.e. $\Lambda^{(i)} = \Omega^{(i)}$) or it may be a subset of $\Omega^{(i)}$ or the sample-based estimates of a set cdf's selected to represent class i. Finally the term sample-based estimate is used to refer to any estimate of a cumulative distribution function or its corresponding density which is based on a random sample from the distribution in question. A number of suitable estimators exist[4] and the present formulation does not restrict the type of estimator. Later attention will be focused on distance measures based on densities. In the parametric case the densities will be estimated by estimating the parameters describing the densities (parametrically estimated pdf's). In the nonparametric case density estimates will be based on histograms (density histogram estimation). To obtain a density histogram estimate of a pdf the observation space is partitioned into square bins and the probability density estimate in any bin is the percent of vectors used to estimate the density which fall in the bin.

A number of special cases of the above formulation are now considered. These special cases are basically a consequence of making different assumptions regarding Ω, and $\Lambda = [\Lambda^{(1)}, \Lambda^{(2)}, \ldots, \Lambda^{(k)}]$. In Type I problems the sets of distribution functions representing the classes are assumed to be known sets. Actually, this problem is not of great interest from a practical point of view, since class distributions are not normally known, but it is interesting from a theoretical point of view because of its relative simplicity.

Type I - The $\Omega^{(i)}$'s are known sets of cdf's

Case (a) The sets $\Omega^{(i)}$ are infinite and $\Lambda^{(i)} = \Omega^{(i)}$

Case (b) The sets $\Omega^{(i)}$ are finite and $\Lambda^{(i)} = \Omega^{(i)}$

Case (c) The sets $\Omega^{(i)} = F^{(i)}$ (single cdf/class) and $\Lambda^{(i)} = F^{(i)}$

Type II problems differ from Type I problems in that the possible distribution functions for each class are known to be q-variate distributions but are otherwise unknown. Consequently, all distributions used in the minimum distance decision rule must be estimated. Since in practice only a finite number of estimated distributions can be utilized this factor must be considered in formulating the problem. If the sets of states of nature (i.e. the $\Omega^{(i)}$'s) are infinite the infinite sets must somehow be replaced by a representative finite set. A similar attitude must be adopted if it is known apriori that the sets $\Omega^{(i)}$ are finite

but it is not known precisely how many distribution functions each $\Omega^{(i)}$ contains (i.e. how many subclasses of wheat are there?); or even if the precise number is known, it may not be known how to obtain a random sample for each distribution function (i.e. how are samples representing different subclasses of wheat selected?). Finally, in the finite case, even if a random sample for each distribution function of interest can be obtained, their number may be so large that for practical reasons it may be desireable to use a smaller number of representative distributions. Thus, the need arises for a method to select a representative set of distribution functions from a larger (possibly infinite) set. To do this assign a distribution $H^{*(i)}$ to $\Omega^{(i)}$, i = 1, 2, ..., k. That is the events to which probability mass is assigned by $H^{*(i)}$ are sets of distributions in $\Omega^{(i)}$. To select a random set of cdf's from $\Omega^{(i)}$ (i.e. to select a random set of training samples for the ith class) is now equivalent to selecting a random sample from $H^{*(i)}$.

The above formulation is rather complicated in that a distribution over a space of functions is involved. This complexity can be avoided by restricting consideration to a parametric family characterized by s real parameters. Making the logical assumption that a one to one correspondence exists between cdf's in $\Omega^{(i)}$ and points in the parameter space $\theta^{(i)}(\equiv E^s)$, it is apparent that assigning a distribution $H^{*(i)}$ to $\Omega^{(i)}$ is equivalent to assigning some other distribution $H^{(i)}$ to the parameter space $\theta^{(i)}$. Consequently, in the parametric case rather than deal with $H^{*(i)}$, which is a cdf on a set of distribution function, only $H^{(i)}$ which is a cdf in E^s need be considered.

It is perhaps worthwhile to restate the above ideas in terms of multispectral data-imagery from an agricultural scene before stating them in a more formal manner. In the interest of simplicity and since it is the case of primary interest assume that the true q-dimensional distribution of the radiance measurements from each field belong to the same parametric family which can be characterized in the paramet-space E^s. This family may have a finite or infinite number of members (i.e. subclasses). Further assume that all the fields in a class (i.e. wheat) can be described by a suitable distribution $H^{(i)}$ over the parameter space. A set of training fields for each class is selected at random. Because of our formulation this is equivalent to selecting a random sample from the parameter space according to the assumed distribution over the parameter space for that class (i.e. $H^{(i)}$). For each of the randomly selected training fields the radiance measurements are used to get an estimated cdf for that field. In this way estimated cdf's for a representative set of training fields are obtained for each class. An unknown field is then assigned to the class that has a training field whose estimated cdf is nearest to the estimated cdf of the unknown field. Since the problem as stated is parametric, one would normally, though not necessarily, use parametrically estimated cdf's.

Type II problems in which the $\Omega^{(i)}$'s are unknown are now formally described. While prime interest is centered in the case where Ω is a parametric family this restriction is not imposed in stating the problem. The description of Type II problems is complicated by the fact that the description of the sets $\Lambda^{(i)}$ is rather involved.

Type II - The $\Omega^{(i)}$'s are Unknown Sets of cdf's
Case (a) - The sets $\Omega^{(i)}$ are infinite in number and $\Lambda^{(i)} = \tilde{\Omega}_{M_i}{}^{(i)}$. The sets $\tilde{\Omega}_{M_i}{}^{(i)}$ are now described. First a set of population cdf's corresponding to a representative set of M_i training fields for class i, i = 1, 2,..., k is selected. Let $\Omega_{M_i}{}^{(i)}$ be this set for the ith class. That is $\Omega_{M_i}{}^{(i)}$ is a random sample of size M_i for $H^{*(i)}$. A sample-based cdf is then obtained for each cdf in $\Omega_{M_i}{}^{(i)}$ for l = 1, 2, ..., k. The resultant set of sample-based estimated cdf's is $\tilde{\Omega}_{M_i}{}^{(i)}$. For the case where parametrically estimated cdf's are used $\tilde{\Omega}_{M_i}{}^{(i)}$ can also be considered to be a random sample of size M_i in the parameter space according to a distribution $H^{(i)}$.
Case (b) - The sets $\Omega^{(i)}$ are finite and $\Lambda^{(i)} = \tilde{\Omega}^{(i)}$ or $\Lambda^{(i)} = \tilde{\Omega}_{M_i}{}^{(i)} \supset \tilde{\Omega}^{(i)}$. If the $\Omega^{(i)}$ are finite sets (i.e. finite number of subclasses) then it is desireable to let $\Lambda^{(i)} = \Omega^{(i)}$, where $\Omega^{(i)}$ is the set of sample-based estimated cdf's for the ith class. In cases where the resultant number of subclasses is impractically large and/or only a random set of M_i training fields is available it is necessary to let $\Lambda^{(i)} = \tilde{\Omega}_{M_i}{}^{(i)} \supset \tilde{\Omega}^{(i)}$ and proceed as in case (a).
Case (c) - The set $\Omega^{(i)} = F^{(i)}$ (Single cdf per class) and $\Lambda^{(i)} = \tilde{F}_N{}^{(i)}$.

Distance Measures

The importance in statistics of distances between cdf's has, of course, long been recognized[5], according to Samuel and Bachi[6] their use appears to fall into two broad categories.

(a) Used for descriptive purposes. For example, as an indicator to quantitatively specify how near a given distribution is to a normal distribution.

(b) Use in hypothesis testing, which is, of course, a special case of decision theory.

There is a tendency for distance functions sufficiently sensitive to detect minor differences in distribution functions (i.e. category (a) use) to be somewhat involved functions of the observations, with the result that their use as test statistics in hypothesis testing has been limited because of the complicated distribution theory. On the other hand, distance functions whose theory is simple enough to be readily used as test statistics often do not distinguish distribution functions sufficiently well. Since in minimum distance classification interest is naturally centered on good discrimination between distribution functions, therefore distance functions that fall into category (b) are normally used. Since the appropriate distribution theory for hypothesis testing is then in general not known it is impossible to theoretically compute probability of error, but it may be possible to establish reasonably tight upper bounds. The approximate probability of error can of course be determined experimentally.

The literature abounds with references to distance measures and no attempt will be made to give a complete bibliography. A representative sample of distance measured is given in Table 1. This Table includes the most widely used distance measures because of their obvious importance, as well as more obscure distance measures whose application to the present problem appears reasonable. In addition a few miscellaneous distance measures have been included to give an indication of the variety of distances that have been suggested. The distances included in this Table are: Cramer-Von Mises[7,8,9,10], Kolmogorov-Smirnov [11,12,9,10], Divergence[13,14,15], Bhattacharyya [15,16], Jeffreys-Matusita[13,14,17], Kolmogorov Variational[15,18,19], Kullback-Leibler[15,20], Swain-Fu[21], Mahalanobis[22,23], Samuels Bachi[6], and Kiefer-Wolowitz[24]. The references cited are by no means comprehensive. In selecting the references the attempt has been made to cite only the original source in addition to survey papers. The paper by Darling[9], Sahler [10] and a certain extent Kalaith[15] fall in this latter category.

Most of the references cited are concerned only with the univariate forms of the distance measure. With the exception of the Samuels-Bachi distance, the extention to the multivariate forms is quite natural. Since it is the multivariate forms that are of interest, these, rather than the more common univariate forms, are given in Table 1. For the Samuels-Bachi distance multivariate forms other than the one presented may be possible.

Table 1 also contains information regarding the metric properties of the distance measures when used in conjunction with three families of distribution functions. The families considered are: C, the family of q-variate absolutely continuous distribution functions; MVN, the family of q-variate normal distribution functions; and MVN_Σ, the family of q-variate normal distribution functions with equal covariance matrices. Since MVN and MVN_Σ are subsets of C it is, of course, true that a metric in C is also a metric in MVN and MVN_Σ. A metric in MVN_Σ need not, however, be a metric in MVN or C.

Because of the importance of the multivariate normal distribution, expressions for the distance between two such distributions are given in Table 2 for each of the distances measured in Table 1 in those instances where the expressions are known.

The distances listed in Table 1 are discussed in the references cited and no attempt will be made to discuss them except for some general comments pertaining to their use in minimum distance classification.

Since a large variety of distance measures is available, the problem naturally arises as to which distance measure to use in a given problem. Unfortunately, no complete answer to this question is presently available, but some general comments are possible. The distribution-free properties* that make the Cramer-Von Mises and Kolmogorov-Smirnov distances so popular in the univariate case do not apply in the multivariate case. Since it is the multivariate case that is of interest these distances lose their special appeal. Intuitively a distance like the Kolmogorov-Smirnov distance does not appear to be as good a distance measure as those involving integration over the whole space. It is also more difficult to compute in parametric situations then some of the integral relations. The Samuels-Bachi distance suffers from a similar computational disadvantage.

The Divergence, Bhattacharyya distance, Jeffreys-Matusita distance, Kolmogorov varia-

* In the univariate case the distribution of the Kolmogorov-Simirnov and the Cramer-Von Mises distances between two estimated distribution functions independent of the underlying distributions being estimated, provided appropriate estimators are used.

Table 1

Multivariate Forms of Distance Measures and Their Metric Properties

Name	Form	Metric in C	Metric in MVN	Metric in MVN_Σ
Cramer-Von Mises	$W = \{\int_{-\infty}^{\infty} (G(\underline{x}) - F(\underline{x}))^2 d\underline{x}\}^{\frac{1}{2}}$	Yes	Yes	Yes
Kolmogorov-Smirnov	$K = \mathrm{Sup}_{\underline{x}} \lvert G(\underline{x}) - F(\underline{x})\rvert .$	Yes	Yes	Yes
Divergence	$J = \int_{-\infty}^{\infty} \mathrm{Ln}(\frac{f(\underline{x})}{g(\underline{x})})(f(\underline{x})-g(\underline{x}))d\underline{x}$	No	No	Yes
Bhattacharyya Distance	$B = -\mathrm{Ln}\int_{-\infty}^{\infty} (f(\underline{x})g(\underline{x}))^{\frac{1}{2}} d\underline{x}$	No	No	Yes
Jeffreys-Matusita Distance	$M = \{\int_{-\infty}^{\infty} (\sqrt{g(\underline{x})} - \sqrt{f(\underline{x})})^2 d\underline{x}\}^{\frac{1}{2}}$	Yes	Yes	Yes
Kolmogorov Variational Distance	$K(p) = \int_{-\infty}^{\infty} \lvert p_g g(\underline{x}) - p_f f(\underline{x})\rvert d\underline{x}$	Yes	Yes	Yes
Kullback-Leibler Numbers	$L_{fg} = \int_{-\infty}^{\infty} \mathrm{Ln}(\frac{f(\underline{x})}{g(\underline{x})})\ f(\underline{x}) d\underline{x}$	No	No	Yes
Swain-Fu Distance	$T = \frac{\lvert \mu_f - \mu_g \rvert}{D_f + D_g}$	No	No	Yes
	Where $D_\cdot = \{\frac{\lvert \underline{\mu}_f - \underline{\mu}_g \rvert^2 (q+2)}{\mathrm{tr}\{\Sigma_\cdot^{-1}(\underline{\mu}_f - \underline{\mu}_g)(\mu_f - \mu_g)^t\}}\}^{\frac{1}{2}}$			
Mahalanobis Distance	$\Delta = \{(\underline{\mu}_g - \underline{\mu}_f)^t \Sigma^{-1} (\underline{\mu}_g - \underline{\mu}_f)\}^{\frac{1}{2}}$			Yes
Samuels-Bachi Distance	$U = \{\int_0^1 [F^{-1}(\alpha) - G^{-1}(\alpha)] d\alpha\}^{\frac{1}{2}}$	No	No	No
	where $F^{-1}(\alpha) = \mathrm{Inf}\{c \mid Q_c \cap Q_\alpha \neq \emptyset\}$			
	and $Q_c = \{\underline{x} \mid \sum_{i=1}^{q} x_i \leq c\}, Q_\alpha = \{\underline{x} \mid F(\underline{x}) \geq \alpha\}$			
Kiefer-Wolfowitz Distance	$V = \int_{-\infty}^{\infty} \lvert F(\underline{x}) - G(\underline{x})\rvert e^{-\lvert x \rvert} d\underline{x}$	Yes	Yes	Yes

Notation

(1) F, G are multivariate cdf's with densities f, g; means $\underline{\mu}_f$, $\underline{\mu}_g$; covariances Σ_f, Σ_g; and prior probabilities p_f, p_g.

(2) $\int_{-\infty}^{\infty}(\ )\,d\underline{x}$ designates a multivariate integral.

(3) For Mahalanobis distance F and G are normal with means $\underline{\mu}_f$ and $\underline{\mu}_g$ and have common covariance Σ.

(4) $|\ |$ designates the absolute value or vector norm.

(5) t designates the transpose

Table 2

Distances Between Two Multivariate Normal cdf's

Name	Distance
Divergence	$J = \frac{1}{2}\,tr[\Sigma_f-\Sigma_g][\Sigma_g^{-1}-\Sigma_f^{-1}]+\frac{1}{2}\,tr[\Sigma_f^{-1}+\Sigma_g^{-1}][\underline{\mu}_f-\underline{\mu}_g][\underline{\mu}_f-\underline{\mu}_g]^t$
Bhattacharyya Distance	$B = \frac{1}{8}(\underline{\mu}_f-\underline{\mu}_g)^t[\frac{\Sigma_f+\Sigma_g}{2}]^{-1}(\underline{\mu}_f-\underline{\mu}_g)+\frac{1}{2}\,Ln\,\frac{\det(\frac{1}{2}[\Sigma_f+\Sigma_g])}{\{\det(\Sigma_f)\det(\Sigma_g)\}^{1/2}}$
Jeffreys-Matusita Distance	$M = [2\{1-\frac{\{\det(\Sigma_f)\det(\Sigma_g)\}^{1/4}}{\{\det(\frac{1}{2}[\Sigma_f+\Sigma_g])\}^{1/2}}\exp(-\frac{1}{8}(\underline{\mu}_f-\underline{\mu}_g)^t[\frac{\Sigma_f-\Sigma_g}{2}]^{-1}(\underline{\mu}_f-\underline{\mu}_g)\}]^{1/2}$
Kullback-Leibler Numbers	$L_{fg} = \frac{1}{2}\,Ln\,\frac{\det(\Sigma_f)}{\det(\Sigma_g)}+\frac{1}{2}\,tr\,\Sigma_f[\Sigma_g^{-1}-\Sigma_f^{-1}]+\frac{1}{2}\,tr\,\Sigma_g^{-1}[\underline{\mu}_f-\underline{\mu}_g][\underline{\mu}_f-\underline{\mu}_g]^t$
Swain-Fu Distance	$T = \frac{\lvert\underline{\mu}_f-\underline{\mu}_g\rvert}{D_f+D_g}$ where $D_\cdot = \{\frac{\lvert\underline{\mu}_f-\underline{\mu}_g\rvert^2(q+2)^{\frac{1}{2}}}{tr\{(\Sigma_\cdot)^{-1}(\underline{\mu}_f-\underline{\mu}_g)(\underline{\mu}_f-\underline{\mu}_g)^t\}}$
Mahalanobis Distance	$\Delta = \{(\underline{\mu}_g-\underline{\mu}_g)^t\Sigma^{-1}(\underline{\mu}_g-\underline{\mu}_f)\}^{\frac{1}{2}}, (\Sigma=\Sigma_f=\Sigma_g)$

Notation

(1) t means transpose

(2) det means determinant

(3) tr means trace

(4) The normal distributions involved have means $\underline{\mu}_f$ and $\underline{\mu}_g$ and covariance matrices Σ_f and Σ_g

tional distance and Kullback-Leibler numbers all belong to a class of distance measures which can be written as the expected value of a convex function of the likelihood ratio*. In fact Ali and Silvey[25] have shown that the expected value of any convex function of the likelihood ratio has properties that might reasonably be demanded of a distance measure. In addition Wacker[4] has shown that in feature selection such distance measures have a weak relationship to the probability of error. Kalaith[15] proved the same relationship for Divergence and the Bhattacharyya distance. Since the class of distance measures under discussion is based on pdf's there is probably a tendency for these distances to reflect differences in pdf's rather than cdf's.

Of the distances based on likelihood ratios the Bhattacharyya distance seems to have been gaining in favor. The prime reason for this is apparently the close relation between probability of error and Bhattacharyya distance, as well as the relative ease of computing Bhattacharyya distance in theoretical problems. Other properties of the Bhattacharyya distance which enhance its prestige as a distance measure have been pointed out by Lainiotis[26] and Stein[27]. A property of considerable theoretical utility is the close relation between the Bhattacharyya distance B, the Jeffreys-Matusita distance M and the affinity ρ namely

$$M = 2(1-\rho)^{\frac{1}{2}} = 2(1-e^{-B})^{\frac{1}{2}} \qquad 8$$

Where

$$B = -\mathrm{Ln}\rho \qquad 9$$

and

$$\rho(F,G) = \int_{-\infty}^{\infty}(f(\underline{x})g(\underline{x}))^{\frac{1}{2}}d\underline{x} \qquad 10$$

Because of the above relationships minimum distance classifications made on the basis of the Bhattacharyya distance, Jeffreys-Matusita distance or affinity all yield identical results, and consequently have identical probability of error.

The Jeffreys-Matusita distance is, however, a metric in a much larger class of distribution (see Table 1). This means that theoretical derivations regarding probability of error can be made using the metric properties of the Jeffreys-Matusita distance in this larger class, and the results are applicable if classification is effected using Bhattacharyya distance or affinity as well. This property has been used extensively by Matusita.

While no strong preference for any distance measure can presently be demonstrated the theoretical properties of the Bhattacharyya distance suggests that it might be a reasonable choice and the experimental results presented later are based on this distance measure.

Minimum Distance Classification And Probability of Error

Considerable literature exists on the minimum distance method with Matusita[28-35] and Wolfowitz[36-39] being the chief contributors. Wolfowitz's work is concerned primarily with estimation while much of Matusita's work deals with the decision problem. Contributions have also been made by Gupta[40], Cacoullous[41,42], Sirvastava[43] and Hoeffding and Wolfowitz[44].

In considering minumum distance decision rules a common requirement is to insist that by using arbitrarily large samples the probability of misclassifying a sample can be made arbitrarily small. This is the motion of consistency and it is a reasonable demand if the pairwise distance between all the sets of distributions associated with each class is greater than zero or

$$d(\Omega^{(i)}, \Omega^{(j)}) > 0$$

for all i, j = 1, 2, ...,k; i≠j 11

In parametric problems in which some distribution is assigned to the parameter space the condition specified by 11 is equivalent to requiring that there is no overlap of regions of the parameter space associated with different classes.

It has been shown[40,34,44] that any minimum distance classification problem for which equation 11 hold is consistent (probability of misclassification approaches zero as sample sizes approach infinity) provided the distance and distribution estimator utilized satisfy certain conditions. These conditions are that the distance used must be essentially a metric (metric property b(2) need not hold) and that for the particular distance measure and estimator used, the probability that the distance between the true and estimated distribution can be made arbitrarily small is one for infinite sample size. Further it is shown that certain distances and estimators satisfy these conditions. In particular in the normal case these conditions are satisfied by using parametrically estimated densities and the Bhattacharyya distance[35]. Similar consistency results are not known for density histogram estimators. The known properties of consistency are summarized more rigorously and in greater detail by Wacker[4].

It is the property of consistency described in

* The likelihood ratio of densities $f(\underline{x})$ and $g(\underline{x})$ is $f(\underline{x})/g(\underline{x})$.

the previous paragraphs which makes the minimum distance decision rule potentially so attractive. In essence consistency says that if the condition specified by 11 is satisfied, and if sufficiently large samples are used then the probability of misclassifying a sample should be very small. Unfortunately in classifying multispectral data-images two problems arise.

(1) The number of distributions associated with any class is very large (perhaps almost infinite) and it is not practical to attempt to store all possible subclass distributions as is essentially assumed in deriving the consistency result described.

(2) It appears that the condition of equation 11 is frequently not satisified, or at least that distributions from different classes are often so nearly alike that the number of samples required to distinguish them is impractically large.

When the condition specified by equation 11 is violated to the extent that $\Omega^{(i)}$ and $\Omega^{(j)}$ overlap on a set of non zero probability then the minimum distance decision rule can obviously no longer be consistent; in this situation the probability of misclassifying a sample will be finite regardless of sample size. Under these circumstances, except for the simple parametric example treated by Wacker[4], essentially no results are available.

RESULTS

Three different classifiers were used to obtain the experimental results. These classifiers are known as LARSYSAA, PERFIELD and LARSYSDC. LARSYSAA is a vector by vector classifier based on the maximum likelihood decision rule[45], while PERFIELD and LARSYSDC are minimum distance classifiers utilizing the Jeffreys-Matusita or equivalent (Bhattacharyya) distance. LARSYSAA and PERFIELD are based on the Gaussian assumption and utilize parametrically estimated pdf's while LARSYSDC utilize density histograms to estimate the pdf's. All three classifiers assume equal subclass probabilities and operate in the supervised mode*.

Two examples are discussed. The first example compares the sample classification accuracy (% samples correct) of a parametric with a nonparametric minimum distance classifier. The second example compares the vector classification accuracy (% vectors correct) of the parametric maximum likelihood classifier LARSYSAA with the parametric minimum distance classifier PERFIELD. The data used in both examples is essentially the same, but as subsequently described the training procedures differ considerably.

The two examples discussed are problems in species identification of agricultural fields. In this context it is usually logical to assume that all the measurement vectors from a given physical field belong to the same class. This assumption was made in defining samples for the minimum distance classifiers and in determining the classification accuracy of the maximum likelihood classifier. In other words, for the minimum distance classifiers each sample to be classified represents a physical field, while for the maximum likelihood classifier all vectors from a field are assumed to belong to the same class.

The data for the examples to be discussed has 13 spectral bands and was collected by the University of Michigan Scanner. For ease in referring to different spectral bands the wavelength channel number correspondence of Table 3 is utilized. The data was collected at an altitude of 3000 ft., between 9:45 and 10:45 a.m. E.D.T., on June 30, 1970, from Purdue University flightlines 21, 23 and 24 respectively. The exact location and orientation of these flightlines, which are located in Tippecanoe County, Indiana, is shown in Fig. 3. The flightlines extend the 24 mile length from the north to the south end of the county and are roughly equally spaced in the east-west direction. Since the scanner geometry is such that at an altitude of 3000 feet the field of view is roughly 1 mile, the area covered by the three flightlines, approximately 72 square miles, is about 1/7 of the total area in the county. The scanner resolution and sampling rate are nominally three and six milliradians respectively. This means that at nadir the scanner "sees" a circle about 9 feet in diameter and that the spacing between adjacent pixels is about 18 feet. Since the scanner resolution and sampling rate are independent of look angle the distance between adjacent pixels is approximately 30% larger at the edge of the scanner's field of view with a corresponding change in the shape and area "seen" by the scanner. At the sampling rate indicated there are 220 samples across the width of a flightline and each flightline contains 5000 to 6000 lines. This means each flightline contains somewhat more than 10^6 pixels of which 10% to 20% are typically used for test purposes.

For both examples four principle ground cover

* Supervised refers to the fact that samples whose classification are know are available to "train" the classifier.

categories are considered; wheat, corn, soybeans and other. Although the other class includes a considerable variety of ground cover most of the agricultural fields in this category are either small grains (other than wheat) or forage crops. There are also some bare soils and diverted-acre fields. Some natural categories such as trees and water are also included in this class. For most of the subcategories for the class other ground cover is fairly complete, but the spectral properties of the ground cover are quite variable from field to field within a subcategory. Most of the wheat in the flightline was mature and ready or nearly ready for harvest. In fact some portion of it had already been harvested. For corn and soybeans the crop canopy at flight time was such that the ground was not covered by vegetation when viewed from above and consequently the radiance is greatly influenced by the soil type. This fact makes it difficult to discriminate corn and soybeans at this time of year and consequently high classification accuracies are not to be expected, especially since corn and soybeans constitute a considerable fraction of the ground cover.

Table 3

Correspondence Between Channel Numbers and Spectral Bands

Channel Number	Spectral Band (Micrometers)
1	0.40-0.44
2	0.46-0.48
3	0.50-0.52
4	0.52-0.55
5	0.55-0.58
6	0.58-0.62
7	0.62-0.66
8	0.66-0.72
9	0.72-0.80
10	0.80-1.00
11	1.00-1.40
12	1.50-1.80
13	2.00-2.60

While the particular training procedure used in each example is different some general observations are possible. It is evident that some of the variables which affect radiance tend to be constant within a physical field, but vary from field to field. Such variables are usually related to farm management practices and include such factors as variety of species, fertilization rates, crop rotation practices, etc. Also the variability in soil type can normally be expected to be greater between fields then within fields. Consequently it is not uncommon for all data from one field to be fairly "uniform" but still be quite different from the data from another field; even though the class (species) is the same in both fields. In terms of probability densities the density from each individual field might reasonably be approximated by a normal distribution; in that it is typical unimodal and reasonably symmetrical, but the data from several fields combined frequently exhibit severe multimodality. Under these circumstances, in order that the Gaussian assumption is approximately satisified (for classifiers making this assumption), subclasses are usually defined for each main class, such that the distribution for each subclass is unimodal. Perhaps if data from a sufficient variety of fields could be combined for a given crop species a unimodal distribution would result for each main class and the definition of subclasses would not be necessary, even for a parametric classifier. The class distribution in this case would naturally be broader than the distribution of any "subclass" of which it is composed. It is presently not known in the above situation whether better classification is achieved with parametric (Gaussian) classifiers by using many subclasses whose distribution are relatively narrow, or using fewer subclasses with broader distribution. In practice there appears to be a tendency toward the definition of many subclasses. In nonparametric classifiers it should of course not be necessary to define subclasses as there is no need for densities to be unimodal.

On the basis of the above discussion a fairly general parametric model which at least qualitatively behaves much like the actual multispectral data results when every field associated with each main class is considered as a potential subclass. The variation in distribution parameters from field to field is accounted for by a distribution over the parameter space. This is precisely the problem previously formulated as problem Type II case (a).

Example 1 - Parametric vs Nonparametric

The classifications performed for this example can be segregated into the four categories shown below.

1) Classifications with the parametric classifier PERFIELD
a) Every training field treated as a subclass.
b) Data from all training fields for each principle class combined (no subclasses).
2) Classifications with the nonparametric classifier LARSYSDC
a) Every training field treated as a subclass.
b) Data from all training fields for each principle class combined (no subclasses).

In the classification procedure each flightline was treated as a separate data set. The training and classification method is described for one flightline with other flightlines receiving similar treatment. Initially test and training data must be defined. Every field of any significant size whose classification had been determined by field observation was included as a possible test or training field. These fields were segregated into the four principle classes. Roughly 10% of the fields in each class were then selected at random to serve as training fields. The remaining fields were used as test fields. Table 4 give a break down of the number of test and training field for each flightline. After the training fields had been selected the subclass or class densities were estimated and stored. The test fields were then classified on the basis of their estimated densities by the minimum distance rule. The computations to estimate a density function for PERFIELD are substantially simpler than for LARSYSDC since for PERFIELD only the mean and covariance need be estimated while for LARSYSDC the density histogram must be generated. A bin size of 5 was used for the density histograms in PERFIELD. (The data ranges was 0 to 256). Only 3 of the 13 channels were used in performing the classifications. These were selected in a more or less arbitrary manner, although it was known that the selected set (1,8,11) were among the better subsets of channels.

Table 4

Number of Test and Training Fields

Flightline	Number of Test(Training) Fields				
	Total	Wheat	Corn	Soybeans	Other
21	218(22)	23(2)	79(8)	57(6)	59(6)
23	141(15)	18(2)	58(6)	55(6)	10(1)
24	156(18)	19(2)	52(6)	43(5)	42(5)

The results of the classification are shown in Fig. 4. Rather than present the classification results for each flightline individually the performance averaged over the three flightlines is given. The results therefore give some indication of the classification accuracy one might expect on the average for this type of data for the training method used. In view of the random nature of the training procedure it is felt that this is a more meaningful presentation than quoting the results for each flightline individually.

Example 2 - Maximum Likelihood vs Minimum Distance Classification

For this example the data from flightlines 21, 22, and 23 was classified using:
a) The parametric maximum likelihood classifier LARSYSAA.
b) The parametric minimum distance classifier PERFIELD.

The training procedure in this case is considerably different than the procedure for Example 1. In this case small areas approximately one acre in size were selected from flightlines 21, 23 24 on this basis of a sampling scheme. The sampling scheme simply used every nth acre in the flightline belonging to the class in question as a "training acre". The data from the acres selected in this manner was used to train the classifier. In this manner 59 wheat acres, 44 corn acres, 23 soybean acres and 46 other acres were selected. The sampling rate n was different for the various principle classes. If every training acre were treated as a separate subclass a total of 172 subclasses result. This number exceeds the capabilities of the classification programs. Consequently it was necessary to reduce the number of subclasses to a reasonable number. This was accomplished by means of a clustering program which groups together the acres within each principle class whose estimated pdf's are similar[4]. As a result of this grouping the number of subclasses defined for the principle classes: Wheat, Corn, Soybeans and Other were 4, 10, 6 and 10 respectively. Density histogram estimates of the resulting 4 wheat subclasses are shown in Fig. 5. Note that even after clustering considerable evidence of multimodality still exists, particularily for the first subclasses. In fact in some channels the contribution of all 4 acres assigned to subclass 1 are clearly evident. It is possible that this data should have been segregated into a greater number of subclasses. After the subclasses had been defined by clustering the statistics (means and covariance) were computed for each subclass. The feature selection capability of LARSYSAA[45] was then used to select the "best" 4 of the 13 channels for classification. This selection is based on the average. Divergence between all possible subclass pairs, excluding subclass pairs from the same class. On this basis channels 2,8,11 and 12 were selected. Using these channels both the training acres as well as the test fields were classified both with LARSYSAA and PERFIELD. The classification results for the training acres are shown in Fig. 6 while the results for the test fields (again averaged over the 3 flightlines) are shown in Fig. 7.

Discussion of Experimental Results

It is suggested that in evaluating a classifier a reasonable index of comparison is the overall average classification accuracy. This performance index has the advantage that it gives an indication of the classification accuracy that might be expected from the classifier for similar data and training procedures. For a relatively small data set, it is usually relatively easy to devise a training procedure or classifier which superficially looks superior but whose apparent superiority disappears when results are averaged over a number of data sets. A disadvantage of the suggested performance index is the necessity to do a reasonable number of classifications.

On the basis of average classification accuracy and the training procedures used there is no evidence that the parametric minimum distance classifier is superior to the nonparametric classifier. Neither is there any evidence that using a relatively large number of subclasses improves classification accuracy on the average. This is contrary to expectations.

Actually when each field is treated as a subclass one would expect the nonparametric classifier to perform better than the parametric classifier only if the Gaussian assumption was seriously violated for the various training or test fields involved. Furthermore, for the nonparametric classifier to exhibit any real advantage the nonnormal structure of the data must bear some resemblence from field to field (e.g. modes must appear in same relative positions). Since the nonparametric classifier does not exhibit any superior performance neither of the above factors apparently occur with any consistency.

When the data from all the training fields is grouped one would expect that the data would be multimodal and that the nonparametric classifier would be much superior. The basic fallacy in this reasoning appears to be that although the class distributions are multimodal the samples to be classified are usually unimodal. In other words the distribution of any sample to be classified is not really a random sample from the distribution of any class. Instead it simply tends to account for one of the modes in the class distribution. Furthermore, there is no apparent way of rectifying this situation within the constraints of minimum distance classification.

The fact that the parametric classifier does so well (comparatively) when no subclasses are considered attests to the robustness* of the Gaussian assumption in minimum distance classification.

It must be recognized that in assessing a classifier factors other than the performance index considered are of importance. One other factor that should be considered is the consistency of the results. That is, how near to the average can one expect to get for any given classification. The variance in the average performance is a measure of this consistency. In this regard, although the number of classifications is small , there is evidence that the nonparametric classifier is better than the parametric version and that for the parametric classifier the variance in average performance is increased by combining the data from many fields. This small advantage hardly warrants the additional complexity of the nonparametric implementation.

The results comparing the minimum distance and maximum likelihood classifiers show fairly conclusively that in general the sample classification accuracy of minimum distance classifiers is higher than the vector classification accuracy of maximum likelihood classifier of the same data. This is true for both the test and training data. It is recognized of course that the quantities being compared are by nature somewhat different but nevertheless they represent the natural method of expressing the classification accuracy of each classifier individually and do afford some measure of comparison. This result agrees with expectations although a greater improvement might have been anticipated.

It is convenient to define the difference between the sample classification accuracy and the vector classification accuracy as the improvement factor. The exact value of the improvement factor depends on the particular data but qualitatively it is obvious that for Type II case (a) problems the improvement will be very small or non existent both when the separation of the parameter space densities for all classes is large (one can't improve a high vector classification accuracy much) as well as when no separation exists (subclasses of different main classes can then not be distinguished by either classifier). The experimental evidence suggest that for moderate overlap of the parameter space densities the improvement factor will be of the order of 5% to 10%.

---

* A robust classifier is relatively insensitive to the underlying assumptions about the distributions involved.

In concluding it should be mentioned that no comparative computation times have been given. The fact that the experiments involved a number of different programs, two computer systems (one in a time sharing mode) and the inherent dependence of processing time on the Classification Parameters and on the manner in which the data is stored (data retrieval time is by no means negligible) makes it virtually impossible to give meaningful comparative times. Suffice it to say that to classify a typical flightline time would be measured in fractions of an hour to hours on an IBM 360 System Model 44, and that PERFIELD is the fastest classifier, followed by LARSYSDC and LARSYSAA in that order.

CLOSURE

Although only two examples have been presented numerous other classifications have been performed on similar data and the results generally support the results presented. Even considering only the classification discussed the volume of data involved is quite substantial and is certainly adequate for a reasonable test.

For the type of data considered two basic conclusions appear reasonable.
(1) The classification accuracy of a nonparametric minimum distance classifiers, utilizing density histograms for estimating pdf's, is on the average not any larger than the classification accuracy of the parametric (Gaussian) classifier based on parametrically estimated pdf's. The variability in performance of the nonparametric classifier appears somewhat smaller. Since the parametric classifier requires less storage and is faster than the nonparametric classifier the latter classifier is not an attractive alternative.
(2) The average sample classification accuracy of a parametric (Gaussian) minimum distance classifier is larger than the average vector classification accuracy of a maximum likelihood vector classifier. Ignoring the problem of sample definition the minimum distance classifier is faster and is an attractive alternative to the maximum likelihood classifier in situations where it can be utilized.

The disparity between test and training results for both minimum distance and maximum likelihood classifiers is much greater than the differences due to classifier type or the specific implementation. This suggests that given the present state of the art greater improvement in classification accuracies will probably result from investigations intended to improve the training procedure than from investigation of classifier types.

ACKNOWLEDGEMENT

The work reported in this paper was supported by the National Aeronautics and Space Administration of the United States under Grant NGR-15-005-112.

REFERENCES

1. D.A. Landgrebe, "Systems Approach to the Use of Remote Sensing", LARS Information Note 041571, Purdue University, Lafayette, Indiana, April, 1971.

2. A.G. Wacker and D.A. Landgrebe, "Boundaries in Multispectral Imagery by Clustering," 1970 IEEE Symposium on Adaptive Processes (9th) Decision and Control, pp. X14.1-X14.8, December 1970.

3. R.L. Kuehn, E.R. Omberg and G.D. Forry, "Processing of Images Transmitted from Observation Satellites," Information Display, Vol.8, No.5, pp. 13-17, September/October 1971.

4. A.G. Wacker, "Minimum Distance Approach to Classification,", Ph.D. Thesis, Purdue University, Lafayette, Indiana, January 1972. Also available as LARS Information Note. 100771, Purdue University, Lafayette, Indiana, October 1971.

5. Z.W. Birbaum, "Distribution Free Tests of Fit for Continuous Distribution Functions," Ann. Math. Stat., Vol.24, pp. 1-8, 1953.

6. E. Samuel and R. Bachi, "Measures of Distances of Distribution Functions and Some Applications," Metron, Vol.23, pp. 83-122, December 1964.

7. H. Cramer, "On the Composition of Elementary Errors," Skand. Aktuarietids, Vol.11, pp. 13-74 and 141-180, 1928.

8. R. Von Mises, "Wahrscheinlichkeitsrechnung," Leipzig-Wien, 1931.

9. D.A. Darling, "The Kolmogorov-Smirnov, Cramer-Von Mises Tests," Ann. Math. Stat., Vol.28, pp. 823-838, December 1957.

10. W. Sahler, "A Survey on Distribution-Free Statistics Based on Distances Between Distribution Functions," Metrika,

Vol.13, pp. 149-169, 1968.

11. A.N. Kolmogorov, "Sulla Determinazione Empirica Di Une Legge Di Distribuzione," Giorn, dell'Insit. degli att., Vol.4, pp. 83-91, 1933.

12. N.V. Smirnov, "On the Estimation of the Discrepancy Between Empirical Curves of Distribution for Two Independent Samples," Bull. Math. Univ. Moscow, Vol.2, pp. 3-14, 1939.

13. H. Jeffreys, "An Invariant for the Prior Probability in Estimation Problems," Proc. Roy. Soc. A., Vol.186, pp. 454-461, 1946.

14. H. Jeffreys, "Theory of Probability," Oxford University Press, 1948.

15. T. Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," IEEE Trans. on Comm. Tech., Vol. COM-15, pp. 52-60, February, 1967.

16. A. Bhattacharyya, "On a Measure of Divergence Between Two Statistical Populations Defined by Their Probability Distributions," Bull. Calcutta Math. Soc., Vol.35, pp. 99-109, 1943.

17. K. Matusita, "On the Theory of Statistical Decision Functions," Ann Instit. Stat. Math. (Tokyo), Vol.3, pp. 17-35, 1951.

18. B.P. Adhikari and D.D. Joshi, "Distance Discrimination et Resume Exhaustif," Pbls, Inst. Stat., Vol.5, pp. 57-74, 1956.

19. C.H. Kraft, "Some Conditions for Consistency and Uniform Consistency of Statistical Procedures," University of California Publications in Statistics, 1955.

20. S. Kullback and R.A. Leiber, "On Information and Sufficiency," Ann. Math. Stat., Vol.22, pp. 79-86, 1951.

21. P.H. Swain and K.S. Fu, "Nonparametric and Linguistic Approaches to Pattern Recognition," LARS Information Note 051970, Purdue University, Lafayette, Indiana, June 1970.

22. P.C. Mahalanobis, "Analysis of Race Mixture in Bengal," J. Asiat. Soc. (India), Vol.23, pp. 301-310, 1925.

23. P.C. Mahalonobis, "On the Generalized Distance in Statistics," Proc. Nat'l. Inst. Sci. (India), Vol.12, pp. 49-55, 1936.

24. J. Keifer and J. Wolfowitz, "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters," Ann. Math. Stat., Vol.27, pp. 887-906, 1956.

25. S.M. Ali and S.D. Silvey, "A General Class of Coefficients of Divergence of one Distribution From Another," J. Roy. Stat. Soc., Ser. B, Vol.28, pp. 131-142, 1966.

26. D.G. Lainiotis, "On a General Relationship Between Estimation, Detection, and the Bhattacharyya Coefficient," IEEE Trans. on Information Theory, Vol.IT-15, pp. 504-505, July 1969.

27. C. Stein, "Approximations of Improper Prior Probability Measures," Dept. of Statistics, Stanford University, Stanford, California, Tech. Report 12, 1964.

28. K. Matusita, "On Theory of Statistical Decision Functions," Ann. Inst. Math. (Tokyo), Vol.3, pp. 17-35, 1951.

29. K. Matusita, "On Estimation by the Minimum Distance Method," Ann. Inst.Stat. Math. (Tokyo), Vol.5, pp. 59-65, 1954.

30. K. Matusita, Y. Suzuki, and H. Hudimoto, "On Testing Statistical Hypothesis," Ann. Inst. Stat. Math. (Tokyo), Vol. 6, pp. 133-141, 1954.

31. K. Matusita and H. Akaike, "Decision Rules Based on the Distance for the Problems of Independence Invariance and Two Samples," Ann. Inst. Stat. Math., Vol.7, pp. 67-80, 1956.

32. K. Matusita and M. Motoo, "On the Fundamental Theorem for the Decision Rule Based on Distance || ||," Ann. Inst. Stat. Math., Vol.7, pp. 137-142, 1956.

33. K. Matusita, "Decision Rule Based on the Distance for the Classification Problem," Ann. Inst. Stat. Math. (Tokoyo), Vol.8, pp. 67-70, 1956.

[34]. K. Matusita, "Distance and Decision Rules," Ann. Inst. Stat. Math. (Tokyo), Vol.5, pp. 59-65, 1954.

[35]. K. Matusita, "Classification Based on Distance in Multivariate Gaussian Case," Proc. 5th Berkeley Symposium on Math. Stat. and Prob., Vol.1, pp. 299-304, 1967.

[36]. J. Wolfowitz, "Consistent Estimations of the Parameters in a Linear Structural Relationship," Skand. Aktuarietids, pp. 132-151, 1952.

[37]. J. Wolfowitz, "Estimation by the Minimum Distance Method," Ann. Inst. Stat. Math. (Tokyo), Vol.5, pp. 9-23, 1953.

[38]. J. Wolfowitz, "Estimation by the Minimum Distance Method in Nonparametric Difference Equations," Ann. Math. Stat., Vol.25, pp. 203-217, 1954.

[39]. J. Wolfowitz, "The Minimum Distance Method," Ann. Math. Stat. Vol.28, pp. 75-88, 1957.

[40]. S. Das-Gupta, "Nonparametric Classification Rules," Sankhya, Indian Jour. of Stat., Series A, Vol.26, pp. 4-30, 1964.

[41]. T. Cacoullos, "Comparing Mahalanobis Distance I: Comparing Distances between Known Populations and Another Unknown," Sankhya, Indian Jour. Stat., Series A, Vol.27, pp. 1-22, March 1965.

[42]. T. Cacoullos, "Comparing Mahalanobis Distances II: Bayes Procedures When the Mean Vector are Unknown," Sankhya, Indian Jour. Stat., Series A, Vol.27, pp. 23-32, March 1965.

[43]. M.S. Srivastava, "Comparing Distances Between Multivariate Populations - The Problem of Minimum Distance," Ann. Math. Stat., Vol.38, pp. 550-556, April 1967.

[44]. W. Hoeffding and J. Wolfowitz, "Distinguisability of Sets of Distributions," Ann. Math. Stat., Vol.29, pp. 700-718, September 1958.

[45]. K.S. Fu, D.A. Landgrebe, and T.L. Phillips, "Information Processing of Remotely Sensed Agricultural Data," Proc. IEEE, Vol.57, pp. 639-654, April 1969.

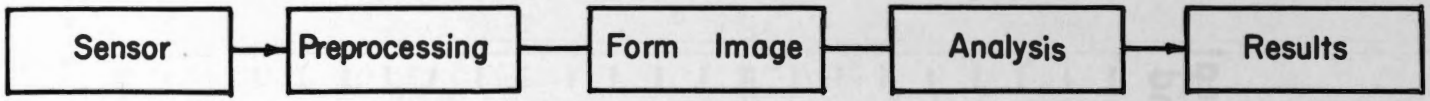
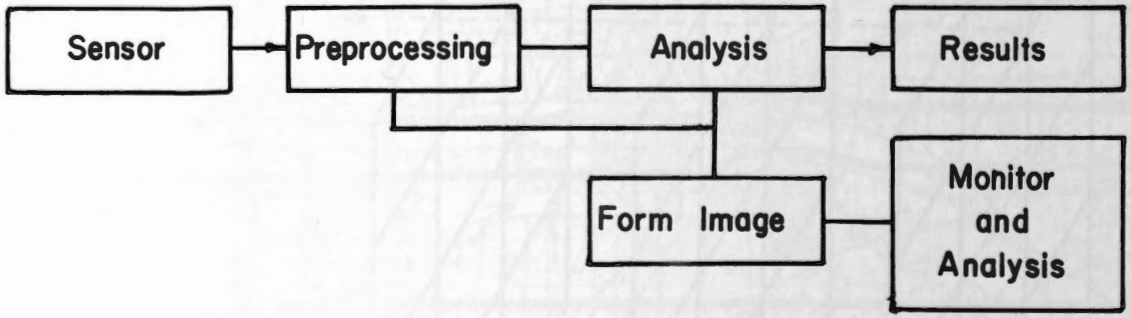


Image Oriented



Numerically Oriented

Fig. 1 Organization of Image and Numerically Oriented Systems

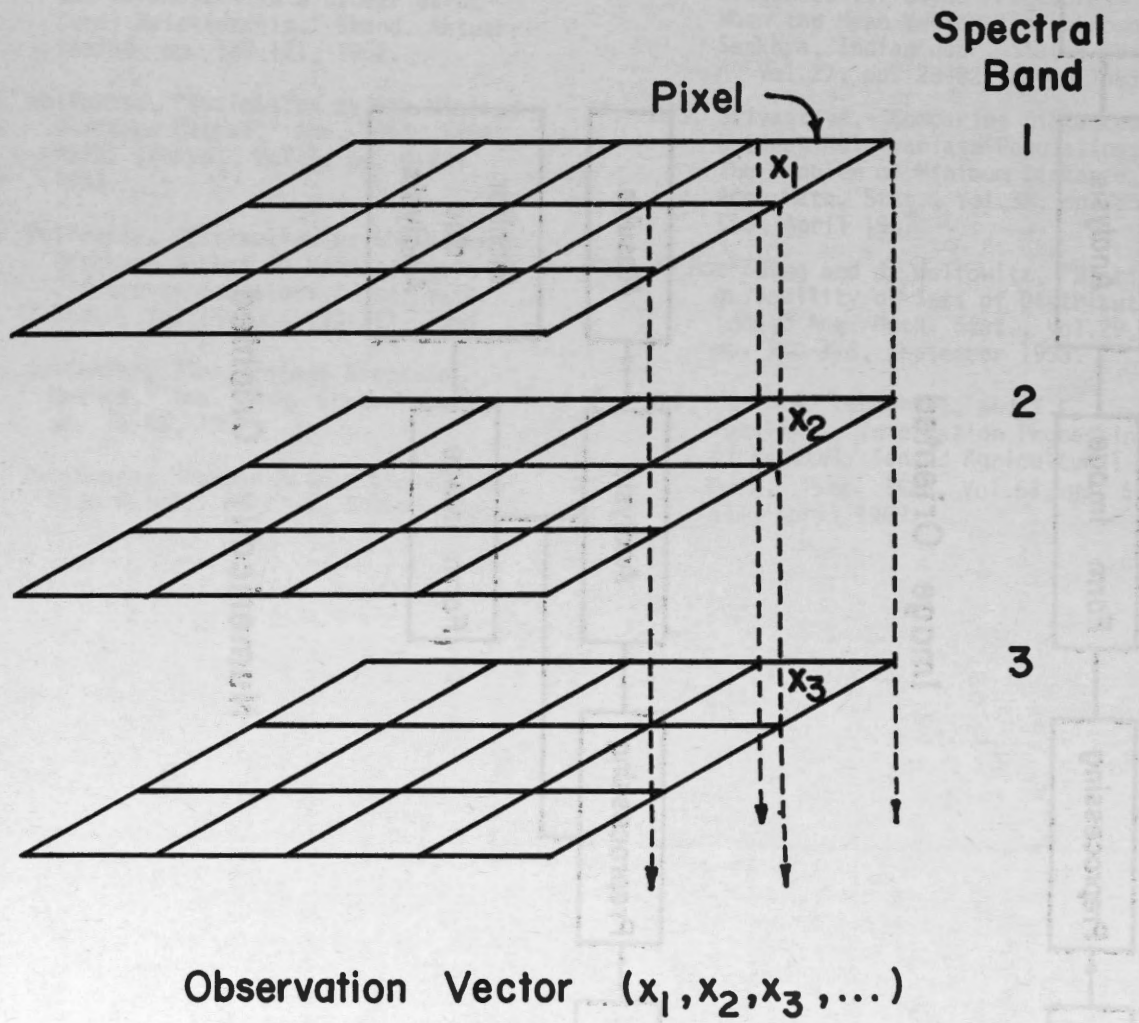


Fig. 2 Formation of Multispectral Data-Image

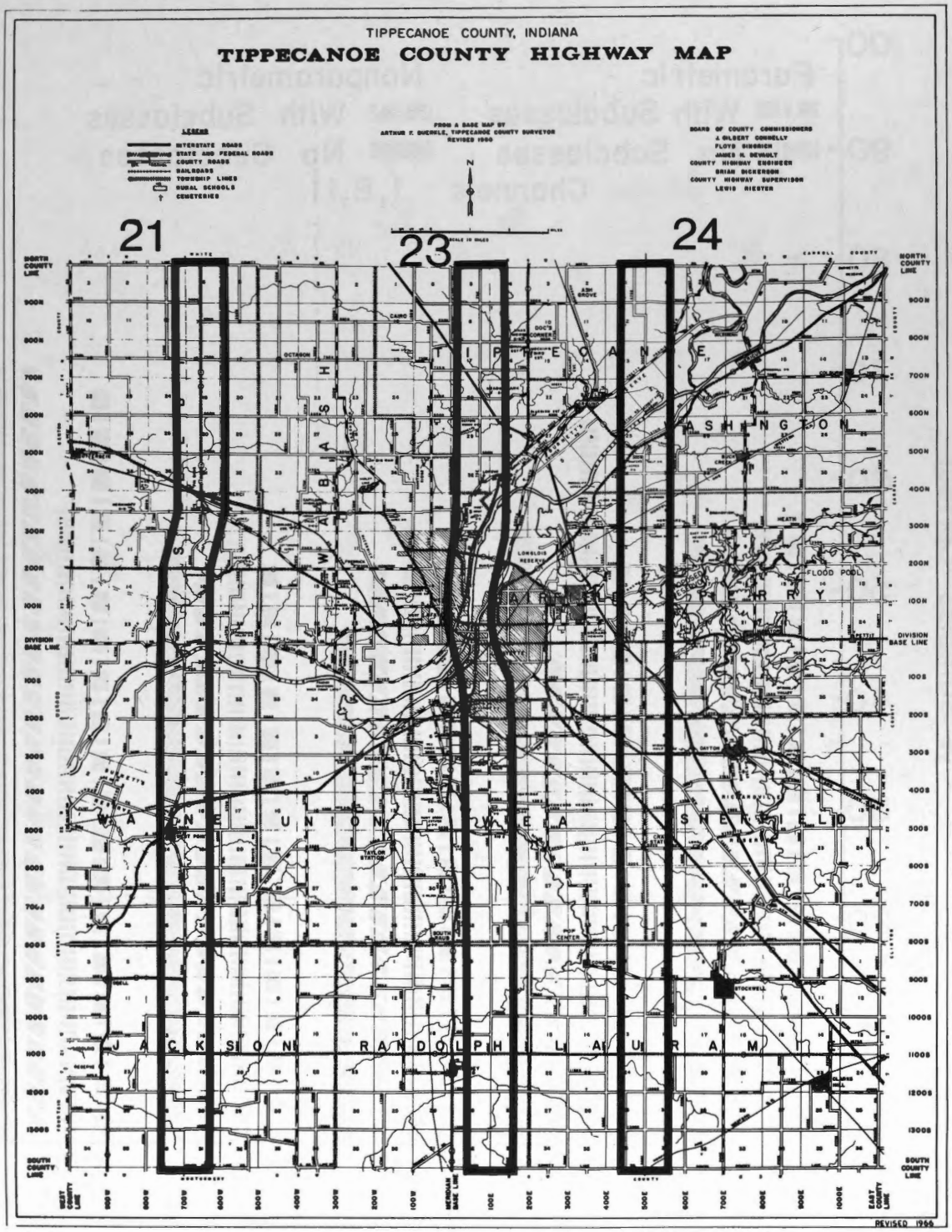


Fig. 3 Location of Tippecanoe County Flight-lines 21, 23 and 24

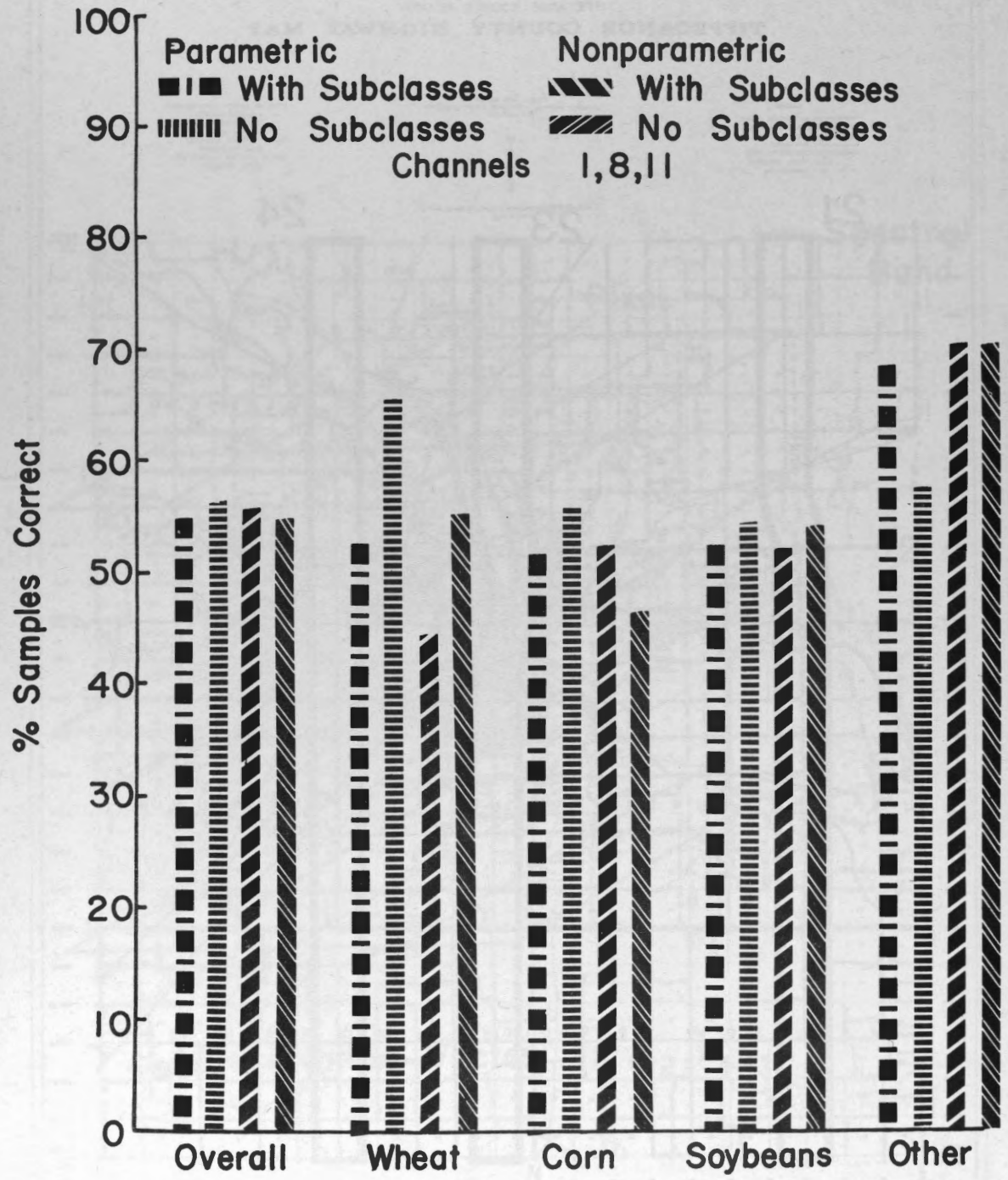


Fig. 4 Comparison of Average Test Performance for Parametric and Nonparametric Minimum Distance Classification Using Bhattacharyya Distance and Random Training

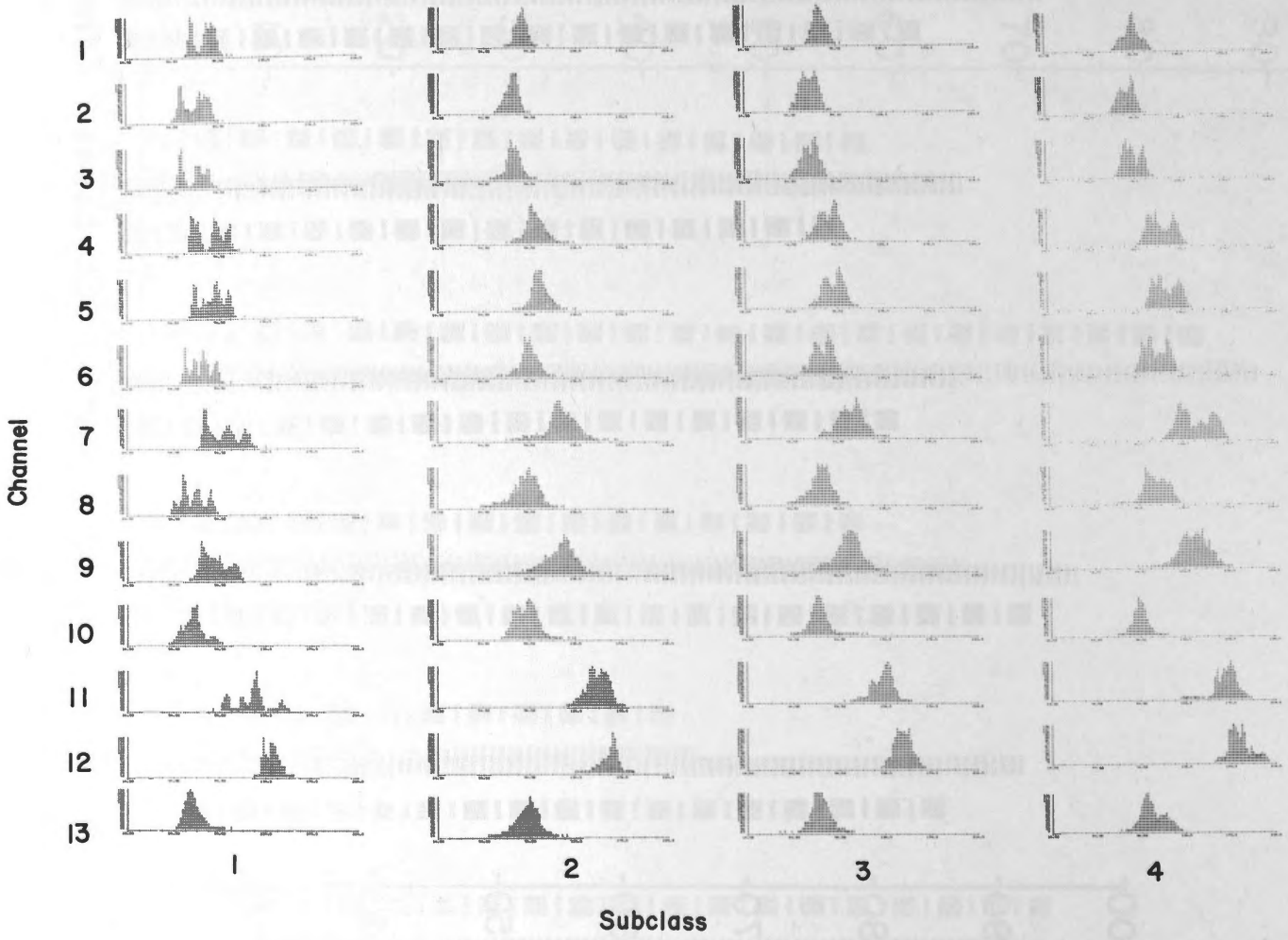


Fig. 5 Histograms for Wheat Subclasses Obtained as Result of Clustering Wheat Acres

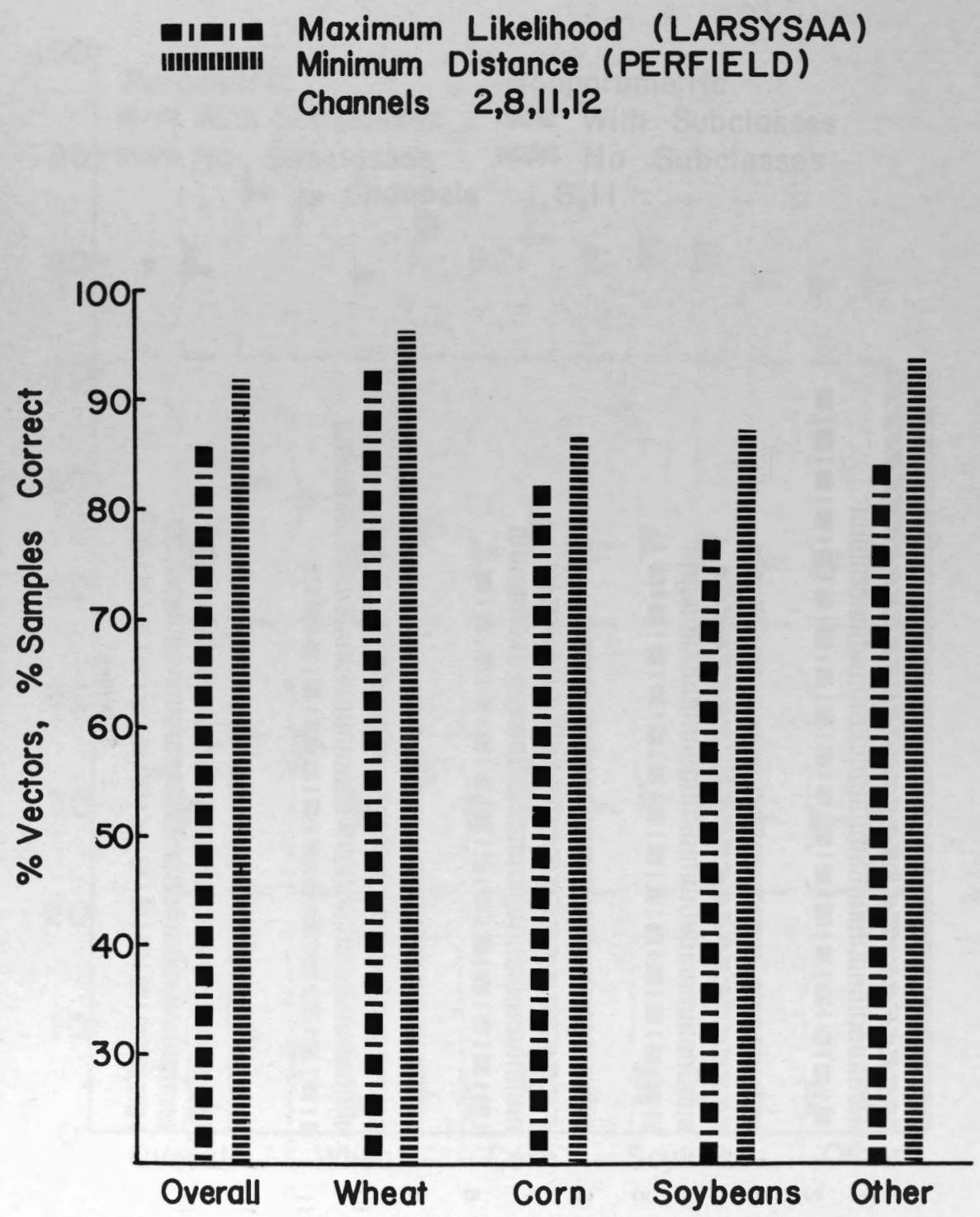


Fig. 6 Comparison of the Training Performance for Minimum Distance and Maximum Likelihood Classification

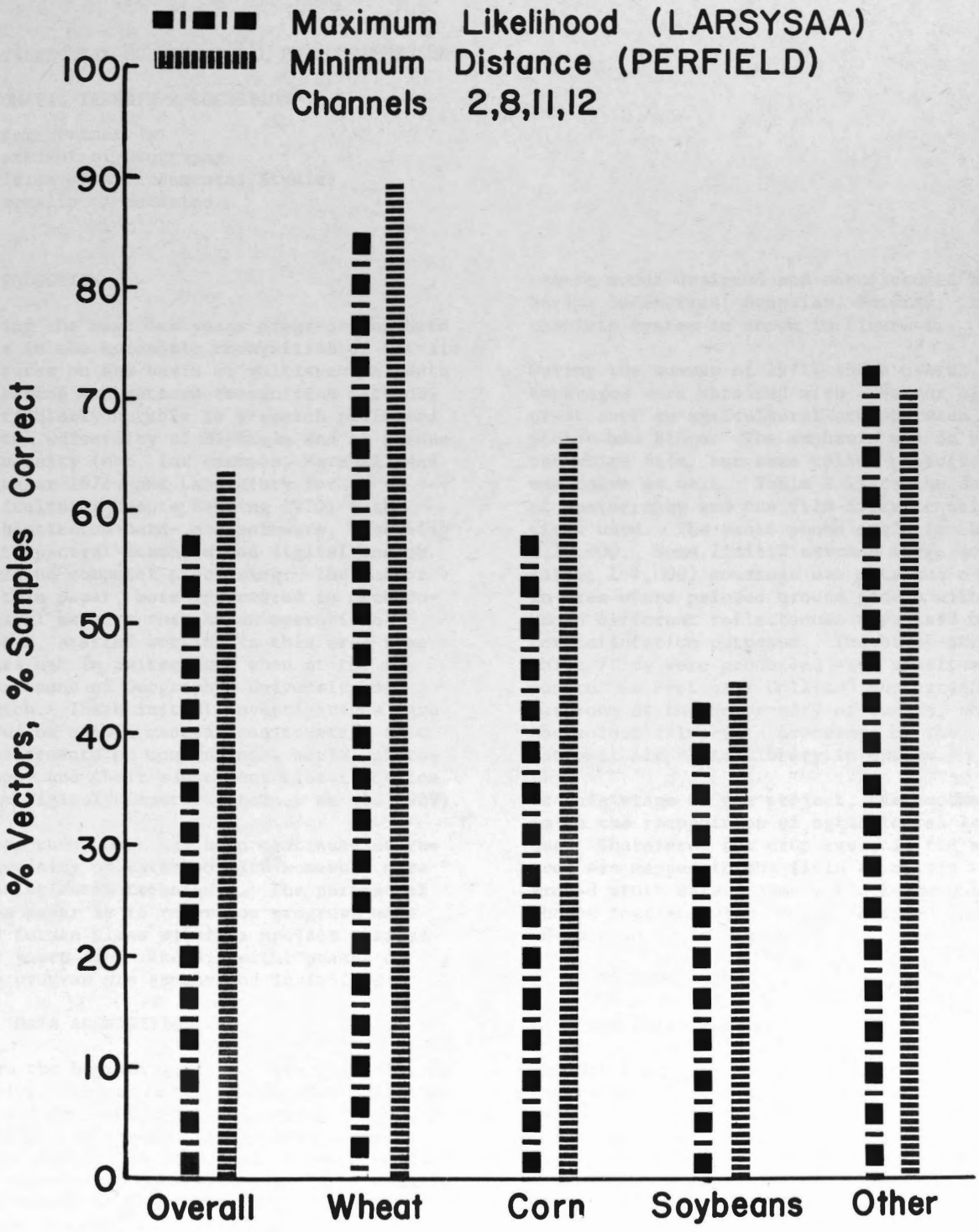


Fig. 7 Comparison of Average Test Performance of Minimum Distance and Maximum Likelihood Classification

MULTISPECTRAL-MULTITEMPORAL PHOTOGRAPHY AND AUTOMATIC TERRAIN RECOGNITION

Dieter Steiner
Department of Geography
Division of Environmental Studies
University of Waterloo

INTRODUCTION

During the past few years progress has been made in the automatic recognition of terrain features on the basis of multispectral data gathering and pattern recognition methods. Particularly notable is research performed at the University of Michigan and at Purdue University (see, for example, Marshall and Kriegler 1971, and Laboratory for Agricultural Remote Sensing 1970) with sophisticated hard- and software, involving multispectral scanners and digital and/or analogue computer processing. The author of this paper, more interested in methodological aspects than in an operational system, started working in this area some years ago in Switzerland when at the Department of Geography, University of Zurich. These initial investigations were based on simple manual densitometric spot measurements on conventional aerial photography and their subsequent classification on a digital computer (Steiner et al. 1969).

Since then, work has been continued at the University of Waterloo with somewhat more sophisticated techniques. The purpose of this paper is to report on progress made and future plans within a project started two years ago. The different phases of the program are summarized in Table 1.

1. DATA ACQUISITION

From the beginning, it was the intention to explore the automatic recognition potential within the photographic spectral range. To this end an experimental multispectral camera setup was developed, drawing partly on experiences made with similar systems at the School of Forestry, University of Minnesota, St. Paul (Ulliman et al. 1970), and at the School of Planning, University of Tennessee, Knoxville (Haddox 1968). It consists of four electric Hasselblad 500 EL 70 mm cameras with 80 mm Zeiss Planar lenses, a timer which permits simultaneous firing of all cameras at regular intervals, a viewer for operational control and a camera mount designed and manufactured by Strite Industries, Hespeler, Ontario. The complete system is shown in Figure 1.

During the summer of 1971, three useful coverages were obtained with a Beaver aircraft over an agricultural area between Guelph and Elora. The emphasis was on black-and-white film, but some colour photography was taken as well. Table 2 lists the dates of photography and the film-filter combinations used. The basic photo scale is about 1:30,000. Some limited extreme large scale (about 1:4,000) coverage was obtained over an area where painted ground panels with known different reflectances were laid out for calibration purposes. The black-and-white films were processed with sensitometric control in Professor Collins' photographic darkroom at the University of Guelph, whereas the colour films were processed by the National Air Photo Library in Ottawa.

At this stage of the project, the emphasis is on the recognition of agricultural land use. Therefore, the crop cover in the study area was mapped in the field to obtain ground truth data. Sample black-and-white photos together with ground observed land use are shown in Figure 2.

2. DATA PROCESSING

2a. Film Digitization

The first data processing step is film digitization. For this purpose an Optronics Photoscan is being employed. This is a high-speed drum scanner which digitizes a transparency to 8-bit accuracy (i.e., 256 discrete gray levels within a density range of 0.0-2.0 or 0.0-3.0) with a spatial resolution of 50, 100 or 200 μ. The measurements are logged on magnetic tape. With the coarsest resolution, a 70 mm frame is digitized to about 90,000 picture elements, and the instrument performs this operation in about one minute. So far, only a few sample frames have been digitized, and no analysis of any kind has been carried out as yet.

The following then is a brief account of plans for future work.

2b. Image Compression

One problem to be considered before any actual data processing can be carried out is the compaction of the image data to alleviate the enormous storage requirements for digital processing. According to Rosenfeld (1969) three methods can be considered and used singly or in combination. These are encoding, sampling and quantization. Efficient encoding may take any of the following forms:

1) Short codes can be used to represent gray levels that occur often, longer codes to represent gray levels that occur infrequently.

2) If blocks of constant gray levels exist the image may be described in terms of sizes of such blocks.

3) If successive gray levels are interdependent, previous levels can be used to predict the following ones, and only deviations from predictions need be encoded.

An acceptable approximation may be obtained by taking the gray levels only at a limited number of sample points. The gaps between these points can then be bridged over by polynomial or sinusiodal interpolation. Obviously, this requires calculations when, for processing purposes, the full image has to be reconstituted. In other words, there will be a tradeoff between storage and processing time requirements. Quantization approximates an image by reducing the original number of gray levels to a smaller one. The gray levels can be unevenly spaced, for example, such that they have about equal frequencies of picture elements, which permits to make most efficient use of a given number of levels.

Still another approach produces an image transform (for example, Fourier transform), and applies compaction techniques such as quantization or sampling in the transform domain. An image can then be reconstructed by applying the inverse transformation (see Silverman 1971).

2c. Image Registration

Another problem is of a geometrical nature: In order that multispectral and/or multitemporal image data may be compared and combined, the images in question have to be in proper register. It is believed that differential distortions due to the camera lenses operating at different wave-lengths (see Slater 1969) and also film dimensional changes are negligible with respect to the degree of resolution used (50 μ at the most). The only sources of geometric distortions to consider are then positional and attitude changes of the photographic platform (i.e., roll, pitch, yaw, altitude and horizontal position). Since photogrammetric accuracy is not intended at this stage, an attempt will be made to register images by simple scaling, translation and rotation, thus disregarding effects due to tilt. To this end the position of a number of clearly defined match points will be determined with a coordinate measuring instrument. Should it turn out that the results are unsatisfactory, then, rather than trying to match image frames as a whole, a fit will be attempted for smaller image regions. Such procedures are discussed in Linstedt (1971).

2d. Radiometric Corrections

Conventional photo interpretation as well as automated techniques of image analysis depend heavily on the quality of tonal information. Ideally, the tone of every resolution element should be a consistent function of the vertical upward reflection of radiation from the corresponding terrain patch. Factors which tend to distort the radiometric values on an image and to increase their variability within a frame or from image to image are the following:

1) Variable surface illumination (incl. shadows) due to gradient and orientation of terrain slopes;

2) Change of terrain radiance with view angle (see Steiner & Haefner 1965);

3) Atmospheric interference (attenuation and haze radiance) along the path from the ground to the camera;

4) Change of atmospheric effects with view angle (as the path through the atmosphere becomes longer);

5) Lens shading;

6) Differences in film gamma due to production variation, age of material and processing.

In order to eliminate these tone variation increasing effects and, consequently, to enhance the chances for a subsequent automatic recognition of terrain features, the introduction of radiometric corrections will have to be considered. The calibrated ground targets mentioned earlier will serve to make adjustments for variable atmospheric effects. Changes in film gamma can be taken care of by processing the material with sensitometric control.

More difficult is to correct for tonal changes as a function of within frame position (factors 2, 4 and 5 above). One possible approach to solving the problem is an analytical one whereby the influence of each component is determined separately first and the combined effect is derived subsequently. However, the necessary data are not readily available. Particularly difficult in this respect is the change of atmospheric interference as a function of view angle. It is very much dependent on the weather situation and the degree of atmospheric pollution at the time of photography and, consequently, available models for an "average" atmosphere are of little use. It is therefore proposed to determine the overall combined effect on a purely empirical basis, i.e., by trend surface analysis carried out on gray tones averaged from several images. The thus derived functions will then allow to apply appropriate corrections. An alternative approach which would reduce the necessary processing time is a table lookup procedure (see Bakis et al. 1971).

The most bothersome of these factors is terrain slope. Its effect cannot be removed by simple means. For the beginning a study area has been chosen with terrain that is not heavily accentuated so that the problem will be minimal and a correction unnecessary. Future investigations should consider this tone distorting factor as well, however. It is possible to determine slope automatically by means of optical-electronic correlation of two photos forming a stereo pair, a procedure already in use in photogrammetry (Konecny and Refroy 1968). Such a system would, of course, require photogrammetric control. The present project will have to resort to a simpler approach. The most promising would seem to be the comparative use of a digitized terrain model obtained from a topographic map of the same area. Such models have been used to automatically generate hill shading for topographic mapping (see, for example, Yoeli 1966). It is obvious that they could also be employed to achieve the opposite, i.e., to remove "hill shading" from aerial photographs. However, the geometrical fitting of photographic images to terrain models will be an additional problem.

2e. Pattern Recognition

The previous operations can be regarded as preprocessing steps. When appropriate procedures have been found for them it will then be possible to proceed with an attempt to recognize terrain cover types portrayed on the photos automatically by combining multispectral and/or multitemporal information. It is obvious though that different preprocessing methods may affect the quality of the results in the recognition phase. Consequently, it may be necessary to go through the whole processing sequence several times and to experiment with different combinations.

Automatic classification techniques fall under the general heading of pattern recognition. A pattern, according to pattern recognition terminology, is a vector of measurements (called features) taken, in the case of remote sensing, on the imaged terrain area to be classified. The situation can be geometrically visualized with each terrain patch having a point location in the multidimensional feature space. The goal of pattern recognition is then to subdivide this space by linear or curved boundaries in such a manner that each ensuing region is associated with a particular class of terrain cover type. When these regions are defined, new observations can then be allocated to classes according to their location in the feature space. The usual procedure is to establish these regions on the basis of samples taken in a key or training area, to evaluate the resulting classification with respect to recognition accuracy, and then to further appraise the quality of the procedure by applying it to test areas outside the original training area. A number of different pattern recognition methods are available and will be experimented with in this project. To review them goes beyond the scope of this paper. They are discussed in recent papers by Fu (1971) and Steiner (1970).

The features used as input to the pattern recognition phase are tonal and/or textural measurements. The simplest approach to automatic terrain recognition consists of taking all the tonal (i.e., multispectral and/or multitemporal) information available

for each successive resolution element and of determining its class membership. A more elaborate technique tries to establish a classification on a per-field basis. This means that an image has to be subdivided first into homogeneous regions (agricultural fields for the purposes of this project). A classification of individual fields is then arrived at on the basis of all resolution elements lying within this field, i.e., a vote is taken according to observed class membership frequencies. This approach is intuitively appealing because it is closer to the procedure used in conventional interpretation and better results in terms of accuracy can be expected. The automatic recognition of field boundaries by means of clustering or edge detection shows some promise (see Anuta 1970), but the problems involved are fairly complex. An easier, though not automatic, procedure is to determine these boundaries on the photo by visual interpretation, to digitize them and use them as a base for subsequent processing (Johnson et al. 1969).

Once the potential of using tonal information for pattern recognition has been evaluated, it is intended to experiment with textural information as well. This is somewhat more complex. The problem is to find adequate parameters which describe local textural properties. There are two principal approaches possible:

1) Analysis in the spatial domain by deriving various statistical measures from the gray levels of the picture elements within the image region in question. An example are the various measures proposed by Maurer (1971).

2) Analysis in the frequency domain. This requires the calculation of Fourier transforms for the image regions under scrutiny. An attempt can then be made to derive significant parameters from these transforms. (Compare with the use of optically produced Fourier transforms as reported, for example, by Palgen 1970). Fourier transformation using digital techniques poses a problem with respect to processing time, but it can be minimized by resorting to a fast algorithm developed recently (Fisher 1970).

3. DATA DISPLAY

The classification results obtained during the pattern recognition phase will be displayed by computer mapping techniques. The use of the line printer for this purpose will be given first priority, because it allows translating the image raster directly into a decision output raster. The necessary software is available at this Department. The use of a XY-plotter for data display will be considered at a later stage.

Computer mapping techniques will also be used in an intermediate phase to analyse radiometric distortions within the field of view and to check on the quality of image registration.

ACKNOWLEDGMENTS

Financial support provided for this project by the National Research Council of Canada is gratefully acknowledged. Thanks are extended to the Ontario Department of Lands and Forests for making a Beaver aircraft available for the aerial photography phase, to Mr. Victor Zsilinsky of the same Department for valuable advice, to Professor Stanley H. Collins and Ron Shillum, School of Engineering, University of Guelph for the processing of the films under sensitometric control, and to Professor John Starkey, Department of Geology, University of Western Ontario, for the possibility of using his microdensitometric drum scanner (Optronics Photoscan). Finally, the following individuals, who were employed during various phases of the project up to now and made valuable contributions, should be mentioned: Robert A. Ryerson, Howard Turner and Gordon E. Dittmer, all graduate students at this Department, Matthias Winiger, graduate student and instructor from the University of Bern, Switzerland, and Otto F. Matt, computer programmer, now at the Department of Geography, University of Zurich.

REFERENCES

1) P.E. Anuta. Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques. IEEE Trans. on Geoscience Electronics, Vol. GE-8 (4): 353-368, 1970.

2) R. Bakis, M.A. Wesley and P.M. Will. Digital Correction of Geometric and Radiometric Errors in ERTS Data. Proc. 7th Int. Symposium on Remote Sensing of Environment, Vol. II: 1427-1436, 1971.

3) J.R. Fisher. Fortran Program for Fast Fourier Transform. Naval Research Lab. Rep. 7041, 20 p. Washington, D.C., 1970.

4) K.S. Fu. On the Application of Pattern Recognition Techniques to Remote Sensing Problems. Purdue Univ., School of El. Engin. Techn. Rep. TR-EE 71-13, 91 p., 1971.

5) J.T. Haddox. The Hasselblad Camera System in Aerial Reconnaissance. NASA Techn. Letter 132 (Mission 73, Summary and Data Catalog): 65-68, Houston, Texas, 1968.

6) C.W. Johnson, L.W. Bowden and R.W. Pease. A system of Regional Agricultural Land Use Mapping Tested against Small Scale Apollo 9 Color Infrared Photography of the Imperial Valley (California). Studies in Remote Sensing of S. California and Related Environments. Techn. Rep. V, 96 p., Dept. of Geography, Univ. of California, Riverside, 1969.

7) G. Konecny and D. H. Refroy. Maps from Digitized Stereomat data. Photogramm. Engin. 34(1): 83-90, 1968.

8) Laboratory for Agricultural Remote Sensing. Remote Multispectral Sensing in Agriculture. Annual Rep., Vol. 4, 112 p., Purdue University, Lafayette, Ind., 1970.

9) C. Linstedt. Preliminary Experiments on Geometric Transformations of Images. IBM Research RC 3314, 42 p., Yorktown Heights, N.Y., 1971.

10) R.E. Marshall and F. Kriegler. An Operational Multispectral Surveys System. Proc. 7th Int. Symposium on Remote Sensing of Environment, Vol. III: 2169-2191, 1971.

11) H. Maurer. Measurement of Textures of Crop Fields with the Zeiss Scanning Microscope-Photometer 05. Proc. 7th Int. Symposium on Remote Sensing of Environment, Vol. III: 2329-2334, 1971.

12) J.J.O. Palgen. Applicability of Pattern Recognition Techniques to the Analysis of Urban Quality from Satellites. Pattern Recognition 2: 255-260, 1970.

13) A. Rosenfeld. Picture Processing by Computer. 196 p., Acad. Press, New York, 1969.

14) H.F. Silverman. An Optimizing Fourier Domain Compaction Method for ERTS Data. IBM Research RC 3277, 24 p., Yorktown Heights, N.Y., 1971.

15) P.N. Slater. Multiband Camera Monograph. Optical Sciences Center Techn. Rep. 44, 74 p., Univ. of Arizona, Tucson, 1969.

16) D. Steiner. Automation in Photo Interpretation. Geoforum 1 (2): 75-88, 1970.

17) D. Steiner and H. Haefner. Tone Distortion for Automated Photo Interpretation. Photogramm. Engin. 31 (2): 269-280, 1965.

18) D. Steiner, K. Baumberger and H. Maurer. Computer-Processing and Classification of Multi-Variate Information from Remote Sensing Imagery. Proc. 6th Int. Symposium on Remote Sensing of Environment, Vol. II: 895-907, 1969.

19) J.J. Ulliman, R.P. Latham and M.P. Meyer. 70-mm Quadricamera System. Photogramm. Engin. 36 (1): 49-54, 1970.

20) P. Yoeli. Die Mechanisierung der analytischen Schattierung (The mechanisation of analytical hill shading). Kartogr. Nachrichten 16 (5): 103-107, 1966.

Table 1: Phases of the Waterloo project on automatic terrain recognition on the basis of multispectral and multitemporal photography

Phases	Purpose
DATA ACQUISITION	
1. Photographic Flights	To obtain black-and-white multispectral coverage at different times of the year with calibrated reflectance ground targets and supplementary colour photography
2. Film Processing	To process black-and-white films with sensitometric control
DATA PROCESSING	
1. Film Digitization	To digitize film frames with drum scanner into machine-readable format
2. Image Compression	To compact digitized image data with a view on computer storage requirements
3. Image Registration	To bring multispectral and/or multitemporal image data into geometrical register so that they can be overlayed
4. Radiometric Corrections	To decrease the variability of photographic tones by correcting for radiometric distortions (lens shading, view angle effects, film gamma changes, etc.)
5. Pattern Recognition	To combine the overlayed image data numerically into decision functions which allow a classification of terrain patches as to the nature of their cover
DATA DISPLAY	To map the recognition results using computer output devices (line printer and plotter)

Table 2: Coverage obtained with multispectral camera setup over Guelph-Elora area, Ontario, 1971

Film	Filter(s)	Dates		
		June 15	July 7	August 12
Kodak Tri-X Pan	Wratten 47B (Blue)	x	x	x
id.	Wratten 58 (Green)	x	x	x
id.	Wratten 29 (Red)	x	x	
Kodak Infrared Aerographic	Wratten 87 (IR)	x	x	x
Kodak Ektachrome	Wratten HF3 + HF4 (Haze-cutting)		x	x
Kodak Aerochrome Infrared (2443)	Wratten 12 (Yellow)		x	x

Fig. 1. Multispectral camera system with four Hasselblads and viewer in the middle. It fits into a regular camera mount.

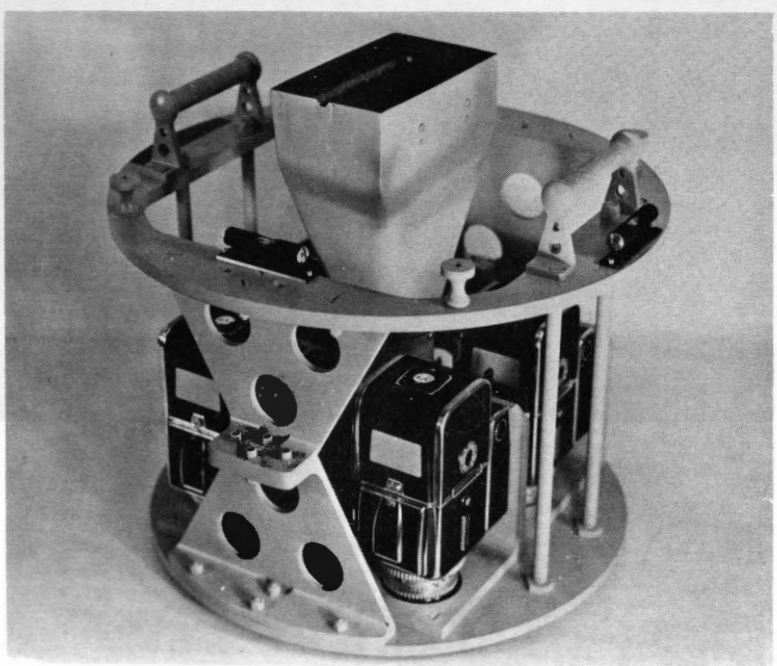


Fig. 2a. Sample photos of coverage obtained on July 7, 1971. Panchromatic/ blue filter image is in the upper left, panchromatic/green filter in the upper right, panchromatic/ red filter in the lower left, and infrared in the lower right hand corner. A frame side corresponds roughly to 1 mile on the ground. Area shown is on Swan Creek, southeast of Elora, Ontario.

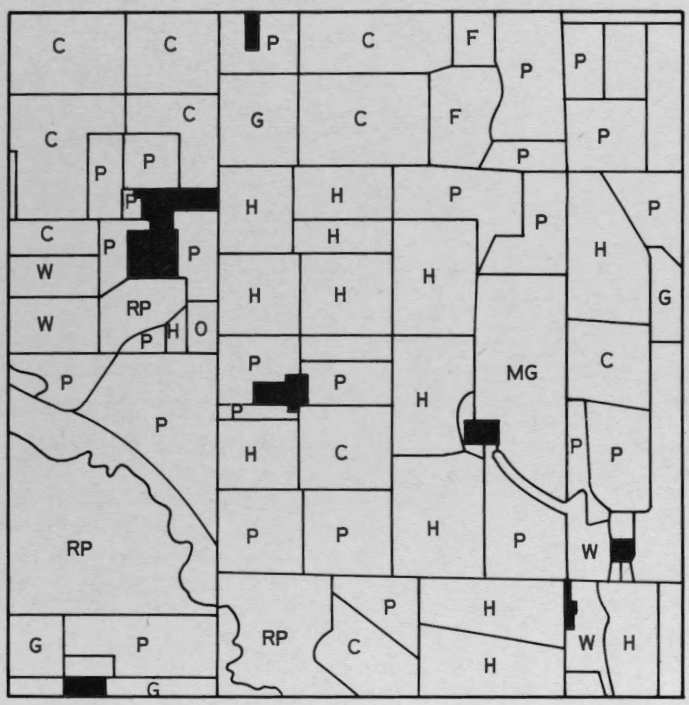


Fig. 2b. Land use sketch map.

G = small grains, undifferentiated
MG = mixed grains
W = wheat
C = corn
H = hay
P = improved pasture
RP = rough pasture
F = fallow
W = woodland

MICROWAVE RADIOMETRY FOR REMOTE SENSING FROM AIRCRAFT AND SPACECRAFT

A.W. Adey,
Communications Research Centre,
Ottawa.

ABSTRACT

The primary aim of this paper is to draw to the attention of those concerned with surveys of the earth and near-earth environment, the potential of the microwave radiometer in the role of a remote-sensing device. It is based on a more extensive review which is being published as a CRC Report*, preprints of which have been made available to attendees at the symposium. This report identifies a number of major application areas and provides information on some primary sources of published literature. The referenced review, while not exhaustive in depth or scope of coverage, should still provide a useful statement of the status of an important and rapidly-developing area of remote-sensing activity. It could also serve as a basis for a more comprehensive study.

The report first presents a short discussion of some basic principles of radiometry, followed by a number of general radiometer design factors based on these principles. The report continues with a review of the status of current programmes of airborne and satellite remote-sensing based on microwave radiometers. It provides details of device capabilities and limitations and outlines some research areas and problems. Studies involving penetration of sub-surface materials are noted. The final section of the report comprises a bibliography. In the choice of the references the emphasis has been on details of programmes, applications, operational factors and results, rather than on specific hardware.

INTRODUCTION

A preliminary review of the status of the remote-sensing area covered by the title of the paper has been prepared and is being published as CRC Report No. 1231, dated February 1972. Preprints have been made available to those attending this symposium.

* Microwave Radiometry for Surveillance from Spacecraft and Aircraft. CRC Report No. 1231, February 1972.

For this paper some examples have been selected, for the most part from the works cited in the bibliography of the report, to illustrate (a) a number of the environmental factors on which the technique is based and which must be considered as facts of life in designing an experiment and interpreting the data; (b) some instrumentation and operational details of the USSR and USA programmes; and (c) a few results to illustrate possible applications of the technique.

RADIOMETER OPERATION

The technique is based on the fact that terrestrial materials have a wide range of emission factors at microwave frequencies, this factor being related to the electromagnetic properties, the internal structure and the boundary geometry. One attempts, on the basis of some prior knowledge of the probable magnitude of the emission factor of various terrestrial bodies and structures, to interpret the level of signal received (generally expressed in degrees Kelvin) in terms of the class of radiator (e.g. ship, water, ice, precipitation cell, etc.) or of some parameter of a particular class (e.g. the salinity of water or the vertical temperature profile of the atmosphere). Some nominal terrestrial emission factors are listed in Table I. One can expect to find essentially the complete range, from close to zero (a perfect reflector) up to unity (a black body).

The total radiation received by the radiometer will be made up of two components - the radiation or emission of the objects or terrain under study, and that from a number of types of spurious or interfering sources. The latter can reach the radiometer either directly through back or side lobes, or indirectly by transmission or reflection, thus complicating an experiment aimed at measuring emission. The interference can be due to (a) cosmic or galactic noise, which increases steeply with decreasing frequency, amounting to the order of 100°K at 300 MHz and becoming negligible above 2 GHz; (b) radiation from extra-terrestrial radio sources such as the sun, some of which

have a variable output; (c) atmospheric radiation, which originates in gases such as oxygen and water vapour and in precipitation cells, clouds and fog; (d) radiation from areas of terrain outside the main beam and suitably situated and oriented; (e) manmade interference, varying with location and time.

Fig. 1 (Hogg, 1959) shows the level of the radiation to be expected, at a ground-based radiometer, from atmospheric oxygen and water vapour for a number of elevation angles. This was calculated for a water vapour content of 10 g/m^3 at ground level, falling to zero at 5 Km., and is representative of a summer day in temperate latitudes. The strong effect of frequency (particularly in the neighbourhood of the 22 GHz water vapour line) and elevation angle is evident.

Figure 2 (calculated from Stogryn, 1971) illustrates an important factor in radiometry. This is the relation between the electromagnetic properties of any radiating material and the depth in that material from which a significant amount of radiation can penetrate to the surface. The strong effect of salinity at the lower frequencies is evident, as is the fact that any radiation received from water, for frequencies above 2 GHz, originates in essentially the upper 1-2 cm.

Theoretical studies suggest that, for frequencies below approximately 2 GHz, the brightness temperature of water should be sensitive to the salinity, with this sensitivity increasing with decreasing frequency. The curves of Figure 3 (from Paris, 1971) illustrate the effect. A recent experiment carried out with a lowest frequency of 1.4 GHz tends to confirm the prediction (Droppleman and Mennela, 1970).

PROGRAMME DETAILS

Tables II-IV summarize the current radiometer programmes of the USSR and of the NASA in the United States. Other smaller-scale U.S. programmes are being conducted under the sponsorship of the Coast Guard (mainly oil pollution studies), the Defence Department and Industry. The U.S. aircraft radiometer studies involve operation in approximately twenty different wavelength bands within the limits noted in Table II.

EXAMPLES OF DATA

A few results have been selected from the publications cited in the bibliography of the C.R.C. Report No. 1231. The examples fall into three areas of remote-sensing application: ice reconnaissance, classification of land features, and detection and classification of targets in a water environment. The absence of a transmitter signal and the possibility of operating during darkness and considerable levels of precipitation are special advantages in many such applications.

The first two examples are in the area of ice reconnaissance. The data in Figure 4 (Basharinov, 1971), presumably from the Cosmos 243 satellite, illustrate a presentation technique for classification of ice types on the basis of multifrequency brightness temperatures. The example in Figure 5 (Adey et al., 1971 and 1972; Hartz, 1971) was selected from results of the initial field tests, off Resolute in February-March 1971, of a multifrequency, UHF radiometer being developed at the Communications Research Centre for application to the measurement of the thickness of sea ice. The recording shown was made during a flight at an altitude of 150 meters and at a speed of 140 knots or 260 Km/hour. The deflection of the trace on passing from ice to water and the recovery over the ice in the lead might well have been somewhat restricted by the combination of aircraft altitude and speed, antenna beamwidth, radiometer time constant and the width of the lead. The data emphasize the sharp radiometric contrast between sea ice and water, and thus demonstrate the potential use of the technique in the areas of transportation and ice reconnaissance.

Figure 6 (Richer, 1969) illustrates the brightness temperature of a vehicle as a function of the viewing aspect. The frequency was 35 GHz. Figure 7 (Richer, 1969) shows a scan taken at 35 GHz, which emphasizes the brightness temperature contrast between metal vehicles and wooded areas and the change in the level of radiometric signatures with distance. The exploitation by the microwave radiometric technique of the contrast between a metal ship and its water environment is illustrated in Figures 8-10 (Copeland, 1969), the frequency again being 35 GHz. The detail present in the barge signature of Figure 9 provides an example of the usefulness of the technique in target identification and classification. The analogy of the data of Figure 10 with a corresponding active radar system analysis is apparent. This is an application area where the possible advantages of combining an active and a passive system are just coming to be identified and investigated.

REFERENCES

(1) A.W. Adey et al. Theory and Field Tests of a UHF Radiometer for Determining Sea Ice Thickness.
Canadian Aeronaut. & Space Journal, 17, p.425, December 1971.
Also, Theory & Field Tests of a Microwave Radiometer for Determining Sea Ice Thickness.
Tech. Note No. 637.
Communications Research Centre, Ottawa, January 1972.

(2) A.E. Basharinov, et al. Features of Microwave Passive Remote Sensing.
Proc. 7th International Symposium on Remote Sensing of Environment, Univ. of Michigan, May 1971.

(3) W.O. Copeland et al. Millimeter Wave Systems Applications.
IEEE 1969 G-MTT International Microwave Symposium, Dallas, Texas, May 1969.

(4) J.D. Droppleman and R.A. Mennela. An Airborne Measurement of the Salinity Variations of the Mississippi River Outflow.
Jour. Geophys. Res., 75, p.5909, 20th October 1970.

(5) T.R. Hartz. A Radiometer Method for Determining the Thickness of Sea Ice.
C.R.C. Report No. 1217.
Communications Research Centre, Ottawa, May 1971.

(6) D.C. Hogg. Effective Antenna Temperatures Due to Oxygen and Water Vapour in the Atmosphere.
Jour. Appl. Phys., 30, p.1417, September 1959.

(7) J.F. Paris. Transfer of Thermal Microwaves in the Atmosphere.
Texas A&M University (Dept. of Meteorology) Report, 2 volumes, May 1971.
NASA Grant No. NASA NGR-44-001-098; ONR Contract No. Nonr 2119(04); Water Resources Inst. Proj. No. 5013 and Dept. of Defence.

(8) K.A. Richer. Near Earth Millimeter Wave Radar and Radiometry.
IEEE 1969 G-MTT International Microwave Symposium, Dallas, Texas, May 1969.

(9) A. Stogryn. Equations for Calculating the Dielectric Constant of Saline Water.
IEEE Trans. on Microwave Theory and Techniques, MTT-19, p.733, August 1971.

Table I.
Emission factors of some terrestrial materials.

Material	Emission Factor
Metals	0
Sea Water - 1 GHz	0.27-0.32
60 GHz	0.5-0.37
Fresh Water - 1 GHz	0.35-0.37
10 GHz	0.38-0.39
Soils (Dep. on moisture type)	0.6-1.0
Thick sea ice	0.8-0.9

Table II.
U.S. (NASA) Microwave Radiometry Programme.
(C240A is no longer in operation).

Satellite Plans (NASA)

Skylab - 1.4 GHz (21 cm.)

Nimbus - 19.3 GHz scanner (1.55 cm.)
- 22.2, 31.4, 53.7, and 58.8 GHz (1.35 - 0.51 cm.)

Aircraft - Lockheed P3A - 30 KFT.
- Convair 240A - 15 KFT.
- Convair 990 - 35 KFT.

Freq. (GHz) - 1.4 - 94

(cm.) - 21 - 0.32

Table III.
USSR Satellite Microwave Radiometry Programme.

Satellite - Cosmos 243

(Cm.) - 0.8, 1.35, 3.4 and 8.5

Beamwidth (deg) - 9 (at 8.5 Cm.)
- 3.5 (others)

Sensitivity (K^o) - 2 (at 0.8 and 1.35 Cm.)
0.7 (at 3.4 and 8.5 Cm.)

Orbit - 319 x 210 Km.
- 71.3 deg.

Table IV.
USSR Aircraft Microwave Radiometry Programme

Aircraft - IL - 18

(Cm.) - (a) 0.8, 1.35, 1.6 and 3.2
(b) 3, 10 and 30

Sensitivity (K^o) - 1-2

Altitude (Kft.) - 30

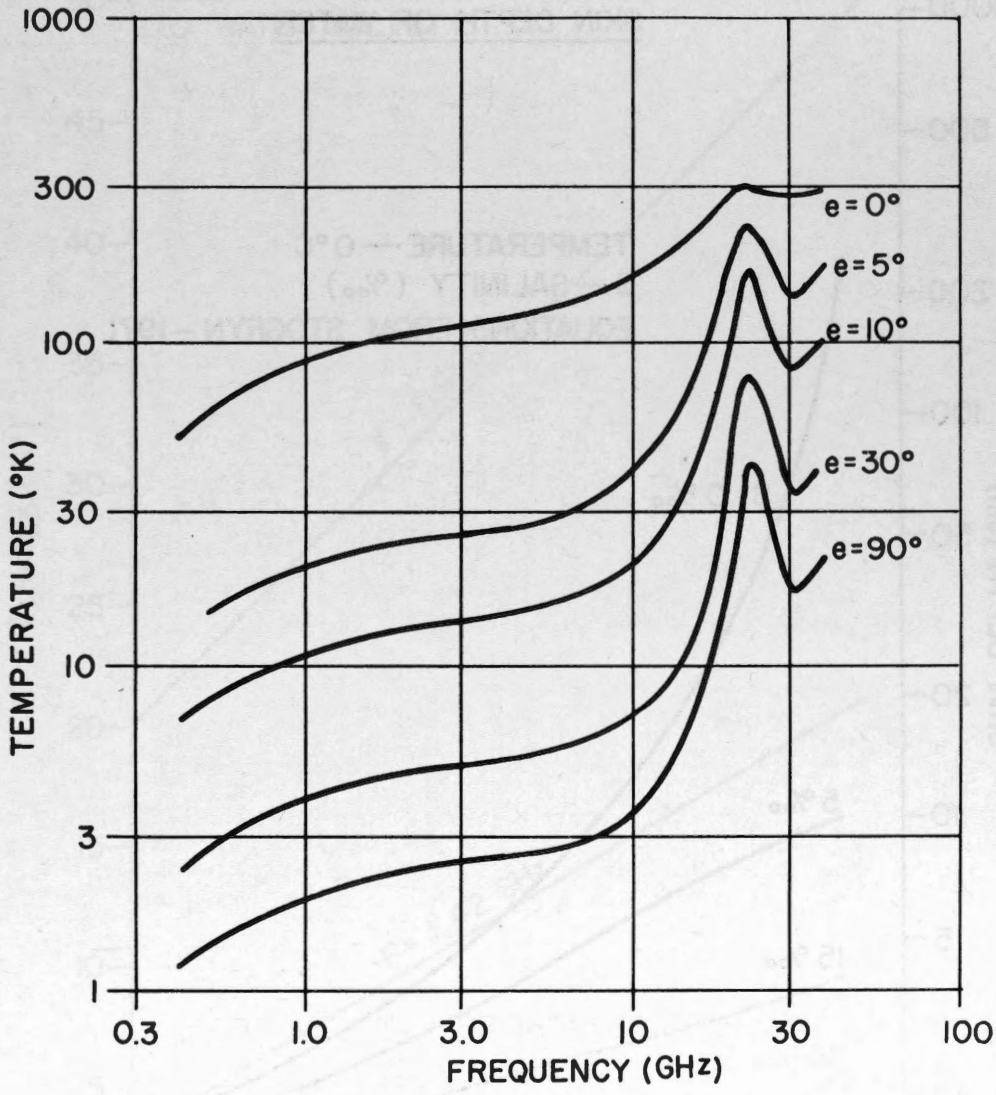


Figure 1. Atmospheric emission (in degrees Kelvin) due to oxygen and water vapour (Hogg, 1959).

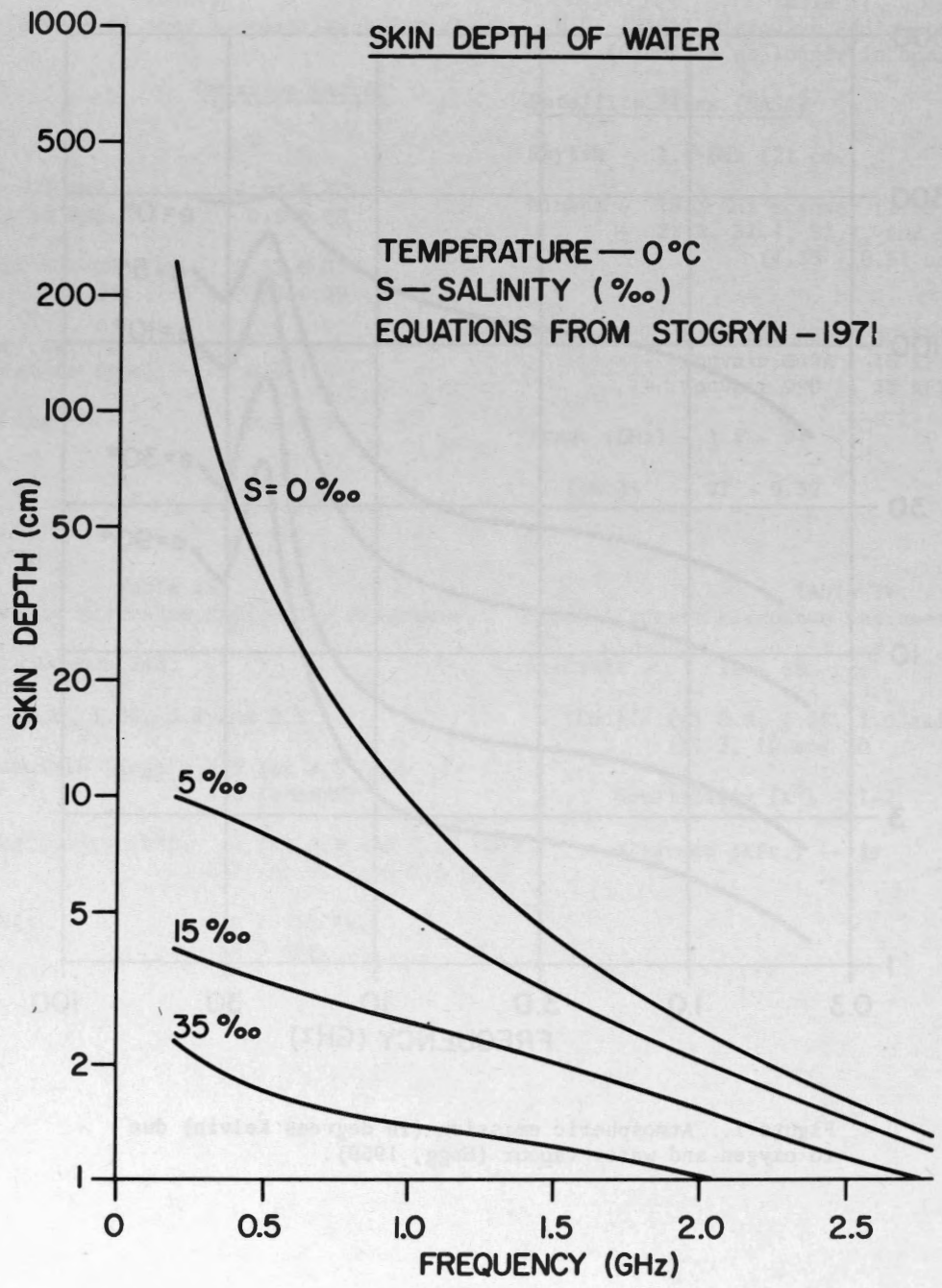


Figure 2. Skin depth of water as a function of frequency and salinity (calculated from Stogryn, 1971).

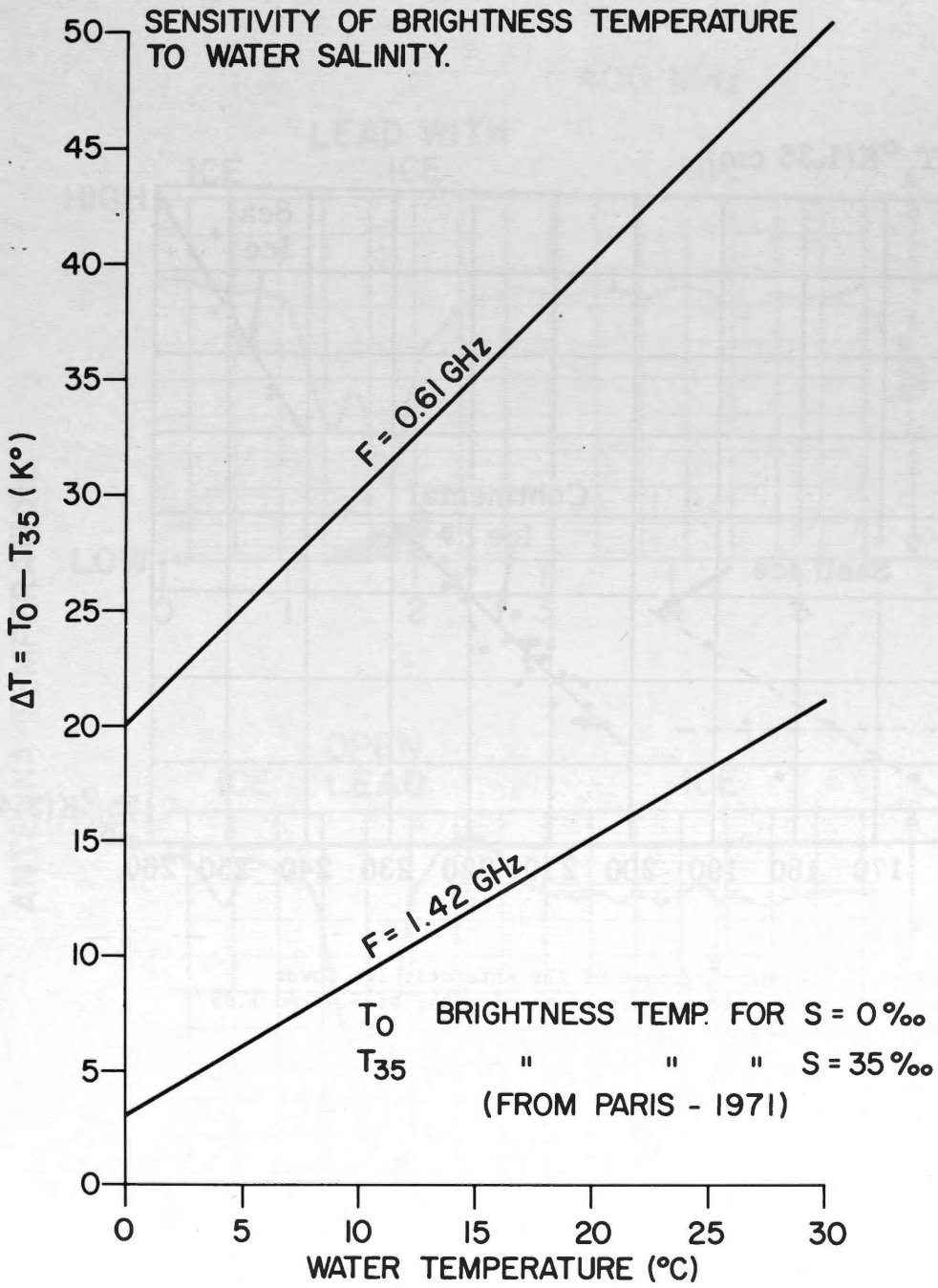


Figure 3. Difference between brightness temperatures of fresh water (S=0 °/oo) and ocean water (S=35 °/oo) as a function of water temperature, for frequencies of 0.61 and 1.42 GHz (from Paris, 1971).

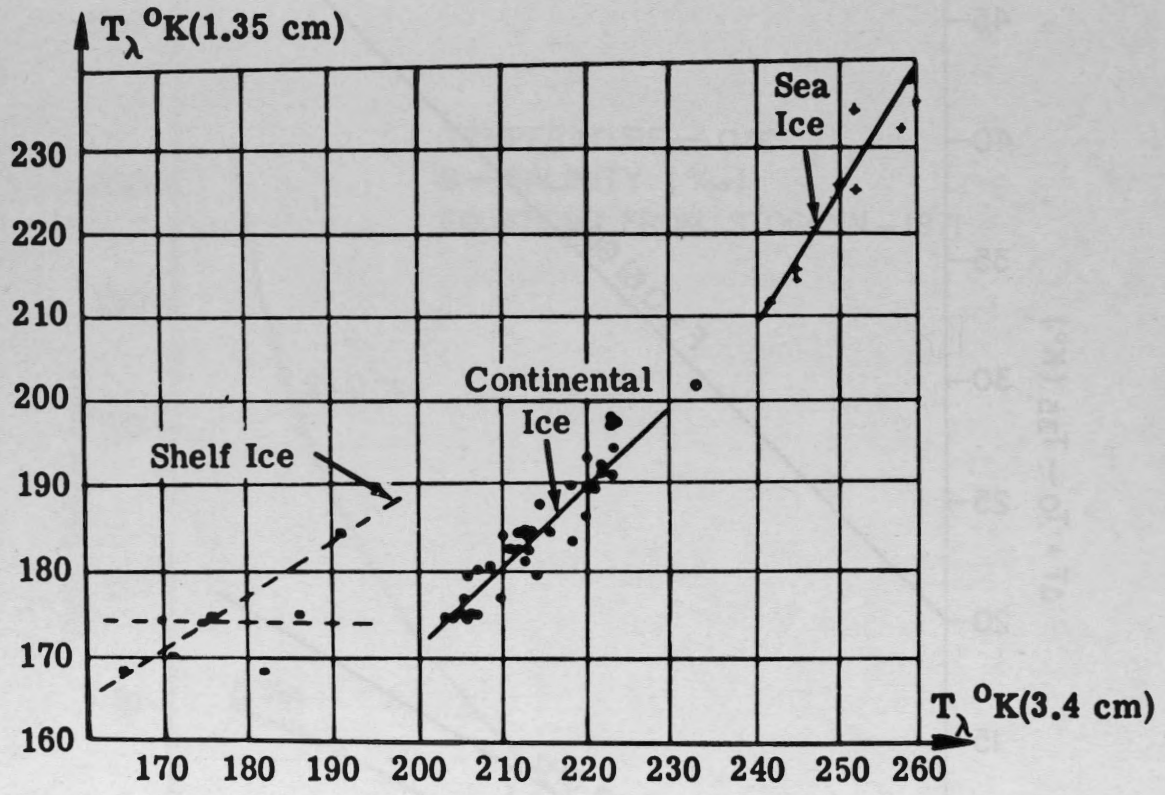


Figure 4. A correlogramme of the Antarctic Ice Cover radiobrightness temperature at wavelengths of 3.4 and 1.35 cm. (Basharinov, 1971).

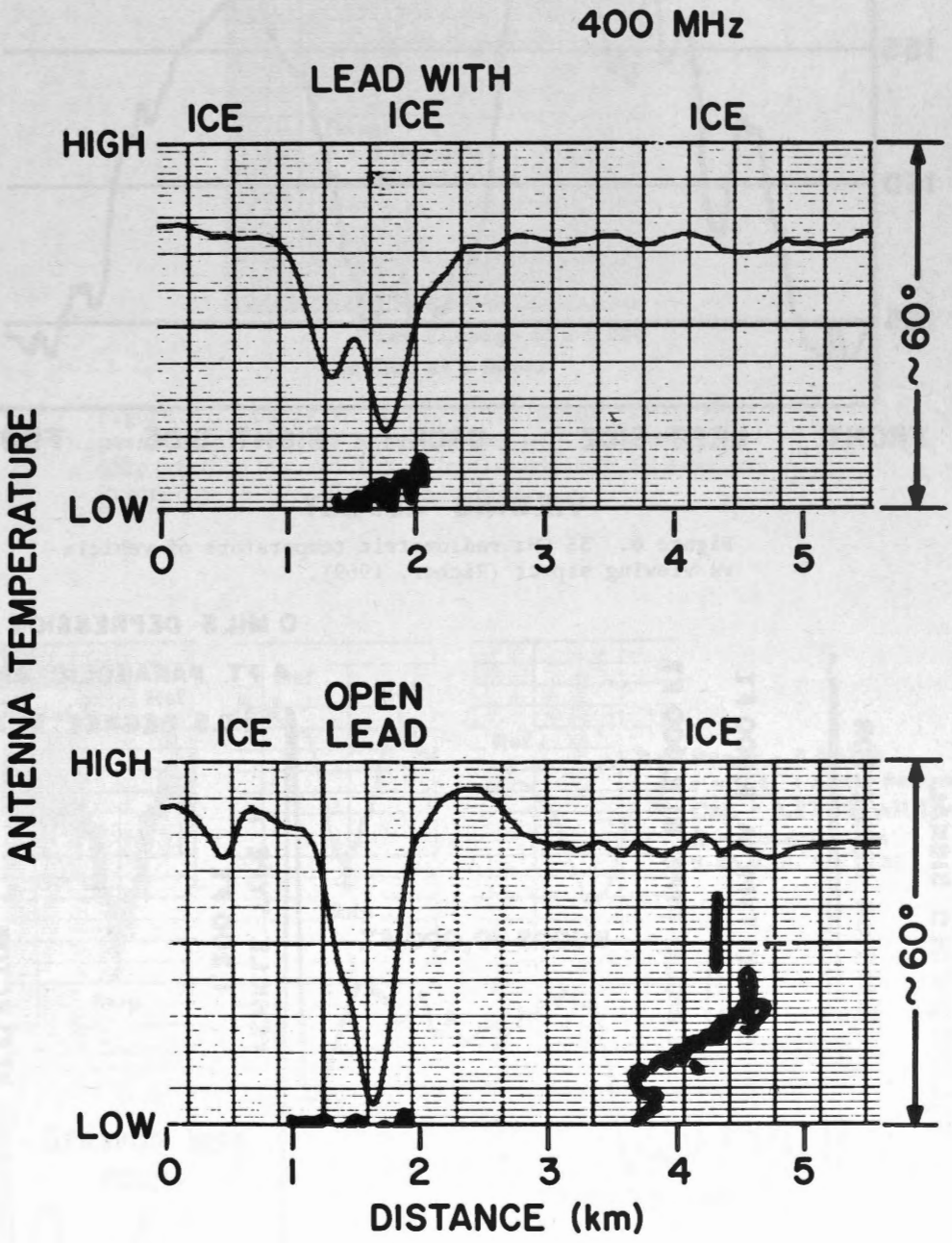


Figure 5. Radiation recorded during two flights over an open lead in first-year Ice. The flights took place in Barrow Strait in February 1971. Frequency: 400 MHz. Top: Lead with ice; bottom: completely open lead (Adey et al, 1971 and 1972).

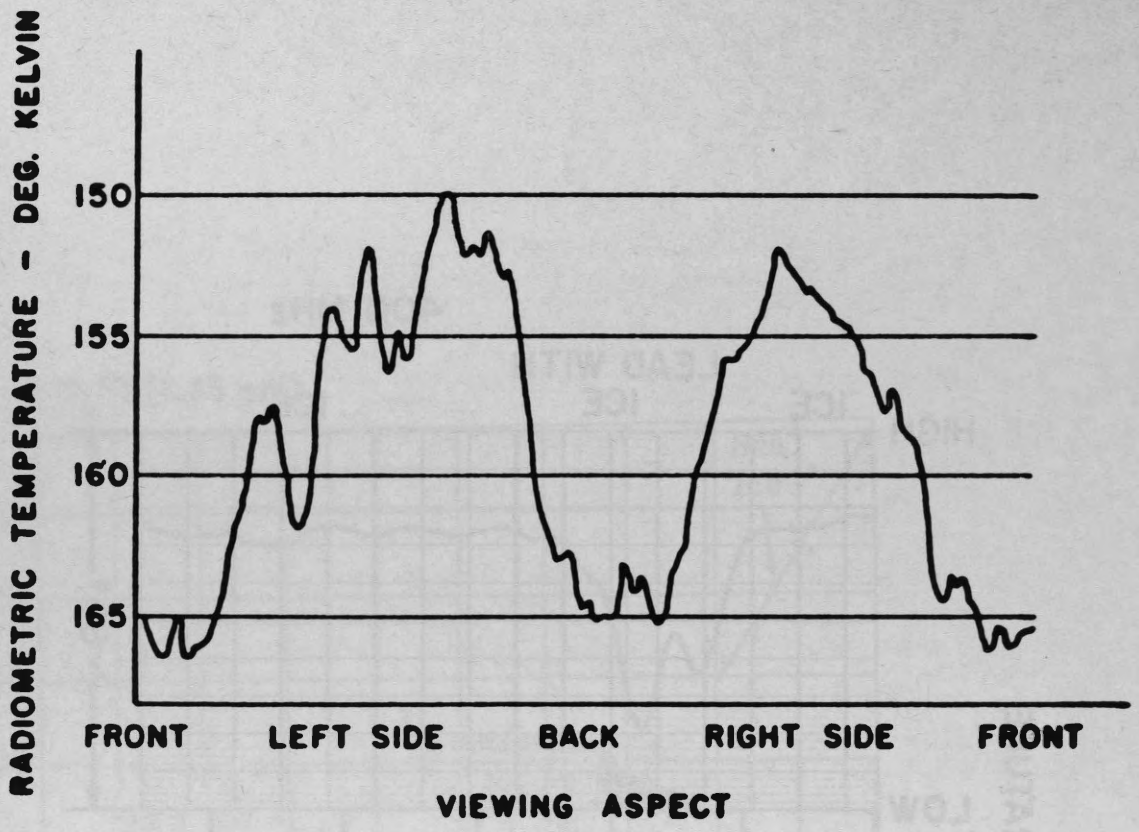


Figure 6. 35 GHz radiometric temperature of vehicle vs viewing aspect (Richer, 1969).

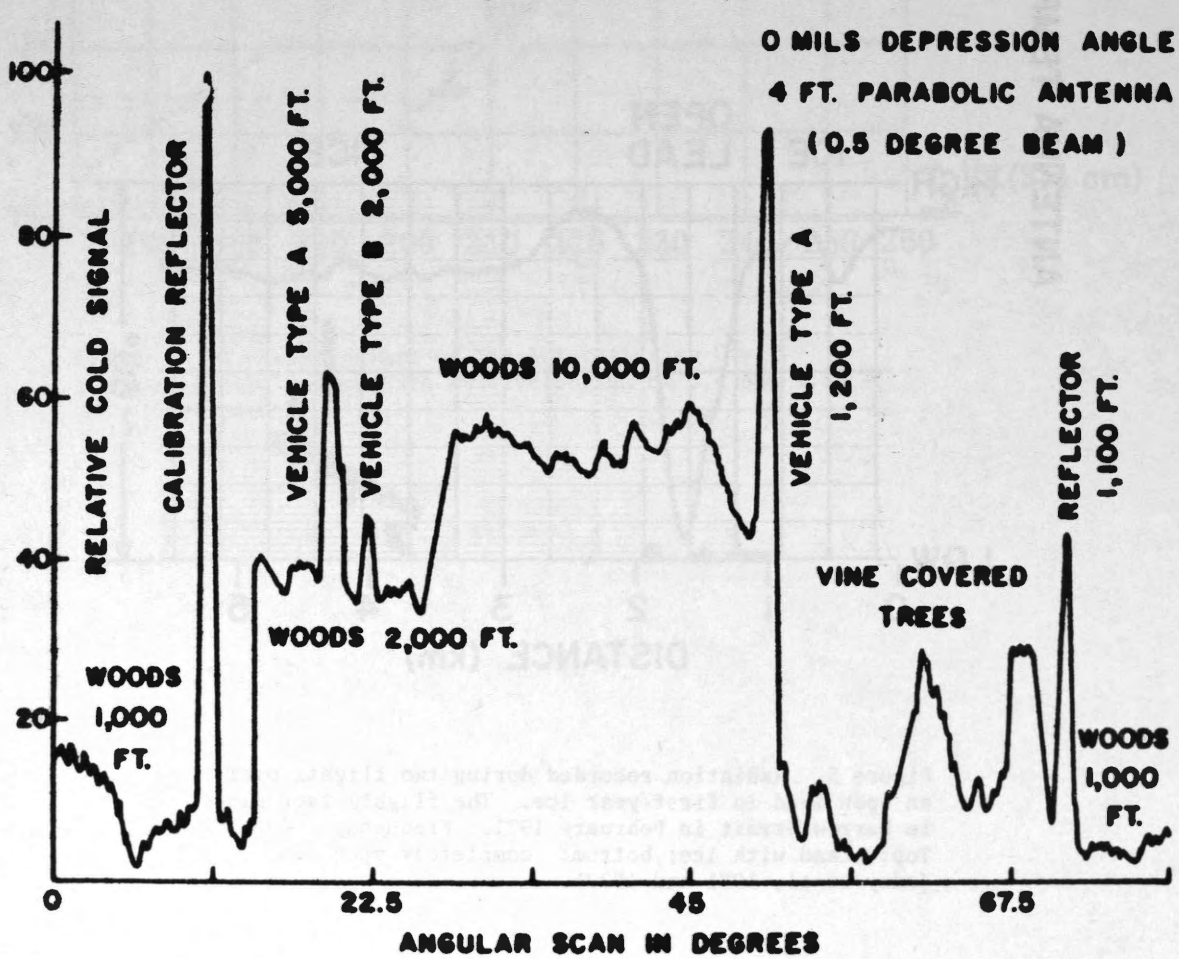
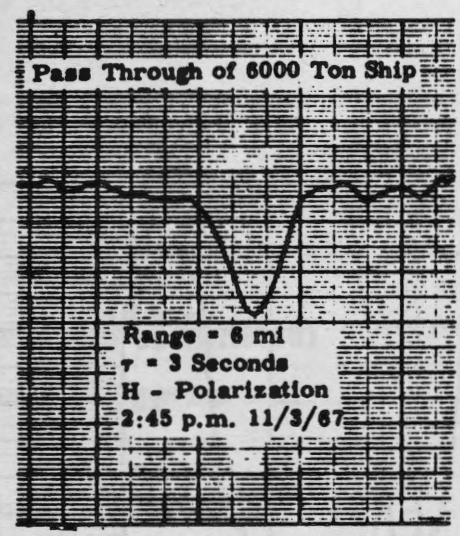
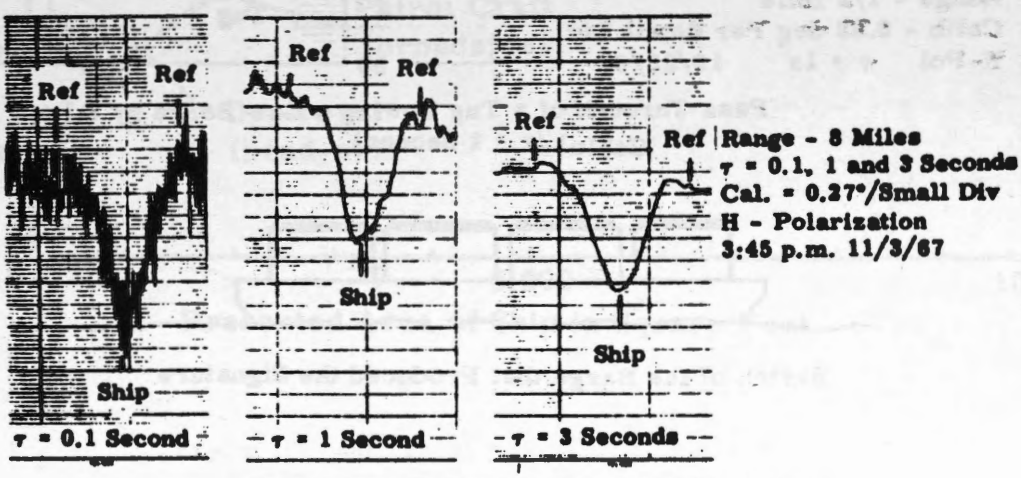


Figure 7. 35 GHz radiometric scan temperature of terrain (Richer, 1969).

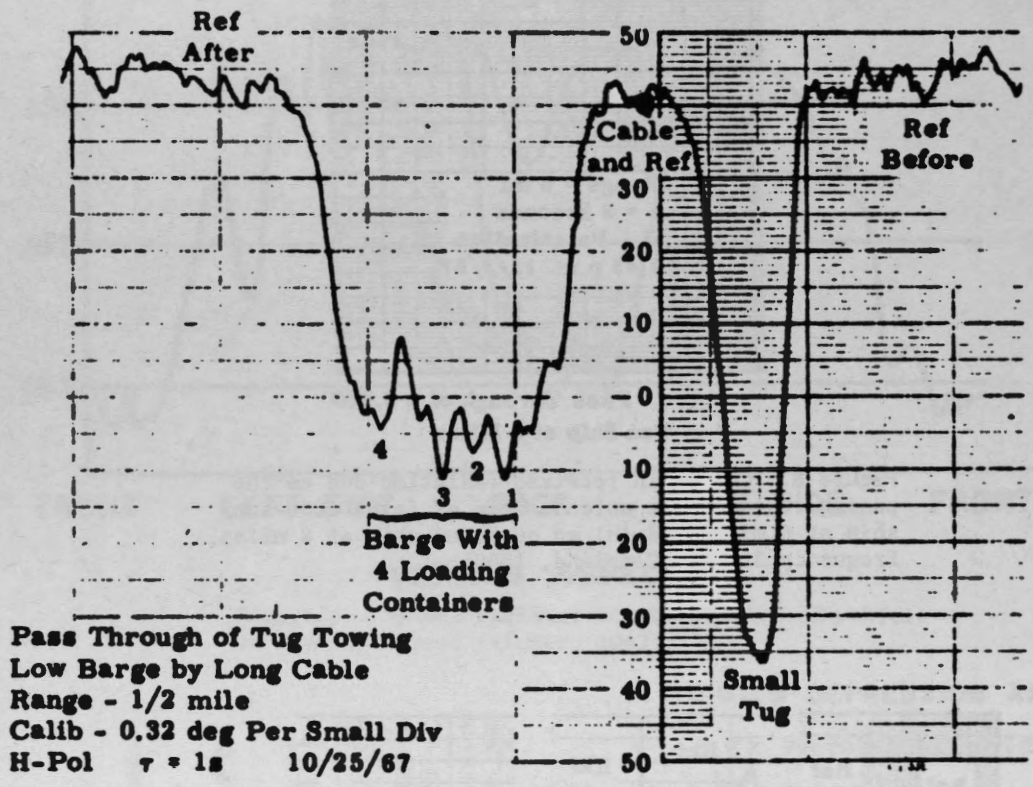


Pass Through of a 6,000 Ton Ship at 6 Miles

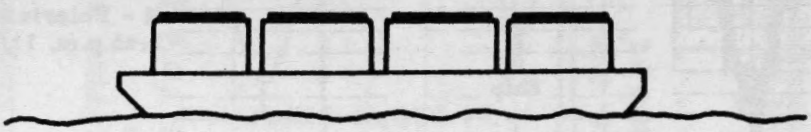
Figure 8. Change in received radiation due to the passage through the antenna beam of (a) a 6000-ton ship at 6 miles and (b) an outbound ship at 8 miles. Frequency 35 GHz (Copeland, 1969).



Pass Through of an Outbound Ship at 8 Miles (τ = 0.1, 1 and 3 Seconds)



Pass Through of a Tug Towing a Low Barge by a Long Cable (τ = 1 Second)



Sketch of the Barge that Produced the Signature

Figure 9. Pass through of a tug towing a low barge with four loading containers. Frequency: 35 GHz (Copeland, 1969).

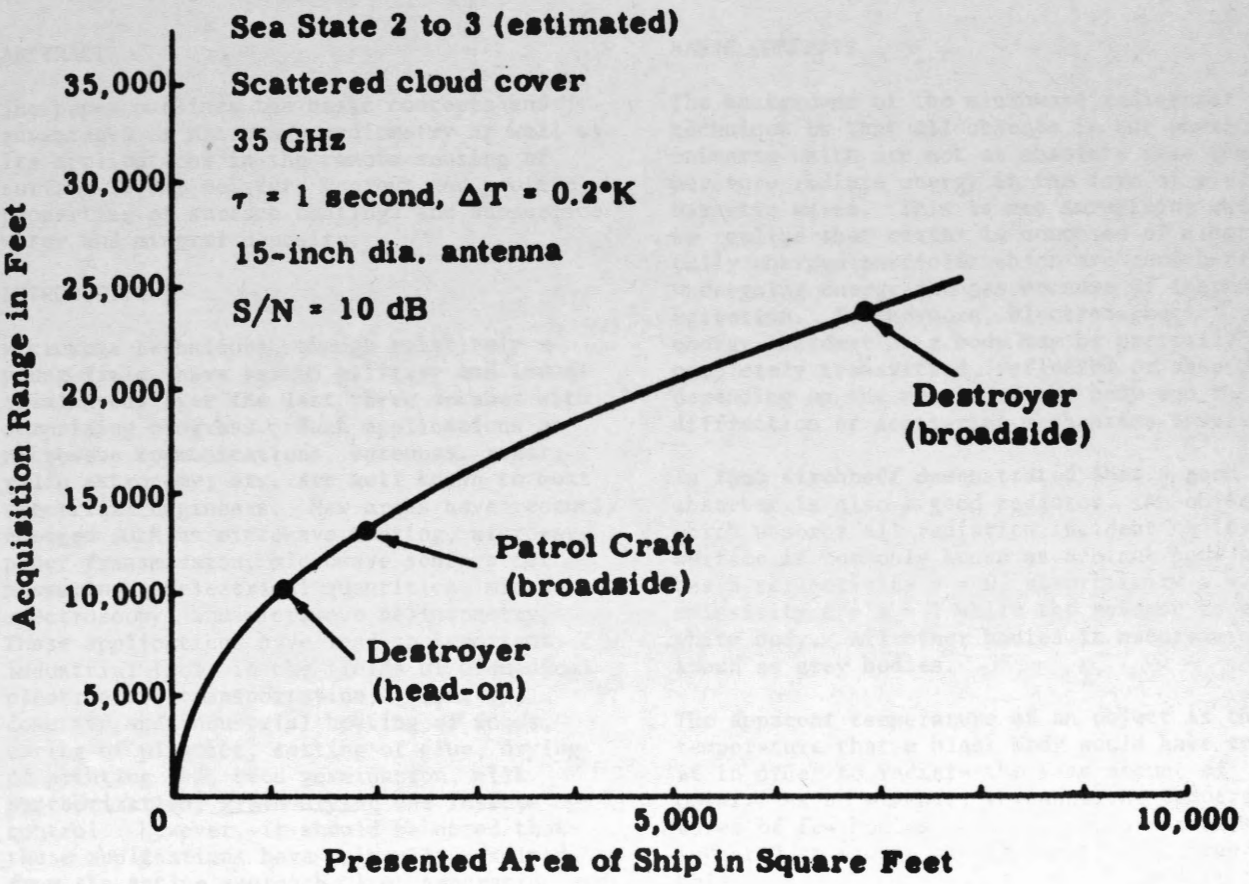
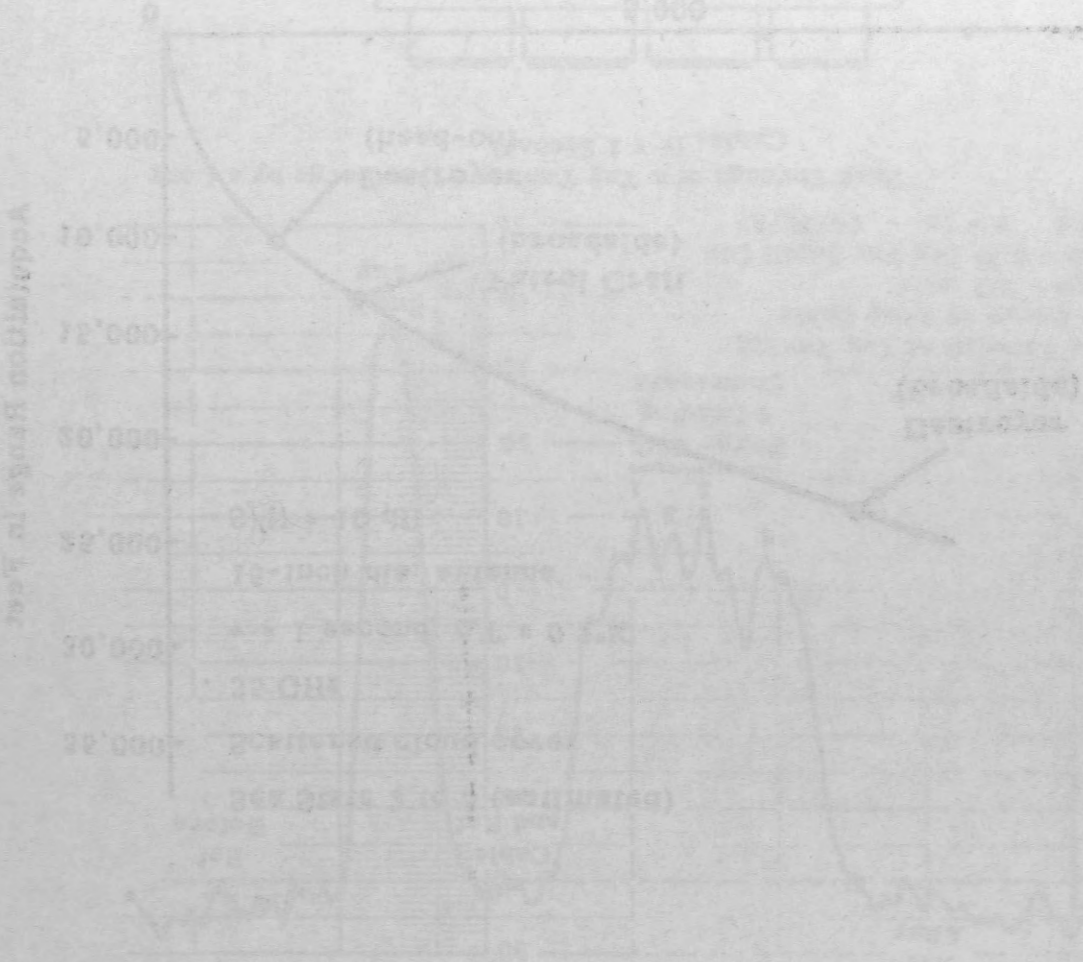


Figure 10. Calculated radiometer acquisition capability. Frequency: 35 GHz (Copeland, 1969).



NEW DIRECTIONS IN MICROWAVE RADIOMETRY FOR REMOTE SENSING

M.A.K. Hamid,
Antenna Laboratory,
Department of Electrical Engineering,
University of Manitoba,
Winnipeg, Manitoba R3T 2N2.

ABSTRACT

The paper outlines the basic concepts and advantages of microwave radiometry as well as its applications in the remote sensing of surface level, moisture content and profile, properties of surface coatings and subsurface water and mineral deposits.

INTRODUCTION

Microwave techniques, though relatively a young field, have served military and industrial needs over the last three decades with surprising progress. Such applications as microwave communications, antennas, radar, radio astronomy, etc. are well known to most electrical engineers. New areas have recently emerged such as microwave heating, microwave power transmission, microwave sensors for measuring nonelectrical quantities, microwave spectroscopy, and microwave tellurometry. These applications have lead to important industrial tools in the fields of biomedical electronics, transportation, telemetry, domestic and industrial heating of foods, curing of plastics, setting of glue, drying of printing ink, seed germination, milk pasteurization, grain drying and insect control. However, it should be noted that these applications have primarily resulted from the active approach, i.e. generation and transmission of waves, modulated or unmodulated, continuous or pulsed, and making use of reflected, transmitted, diffracted, dissipated or scattered power to monitor, control or initiate the process involved. In contrast, relatively little has been achieved with the passive radiometer approach which may be superior than the active approach for certain applications and complementary for others.

The purpose of this paper is to outline some of the areas where microwave radiometry has been applied for remote sensing and the salient applications which either resulted or could be achieved with further investigations.

BASIC CONCEPTS

The background of the microwave radiometer technique is that all objects in the physical universe which are not at absolute zero temperature radiate energy in the form of electromagnetic waves. This is not surprising when we realize that matter is composed of electrically charged particles which are constantly undergoing energy changes because of thermal agitation. Furthermore, electromagnetic energy incident on a body may be partially or completely transmitted, reflected or absorbed depending on the nature of the body and the diffraction or scattering mechanisms involved.

In 1860 Kirchhoff demonstrated that a good absorber is also a good radiator. An object which absorbs all radiation incident on its surface is commonly known as a black body and has a reflectivity $r = 0$, absorptivity $a = 1$, emissivity $\varepsilon = a = 1$ while the reverse is a white body. All other bodies in nature are known as grey bodies.

The apparent temperature of an object is the temperature that a black body would have to be at in order to radiate the same amount of power. As an example, the apparent temperatures of few bodies in the microwave region, say at 3 cm wavelength (X-band), are given below:

	Emissivity	Apparent Temp. at X Band
Sky	<0.1	25°K
Sea	0.3	125°K
Iceberg	0.9	225°K

Since the emissivity of metals is below that of most materials, and certainly below the emissivity of water, soil, ice, or open sky, there is obviously strong contrast between metallic and background dielectric bodies in general.

Apart of the emissivity of materials, expressed in terms of radiation temperature, there are two other significant factors in microwave radiometry. These are the background noise temperature and the receiver parameters

(antenna gain and beamwidth, receiver time constant, noise figure and sensitivity). Considerable information about these parameters is available in the literature [1-32].

From the classical theory of black body radiation, any perfectly absorbing body emits radiation at a wavelength λ into an isothermal enclosure in accordance with Planck's radiation law:

$$E_\lambda = \frac{8\pi ch}{\lambda^5\left[e^{\frac{ch}{\lambda KT}} - 1\right]} \text{ joules/m}^3 \qquad (1)$$

where
h = Planck's constant,
c = velocity of light,
K = Boltzmann's constant, and
T = absolute temperature of the body in degrees Kelvin.
If at microwave frequencies we assume that

$$\frac{ch}{K} \ll \lambda T$$

then we have the approximate relation

$$E_\lambda \simeq \frac{8\pi KT}{\lambda^4} \text{ joules/m}^3 \qquad (2)$$

which is known as the Rayleigh-Jeans equation. Since E_λ is the rate at which energy is radiated in a given direction per unit solid angle per unit area, i.e.

$$E_\lambda = \frac{4\pi}{c} i_\lambda \qquad (3)$$

where
i_λ = intensity of radiation,
then the power P_λ radiated over the range $\Delta\lambda$ is given by

$$P_\lambda = i_\lambda \Delta\lambda = \frac{2KTC\Delta f}{\lambda^2} \text{ watts/m}^2\text{/steradian} \qquad (4)$$

where
Δf = frequency range corresponding to $\Delta\lambda$.
Thus it may be shown that power received by a plane polarized antenna with gain $G(\theta,\phi)$ is given by

$$P_R = \frac{KB}{4\pi}\int_\Omega T(\theta,\phi)\, G(\theta,\phi)\, d\Omega \quad \text{watts} \qquad (5)$$

where
B = bandwidth of the system,
$T(\theta,\phi)$ = temperature distribution across the body,
$d\Omega$ = element of the solid angle Ω subtended by the body at the antenna, and
r,θ,ϕ = spherical coordinate system with origin at the radiometer antenna.
Similarly, the power P_G received from a grey body is given by

$$P_G = \frac{K}{4\pi}\int_{\Delta\lambda}\int_\Omega T_e(\lambda,\theta,\phi)\, G(\theta,\phi)\, d\Omega\, d\lambda \qquad (6)$$

where T_e is the effective temperature defined by

$$T_e(\lambda,\theta,\phi) \triangleq \varepsilon(T,\lambda,\theta,\phi) \cdot T(\theta,\phi) \qquad (7)$$

It should be noted that, unlike P_R, P_G is dependent on wavelength due to the frequency dependence of the emissivity. It is obvious that, in the event the body of interest is surrounded by a radiating surface of different effective temperature, the quantity P_G may be written as the sum of a constant term plus another term of the form given in (6) except that T_e is replaced by the difference between the body and background effective temperatures.

WHAT IS A MICROWAVE RADIOMETER?

A microwave radiometer is analogous to the human eye and is, in effect, an extremely sensitive radio receiver. This receiver is provided with a highly directional antenna, usually a parabaloidal reflector, and the technique of observation consists of directing the antenna at the object or region of space under observation and reading the power picked up by the antenna, as indicated through the circuits and equipment associated with the output display of the receiver.

BASIC ADVANTAGES OF MICROWAVE RADIOMETRY

- High detection capability: detection signal is determined by (a) material emissivity relative to background, geometrical dimensions and position relative to ground level and antenna, (b) antenna gain, beamwidth and bandwidth, and (c) receiver gain, time constant, noise level and sensitivity.
- Simple, passive, continuous, contactless, low power and compact particularly with solid state receivers and active antennas.
- Permits scanning, multifrequency operation, data storage, correlation and display at a center location.
- Sensitivity improves with sea-state.
- Has an all-weather capability (particularly at lower microwave frequencies).

TYPICAL APPLICATIONS OF MICROWAVE RADIOMETRY

- Environmental Protection: measurement of snow and ice thickness, flood levels, hail

and rainstorms; oil pollution surveillance in lakes, rivers and high seas; monitoring of air and electromagnetic pollution in residential areas.

- Agricultural Industry: monitoring of permafrost and subsurface water; monitoring of flood conditions, hail and rainstorms and moisture profiles; optimum frequency for heating, pasteurization, insect control of agricultural products; counting and grading of agricultural products.

- Electrical Industry: measurement of water level in dams, lakes, and rivers and fault diagnosis in cables and waveguides.

- Oil Industry: measurement of oil level in tanks and reservoirs; monitoring of void fraction, fluid density and speed in pipelines.

- Paper and Textile Industry: measurement of moisture content and thickness of web materials.

- Aviation Industry: detection of wind and rainstorms; detection of airborne aircraft, landing strips and coastal lines.

- Mining Industry: detection of oil, gas and mineral deposits.

There are other applications in the fields of hydrology, oceanography and sediments which are discussed in detail elsewhere [16].

SPECIFIC APPLICATIONS OF MICROWAVE RADIOMETRY

Monitoring of Surface Level:

Adey [13] showed that the radiation temperature of ice is the integral contribution from layers through which the emission is attenuated by a certain parameter. If instead of ice we have a liquid (e.g. water, oil) in a metal tank of reservoir, the contribution from the walls may be neglected while the received signal in a radiometer located at the ceiling is directly proportional to the type of liquid and the height. It is therefore possible to relate the reading directly to the height of the fluid. It is obvious that this technique can be employed to monitor the water level in lakes and dams using airborne microwave radiometers.

Moisture Content and Profile:

Since the emissivity of web materials (e.g. paper, textiles) is proportional to the dielectric permittivity (and hence the moisture content), it is obvious that the variations in the receiver temperature readings are directly related to the variations in moisture content. With a narrow beam antenna, high signal to noise ratio, short receiver time constant and proper calibration with a reference noise source, it is possible to construct radiometers to within 0.2°K resolution and to monitor the moisture content versus position along the web to better than 0.1% accuracy. This accuracy is considerably better than the 0.25% now achieved at a higher cost by gamma ray and optical monitors. It is obvious that this technique can also be employed for monitoring the moisture profile (e.g. soils, stored granular materials, etc.) in the same manner as time domain reflectometery. Here the moisture profile is retrieved once the radiometer is extended to operate as a multifrequency radiometer.

Monitoring of Surface Properties:

The properties of a coating on a metal substrate can be characterized in terms of radiation temperature measured by a sensitive microwave radiometer. Since the contribution from the metal is negligible, it is possible to calibrate a particular coated panel in terms of a desirable or reference temperature and examine similar panels by comparison with the reference temperature. The technique is adequate for retrieving the thickness, moisture (or solvent) content, inhomogeneity and roughness of the finished surface. The accuracy of the technique is significantly improved when a second radiometer is employed such that the two form a short base interferometer system using correlation receivers.

Monitoring of Subsurface Water and Mineral Deposits:

A narrow-beam high-resolution airborne radiometer may detect at low altitudes the depth and extent of subsurface water and mineral deposits. This is basically due to the difference in the effective temperatures of the deposit and the surrounding soil which leads to a sudden variation in the radiometer signature. The extent of the deposit can only be determined from multiple passes, while the depth below the ground level requires reference to a catalogue of data obtained through rather complicated multifrequency measurements.

CONCLUSIONS

It is anticipated that single or multiple airborne multifrequency scanned microwave radiometers will be employed for the remote sensing of natural resources and continuous monitoring of nonelectrical parameters in a wide variety of industrial processes.

ACKNOWLEDGMENTS

The author wishes to thank Mr. O.A. Sandoz, L. Rowlandson, Dr. A.W. Adey and Dr. J.D. Hunter for many profitable discussions throughout this work.

This research was supported by the Defence Research Board of Canada under Grant 3801-42.

REFERENCES

[1] G.C. Southworth. Microwave Radiation from the Sun: Journal of the Franklin Institute, vol. 239, no. 4, April 1945.

[2] R.H. Dickie. The Measurement of Thermal Radiation: Review of Scientific Instruments, vol. 17, July 1946.

[3] E.M. Sparrow and R.D. Cess. Radiation Heat Transfer: Brooks/Cole, California, 1966.

[4] H.C. Hottel and A.F. Sarofim. Radiative Transfer: McGraw-Hill, N.Y., 1967.

[5] T.P. Merritt and F.F. Hall. Blackbody Radiation: Proc. IRE, 1959, 1435-1411.

[6] D.B. Harris. Microwave Radiometry-I: Microwave Journal, 3, 1960, 41-46.

[7] D.B. Harris. Microwave Radiometry-II: Microwave Journal, 3, 1960, 47-54.

[8] R. Meredith and F.L. Warner. Superheterodyne Radiometers for use at 70 Gc and 140 Gc: IEEE MTT-11, 1963, 397-411.

[9] M.W. Long and J.C. Butterworth. New Technique for Microwave Radiometry: IEEE MTT-11, 1963, 389-397.

[10] J. Dijk, M. Jenken and E.J. Maanders. Antenna Noise Temperature: Proc. IEE, 115, 1968, 1403-1410.

[11] K. Voss. Coast Guard Discovers that AN/AAR-33 Sperry Iceberg Tracker System Aids Mapping: Technology Week, 1967, 38-40.

[12] H.L. Clark. Some Problems Associated with Airborne Radiometry of the Sea: Applied Optics, vol. 6, no. 12, Dec. 1967, pp. 2151-2157.

[13] A.W. Adey. Theory and Field Tests of a Microwave Radiometer for Determining Sea Ice Thickness: presented at AGARD EM Wave Prop. Panel XVII Am. Symp. on Prop. Limitations in Remote Sensing, Colorado, 1971.

[14] A.W. Adey. A Survey of Sea-Ice-Thickness Measuring Techniques: Dept. of Communications, CRC Rept. 1214, Ottawa, 1970.

[15] J.W. Dees and J.C. Wiltse. An Overview of Millimeter Wave Systems: Microwave J., 12, Nov. 1969, 42-49.

[16] A.T. Edgerton, et al. Microwave Emission Characteristics of Natural Materials and the Environment: (a summary of six years research), Contract No. N00014-70C-0351, Final Tech. Rept. 9016R-8, Aerojet-General Corporation, El Monte, California, Feb. 1971.

[17] J.K. Schindler. The Electromagnetic Scattering of Broad Bandwidth Random Signals by a Discrete Target: AFCRL-67-0602, Nov. 1967, DSIS Acc. No. 68-01838.

[18] Born and Wolf. Principles of Optics: Pergamon Press, 1959, p. 621.

[19] P. Beckmann. The Depolarization of Electromagnetic Waves: The Golem Press, Boulder, Colorado, 1968.

[20] D. Friedman. Infrared Characteristics of Ocean Water (1.5 - 15μ): Applied Optics, vol. 8, no. 10, Oct. 1969, 2073-2078.

[21] L.J. Pierre, Jr. An Exploratory Study of the Infrared Characteristics of Surface Slicks: Naval Research Lab., Washington, D.C., NRL Rept. 7235, Jan. 29, 1971.

[22] Special Issue of Microwaves on Electronic Wavefare: vol. 10, no. 10, Oct. 1971.

[23] J.M. Flaherty. Microcircuit Phased-Array Electronic Countermeasures System: Microwave Journal, vol. 14, no. 3, 1971.

[24] H.C. Ko. Coherence Theory of Radio-Astronomical Measurements: IEEE Trans. on AP, vol. AP-15, no. 1, Jan. 1967, pp. 10-20.

[25] R.H. MacPhie. A Compound Intensity Interferometer: IEEE Trans. on AP, vol. AP-14, no. 3, May 1966, pp. 369-374.

[26] T.V. Seling and D.K. Nance. Sensitive Microwave Radiometer Detects Small Icebergs: Electronics, vol. 34, no. 19, May 1961, pp. 72-75.

[27] D.R.J. White, R.W. Hanford and H. Ihara. Passive Detection with Radiometers: Electronic Industries, vol. 21, no. 11, Nov. 1962, pp. 113-117.

[28] C.D. McGillem and T.V. Selling. Influence of System Parameters on Airborne Microwave Radiometer Design: IEEE Trans. on Military Electronics, vol. MIL-7, no. 4, Oct. 1963, pp. 296-302.

[29] K. Maerz. Messung der thermischen Mikrowellenstrahlung des Erdbodens vom Flugzeug aus: Deutsch Luft-u Raumfahrt Forschungsbericht, no. 64-30, Oct. 1964, p. 42.

[30] D.L. Dahl and D.R. Hornbaker. Radiometer View Angle, Its Meaning with Respect to Instrument Applications and Specifications: ISA 20th Annual Conference, Phys. and Mech. Measurements Instrumentation, Proc. Reprint 17-12-3 for meeting October 4-7, 1965, p. 6.

[31] K. Kuenzi and E. Schanda. Microwave Scanning Radiometer: IEEE Trans. Microwave Power and Techniques, vol. MTT-16, no. 9, Sept. 1968, pp. 787-789.

[32] M.A.K. Hamid and J.D. Hunter. Examination of Surface Coatings by Microwave Radiometry: presented at NATO Surface Coatings Meeting, April 9-14, Imperial College London, England.

THE MEASUREMENT OF SNOW WATER EQUIVALENT USING NATURAL GAMMA RADIATION

R.L. Grasty and P.B. Holman
Geological Survey of Canada

ABSTRACT

The natural gamma radiation emitted by potassium, uranium and thorium is attenuated by snow. This attenuation depends on the water-equivalent depth of the snow layer. Air absorption coefficients were determined by flying at different altitudes over a uniform test strip and used to calculate the absorption coefficients for water. Preliminary results using the Geological Survey of Canada high sensitivity airborne gamma-ray spectrometer indicated that a water-equivalent snow depth of 18 cm could be measured to an accuracy of 2 cm over suitable terrain.

INTRODUCTION

Since 1967 the Geological Survey of Canada in conjunction with Atomic Energy of Canada Ltd. has been developing a high sensitivity airborne gamma-ray spectrometer for the purpose of mapping surface concentrations of potassium, uranium and thorium from the air. During the past three summers, the system has been used to fly many thousands of line miles of experimental survey and reconnaissance profiles. In the winter the survey work is discontinued owing to the blanketing effect of the snow water layer. Parts of Canada, however, receive comparatively little snowfall and it may still be possible to fly these areas and achieve satisfactory results. Experiments carried out in the winter of 1970-71 on the absorption effect of snow indicated that the water-equivalent depth of a snow layer could be determined from a comparison of the radiation pattern received with and without the snow layer. In Canada, snow-water surveys are carried out by a variety of agencies for the purpose of flood control, regulation of water to hydro-electric power plants and reservoirs etc. These surveys generally consist of measurements of snow density using snow gauges over a network of points. The accuracy of these surveys depends on the variation in snow depth over the area concerned and the number of point measurements taken. An airborne technique which could provide a quick and accurate measure of the water equivalent of snow would be invaluable. Research on the use of natural gamma emissions from the ground to measure snow-water equivalent has been carried out in the Soviet Union since 1965 (Kogan et al., 1965; Zotimov 1968) and is now reported to be used operationally over large areas. Similar work has been carried out in Norway using total radiation (Dahl and Odegaard, 1970) with promising results. Recently the U.S. National Weather Service using a high sensitivity gamma spectrometer showed that snow-water equivalent could be measured to an accuracy of better than 0.5 inches over flat terrain (Peck et al., 1971).

GAMMA-RAY SYSTEM

The principal gamma rays emitted by rocks and soils orginate from radioactive potassium-40, bismuth-214 and thallium-208. Bismuth-214 and Thallium-208 are decay products in the uranium and thorium decay series and both emit a complex pattern of gamma radiation, whereas potassium-40 emits mono-energetic gamma rays at 1.46 Mev. Potassium-40 gamma radiation and the high energy gamma radiation of bismuth and thallium travel several hundred feet through the air and can be monitored using an airborne gamma-ray spectrometer to measure surface concentrations of potassium, uranium and thorium. The airborne gamma-ray system operated by the Geological Survey of Canada has been described by Darnley (1970). Briefly it consists of twelve 9" x 4" NaI (Tl) crystals kept at a constant temperature. The pulses from these twelve detectors are fed via a summing amplifier into a 128-channel analyzer. Plug-in program cards are used to select four windows of the spectrum for the measurement of potassium, uranium and thorium, and total radiation, from 0.4 to 2.8 Mev. Table 1 shows the windows selected and Fig. 1 a typical laboratory gamma-ray spectrum. A separate plug-in program calibration card monitors the position of the prominent potassium peak and the drift of the spectrum. Because the crystals are maintained at a constant temperature, the drift is minimal and no adjustment

is normally made during the course of a flight. The counts obtained in the four channels are fed directly on to computer-compatible magnetic tape together with height, long and cross-track doppler position, time, date, line number and operating code. For normal survey work total count information is recorded every 0.5 seconds and the three photo peaks every 2.5 seconds. A radar altimeter is used to register the terrain clearance over the 2.5 second counting period.

Due to the finite size of the sodium iodide detectors, a certain fraction of the thallium gamma-rays impinging on the detectors are incompletely absorbed and appear as counts in the lower energy uranium and potassium windows. Similarly, bismuth gamma-rays can appear as counts in the potassium window. The fraction of the higher energy gamma-rays appearing in the lower energy windows are known as stripping ratios or Compton scattering coefficients and have been experimentally determined (Grasty and Darnley, 1971).

THEORY

When an absorber of thickness D is placed between a detector and a point source of mono-energetic gamma-rays, the number of gamma-rays per second, N, received by the detector is given by the well-known expression

$$N = N_o e^{-\mu D}$$

N_o is the number of gamma-rays per second received by the detector with no absorber and μ is the linear absorption coefficient of the absorbing material.

In the case of a detector in an aircraft which is receiving mono-energetic gamma radiation from the ground, the absorbing material is the intervening layer of air. If the source can be considered as a homogeneous infinite half-space, the relationship between count rate N and height H of the aircraft is given by (Godby et al., 1952)

$$N = N_o \int_1^\infty \frac{e^{-\mu_a H x}}{x^2} dx = N_o E_2 (\mu_a H) \quad \ldots 1$$

where N_o is the counts per second at ground level and μ_a is the absorption coefficient of the air corresponding to the gamma energy. The function $E_2(x)$ is known as the exponential integral of the second kind and is tabulated (Placzek, 1953). When the ground is shielded by a snow layer of water-equivalent depth D, the count rate N may be expressed in a similar form (Godby et al., 1952) and is given by

$$N = N_o E_2 (\mu_a H + \mu_w D) \quad \ldots 2$$

μ_w is the absorption coefficient of the water and N_o is the count rate at ground level when no snow is present.

The standard technique of measuring snow-water depth is to fly a preset course initially in the absence of snow and then with snow. The ratio of count rate N before snow and N_s with snow for any particular energy photo-peak will be given by the expression

$$\frac{N}{N_s} = \frac{E_2 (\mu_a H)}{E_2 (\mu_a H + \mu_w D)} \quad \ldots 3$$

If μ_a and μ_w can be determined and H, the height of the aircraft is known, then the snow-water depth can be calculated.

Kogan et al., (1965) have given a similar expression for the variation of total count with snow depth. However, the situation in this case is rather more complicated because gamma-rays originating in the ground can be Compton scattered in the air and arrive at the detector with a lower energy, causing a build-up in the low energy region.

An alternative technique of snow-water measurement for which a pre-snow flight is not required is demonstrated in Fig. 2. This figure shows that with different snow-water depths the gamma gradient at any altitude is dependent on the snow-water depth. By flying a course at two different altitudes H_1 and H_2 the ratio of the count rates N_1/N_2 is given by the equation

$$\frac{N_1}{N_2} = \frac{E_2 (\mu_a H_1 + \mu_w D)}{E_2 (\mu_a H_2 + \mu_w D)} \quad \ldots 4$$

for which the snow-water depth D can be calculated. Fig. 3 shows the ratio of count rates at heights of 200 feet and 600 feet that will be obtained with different snow-water depths using experimentally determined values of the absorption coefficients of water and air for the thallium-208 photopeak.

EXPERIMENTAL METHODS

1. Air Absorption Coefficients:

In order to determine the air absorption coefficients for the three photopeaks of potassium, bismuth and thallium, and also total count, it is necessary to have a test strip of homogeneous surface composition which is flat and easily accessible to ground measurements and also close to a body of water for accurate background radiation measurements. A suitable test strip was chosen on the flat terrace of the Ottawa River, the exact location is shown in Fig. 4. Ground measurements with a portable gamma-ray spectrometer have shown that the surface composition is relatively uniform (Charbonneau and Darnley, 1970). This test strip is used routinely by the Geological Survey of Canada in order to relate airborne gamma-ray spectrometer readings with ground concentrations.

All gamma-ray spectrometric data have to be corrected for background effects. This background originates from cosmic radiation, radiation due to the activity of the aircraft and its equipment and also the presence of bismuth-214 in the air. Radon, a gas which is a decay product in the uranium decay series can diffuse through the ground into the air as it has a comparatively long half-life of 3.8 days. The radon decays by successive alpha emissions into bismuth-214 which being a charged particle is attracted to and attaches itself to any dust that may be present in the atmosphere. In routine mapping for geological purposes this background is determined by flying over a large body of water so that ground radiation is reduced to negligible proportions. The height chosen is the nominal survey height of 400 feet.

For measurement of air absorption, twelve flights were carried out over the 7-mile test strip during the summer of 1971 from 250 to 800 feet at 50 foot intervals. In order to correct the data for background, flights were taken over the Ottawa River at several different heights. The results of the twelve flights corrected for background and Compton scattering in the crystals, are presented in Table 2.

2. Airborne snow-water measurements:

In February 1971 the test strip was flown at eight different altitudes between 400 feet and 800 feet at fifty foot intervals. These results are tabulated in Table 3. The following day after only a light sprinkling of snow, density and depth measurements were taken along three 1000-foot profiles cutting across the flight lines. Twelve density and fifty-six depth measurements were taken.

EXPERIMENTAL RESULTS

1. Background:

The variation of background with height is shown in Fig. 5, together with errors of one standard deviation. The lines plotted are the computed 'least-squares' best fit, and indicate no significant variation in the measured range of 200 to 600 feet for thorium and potassium. Both the total and uranium count show a slight increase with altitude, but this increase is insignificant compared to the variation in count-rate over the test-line at all twelve altitudes, and the background taken at 400 feet can be assumed constant.

2. Absorption coefficient:

From equation 1, the count rate N_H at height H is related to the air-absorption coefficient μ_a by the relationship:-

$$N_H = N_o E_2 (\mu_a H)$$

where N_o is the count rate at ground level and is to be determined. The optimum values of N_o and μ_a were computed using a 'least-squares' regression technique. The computed curves relating count rate and height are plotted for the three radio-element and total count in Figures 6, 7, 8 and 9. All count rates have been normalized to ground level.

Water absorption coefficients were calculated from the air-absorption values using the relation that a fixed mass of water is 1.11 times as effective in attenuating gamma-rays, of energy 0.4 to 3.0 MeV, as the same mass of air (Davisson, 1965). The absorption coefficients of air and water are presented in Table 4.

SNOW DEPTH CALCULATIONS

1. Standard Technique:

This technique requires a survey flight in the absence of snow. A comparison of the count rates before snow and in the presence of snow can be used to compute the water-equivalent snow depths (Equation 3). Table 5 shows the results of the calculated and ground measured values using count rates obtained at the standard survey altitude of 400 feet.

Because of fluctuations in the ground density and depth measurements, the standard error of the mean depth for the three ground profiles is also indicated.

2. Gradient Technique:

This technique requires no pre-snow flight and depends on the results of measurements taken at two different altitudes. The water equivalent depth can be calculated from these two measurements (equation 4).

From the results of the nine flights carried out in the presence of snow, a minimum water depth of 15 cm. was indicated compared to the actual depth of 17.8 cm. as measured on the ground. It is apparent from the curves plotted in Figure 2 that for water depths greater than 7 cm. the gamma gradient shows little variation for flights carried out over 400 feet. The statistical errors occurring because of low signal to background noise, do not allow the gradients to be measured to a sufficient accuracy for reliable depth determinations at these heights.

This technique may prove to be successful for shallow water depths, flown at low altitude, where the gamma gradient shows the greatest variation and the accuracy of the measurements is increased considerably.

DISCUSSION

From the results listed in Table 5, it can be seen that the airborne and ground results agree within the limits of experimental error for all four energy windows.

It should be pointed out that the total absorption coefficient will be dependent on the thickness of the absorbing layer. In any total energy spectra there exists a high proportion of low energy peaks. These low energy peaks are almost completely absorbed by 400 feet of air and the total absorption coefficient will decrease as the high energy peaks become more predominant. The total absorption coefficient also depends on the spectral composition of the gamma radiation emitted by the ground. Fortunately over most of the Canadian shield the elements potassium, uranium and thorium exist in approximately constant proportions.

Absorption in the uranium window is governed predominantly by the scattering of the higher energy thallium-208 gamma-ray photons. This is well illustrated by its unusually low absorption compared to that of thorium or potassium. The higher proportion of thallium-208 gamma-rays compared to those of bismuth-214 results in a build-up of Compton scattered radiation in the uranium window partially compensating for the higher absorption of the lower energy peak.

Scattering also occurs from the uranium and thorium channels into the potassium window. This scattering can be neglected except in exceptional circumstances of abnormally high uranium or thorium concentrations, because the uranium and thorium counts are normally small in comparison to the potassium counts. Because the thorium window is receiving gamma rays of the highest energy emitted by the ground, no problem of scattering can occur. Provided the spectral characteristics of a test line are known, this scattering phenomena should be no problem to the snow depth determinations.

In the past three summers' mapping work the average thorium over water background has remained constant within a very few percent. The uranium background showed a much greater variation and can change by as much as 300%. Because of the low uranium count rate normally received, it is extremely important to obtain an accurate background prior to each flight. In most parts of the Canadian shield this is no problem as lakes of sufficient size for accurate background determinations are abundant. The total and potassium background show an intermediate variation between that of thorium and uranium and is found to be directly related to fluctuations in the bismuth-214 window.

Air absorption coefficients vary with air temperature and pressure, because of changes in density. For large snow water depths this variation can be neglected as the absorption in the snow will be much greater than in the air. For shallow water depths the air absorption coefficients should be corrected for this density change.

Errors due to random counting statistics are generally insignificant over a course of more than one mile. Experiments carried out over a one-mile section of the test strip in the absence of snow, show that the average uranium count could be reproduced to better than 5% and the thorium to better than 2.5%. Results for both the integral and potassium channels showed variations of less than 1%. As the snow depth increases, however, the signal from the ground decreases and the background noise stays at the same level resulting in reduced accuracy. This is particularly noticable in

the uranium channel where the signal to noise is low. However, even at an equivalent water depth of 17 cm. the consistency of the results from all four energy windows indicates that appreciably greater snow depths could be determined.

CONCLUSIONS:

The high sensitivity gamma-ray system operated by the Geological Survey of Canada has been shown to measure a water equivalent snow depth of 18 cm. to an accuracy of 2 cm. by monitoring the absorption in the potassium, uranium, thorium and total count channels. A gradient technique which does not require a pre-snow flight may prove useful at shallow depths and low survey altitude where the gamma-gradient changes appreciably with altitude.

ACKNOWLEDGMENTS:

The success of this work is largely due to the design and excellence of the equipment for which a large number of people have been responsible. In particular Dr. A.G. Darnley of the Geological Survey of Canada and members of the Atomic Energy of Canada, Commercial Products Division have contributed a great deal. P. Fohn and H. Loijens of the Inland Waters Branch of the Department of the Environment assisted in the ground measurements. We would also like to thank Mrs. Barbara Elliott and Miss Dina Bishop for computerizing the drafting of the data.

REFERENCES

(1) B.W. Charbonneau and A.G. Darnley. A test strip for calibration of airborne gamma-ray spectrometers. Geol. Surv. Can., Paper 70-1, Part B, pp. 27-32, 1970.

(2) J.B. Dahl and H. Odegaard. Areal measurement of water equivalent of snow deposits by means of natural radio-activity in the ground. 'Isotope Hydrology 1970', Vienna, Austria, 191-210, 1970.

(3) A.G. Darnley. Airborne gamma-ray spectrometry. Can. Mining Met. Bull. vol. 73, pp. 20-29, 1970.

(4) C.M. Davisson. Gamma-ray attenuation coefficients, in α, β, γ-ray spectroscopy, vol. 1, edited by K. Siegbahn, pp. 827-843, North Holland, Amsterdam, 1965.

(5) E.A. Godby, S.H.G. Connock, J.F. Steljes, G. Cowper and H. Carmichael. Aerial prospecting for radio-active minerals. Atomic Energy Rept. 13, 1952.

(6) R.L. Grasty and A.G. Darnley. The calibration of gamma-ray spectrometers for ground and airborne use. Geol. Surv. Can. Paper 71-17, 1971.

(7) R.M. Kogan, M.V. Nikiferov, Sh.D. Fridman, V.P. Chirkov and A.F. Yakovlev. Determination of water equivalent of snow cover by method of aerial gamma-survey. Soviet Hydrology: selected papers, No. 4, 183-187, 1965.

(8) E.L. Peck, V.C. Bissell, E.B. Jones and D.L. Borge. Evaluation of snow water equivalent by airborne measurements of Passive Terrestrial Gamma Radiation. Water Resources Research, vol. 7, no. 5, pp. 1151-1159, 1971.

(9) G. Placzek. The functions $E_n(x)$. Nat. Res. Council, Atomic Energy Rept. MT-1, 1953.

(10) N.V. Zotimov. Investigation of a method of measuring snow storage by using the gamma radiation of the earth. Soviet Hydrology: Selected Papers, No. 3, 254-266, 1968.

TABLE 1

SPECTRAL WINDOW WIDTHS

ELEMENT	ISOTOPE	γ-RAY ENERGY	ENERGY WINDOW (MeV)
K	K-40	1.46	1.37 - 1.57
U	Bi-214	1.76	1.66 - 1.86
Th	Tl-208	2.62	2.40 - 2.80
ΣK + U + Th			0.40 - 2.80

TABLE 2

VARIATION OF COUNT RATE WITH HEIGHT OVER TEST STRIP (NO SNOW)

HEIGHT (FEET)	INTEGRAL COUNT/0.5 SEC	THORIUM COUNT/2.5 SEC	URANIUM COUNT/2.5 SEC	POTASSIUM COUNT/2.5 SEC
250	1212	133	41.4	553
300	1072	118	37.7	474
350	950	103	35.0	406
400	888	95	35.7	360
450	818	84	33.6	329
500	732	77	28.0	281
550	687	73	29.4	262
600	615	64	27.6	220
650	556	55	22.7	204
700	506	51	23.1	180
750	456	42	22.3	150
800	399	39	20.3	116

TABLE 3

VARIATION OF COUNT RATE WITH HEIGHT OVER TEST STRIP (17.8 CM. WATER)

HEIGHT (FEET)	INTEGRAL COUNT/0.5 SEC	THORIUM COUNT/2.5 SEC	URANIUM COUNT/2.5 SEC	POTASSIUM COUNT/2.5 SEC
400	362	40	19.0	125
450	332	38	15.7	111
500	298	34	16.1	97
550	285	32	16.8	89
600	268	30	16.5	83
650	261	28	16.2	72
700	248	25	19.2	69
750	236	25	16.2	62
800	215	24	18.6	56

TABLE 4

EXPERIMENTAL LINEAR ABSORPTION COEFFICIENTS

	AIR	WATER*
INTEGRAL	0.00113/Ft.	0.0321/cm.
POTASSIUM	0.00162/Ft.	0.0460/cm.
URANIUM	0.00066/Ft.	0.0187/cm.
THORIUM	0.00130/Ft.	0.0368/cm.

*Calculated from Air-Absorption Results (see text)

TABLE 5

GROUND AND AIRBORNE WATER-EQUIVALENT DEPTHS

ELEMENT	COUNTS PER INTERVAL (400 ft.,no snow)	COUNTS PER INTERVAL (400 ft., snow)	AIRBORNE CALCULATED	GROUND MEASURED
INTEGRAL	888	362	17.6 cm	17.8 ± 3.0 cm
POTASSIUM	360	125	15.5 cm	
URANIUM	35.7	19	19.0 cm	
THORIUM	95	40	15.1 cm	

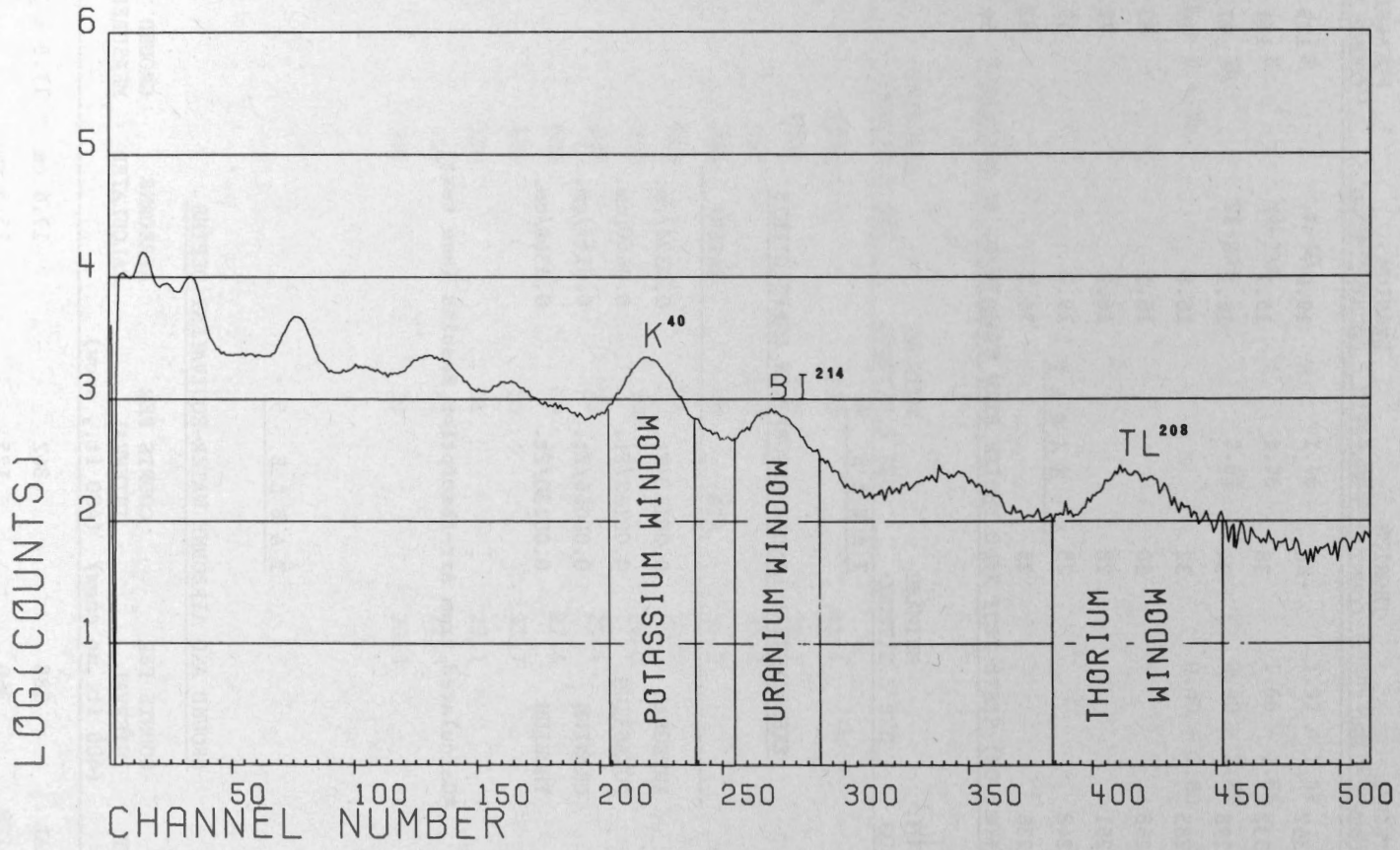


Figure 1. Laboratory Gamma-Ray Spectrum

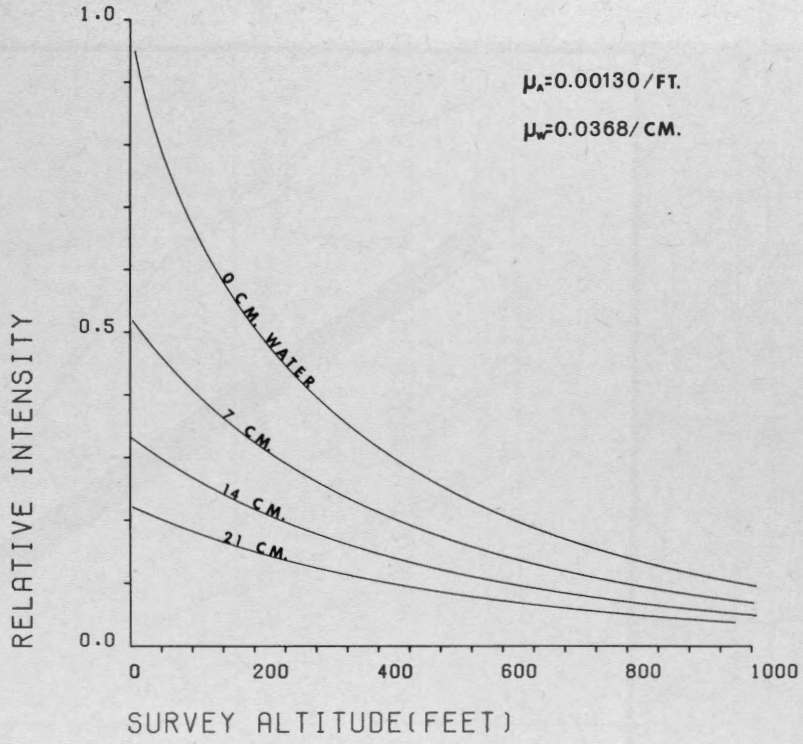


Figure 2. Variation of Count Rate with Height and Water Depth (Thorium)

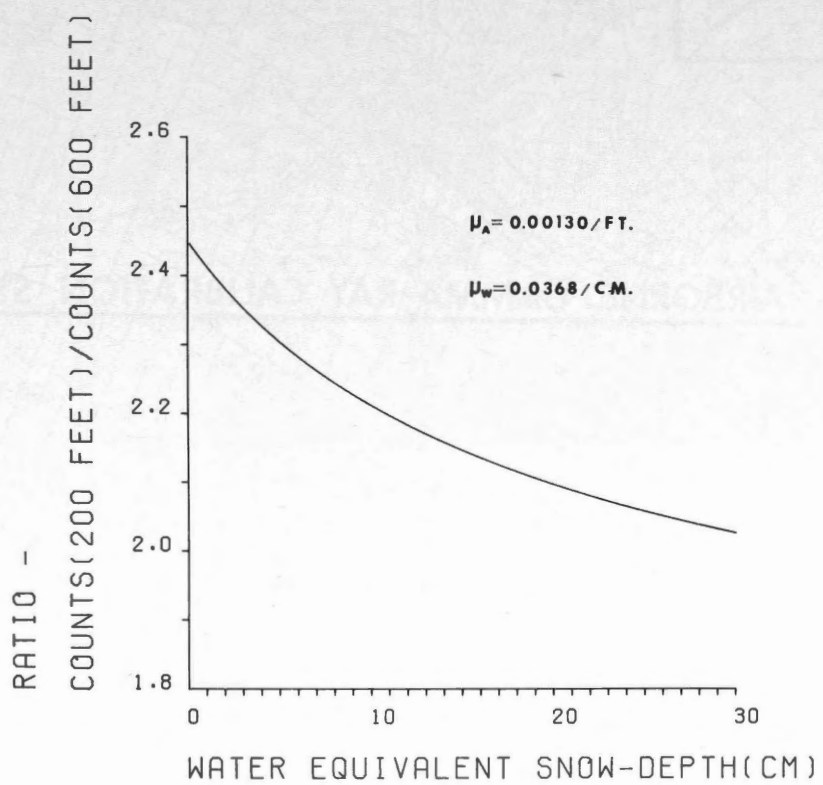


Figure 3. Ratio Variation of Counts at 200 and 600 Feet with Water Depth (Thorium)

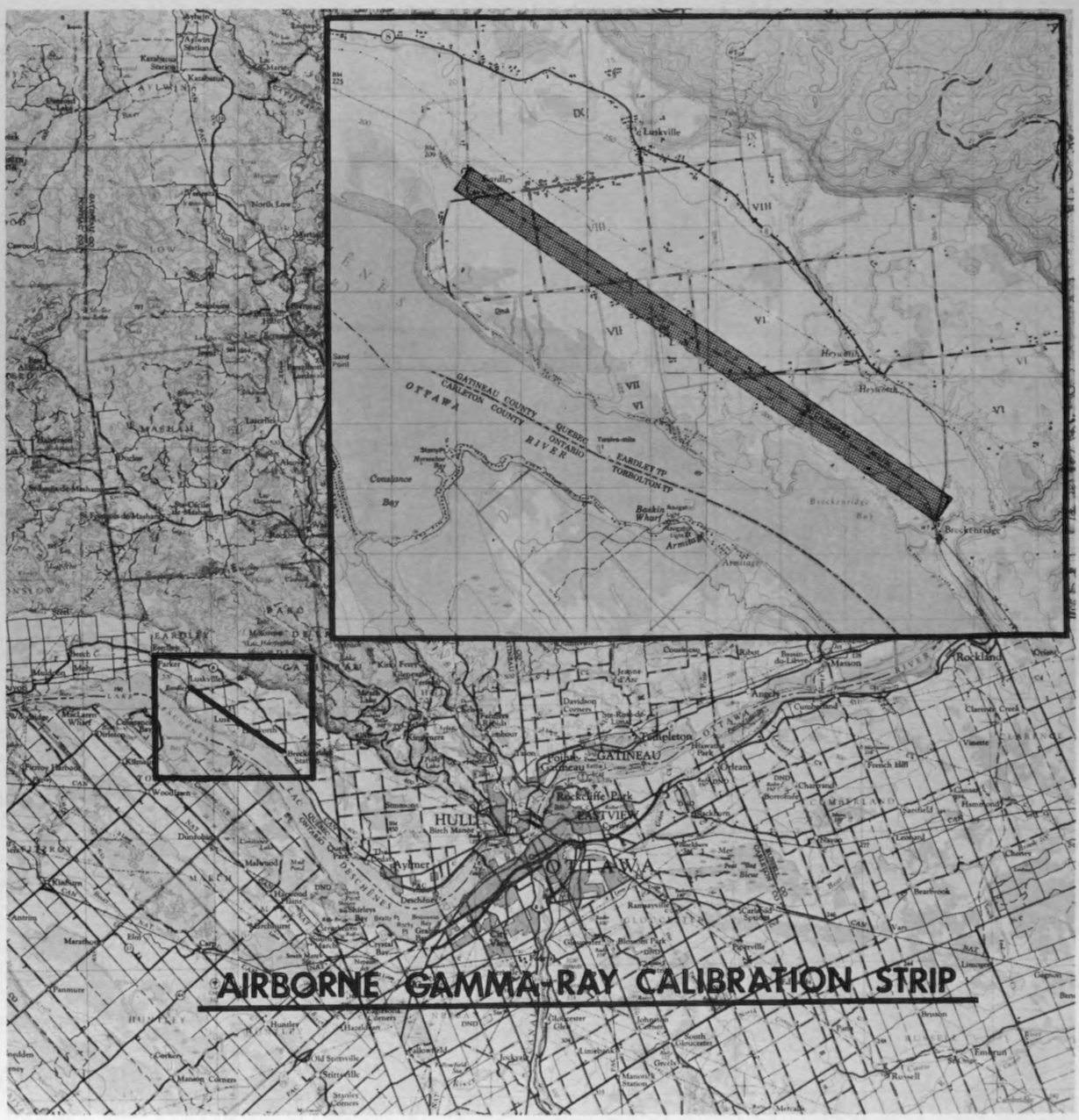


Figure 4. Location of Test Strip

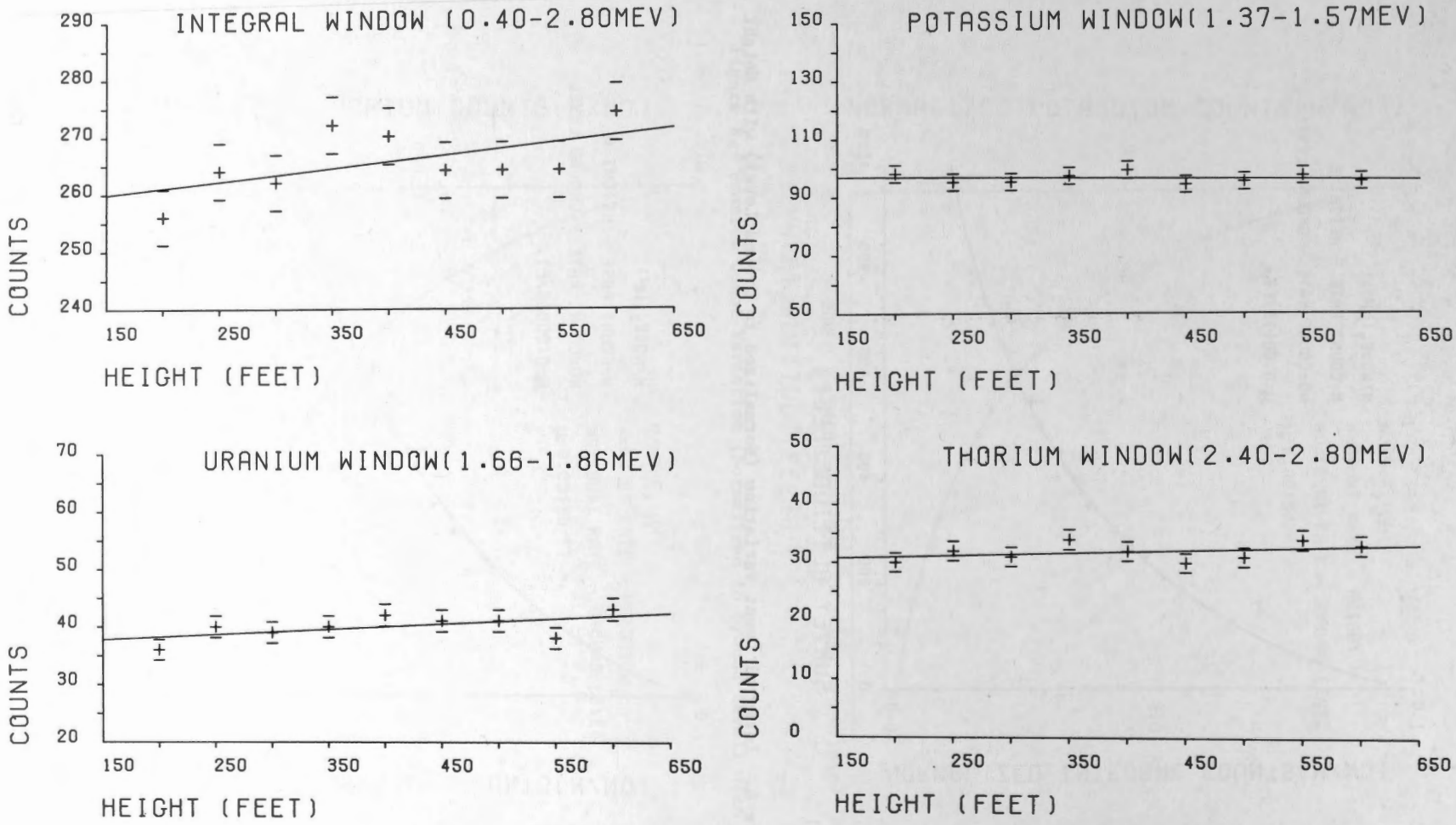


Figure 5. Background Variation with Height

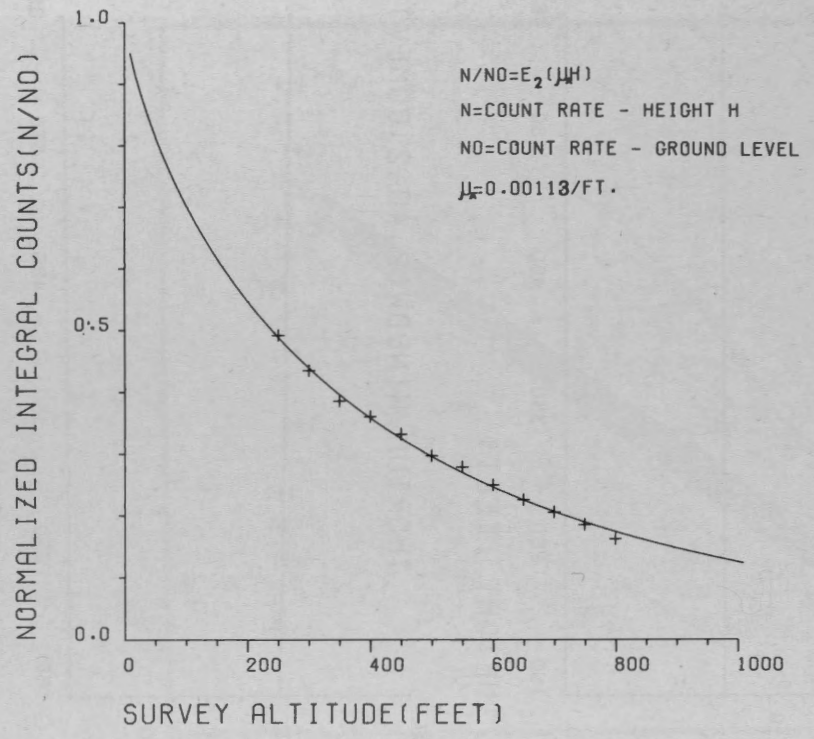


Figure 6. Integral Count Variation (Normalized to Ground Level) with Height

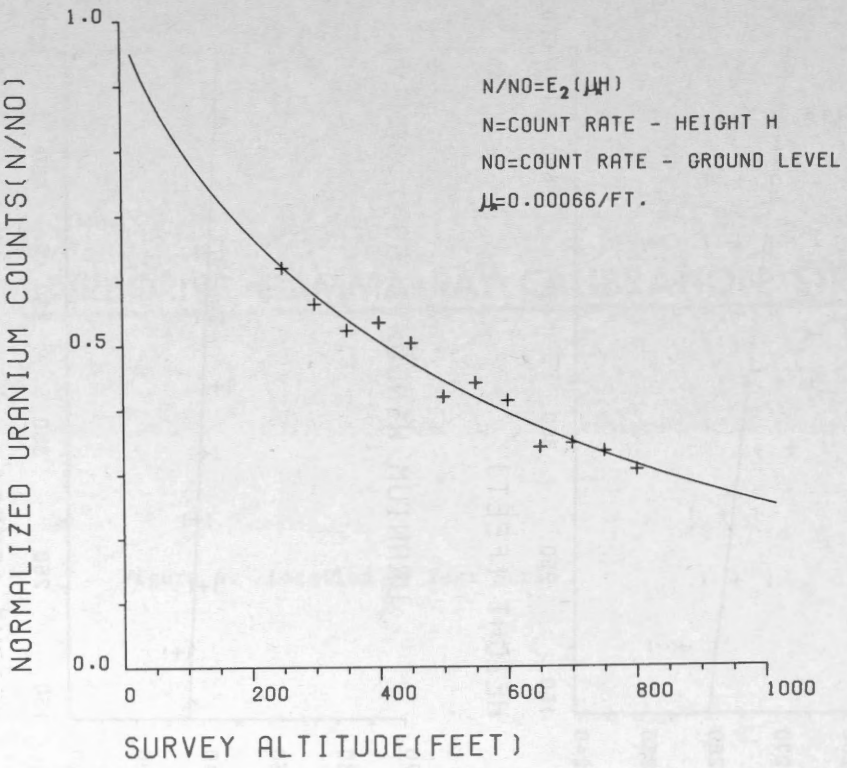


Figure 7. Potassium Count Variation (Normalized to Ground Level) with Height

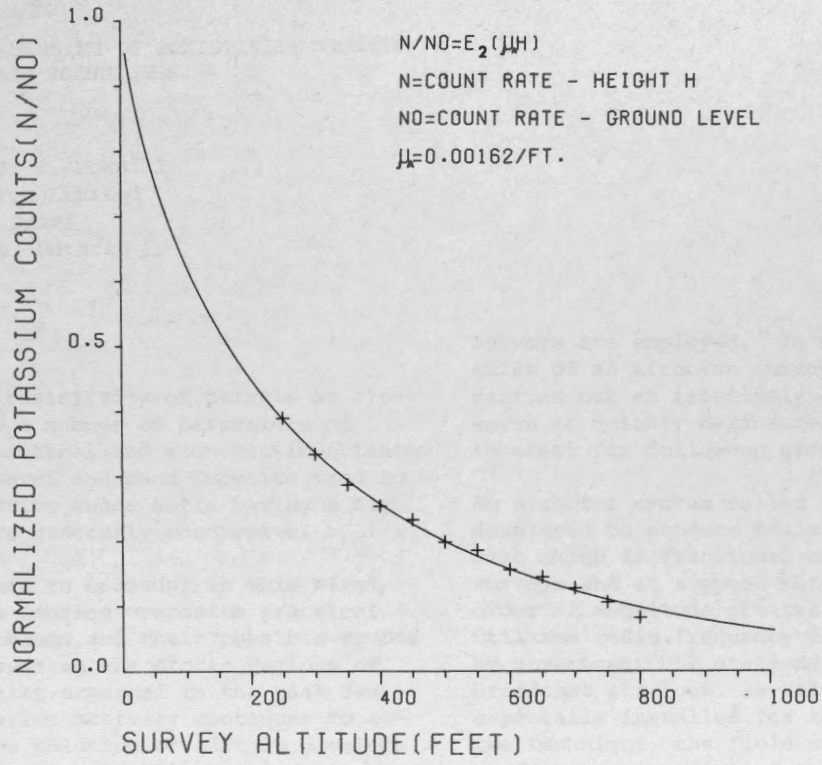


Figure 8. Uranium Count Variation (Normalized to Ground Level) with Height

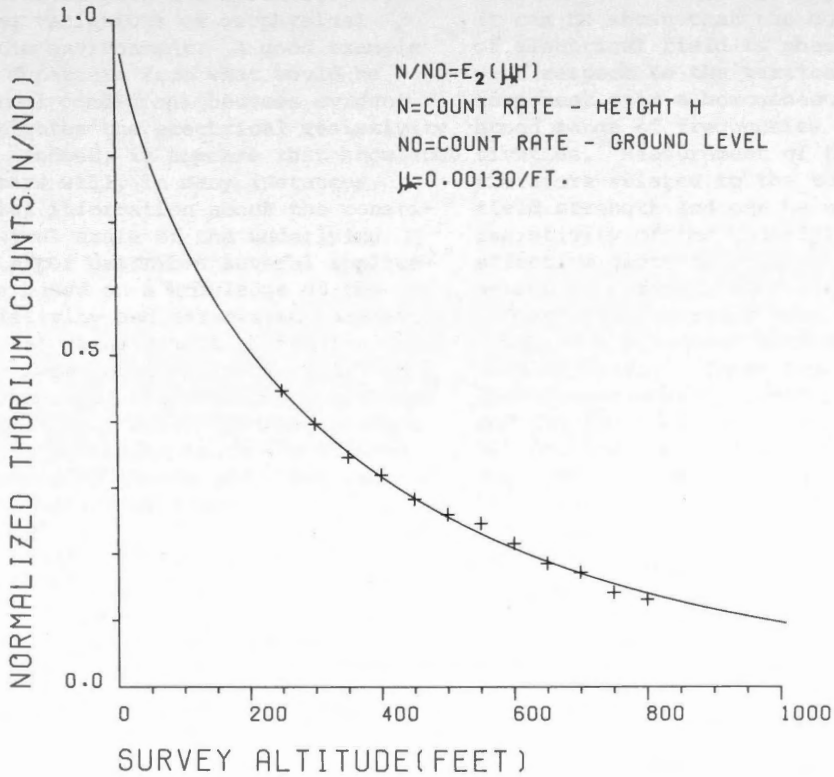


Figure 9. Thorium Count Variation (Normalized to Ground Level) with Height

AIRBORNE REMOTE SENSING OF RESISTIVITY THROUGH THE USE OF E-PHASE TECHNIQUES

J. H. Davies & J. D. McNeill
Barringer Research Limited
304 Carlingview Drive
Rexdale, Toronto, Ontario

INTRODUCTION

The electrical resistivity of terrain is closely related to a number of parameters of considerable practical and economic importance. For example, gravel and sand deposits tend to be highly resistive where soils having a high clay content are generally conductive.

It is appropriate to consider in this First Canadian Remote Sensing Symposium practical measurement problems and their possible remote sensing. Interest in the Arctic regions of Canada has greatly expanded in the past few years. Exploration activity continues to accelerate and the solution of certain problems which are unique to cold climates has become a matter of urgency. Fortunately the extreme conditions associated with the north lead to some interesting variations of geophysical properties of the environment. A good example of the radical departure from what would be regarded as normal conditions becomes evident when one investigates the electrical resistivity of the ground. Indeed, it appears that knowledge of these parameters will, in many instances, provide essential information about the constitution and physical state of the underlying terrain. This paper describes several applications which are based on a knowledge of the subsurface resistivity and associated parameters They include: (a) measurement of sea-ice thickness, (b) location of permafrost in the discontinuous zone, (c) location of permafrost voids in the continuous zone, (d) measurement of the permafrost thickness in the continuous zone, (e) location of gravel and other aggregates in the discontinuous zone.

A feature of special interest is that it should be possible to remotely sense all of the quantities listed above from an airborne platform, either a fixed-wing aircraft or helicopter. This leads to two important advantages which have been recognized by the mineral exploration geophysicist; first, an enormous reduction in the length of time and cost required to survey a given area, and secondly the overall picture that arises from viewing the phenomenon in question over a large region, yielding a perspective which is not possible when highly localized surveys are employed. In many cases the results of an airborne survey, even when carried out at relatively coarse line spacing, serve to quickly delineate areas of enhanced interest for follow-up ground surveys.

An airborne system called E-PHASE has been developed to produce resistivity maps at a cost which is fractional compared with ground surveys and at a speed which is at least an order of magnitude greater. The system utilizes radio frequency fields transmitted by government VLF stations and commercial broadcast stations, as well as transmitters especially installed for the survey. In the technique, the field strength of the quadrature component of the horizontal electric field of the propagated wave is measured with respect to the vertical electric field. It can be shown that the horizontal component of electrical field is phase shifted by 45° with respect to the vertical electrical component over a homogeneous earth for a broad range of frequencies and ground resistivities. Measurement of this component is therefore related to the total horizontal field strength and can be used to derive the resistivity of the underlying terrain. The effective depth to which the resistivity is sensed is a function of the frequency of the radiation and surveys have been carried out using both broadcast band and VLF frequencies simultaneously. These two signals provide penetration depths varying between 10 feet and 100 feet for the broadcast band and 50 to 500 ft. for VLF. Simultaneous use of two or more frequencies enables layering effects to be studied and show considerable potential for the mapping of permafrost distribution and the location of gravel deposits.

THEORY OF E-PHASE SYSTEM

Consider the geometry shown in Figure 1; if we assume that the earth is perfectly conducting a measurement of the electromagnetic field components at a distance from the transmitter of at least five wavelengths (i.e. greater than 100 kilometers in the case of VLF transmitters) will show that there is a vertical electric field denoted by E_z, and a horizontal

magnetic field denoted by H_y. However, as the electrical resistivity of the earth increases from zero a horizontal component of electric field is generated which accounts for power flow into the now lossy earth. The field components are now as shown in Figure 2 where the horizontal electric field is denoted as E_x. For the case of a homogeneous earth, the range of resistivity normally encountered in geophysical surveys (1 to 10,000 ohm-meters), and if the measurements are made at VLF or broadcast frequencies, it can be shown that there is a very simple relationship known as the wave tilt between the horizontal electric field and the resistivity of the underlying terrain, as shown in equation 1, from Maxwell's equations to be:

$$\frac{E_x}{E_z} = (iw\varepsilon_o\rho)^{1/2} \qquad (1)$$

w = $2\pi f$, f = frequency (Hz)
ε_o = permittivity of free space
= 8.85×10^{-12} F/m.
ρ = ground resistivity (ohm meters)
i = $(-1)^{1/2}$

Equation (1) shows that the ratio of horizontal vertical electric field strength is directly proportional to the square root of the ground resistivity. The factor $i^{1/2}$ in the equation indicates that the electrical phase angle between the horizontal electric field and the vertical electric field is 45^o. Thus, the resultant electric field is in general elliptically polarized. Measurements of the polarization ellipse (called "wave tilt" measurements and basically measurements of the horizontal electric field) have been used for many years to determine the electrical properties of the ground and the most recent surveys have been carried out by the Norwegian Defence Research Establishment by Eliassen (1). Measurements of wave tilt cannot, however, be made from a moving platform since they require accurate alignment of the sensing antennas with respect to the surface of the earth. The horizontal electric field that is detected at VLF frequencies is of the order of 1% of the vertical electric field and any rotation of the horizontal electric field antenna results in serious errors due to picked-up vertical electric field. The problem is made more difficult by the fact that the vertical electric field is slightly tilted over topographic features and it is necessary for the horizontal antenna to follow the surface relief.

Development of these electromagnetic equations (4) shows that measurement of W, the wave tilt is in effect a measurement of the surface impedance of free space (η_o):

$$\text{Wave tilt } W = \frac{iw\varepsilon_o}{\gamma_o} Z_1 = \frac{Z_1}{\eta_o} \qquad (2)$$

A technique has been developed which surmounts this difficulty and which using equations (1) and (2) is the basis of the Barringer airborne E-PHASE system (2) which directly senses ground resistivity. Briefly, in this instrument the vertical electric field is used as a phase reference against which to measure the quadrature phase (i.e. that component which is 90^o out of phase with the phase reference) component of the horizontal electric field. As equation 1 shows the quadrature phase component of the horizontal electric field will be reduced from the total value by the factor $(2)^{-1/2}$ but this is of no importance since that factor can be included in the calibration. Because the measurement is made of the quadrature phase component of the horizontal electric field the system is essentially insensitive to small rotations of the sensing antenna and also of the vertical electric field, such as might be caused by topographic relief. Such a system has been flown in several areas in Canada, and the United Kingdom during the summer of 1970 and typical results from such a survey are shown in Figure 3 where the horizontal electric field has been converted to apparent resistivity using equation 1, and contoured. subsequent measurements of the ground resistivity carried out using the conventional Wenner array have confirmed the accuracy of the airborne measurements.

When the homogeneous half space is replaced by a two-layered situation the phase and amplitude of the horizontal electric field becomes much more variable. Under certain circumstances however a considerable simplification takes place. If, for example, we are dealing with two-layered case in which the lower layer is highly conductive so as to be virtually a perfect reflector at the frequency in question and if in addition the upper layer is lossy and/or has a large dielectric constant it transpires that the horizontal electric field measured at the surface becomes basically a measure of the thickness of the upper layer, virtually independent of the resistivity of dielectric constant of this layer. This fact suggests that the technique should be very useful for measuring the thickness of sea-ice.

SEA ICE MEASUREMENTS WITH E-PHASE

The results of the theoretical calculations, plotted in Figure 4, illustrates the resulting horizontal electric field (quadrature phase) as a function of thickness for sea ice of various resistivity and dielectric constant. The calculations given here are exact; it is seen that for high resistivities the effect of varying dielectric constant is not signifi-

cant until very large thickness (of the order of 20 meters) arise, and that the most serious limitation occurs for low ice resistivities when the electrical thickness of the ice becomes large, even for relatively small physical thickness, and the ice becomes a significant fraction of a skin depth thick. This limitation should not prove serious since in reality thicker ice (generally multi-year) tends to be resistive. Thus measurement of the horizontal electric field (quadrature phase) to high accuracy should yield a measurement of sea-ice thickness to equally high accuracy.

There is a certain dependence of the horizontal electric field on the temperature of the sea water. This temperature dependence arises because a small fraction of the horizontal electric field arises from the sea-water itself, the resistivity of which a function of temperature. The current airborne E-PHASE system has the ability to detect a quadrature phase horizontal electric field as small as 0.1% of the primary vertical electric field. In order to achieve an overall resolution of one-half meter in the ice thickness the detection resolution should be improved to approximately 0.02% of the primary field and this presents no problem in a specially designed unit. For the performance of an actual survey the E-PHASE equipment is installed in a suitable aircraft (either fixed wing or helicopter which has a radar altimeter to enable the pilot to maintain approximately constant distance from the ice surface. It is not necessary for this distance to be maintained within close tolerance. The type of navigation employed would depend on the purpose of the survey and the usefulness of the ice thickness data will, of course, depend on how accurately the position of the aircraft was known at the time. The actual ground resolution which is achieved in the direction of the line of flight is determined essentially by the flying height, the greater the resolution. Thus, for highly resolved surveys a flight altitude of about 150 ft. would be selected resulting in ground resolution of about 75 ft. In the situation where the system was being used to direct the passage of a ship through areas of minimum ice thickness it is possible that an onboard helicopter would be used and that the helicopter navigation system would consist of a radar link with the ship which supplied polar coordinates of the helicopter with respect to the ship. It would be quite feasible to telemeter the data back to the ship and to plot profiles of the ice thickness in real time so as to have up-to-date information about actual ice conditions. The annual mean maximum and minimum ice conditions are shown in Figures 5 and 6, where Figure 7 is a locating map with place names (3).

DETECTION OF PERMAFROST IN THE DISCONTINUOUS ZONE WITH E-PHASE

Physical Properties of Permafrost

In Canada the permafrost region which underlays about the total land area is divided into two zones, based on the arbitrary selection by Russian permafrost investigators of the -5°C isotherm of mean annual ground temperature, measured just below the zone of annual variation. Regions to the north of this imaginary line are referred to as regions of continuous permafrost, those to the south as regions of discontinuous permafrost. In the continuous zone permafrost occurs everywhere beneath the ground surface except in newly deposited unconsolidated sediments where the climate has just begun to impose its influence on the ground thermal regime. The thickness of permafrost is about 200' at the southern limit of the continuous zone increasing steadily to more than 1,000' on the northern part of the zone as Figure 8 shows. The active layer generally varies in thickness from about 1 1/2 to 3 feet and usually extends to the permafrost table. The temperature of the permafrost in this zone, at the depth at which annual fluctuations become virtually imperceptible (i.e. less than about 0.1°F), known as the level of zero annual amplitude, ranges from about 23°F in the south to about 5°F in the extreme north.

In the discontinuous zone, frozen and unfrozen layers occur together. In the southern fringe of the zone, permafrost occurs in scattered islands a few square feet to several acres in size and is confined to certain types of terrain, mainly peatlands as in Figure 9. Other occurances are associated either with the north-facing slopes of east-west oriented valleys, or with isolated patches in forested stream banks, apparently in combination with increased shading from summer thawing and reduced snow cover. Northward it becomes increasingly widespread in a greater of variety of terrain types. Permafrost varies in thickness from about two feet to ten feet or more depending on local climatic and surface terrain conditions. The active layer does not always extend to the permafrost table. The temperature of the permafrost in the discontinuous zone at the level of zero annual amplitude generally ranges from a few tenths of a degree below 32° F at the southern limit to 23°F at the boundary of the continuous zone.

Civil Engineering Constraints of Permafrost

The hazards associated with building structures such as pipelines, buildings, roads etc. through regions of discontinuous permafrost are dis-

cussed in detail by Brown (5). Under low temperature conditions permafrost exhibits relatively high strength and hardness, which can be attributed in part to the cementing action of the ice which binds the soil particles into a solid mass. The mechanical properties of frozen ground, in which ice fills some or all of the interstitional space between soil grains, therefore tend to approach those of ice. The strength of frozen ground increases with decrease of temperature and, in general, with increase in moisture (ice) content. For some soils, such as clays, the increase in strength is relatively small at temperatures just below freezing, mainly because of the amount of unfrozen water in the material. Frozen sands that are well cemented by ice usually have considerably greater strength than fine grained materials, particularly at temperatures nearing thawing. Although frozen soil provides excellent bearing for a structure, its strength properties are greatly reduced with increase in temperature and, if thawed, may be lost to such an extent that it will not support even light loads. The most serious difficulties arise with those soils, usually fine grained, that have large moisture (ice) contents. When thawed these materials turn to a slurry with little or no strength, and large settlements or perhaps failure of a structure may occur.

It is obvious that, as the north is developed, techniques must be established for quickly and accurately surveying the distribution of permafrost in the discontinuous zone. Fortunately, one of the useful properties of permafrost is that its electrical resistivity is usually substantially higher than the resistivity of the corresponding unfrozen ground.

Electrical Resistivity of Permafrost

The magnitude of the resistivity change is variable and is the subject of considerable study. For temperature above freezing the ionic conductivity of soils depends upon the mineral composition, the amount of water and the salinity of the water. As the temperature of the soil is decreased below freezing, the fraction of free water to frozen water decreases with a consequent increase in the sample resistivity which depends on this ratio. At a given temperature the fraction of unfrozen to frozen water depends on the porosity, pore dimensions, salinity, and the mineral compositions of the soil. The resistivity of pure ice is very high. Figure 10 from Ogilvy (6) summarises this behaviour; it is seen that the relative change in resistivity upon freezing is much greater for saturated sand than for clay, and that the change of resistivity for the saturated sand occurs over a much smaller temperature range. The change of resistivity of the clay sample illustrated in this Figure is in agreement with the results shown by Hoekstra (7), whose measurements on frozen silt are shown in Figure 11, where it is seen that the effect of freezing is to reduce the conductance by approximately a factor of 10.

Thus, although for very fine materials such as clay and silt the resistivity change on freezing is not extremely large, for the average soil it appears that the resistivity change upon freezing will be moderately great and that the measurement of this resistivity will assist in determining the existence of permafrost.

It should be noted that there will always be some ambiguity since in some cases the resistivity of frozen soils consisting of very fine particles will overlap the resistivity of coarser material in the unfrozen state, especially if the latter is in a region where the ground water is relatively conductive (although the Russian results shown in Figure 11 do not indicate such an overlap). It is, at any rate, probable that a combination of airborne resistivity mapping and aerial photograph interpretation will suffice in many cases of interest to outline areas of permafrost in the discontinuous region.

Application of E-PHASE to Location of Permafrost

The use of the Barringer E-PHASE system to measure ground resistivity has been previously described and the results of a typical survey carried out at VLF frequency are shown in Figure 3, where the survey data has been presented in the form of contours of apparent resistivity.

Now at VLF frequencies (15 to 25 kilohertz) the penetration of radio waves into earth of average resistivity is relatively large and increases with resistivity as shown in Figure 12. Thus, for soils having resistivity of the order of 1000 ohm-meters the penetration depth of radiation is 100 meters and the resistivity is actually sensed to approximately that depth. Since the resistivity increases with freezing, the effective penetration will be even larger for those areas where there is permafrost. For many applications penetration of this order would be considered too large. If for example, the reason for determining the permafrost distribution is connected with the construction of highways, building sites, pipelines etc. it is unlikely that the average depth of penetration should exceed 40 to 50 ft. Fortunately, the penetration depth for plane waves is a function of the frequency of the radiation and by utilizing a radio station

having an appropriate frequency it is possible to control the effective depth. The obvious choice to achieve shallower penetration are radio stations operating at standard broadcast frequencies. For example, a station operating at the top end of the broadcast band would have, for a given resistivity, a penetration of about 10% of that which would occur at VLF frequencies. Thus, at 1000 ohm-meters this depth would be of the order of 10 meters or 30 or 50 ft. This is obviously of the right order of magnitude and such broadcast stations would make ideal transmitters for location of permafrost in the discontinuous zone. In fact, one would, in all probability carry out the survey flying two receivers simultaneously, since this is quite feasible, and a good deal of additional information would be yielded by such a technique. In essence the two layer resistivity case, so familiar to geophysicists, would become partly resolved by this approach. Resistivity highs which appeared only on the broadcast band E-PHASE would have a small probability of exceeding a certain depth, whereas those which appeared only on the VLF system occur at a considerable depth and do not reach the surface. An attractive feature of E-PHASE is that, having presented the data in the form of apparent resistivity, it is possible to use the results of Figure 13 to give some indication, given the apparent resistivity as to the depth of penetration and thus the depth to which the resistivity is being sensed.

There is one disadvantage associated with the use of broadcast band transmitters in the far north and that is, of course, that they are relatively sparsely located and unlike VLF stations described in earlier sections which can be received at distances of thousands of miles from the transmitter, the radiation from broadcast band transmitters falls to small values at relatively short distances from the transmitting antenna, of the order of 100 miles or so, depending on the transmitter power and ground resistivity. This should not present a serious problem for it is quite feasible to supply one's own transmitter, situated at the edge of the survey site. By placing the transmitter close to the survey area a simple, temporary antenna will suffice, and the radiated power need not be large. An additional advantage of such a technique is that, within the limits imposed by the Dept. of Transport, the frequency can be chosen to optimize the penetration depth. From such a simple antenna and low power transmitter it would still be possible, at frequencies within the vicinity of the broadcast band, to survey many hundreds of square miles without having to relocate the transmitter site.

Table 1 lists the transmitters in the Yukon Territory and Northwest Territories (similar information is available for the western provinces and presumably also for the State of Alaska), giving their frequency, call sign, power output and an estimated figure for the maximum range from the station at which the survey can be performed. This last quantity is determined by the field strength which is in turn a function of the transmitter power, transmitter antenna configuration, frequency and the average conductivity of the soil between the transmitter antenna and the survey site. Because of the influence of these variables it is impossible to give an accurate figure for the field strength and for this reason the range is given in Table 1 and are estimates only. Essentially all of the transmitters listed in Table 1 are operated by the Canadian Broadcasting Corporation as remote relay transmitters and measured field strength data is available once a specific site has been selected. It will be observed that there is a good selection of broadcast transmitters down the McKenzie River and that a 1 kilowatt station is located in Inuvik. This transmitter will provide more than adequate coverage of the McKenzie River delta.

LOCATION OF CONSTRUCTION MATERIALS USING E-PHASE

Physical Properties

The normal electrical properties of soil encompass a wide range of resistivity. The most conductive materials usually have a large clay content wherein ion exchange effects tend to enhance the conductivity, resulting in resistivities of only tends of ohm-meters. On the other hand under ordinary circumstances the most resistive materials are sands and gravels. For these materials the actual resistivity of the substance itself may be very high indeed, in which case the apparent resistivity of sand or gravel in its matrix is determined by the resistivity of the included water, which in turn is a function of the dissolved salts etc. Nevertheless, sand and gravel usually appear at the most resistive part of the spectrum, as shown in Figure 14 from Eliassen (1) which shows gravel to generally be in excess of 1000 ohm-meters. This figure is extremely interesting because it represents a statistical summary of a large number of "wave-tilt" measurements carried out in Norway in order to determine the resistivity of various environments. The results are particularly pertinent to this paper for several reasons: first, the environment in Norway is probably somewhat representative of that to be found in many regions of Canada. Secondly, the resistivities were actually determined using the "wave-tilt" method and as already described the E-PHASE

technique is essentially a measurement of wave-tilt. From the figure it is seen that sand and gravel show the lowest conductivity of the unconsolidated sediments, although there is, as is always the case with geophysical measurements, some overlap. It is apparent from this table that measurement of electrical resistivity will materially assist in the location of sand and gravel deposits. Indeed the use of resistivity to locate gravel deposits in the American Mid-West has been reported by Wilcox (8) nearly twenty years ago. As in the case of permafrost it is probable that data from an airborne resistivity survey would be enhanced considerably by the addition of aerial photography.

The resistivity of sand and gravel depends to a certain extent on the environment in which the aggregate is located. Thus, for regions which are essentially clay-free and in which the ground resistivity is very high, one would expect the resistivity of aggregate deposits to be very high, probably in excess of several thousand ohm-meters. In contrast Wilcox's measurements were made in lacustrine clay; in this environment apparent resistivity values in excess of a few hundred ohm-meters are an almost certain indication of sand and gravel. This is yet another example of the situation where geophysical airborne measurements must originate from a region about which a certain amount is known in order to lend some guidance to the interpretation of surveys carried out in completely new areas. Once confidence in the technique has been built up it is possible to reduce the required ground truth to a minimum.

E-PHASE Survey Results

The results of a photogeological interpretation of an area flown with a VLF E-PHASE system (survey results shown in Figure 3) are shown in Figure 13. The two figures illustrate a small section of a survey carried out in southern Ontario. The legend for the photogeology is as follows:

1.. Gravel observed on the surface; 2+.
2. Most probable areas for gravel deposits; 2.
3. Intermediate unit between 2 and 3; 3+.
4. General morainic deposit; 3.
5. General morainic deposit with a possibility of the inclusion of outwash granular deposits, 4.
6. Till and glacio-lacustrine sand and silt; 5.

The E-PHASE contours are those of "apparent resistivity" in ohm-meters where "apparent resistivity" is defined as the resistivity that a uniform half-space would have to have yielded the same value of electric field that was actually measured. It is to be expected, from measurements described in the literature that in this environment the resistivity of granular materials such as sand and gravel would exceed 2000 ohm-meters. Therefore, in the figure all those anomalies which have an apparent resistivity of 2000 ohm-meters or greater and which coincide which photogeological areas interpreted as being of rank 2 or greater are designated as being of rank I, IIa, or rank IIB.

In comparing the two figures it is immediately apparent that the anomalous E-PHASE regions coincide in a great many cases with photogeologically favourable regions. Conversely, the entire southwestern portion of the region is photogeologically uninteresting as is corroborated by the contours. E-PHASE in many instances connects isolated regions which have been designated as being of rank 2 or 2+ photogeologically. See, for example, the extreme northeastern area where a predominant regional strike of approximately northwest-southeast is suggested by E-PHASE, and the center of the area where various photogeological areas of rank 2 are indicated by E-PHASE to be part of a structure striking roughly east-west.

Seen overall, it is evident that E-PHASE, by virtue of the fact that resistivity is being sensed down to depths of the order of hundreds of feet tends, to tie in the physical characteristics of the region in a manner that is not achieved with the photo-geology, and that application of E-PHASE to a photo-interpretation adds a new dimension to the interpretation.

Again, in view of the fact that at VLF frequencies the effective penetration depth is of the order of hundreds of feet and it is probable that sand or gravel found at this depth is of no economic value, it is obvious that such surveys should be carried out with the simultaneous application of E-PHASE systems operating at standard broadcast and VLF frequencies. The former will actually delineate the areas that are most likely to be sand or gravel of economic interest, the latter will give continuity to these results and will indicate the much deeper structures that are in many cases controlling the shallow surficial deposits.

In this application the depth of penetration of the VLF waves would have been too great to detect surficial deposits. Since the depth of penetration of plane waves in a half-space is inversely proportional to the square root of the frequency the obvious way to overcome such a limitation is to the E-PHASE technique at higher frequencies. The use of Prairie broadcast band AM radio stations

results in penetrations only one-eighth of those achieved with VLF frequencies. Consequently the Saskatchewan survey (and others subsequent) used a dual band VLF-Broadcast E-PHASE receiver, to obtain complementary near surface deep logic information.

Figure 15 is a block diagram of the dual frequency E-PHASE receiver concept. The antenna mount contains the vertical dipole antenna for the E_z signal pick-up and preamplification for both VLF/Broadcast frequencies. The mount also contains the horizontal dipole for the E_x signal pick-up and preamplification at both VLF/Broadcast frequencies. The antenna mount is fabricated of some suitable insulating material such as fibre-glass.

Within the main receiver console the E_z reference signal is amplifier and supplied as the reference to a phase lock loop which locks on to the signal. The E_x signal is detected in-phase quadrature with respect to E_z. Outputs of E_z (VLF and Broadcast) and E_x (VLF and Broadcast) quadrature are provided for subsequent analog recording.

In all of these test surveys carried out in the mid-west and Ontario the E-PHASE results demonstrated the importance of looking below the surface by depicting and delineating structures, trends and gravel deposits (subsequently proved out by ground follow-up coring, not recognized in the initial normal photo-interpretation.

REFERENCES

1. Eliassen, K. E. A Survey of Ground Conductivity and Dielectric Constant in Norway Within the Frequency Range 0.2-10 Mc/s. Geofysiske Publikasjoner (Oslo) Vol. XIX No. 11, pp 1-30.

2. Barringer, A. R., McNeill, J. D. The Airborne RADIOPHASE™ System - A Review of Experience. Paper presented at the Canadian Institute of Mining and Metallurgy. 72nd General Meeting Toronto, April 20-22, 1970.

3. Breslau, L. R., Johnson, J.D., McIntosh, J. A., Farmer, L.D., 1970. Environmental Research Relevant to the Development of Arctic Sea Transportation. Marine Technology Society Journal. Vol. 4, No. 5, pp 19-43.

4. Wait, J. R., 1962. Electromagnetic Waves in Stratified Media. Pergamon Press, 1962.

5. Brown, R. J. E. Permafrost in Canada, Chap. 1, 2. University of Toronto Press.

6. Ogilvy, A. A. Geophysical Studies in Permafrost Regions in the USSR. Geological Survey Canada Economic Geology Report No. 26, Mining and Groundwater Geophysics/1967. p. 641.

7. Hoekstra, P. The Physics and Chemistry of Frozen Soils. U.S. Army Terrestrial Sciences Center, Hanover, New Hampshire, Oct. 1968 (Conference Preprint).

8. Wilcox, S. W. Sand and Gravel Prospecting by the Earth Resistivity Method. Geophysics (1944) Vol. 9, pp. 36-46.

TABLE 1

LOCATION		CALL SIGN	POWER (WATTS)	FREQUENCY (KHz)	USABLE RANGE *(MILES)
Beaver Ck.	YT		40	690	10 - 25
Carmacks	YT		40	990	10 - 20
Clinton Ck.	YT		40	990	10 - 20
Dawson	YT	CBDE	40	560	10 - 25
Destruction Bay	YT		40	940	10 - 20
Elsa	YT	CBDD	40	560	10 - 25
Haines Jct.	YT	CBDF	40	860	10 - 20
Mayo	YT	CBDC	40	1230	10 - 15
Swift R.	YT		40	970	10 - 20
Teslin	YT		40	940	10 - 20
Watson L.	YT		40	990	10 - 20
Whitehorse	YT	CFWH	1000	570	50 - 100
Fort Good Hope	NWT		40	920	10 - 20
Fort Norman	NWT		40	920	10 - 20
Fort Providence	NWT		40	1230	10 - 15
Fort Resolution	NWT		40	1150	10 - 20
Fort Simpson	NWT	CFMR	25	1490	10 - 15
Fort Smith	NWT	CBDI	40	860	10 - 20
Frobisher	NWT	CFFB	250	1200	25 - 50
Hay R.	NWT	CBDJ	40	1490	10 - 15
Inuvik	NWT	CHAK	1000	860	50 - 100
Norman Wells	NWT		40	990	10 - 20
Pine Point	NWT		40	880	10 - 20
Wrigley	NWT		40	1280	10 - 15
Yellowknife	NWT	CFYK	1000	1340	25 - 50

*as discussed in the text the value for usable range is a very coarse estimate

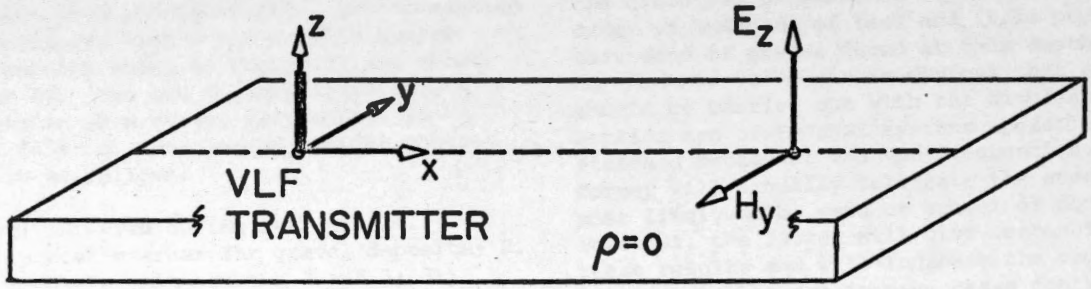
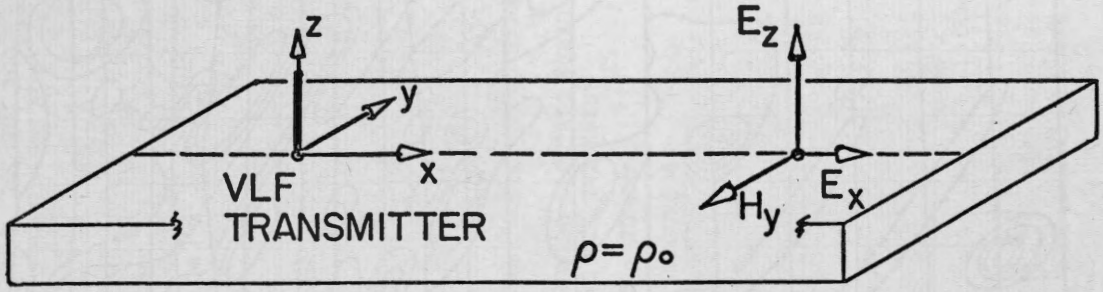


Figure 1 Field Vectors-Lossless Earth



$$\frac{E_x}{E_z} = (iw\epsilon_o\rho)^{\frac{1}{2}} \qquad (1)$$

$w = 2\pi f$ where f = frequency (Hz)

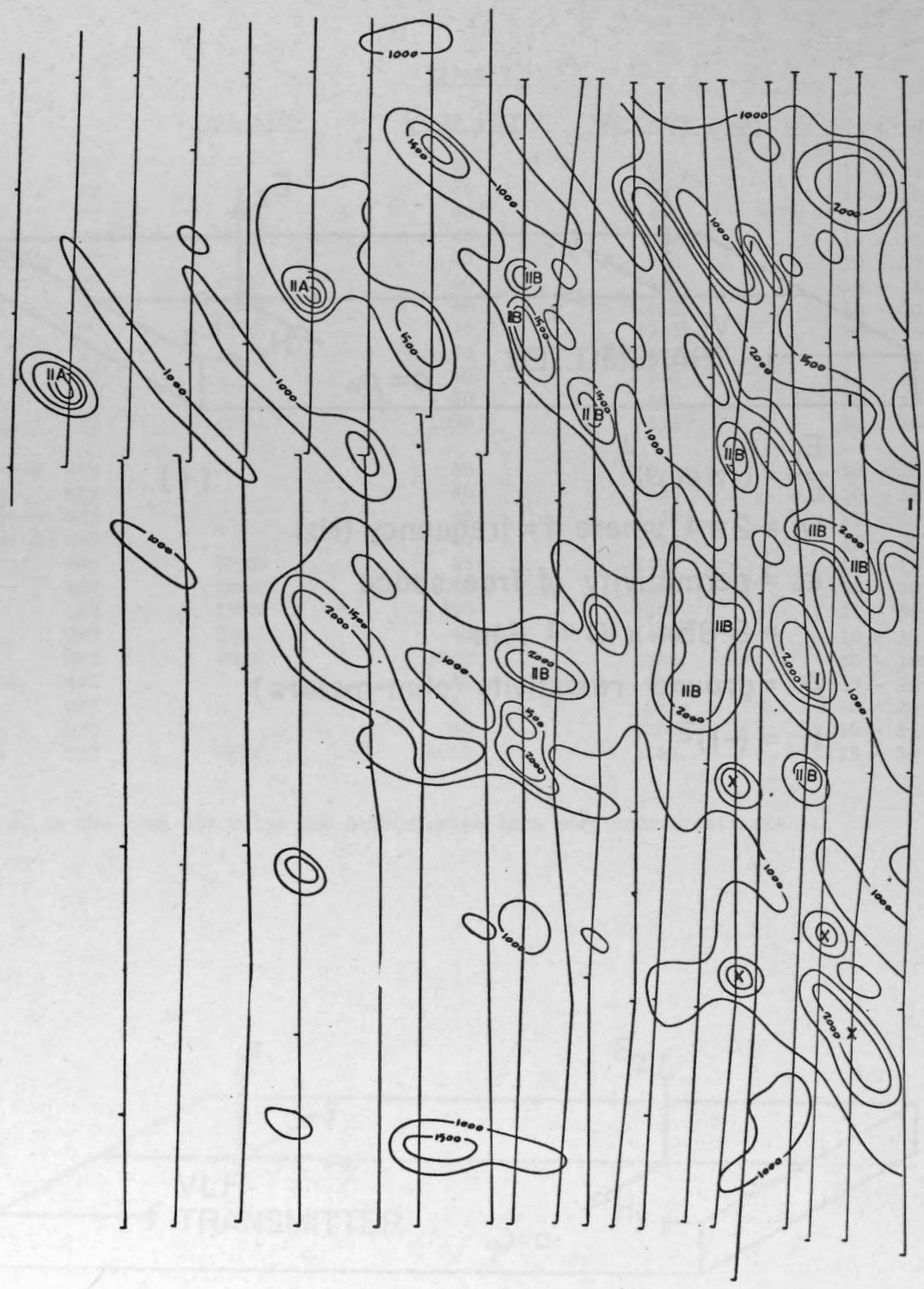
ϵ_o = permittivity of free space

$= 8.854 \times 10^{-12}$ F/m

ρ = ground resistivity (ohm-meters)

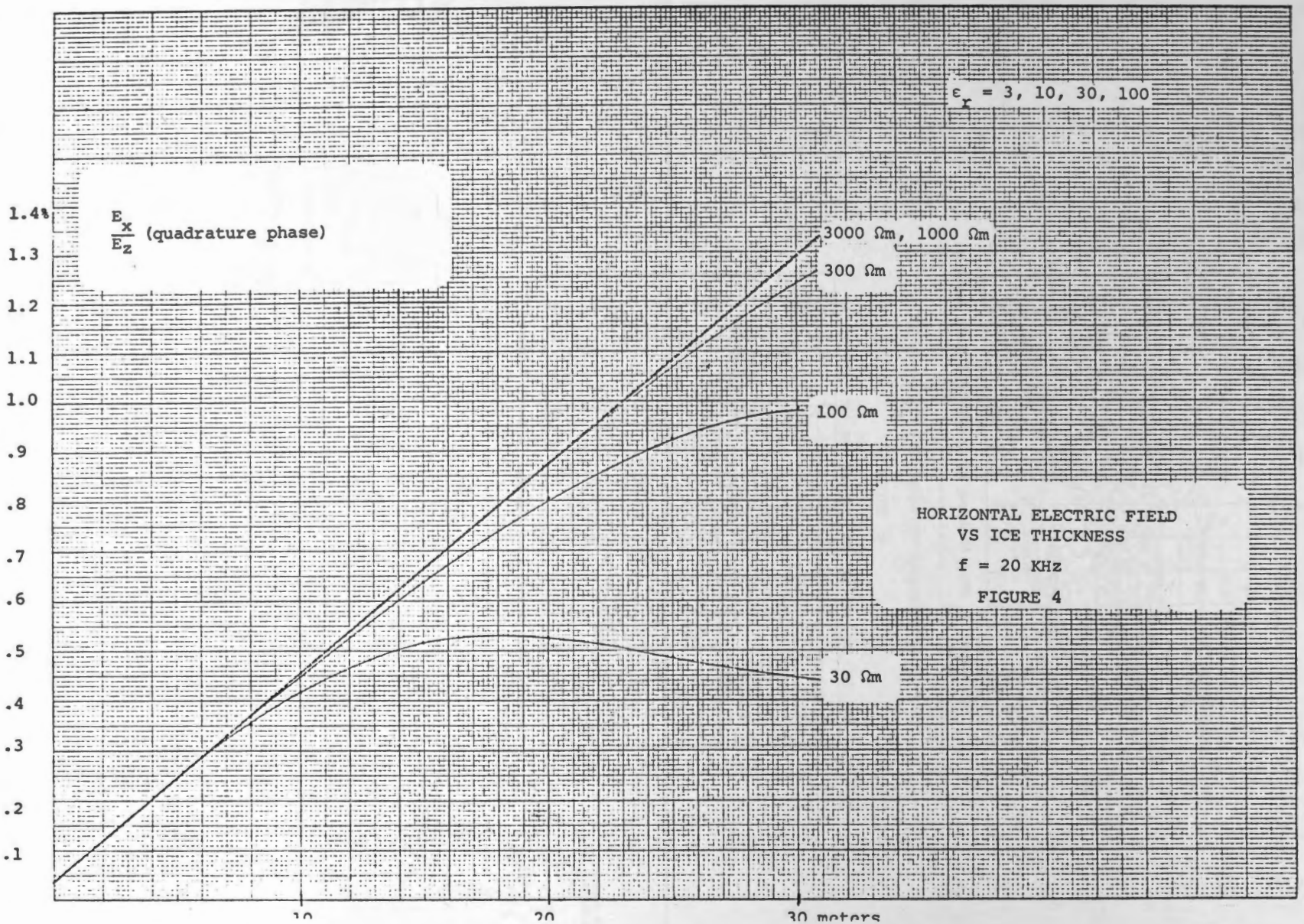
$i = (-1)^{\frac{1}{2}}$

Figure 2 Field Vectors-Lossy Earth



E-PHASE™ SURVEY CONTOURS - OHM METERS

Figure 3



HORIZONTAL ELECTRIC FIELD
VS ICE THICKNESS

f = 20 KHz

FIGURE 4

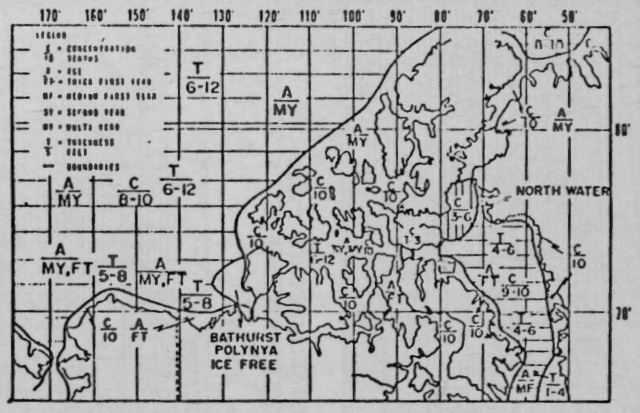


Figure 5 Mean Maximum Ice Conditions in North American Arctic. (3)

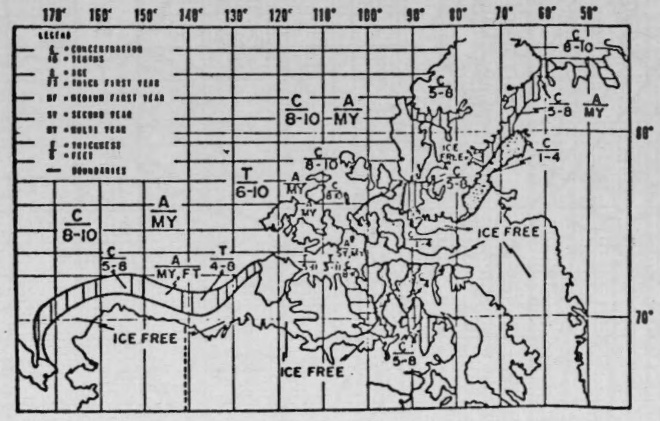


Figure 6 Mean Minimum Ice Conditions in North American Arctic. (3)

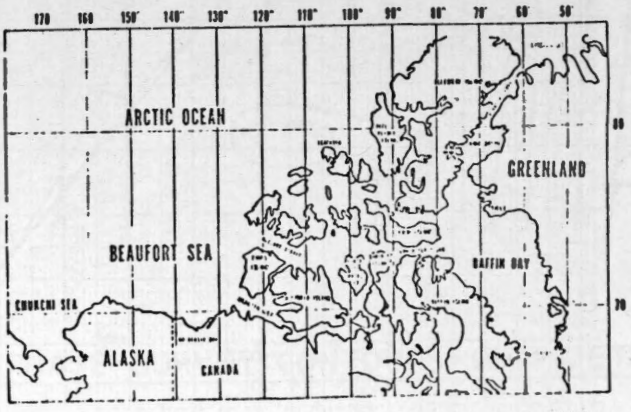
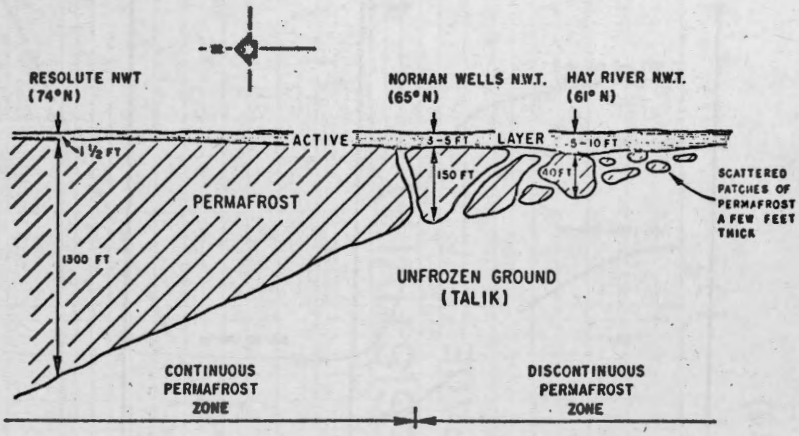
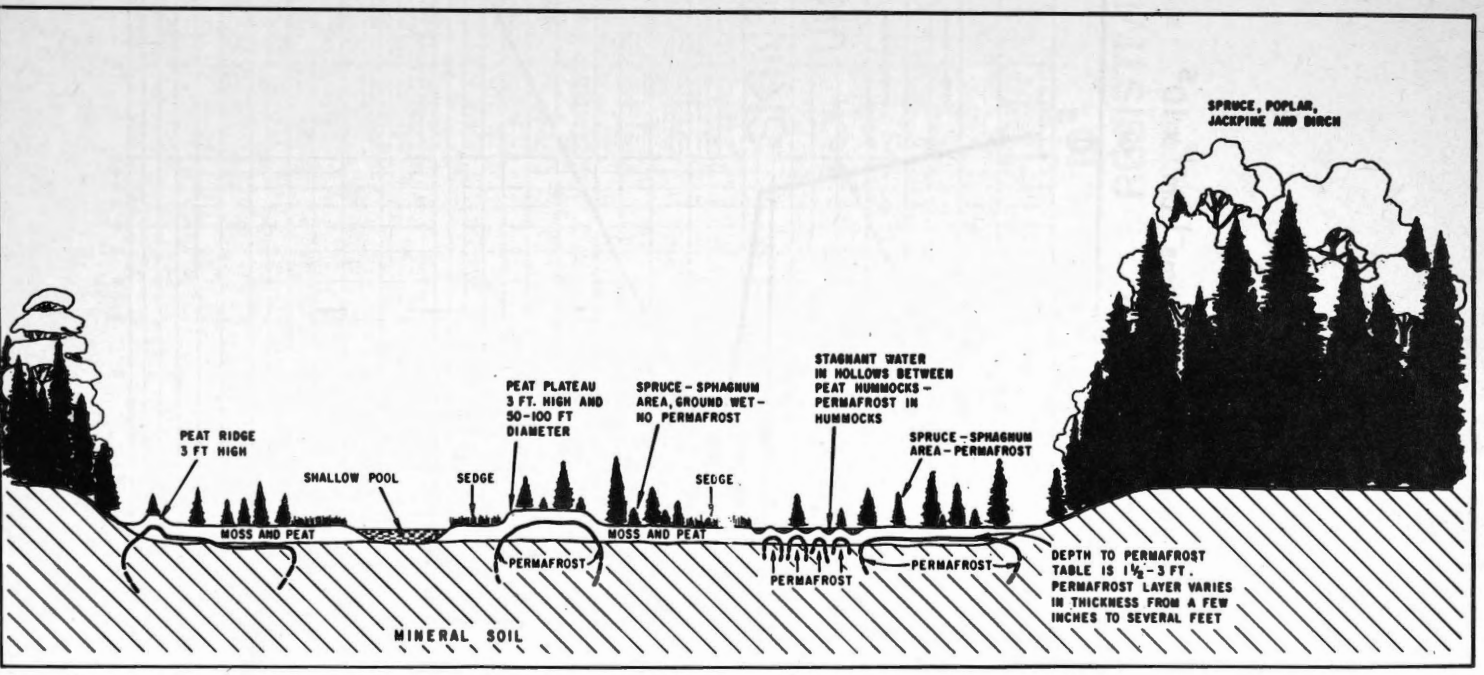


Figure 7 Place Name Chart of North American Arctic. (3)

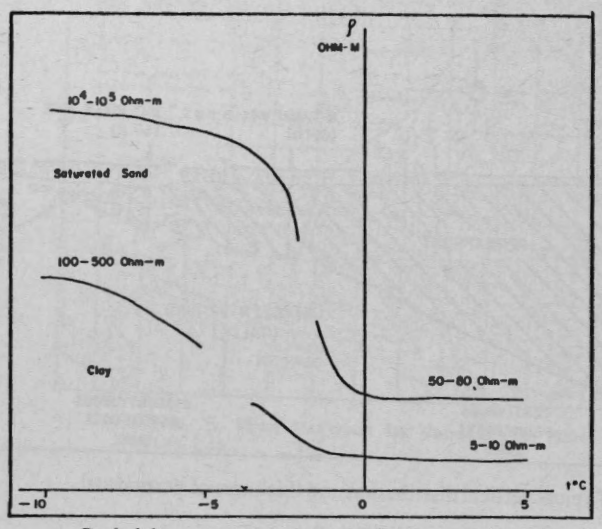


Typical vertical distribution and thickness of permafrost.

FIGURE 8



GURE 9 Profile through typical peatland in southern fringe of discontinuous zone showing interaction of permafrost and terrain factors.



Resistivity versus temperature for saturated sand and clay. (6).

FIGURE 10

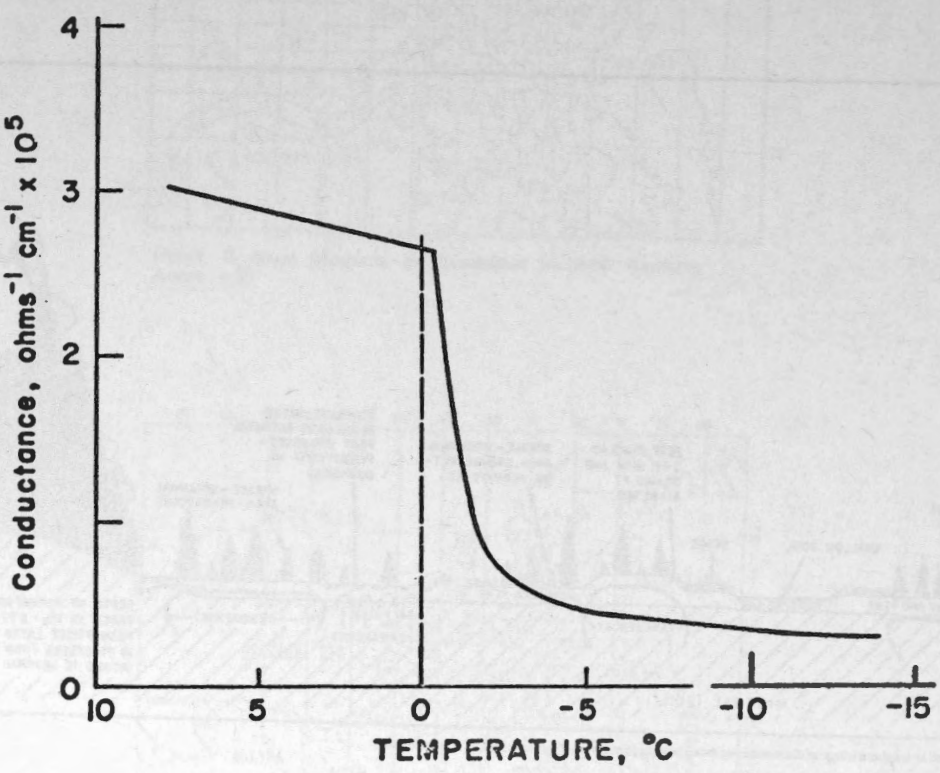


FIGURE 11

The conductance of a frozen silt as a function of temperature. (7)

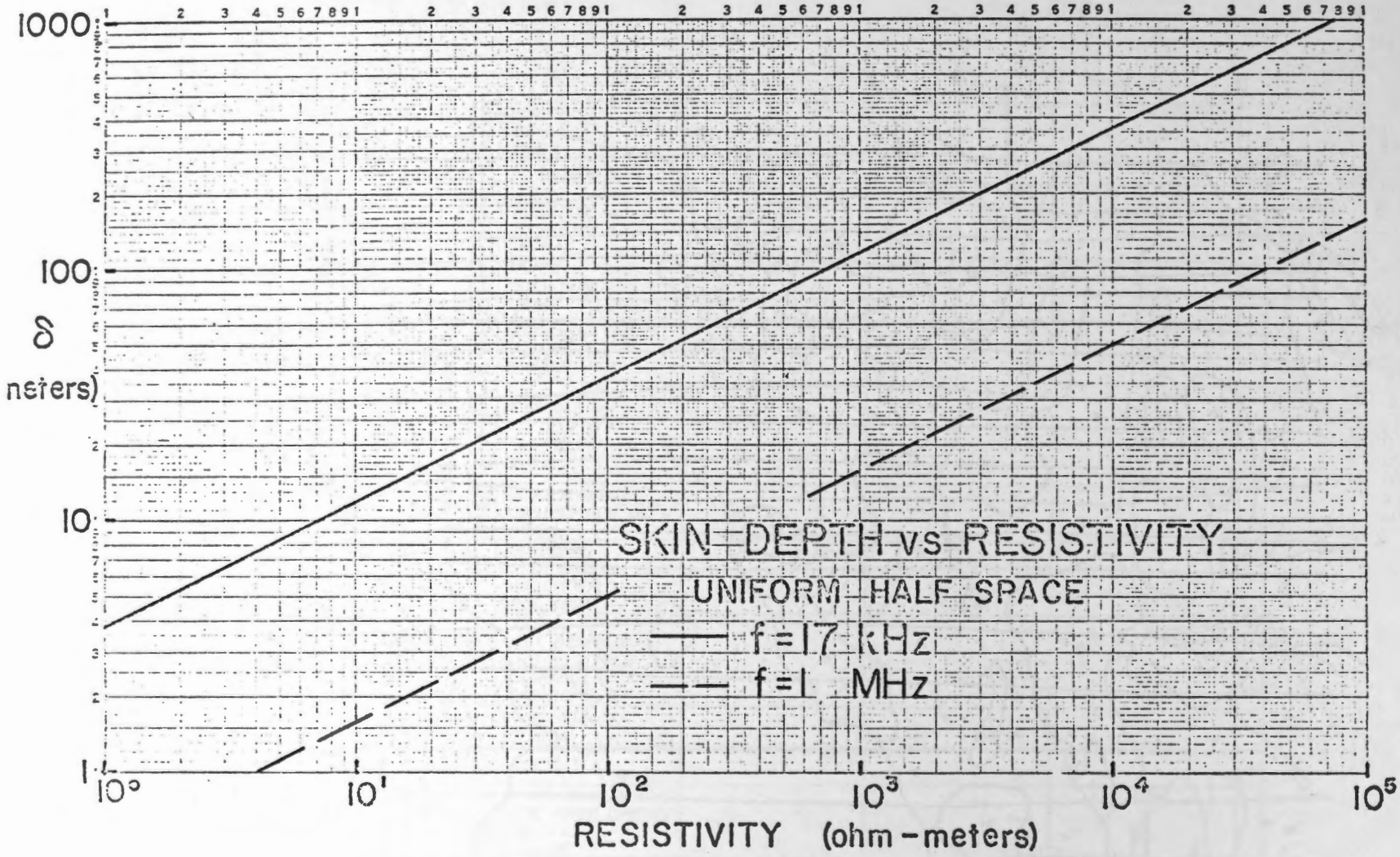

1000
100
δ
meters)
10
1
SKIN DEPTH vs RESISTIVITY
UNIFORM HALF SPACE
f = 17 kHz
f = 1 MHz
10⁰
10¹
10²
10³
10⁴
10⁵
RESISTIVITY (ohm-meters)

FIGURE 12

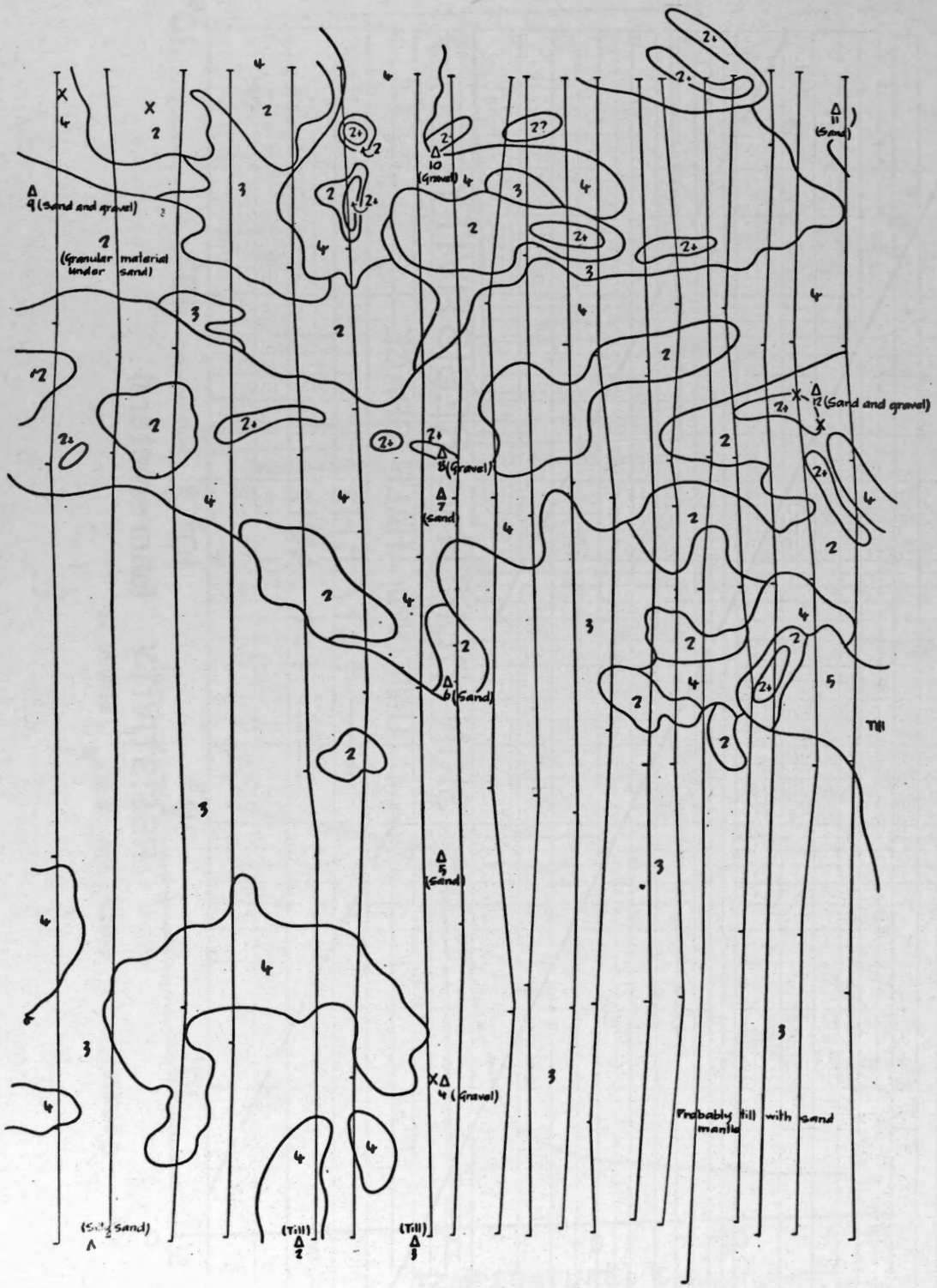
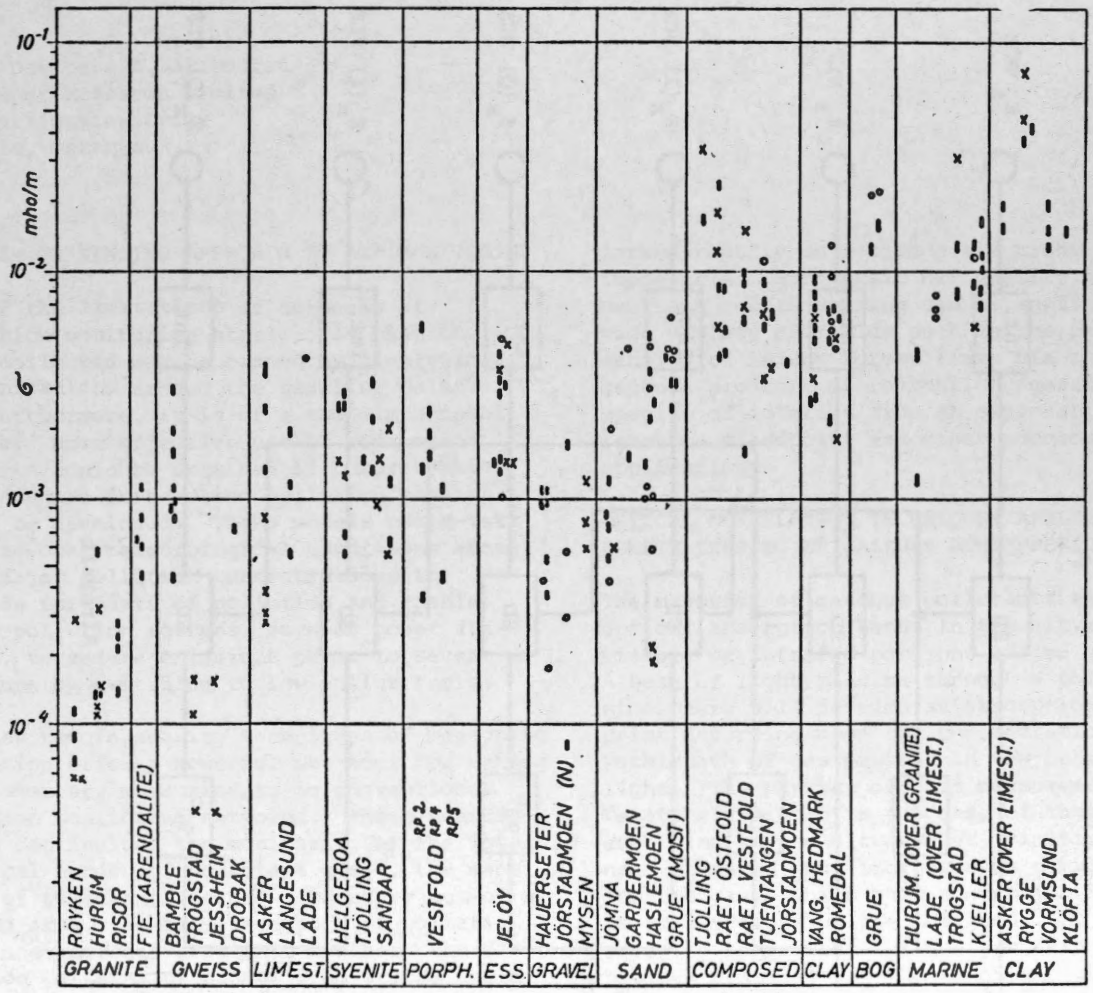


PHOTO-GEOLOGICAL INTERPRETATION OF E-PHASE™ SURVEY AREA

Figure 13



Plots of effective conductivity measured within fairly uniform areas of different ground.

Figure 14

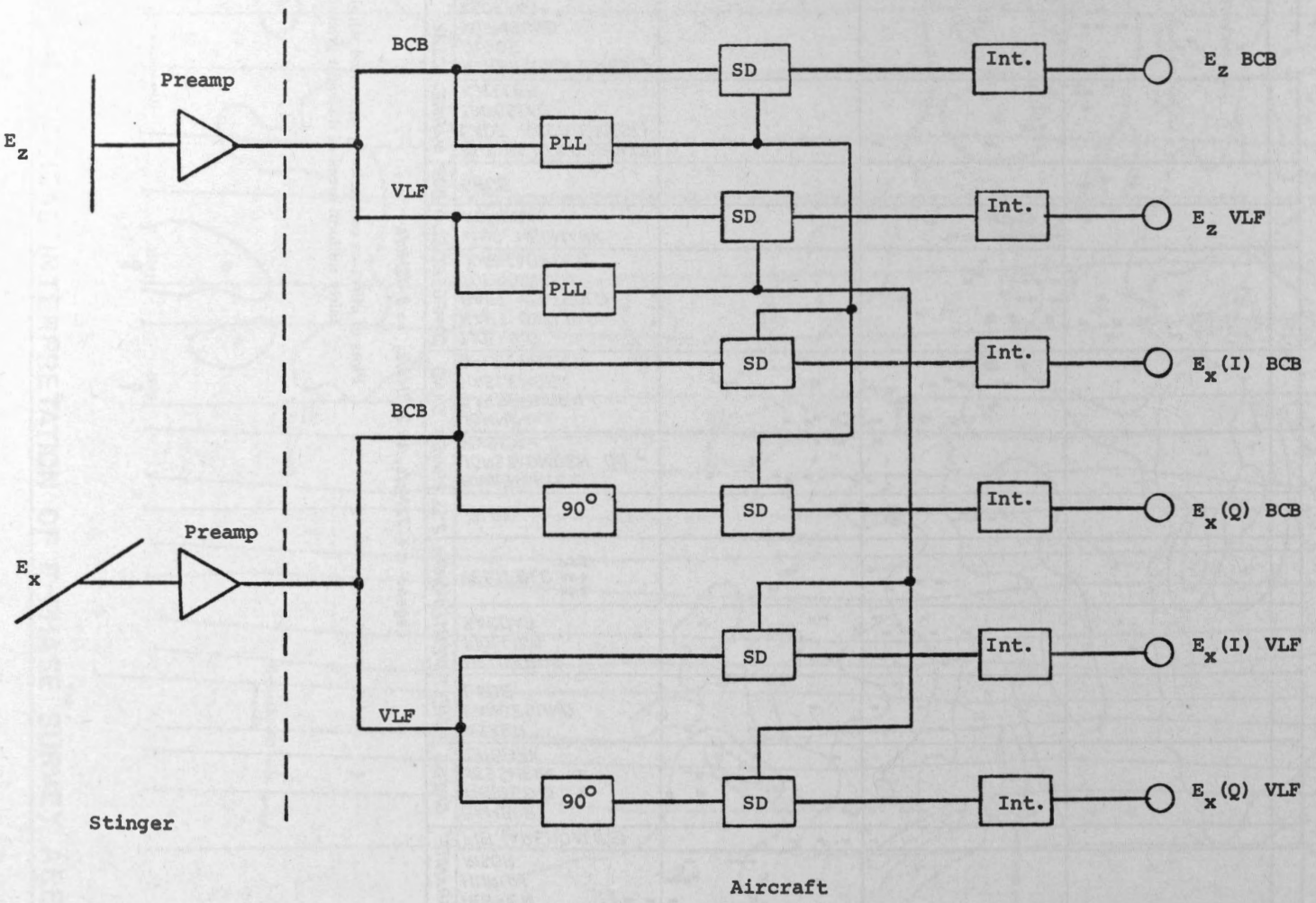


FIGURE 15

MULTI-BAND E-PHASE BLOCK DIAGRAM

DETECTION OF ATMOSPHERIC CONSTITUENTS

J. H. Davies & A. J. Moffat
Barringer Research Limited
304 Carlingview Drive
Rexdale, Ontario

THE REMOTE SENSING APPROACH TO AIR MONITORING

One of the limitations of networks of pollution monitoring stations is that the data collected may be biased by highly local concentrations around the sampling point, and furthermore, it is of a two-dimensional nature. More effective use of monitoring networks could be obtained if comprehensive and accurate forecasting pollution models could be developed. These models could take into account meteorological conditions along with major pollution concentrations to provide forecasts of pollution and enable major pollution sources, such as power stations, to reduce emissions prior to severe buildups by switching to low sulfur fuels.

The new remote-sensing techniques of measuring pollution offer a powerful new tool for making complementary measurements to conventional pollution monitoring networks. These measurements can include the monitoring of the total vertical burden of pollutant gases, the mass flow of pollutants across a boundary out of a closed area, the mass transport of pollution from one regional area into another, the dispersion patterns of plumes from stacks and a variety of other measurements of interest. Many of these measurements have three-dimensional characteristics and are particularly well suited for integration with point sampling measurements to assist in the development of pollution models which can be used in forecasting. Furthermore, remote-sensing techniques are not only applicable to use on the ground in both stationary and mobile fashion, but also to airborne surveys of pollution and even to spacecraft monitoring. It is therefore suggested that there is an important role to be played by remote-sensing techniques in the control of pollution and that the study and further development of these techniques is very worthwhile.

The development of such remote-sensing techniques for detection of air pollution, renders the whole field of atmospheric constituent and air quality open to new measurements performed rapidly on a time scale hitherto impossible. Such basic air quality measurement and remote sensing can be applied to a wide variety of fields such as the remote sensing of latent forest fires via their gaseous products of combustion, gaseous species of interest from an exploration earth resource standpoint and other numerous applications.

OPTICAL CORRELATION TECHNIQUES APPLIED TO REMOTE SENSING OF GASEOUS ATMOSPHERIC POLLUTION

The majority of gaseous pollutants exhibit optical absorption bands in the ultraviolet, visible or infrared portions of the spectrum. A beam of light passing through a polluted atmosphere will develop an absorption fingerprint depending upon the concentration and Pathlength of gas present in the beam of light. The physics of this measurement is relatively simple in the case of those gases absorbing in the ultraviolet, visible or near infrared, but increases in complexity for the gases which show absorption in the thermal regions of the infrared. In the latter case the temperature of the gas being measured becomes critical and when high-altitude measurements are being made, the effects of pressure also have to be taken into consideration.

It is possible to measure concentrations of a gas by noting the absorption obtained at one wavelength corresponding to a strong absorption band in the gas and comparing it with the absorption at an adjacent wavelength where the gas does not absorb. This simple technique, however, tends to be subject to interferences due to the fact that absorption bands are seldom unique and it is generally impossible to pick a pair of wavelengths which will not suffer from some differential absorption due to the presence of gases other than the one being sought.

An effective method of combating this interference problem is to correlate a substantial portion of the absorption spectrum of the gas being measured against a stored replica

or mask of the spectrum. The term correlation spectroscopy has been coined to describe this technique and the method will be described in greater detail below.

While correlation spectroscopy is convenient to apply in the ultraviolet, visible and near infrared portions of the spectrum, it becomes more convenient to use interferometric methods in the middle and far infrared. Detectors are far less sensitive in this region and there is a much greater requirement for large light throughput in the detection device. This condition is satisfied by an interferometer and by scanning the path difference in the interferometer and interferogram can be generated which may be converted into a spectrum by means of a fourier transform operation in a computer. In the correlation interferometer, however, the gas correlation is carried out directly in the interferometer against a stored replica of the interferogram of the gas being detected. The intermediate fourier transform step is omitted thereby greatly simplifying the approach. This is possible due to the fact that the fourier transform of a gas spectrum is as unique as is the spectrum itself.

CORRELATION SPECTROSCOPY

The techniques and applications of correlation spectroscopy have been under development for several years and have been summarized in several publications (ref. 1,2,3,4).

The name of correlation spectroscopy arises from the fact that the main observable chosen as originator of the electrical output is the variation of the correlation function between a mask or array of exit slits situated in the exit plane of a spectrometer and the dispersed power spectrum produced in the spectrometer. The correlation function between the dispersed power spectrum of the incoming light and the array of slits is nothing else but the power in watts passing through the mask from the dispersed power spectrum to a photo-sensitive device situated behind the mask.

The technique is based on the Beer-Lambert Law of absorption;

$$H(\lambda) = H(\lambda) \exp(-\sigma(\lambda) n.x)$$

where

$H_o(\lambda)$ spectral irradiance of the light source ($W\ cm^{-2}\ .nm^{-1}$)

$\sigma\ \lambda$ absorption coefficient of the gas of interest (cm^2 $molecule^{-1}$)

n - number of absorbants per cubic centimeter (cm^{-3} molecule)

x - pathlength (cm)

The first correlation spectrometer constructed termed COSPEC I utilizes a f/3.60 Ebert Fastie spectrometer configuration of 1/4 meter focal length.

The sensor optical system is shown schematically in Figure 1. The field defining fore-optics are comprised of the mirrors M_1 and M_2, a field lens and the spectrometer entrance slit. The dispersed (power) spectrum of the incoming radiation is focussed at the exit mask. The angular position of the grating determines which portion of the spectrum is incident on the exit aperture. In order to obtain a time varying signal from the photodetector, the incoming spectral radiant power has to be spatially and/or time modulated; in this system both can be used.

The spatial modulation is achieved by means of two fork-driven refractor plates situated immediately after the entrance slit. Although these plates have continuous motion, there are effectively only two positions of the plates seen by the incoming light beam. For fifty percent of the time each refractor plate sits at a fixed angle to the beam; thus, the beam "jumps" discontinuously between two different paths, producing a predetermined two-position cyclical displacement of the spectrum with respect to the mask. In this system then, the mask in the exit plane remains stationary and the spectrum jumps back and forth an amount. The refractor plates always introduce spatial modulation or jump. Time modulation of the light source is sometimes introduced. Also used, with or without time modulation of the light source, is another system employing spatial modulation. Superimposed on the sudden jumping of the spectrum across the exit mask is a smoothly drifting motion of the spectrum introduced by smooth rotation of the grating. This method is called spectral scanning.

The response of the photodetector in amperes per watt of incident radiation is a function of the voltage across its dynode chain. By maintaining a constant photodetector response in the presence of source intensity variations via an AGC system, an accurate determination of C (concentration) is achieved.

The tuning fork and associated oscillator circuit is located in the viewing unit and provides the basic system clocking standard. The fork determines the rate at which the spectrum is sampled by virtue of the tyne mounted refractor plates and provides all the reference signals to the AGC circuit.

In the most general case, this response can be expressed as:

$$R = \underbrace{\alpha I_1 \{1 - \psi(\xi)\}}_{\text{no-gas offset}} + \underbrace{\alpha I_1 \psi(\xi) \{1 - e^{-(a_2 - a_1)cL}}_{\text{gas signal}}$$

where a_2, a_1 are the average absorption cross sections per ppm-m of the chosen gas determined by the number of slits n of the mask when in position (2) and (1) with respect to the spectrum. $\psi(\xi)$ is the ratio between the power passing through the mask in position (2) to that of position (1) if the chosen gas had not been present. ξ is the coordinate of a representative feature of the mask (completely arbitrary; i.e., it could be the beginning wavelength of the mask in position (1), and is used as the scan coordinate. In this way the factors affecting the response are divided into two main categories, one related to the chosen gas (a_2-a_1), and the other $\psi(\xi)$ dependent on the background, spectral radiance, filter function, and interfering gases.

In normal operation, the no-gas offset is zeroed out, and a gas measurement can be made using the gas-signal term as soon as a suitable calibration is performed.

SO_2/NO_2 POLLUTION MEASURING BALLOON FLIGHT

The bistable correlation spectrometer for SO_2, NO_2 and I_2 was successfully developed and employed in various field measurement programs with such encouraging results that we speculated and subsequently proved the capability of the technique to undertake measurements of SO_2 and NO_2 from aircraft altitudes. An Aerocommander aircraft was specially outfitted with a correlation spectrometer and undertook a series of successful flights over Toronto, Washington, Chattanooga, Los Angeles and San Francisco. Following upon the success of these airborne flight experiments and concurrent theoretical work evaluating the concept of operating correlation spectrometric equipment from high-altitude platforms, it appeared that the easiest way to simulate a satellite flight would be to undertake a high-altitude-balloon flight experiment over a major pollution target.

Thus a balloon project was sponsored jointly by the NASA Manned Spacecraft Center and the Canadian Department of Energy, Mines and Resources and Chicago was selected as the ideal site because of its large population, heavy industrial activity, excellent ground monitoring network, and its proximity to Lake Michigan as a large background area. The prime aim of this experiment was to see how large the SO_2 and NO_2 signals would be when viewed from high altitude, namely 114,000 feet. At this altitude the balloon would be above most of the ozonosphere, which is the upper atmospheric layer of ozone which acts as a powerful absorber of ultraviolet light. When using the balloon platform the signals of the polluting gases are impressed upon the reflected light from the earth's surface, the light having made two passes through the air layer. Therefore, apart from its normal attenuation, the signal is diluted by atmospheric scattering.

For the balloon flight two correlation spectrometers were flown, one measuring SO_2 in the ultraviolet region, the other measuring NO_2 in the blue visible. These two gases had two critical problems in common, Fraunhoffer line interference in the solar spectrum and dilution of their spectral signatures due to atmospheric absorption and scattering. Sulphur dioxide measurements had the added disadvantage caused by the strong absorption of the ultraviolet radiation by the ozonosphere, and the greater scattering of the shorter wavelengths compared to that which takes place in the visible region where the NO_2 was measured. Mathematical models had been developed and computer programs generated to model the instrument's performance to a variety of outside interferences and calculate the optimum mask designs for the balloon spectrometers.

The results obtained during the balloon flight over Chicago prove conclusively the viability of the correlation technique to monitor SO_2 in the ultraviolet spectral region and NO_2 in the blue visible by clearly demonstrating its efficient solar reflected radiation, modified by target gas signatures impressed at the earth's surface, can be obtained at satellite altitudes. Also the characteristics of the signals obtained at the balloon were identical to those theoretically produced by mathematical modelling. Figure 2 shows the data analysis for the SO_2/NO_2 balloon measurements.

All of the work described so far was undertaken using the bistable correlation spectrometer technique which only allowed the monitoring of one gas at any given instance. Obviously greater benefit could be obtained from an instrument which was capable of measuring both SO_2 and NO_2 simultaneously, and after various design studies were completed, a rotation disc dual gas correlation spectrometer was developed (COSPEC II).

COSPEC II PASSIVE SYSTEM

When the instrument is used as a true remote sensing instrument (i.e., using available natural radiance rather than any controlled light source), the value of $\psi(\xi)$ changes and so does $\bar{a}\ (\xi) = a_2(\xi) - a_1(\xi)$.

This fact, which is common to all passive remote sensing techniques regardless of method or spectral region, determines the direction to follow in order to obtain meaning-

ful results from any instrument technique. In our particular instrument, whose output is given by equation (1), it can be seen that by appropriate modelling and instrumental parameter selection, two positions of the grating ξ_1 and ξ_2 can be found for which $\psi(\xi_1) \simeq \psi(\xi_2)$, while the gas signals are different. The difference of readings from the first to the second is used as the physical observable.

The technique is implemented by using directly the two mask pairs involved in the process, thus taking up the grating scan by masks positioning. For each pair, subtracting the value of the output in the two end positions would cancel the no-gas response leaving a component which depends on the characteristics of the chosen gas and on the value of $\psi(\xi)$. By doing this the effect of $\psi(\xi)$ variations is reduced to one term only, and is a scaling factor in this term rather than an additive one. By engraving the mask pairs needed for the obtainment of the response on a disc, the possibility arose of producing a multi-gas sensor as shown in Figure 3. The multi-stable technique has been incorporated in a commercial instrument known as COSPEC II.

This instrument is capable of also monitoring two gases at the same time. Using the principle just described, a split grating and a disc with four pairs of masks is used (two pairs per gas). The instrument is a folded Ebert spectrometer and uses two cassegrain foreoptical telescopes for light gathering. One telescope is a direct recording and the other a maxwellian system thereby providing narrow and wide angle fields of view. The rotational position of the disc (masks) is monitored by logic diodes, and two PM tubes are used to detect the signals. Internal calibration cells can be positioned in front of the PM tubes to provide calibration references.

COSPEC III LONG PATH MONITORING WITH ACTIVE LIGHT SOURCES

One of the problems of conventional monitoring of pollution with equipment that injects air samples, is that the measurement is highly localized and refers to concentrations in the immediate vicinity of the sensor. A new technique which has been developed involves the use of a high intensity modulated beam of light that can be projected over distances of one or two kilometers and received by a specially adapted correlation spectrometer. Using this approach it is possible to measure the average ambient concentration of a pollutant gas along the beam of light as shown in Figure 4. A commercial version of of this system, known as COSPEC III employs a 75 watt high intensity xenon source which is modulated at 2.5 kilohertz. The electronics circuitry in the COSPEC III equipment when used in the active mode is capable of separating out the component of light that is modulated and detecting only the gas signals which are carried on the modulated light.

One of the problems in developing the COSPEC III has been to achieve an extended range in order to make the system truly different from point sampling techniques. This has been achieved with considerable success by the use of phase locking techniques. The light beam is electronically modulated at a highly stable frequency and an auxillary detector is used in the spectrometer which detects the entire spectrum of light being radiated by the xenon source. A local oscillator which runs freely at a frequency which is approximately that of the modulated light source is locked onto the modulated light source using phase lock techniques. Once this frequency has been acquired by the phase lock loop, the local oscillator signal is employed in the signal processing to separate out the modulated component of light from the unmodulated daylight component. This technique of signal processing achieves an effective filtering bandwidth in the system of a fraction of a hertz and enables far greater ranges to be obtained with a given power of light source than would otherwise be possible.

Figure 4 illustrates schematically the long path monitoring approach. Its great advantage is that average concentrations over large areas can be monitored with a single instrument. For example, it is possible to have four remote light sources separated from the monitor in four different directions. The monitor can be used for sequentially scanning each of the light sources, thereby measuring concentrations with a single instrument over an areas which may be four kilometers in diameter. An added feature of the long path monitoring approach is that since an active light source is used, a twenty-four hour monitoring capability results.

CORRELATION INTERFEROMETRY

Correlation interferometry like correlation spectrometry is based upon cross-correlation of incoming signal against a stored replica - in this instance we work with interferograms which are the Fourier transforms of the input spectra. (Ref. 5 & 6) A basic Michelson correlation interferometer is shown in Figure 5. The beamsplitter B provides amplitude division of the input spectra from F which suffers reflections from mirrors M_1

and M_2 to recombine at the detector D. Here C is the compensator plate added by Michelson to balance the two optical arms. By suitable selection of the position of movable mirror M_2, the two beams can be caused to recombine at D in-phase and hence be additive. If M_2 is now moved one quarter wavelength, a path change of half a wavelength results causing the beams to arrive in-phase and hence a minima occurs at B. If the compensator plate is now oscillated about its central position, a cyclic delay is introduced into the one arm thereby unbalancing the interferometer to generate the well known interferogram. Figure 5 shows the interferogram resulting from a single wavelength input. Figure 6 depicts the forms of various interferograms resulting from several spectral inputs - namely a single discrete wavelength, two discrete wavelengths and finally a series of equispaced spectral lines.

Recent progress in correlation interferometry has resulted in the development of the COPE field widened scanning Michelson correlation interferometer for General Electric's NASA-LRC COPE program[7]. This instrument has been developed to remotely measure CO burdens to detect CO anomalies associated with the CO sink mechanism. Figure 7 shows the first overtone absorption spectrum of CO in the 2.3 micron region and its interferogram while Figure 8 shows a block diagram of the general signal processing involved. The interferogram centered on the delay region characteristic of CO is scanned by the oscillating refractor plate. The interferogram is heterodyned down to remove the high frequency interferogram carrier by mixing with a reference signal. The heterodyned signal is sampled and A to D converted for finally processing within a minicomputer. At this final stage correlation functions are applied to reduce the effects of spectral interferences which of course show up as interferogram interferences in the delay domain of the interferometer. The process of correlation can best be visualized as the application of fixed amplitude digits cross-multiplied with the interferometer's output interferogram to normalize its output to represent zero CO gas output when no CO is present within the field of view, regardless of any interfering gases.

Theoretical modelling and atmospheric radiative transfer studies enable the weighting functions to be calculated for various model atmospheres. Subsequently, the instrument can have its in-program weights continually updated through actual field measurements to ensure no-gas output for no-gas input. In our particular COPE breadboard model the interferogram is A to D converted and the weights held within the minicomputer as digital numbers applied to the digitized interferogram.

The advantages afforded by interferometers result from their large throughput, the spectral multiplex advantage, compact yet flexible design possibilities and ready means for incorporation of correlative techniques for electronic processing, the latter obviating one of the major disadvantages of Fourier transform spectrometers, namely that of transforming the Fourier output back to its original spectral form for analytical interpretation.

The technique of correlation interferometry has been considered as a suitable instrumental approach for the detection of latent forest fires by detecting the trace gaseous products of combustion associated with such small fires[8]. Such considerations require a study of the chemical combustions of the various wood species, the chemical thermodynamics of the combustion and its related oxygen supply. For latent fires it is reasonable to assume low temperature combustion processes are those of major concern, and that such latent fires and/or grass fires would develop in reasonably open spaces, albeit covered by a treed canopy. Several workers have evaluated the gaseous products from wood combustion and even experimental forest fires. The two most copious gases emitted are CO and CO_2, which is naturally reasonable to expect and hence close in direct applicability to the same detection technologies we are working with in the COPE program. While never considered before the detection of such latent fires by their gaseous products is not only theoretically feasible but proves on closer examination to be both desireable and pratical.

ATMOSPHERIC PHYSICS

For remote sensing of air pollutants a natural radiation source must be available to provide the background energy upon which the characteristic absorption spectra will be superimposed. This radiation from the earth as a blackbody at 300° and Figure 9 shows this distribution of energy.

Another important consideration in the remote sensing operation is the transmission of the atmosphere in the spectral region of interest and several recognized windows occur. There are advantages and disadvantages to working in the ultraviolet region and this applies also to the infrared region, and dependent upon the type of measurements to be made, tradeoffs have to be performed to ascertain which wavelength

region offers the greater benefits. While a detailed analysis of the advantages and disadvantages of the ultraviolet versus infrared sensing is not possible in this paper, the salient features are:

1. Ultraviolet sensing has the advantage of high solar flux available as a natural energy source, combined with the ability to penetrate the earth's atmosphere which is fairly transmissive in this region, right down to ground level. However, daylight sensing only is possible.

2. The disadvantages associated with ultraviolet remote sensing are the greater atmospheric scattering problems - Raleigh Mie and aerosol. The absorption of ultraviolet radiation by the ozonosphere, and of course loss of ultraviolet radiation if one penetrates too far into the UV because of the fall-off in solar radiation available from the solar photosphere. Below 2800 Å.

3. The advantages of working the infrared region are the greatly reduced atmospheric scattering problems since IR wavelengths are now large in comparison to small particle and molecule diameters. Also IR operations using the earth's thermal radiation as the source give 24 hour capabilities.

4. The disadvantages of the infrared region are the strong interference effects from natural atmospheric species such as water vapor, CO_2, ozone, etc. Also beyond 3.5 microns the thermal emission from the earth is greater than that energy caused by reflected and scattered sunlight and this thermal emission is modified by the thermal structure of the atmosphere. Consequently infrared probing of the atmosphere down to ground level is difficult and in several cases impossible. This severely affects the applicability of these techniques for monitoring air pollutants in the thermal IR region for several important air pollutant species have characteristic spectra (CO_2, SO_2, O_3, etc.).

The ability of any remote sensor to monitor effectively atmospheric constituents demands the capability of monitoring the burdens close to the surface of the earth, which in turn entails the developing of techniques to probe the atmosphere down to ground level. One technique which provides this capability is the ground chopper concept[9].

THE APPLICATION OF GROUND, AIRBORNE AND SATELLITE MOUNTED REMOTE SENSORS TO POLLUTION MONITORING

One of the principal types of measurement that can be made using remote-sensing techniques for pollution monitoring is the estimation of total vertical burden. This measurement refers to the amount of gas that lies in a column above a unit area of ground and can be expressed for example in milligrams per square meter indicating the mass in milligrams of the gas which lies in a vertical column of air above a square meter of ground. Measurements of vertical burden can be made with the remote sensor placed in a vehicle looking vertically upwards or can be made from an aircraft or spacecraft looking vertically downwards using light reflected from the ground.

An important derivation that can be made from vertical burden measurements is an estimation of the mass transport of a pollutant. Thus when a traverse line of vertical burden measurements is made from a vehicle, it is possible to combine these measurements with information on wind velocities in order to compute the mass of gas being transported across the traverse line. This concept can be made with vehicle mounted equipment looking vertically upwards beneath a smoke plume. The same principle can be applied to an entire plant such as a refinery. A circuit around the plant with upward looking equipment can be used to calculate information on the mass production of the refinery of sulfur dioxide and nitrogen dioxide. The pollutant output from localized sources can be quoted in figures such as tons per hour or tons per day, while mass transport across a boundary can be given in tons per kilometer of traverse line. Examples of mass burden and computed mass transport measurements are shown in Figure 10.

The illumination used in vertical upward looking mode is derived from Raleigh backscatter in the atmosphere. The overhead hemisphere of scattered daylight is of sufficient intensity to make possible high quality measurements of sulfur dioxide and nitrogen dioxide. This technique is unfortunately restricted to use in the visible and ultraviolet spectrum down to 3,000 Å and is not applicable in the infrared due to an insufficiency of scattering in the atmosphere at the longer wavelengths.

The downward looking mode of operation from aircraft and spacecraft is applicable in the infrared spectrum as well as the ultraviolet and visible, but is somewhat less precise to the decrease in intensity of light and the inevitable noise introduced into the measurements by rapid fluctuations in the albedo of the earth.

The satellite survey potential of the method is nevertheless very substantial and it is potentially feasible to scan vast areas in a synoptic fashion and to produce totally new kinds of data on the regional distribution of pollution. It is considered that the satellite monitoring of global and regional air pollution has an important role to play in studying the possible long term buildup of pollution, the interaction between land sources of air pollution and the oceans, the transfer of pollution across international boundaries, the development of predictive models for pollution forecasts and many other problems which are difficult to study solely by ground based techniques. Furthermore, measurements of air pollution from satellites can be interpreted in conjunction with photographic and other optical methods for monitoring water pollution and vegetation damage. When all of these data are added to meteorological and pollution data from ground station networks it becomes possible to provide a uniquely comprehensive picture of the effects of pollution on the ecology and on the environment in general.

REFERENCES

1. Barringer, A. R. and Schock, J. P., Progress in the Remote Sensing of Vapours for Air Pollution, Geologic and Oceanographic Applications, Proceedings 1966 Symposium on Remote Sensing of Environment, University of Michigan, pp. 779-792.

2. Williams, D.T., and Kolitz, B.L., Molecular Correlation Spectroscopy, Applied Optics, Vol. 7, pp 607-616, April 1968.

3. Barringer, A. R., Chemical Analysis by Remote Sensing, 23rd Annual ISA Instrument Automation Conference, New York, New York, October 1968.

4. Millan, M.M., Townsend, S.J., and Davies J. H., Study of the Barringer Remote Sensor Correlation Spectrometer, University of Toronto, Institute of Aerospace Studies, Toronto, Report 146, 1969.

5. Barringer, A. R. and McNeill, J. D., Advances in Correlation Techniques Applied to Spectroscopy, National Analysis Instrumentation Division of ISA Symposia, New Orleans, May 1969.

6. Dick, R. and Levy, G., Correlation Interferometry, Aspen,International Conference on Fourier Spectroscopy, 1970, AFCRL 71-0019, January 1971.

7. Bortner, M. H., Davies, J. H., Dick, R., Grenda, R. N., and LeBelle, P. J., A Solution to the Carbon Monoxide Pollution Experiment, A Solution to the Carbon Monoxide Sink Anomaly. Joint Conferences on Sensing of Environmental Pollutants, AIAA, Palo Alto, Nov. 1971.

8. Levy G., Davies J. H. Report on Remote Sensing of Forest Fires by their Gaseous Products of Combustion. BRL/NRC Dec. 1970.

9. Zwick, H. H., Millan, M. M., (1971. Feasibility Study of Using Ground Modulator to Discriminate between Ground and Atmospheric Scattered Radiation.

ACKNOWLEDGEMENTS

The authors acknowledge the support received from many agencies contributive to the development of correlation spectroscopy and interferometry including NASA, MAS, MSFC, HQ; EPA; DEMR, DRB and NRC Ottawa; USAF and U.S. Army, NASA-LRC and General Electric and NRC Associate Committee for Forest Fire Protection and Forest Fire Research Institute.

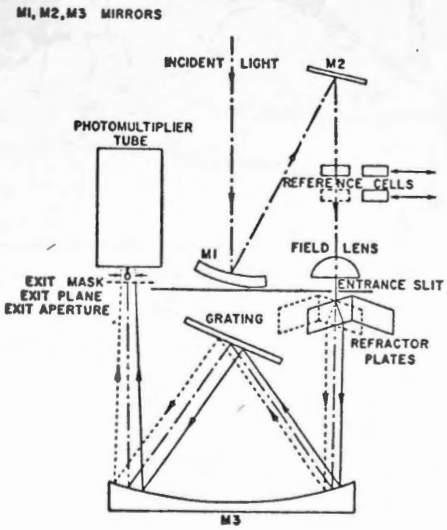
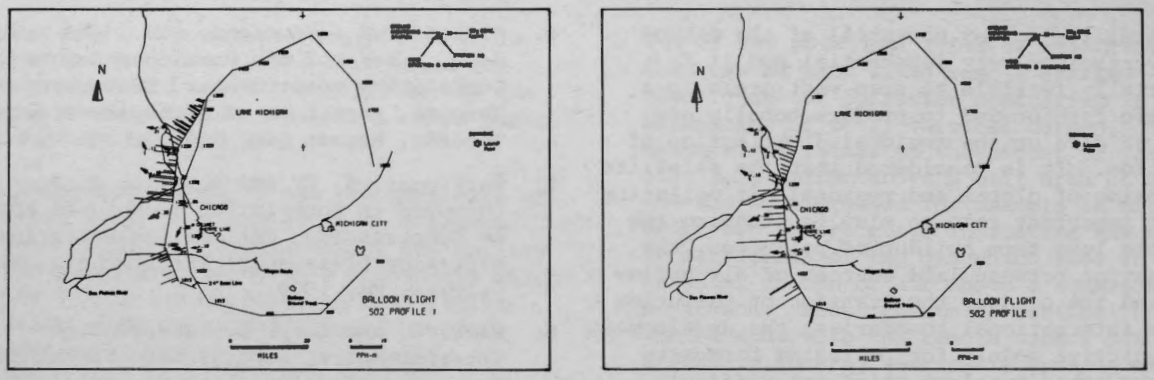


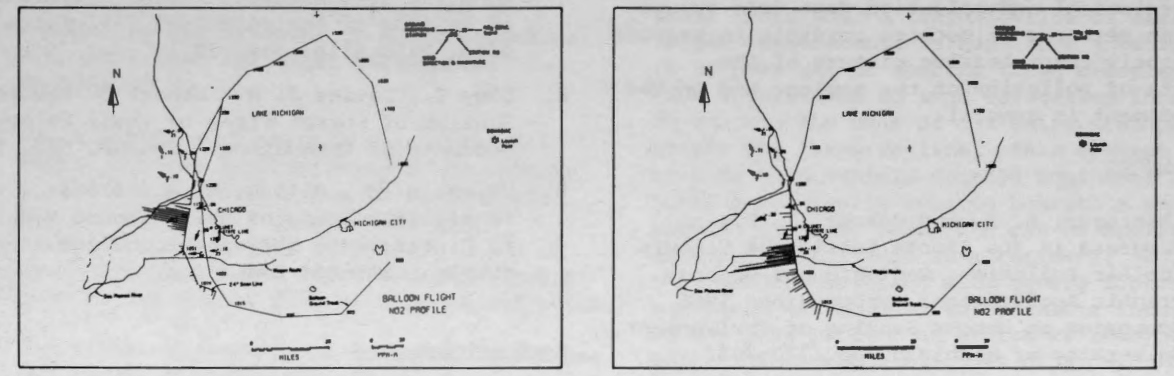
FIGURE 1

SCHEMATIC LAYOUT OF THE COSPEC I

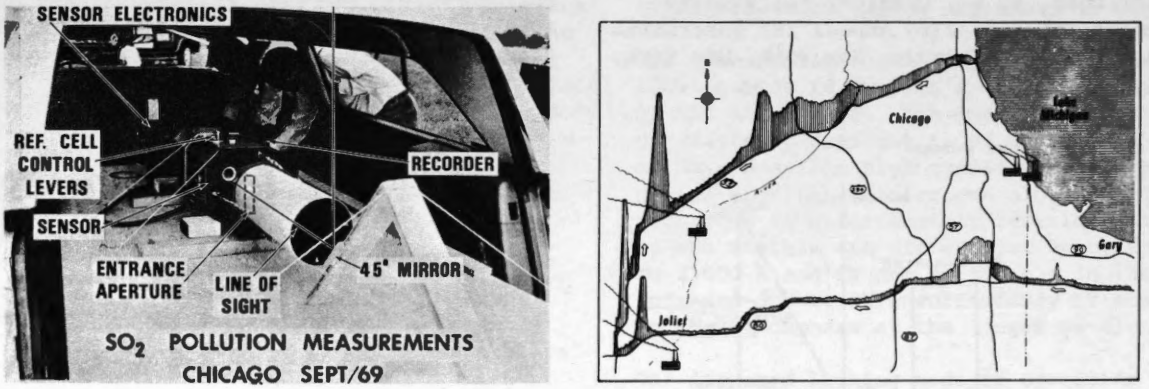
CORRELATION SPECTROMETER



These computer plotted maps show the course of the balloon float over the Chicago Metropolitan Area and Lake Michigan. The strength of the signal in ppm-m is given as a bar to the left of the route. The map to the left depicts the data obtained when the spectrometer viewed 24° off the vertical; the map on the right shows the data directly beneath. These maps show the total sulfur dioxide in the atmosphere measured from the float altitude of 22 miles.



These maps show the computer plot of nitrogen dioxide measured from the balloon over Chicago. On the left is the offset view of NO_2; on the right is vertically beneath. A cloud of this brownish-gas is distributed over the city. The blank spaces in both sets of maps result from partial cloud cover which blocked the field of view.



The station wagon, though it missed complete measurements under the balloon as originally planned, charted a continuous and very revealing profile of the total burden of SO_2 in the atmosphere over the Chicago area. Moderate offshore lake winds blew the gases west from the major industrial complex of South Chicago.

FIGURE 2

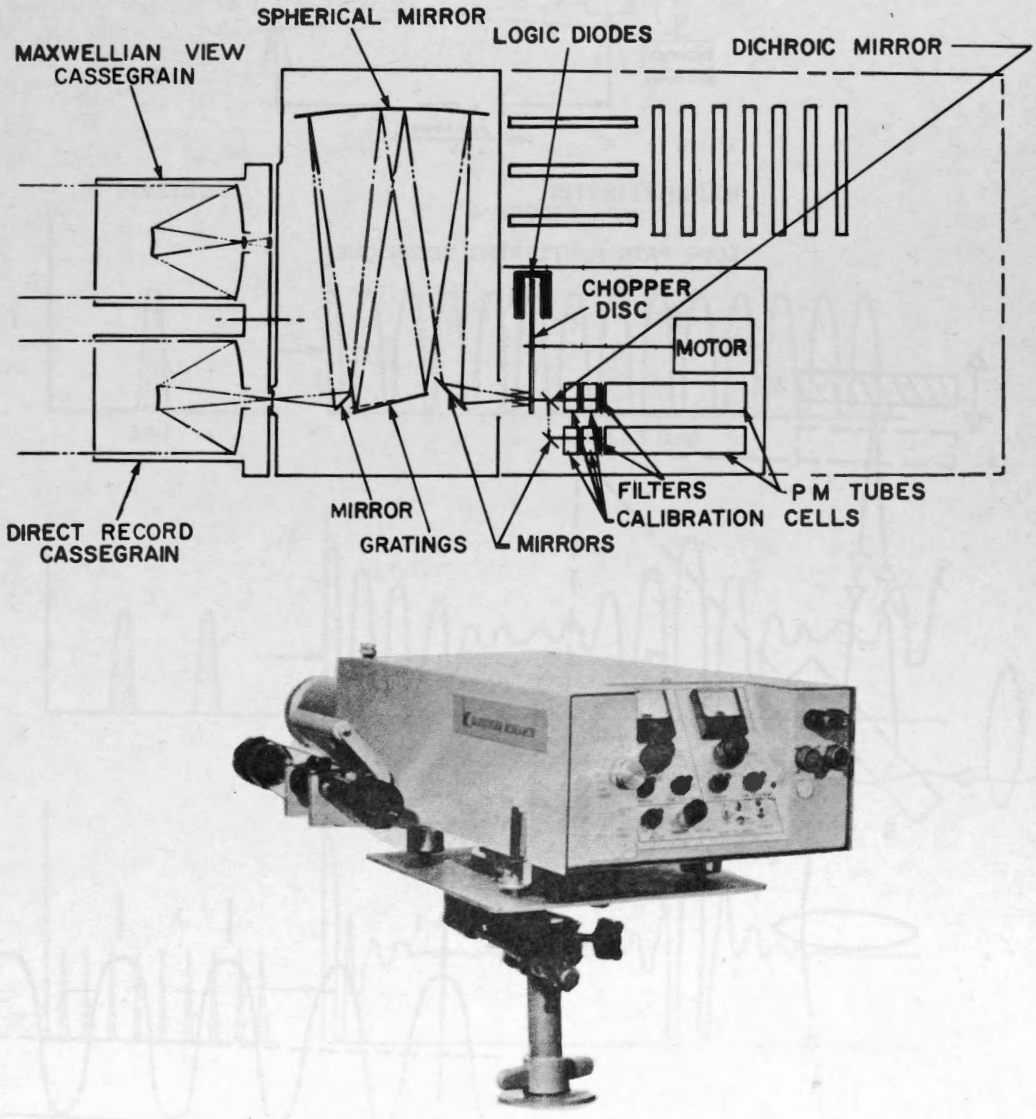


FIGURE 3

DUAL GAS CORRELATION SPECTROMETER COSPEC II

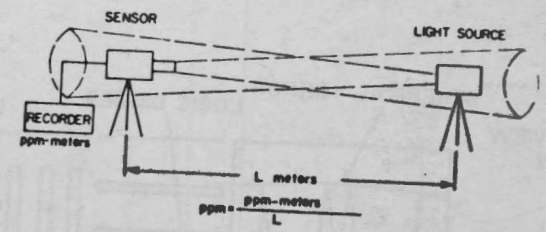


FIGURE 4

LONG PATH MONITORING TECHNIQUE

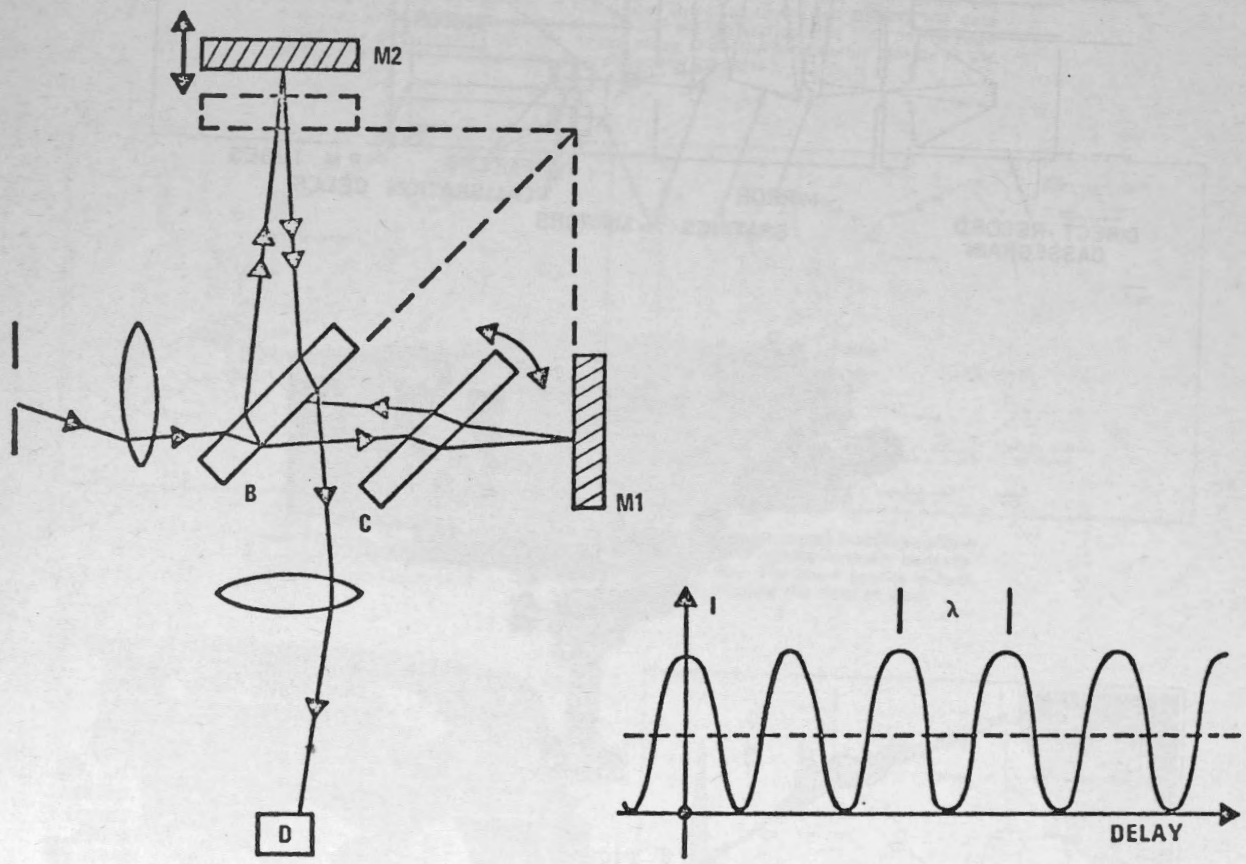


FIGURE 5

CORRELATION MICHELSON INTERFEROMETER

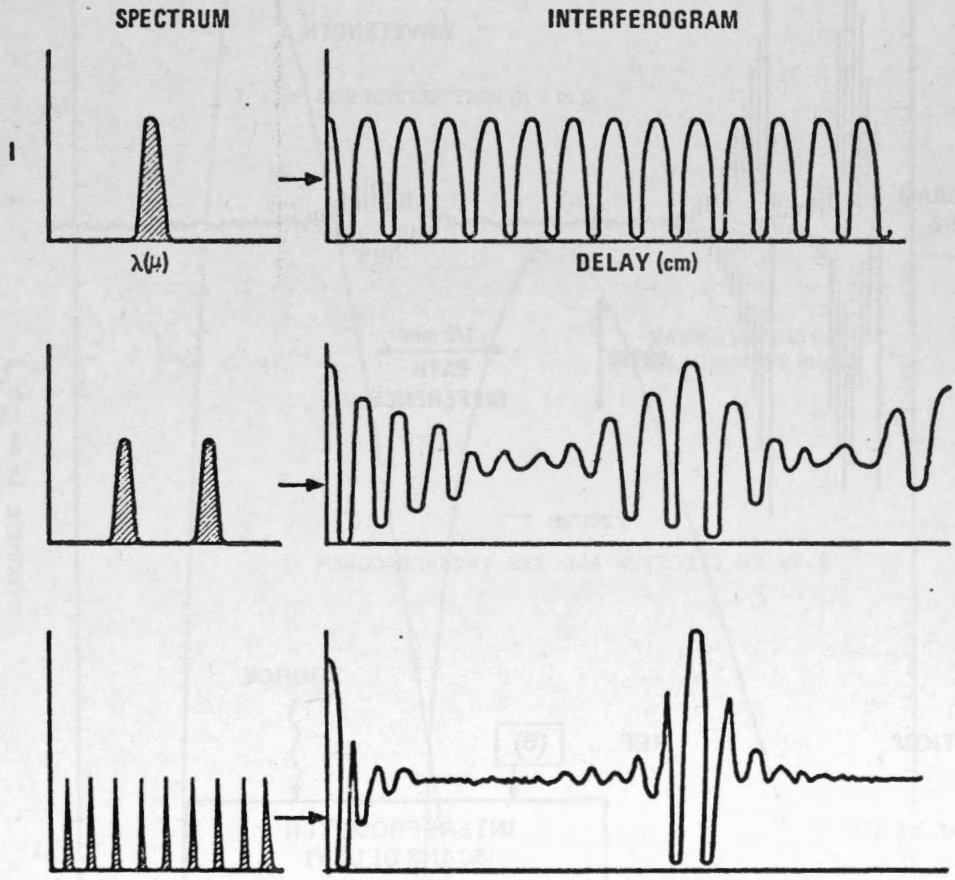


FIGURE 6

FOURIER TRANSFORMS

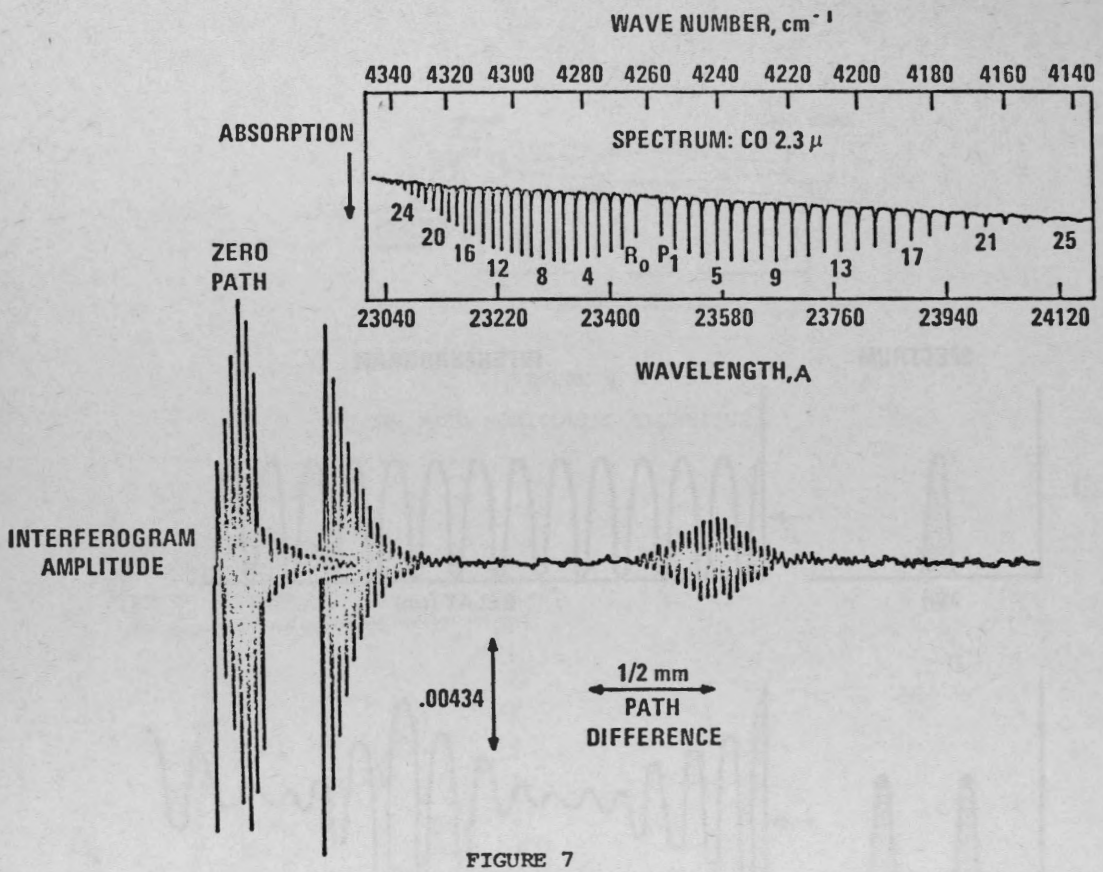


FIGURE 7

2.3μ CO SPECTRUM AND ITS INTERFEROGRAM

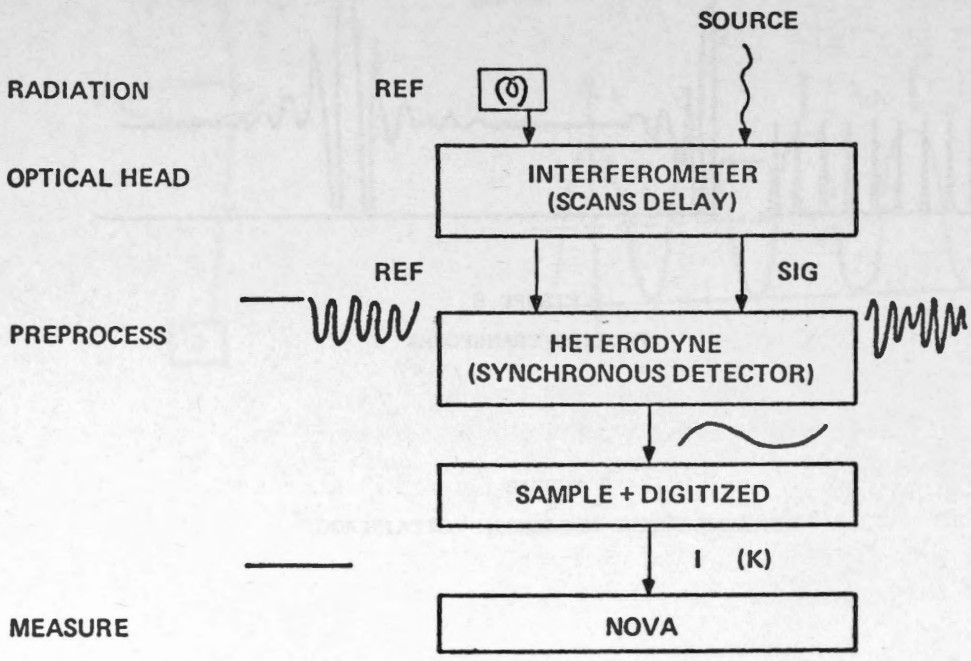
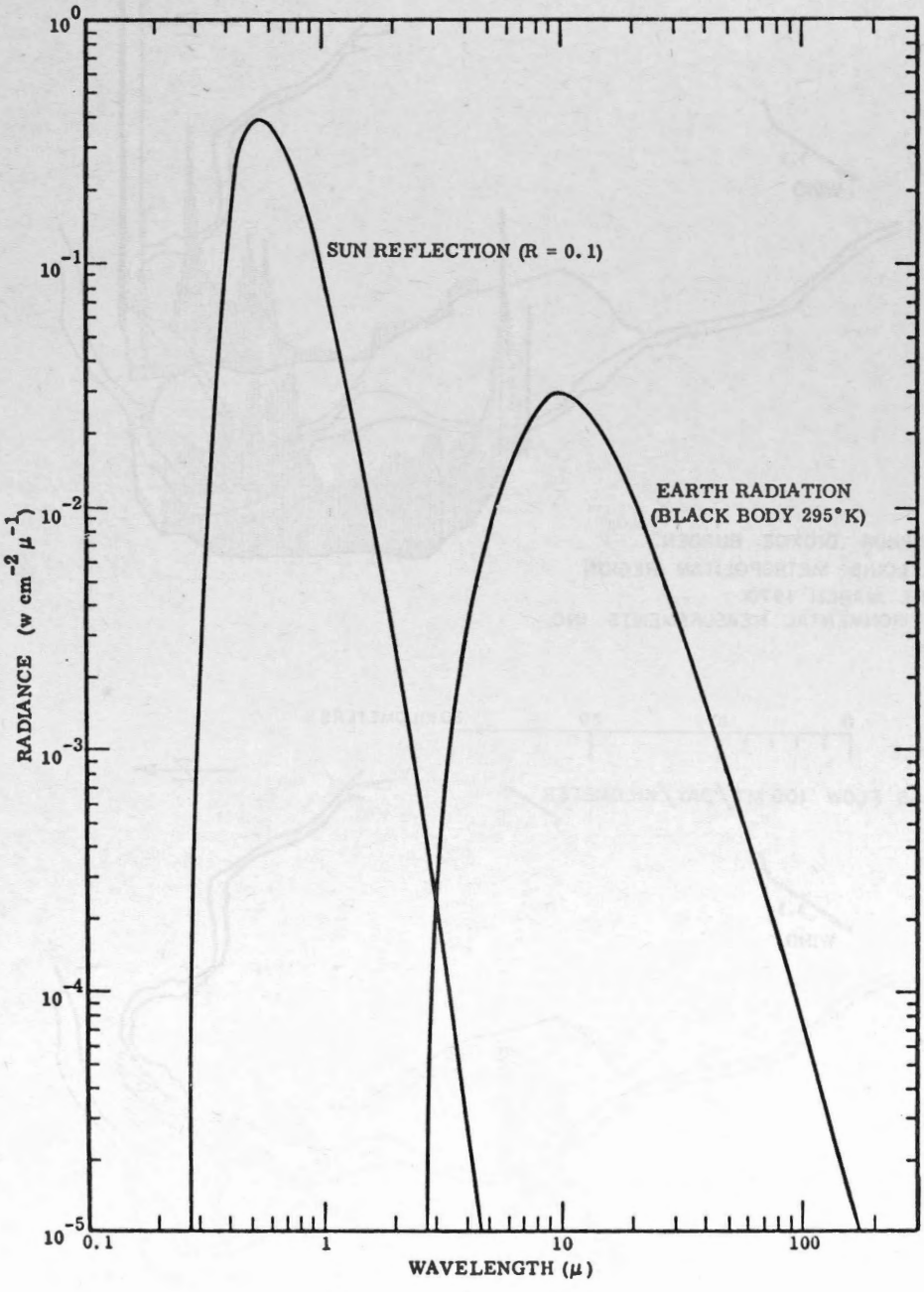


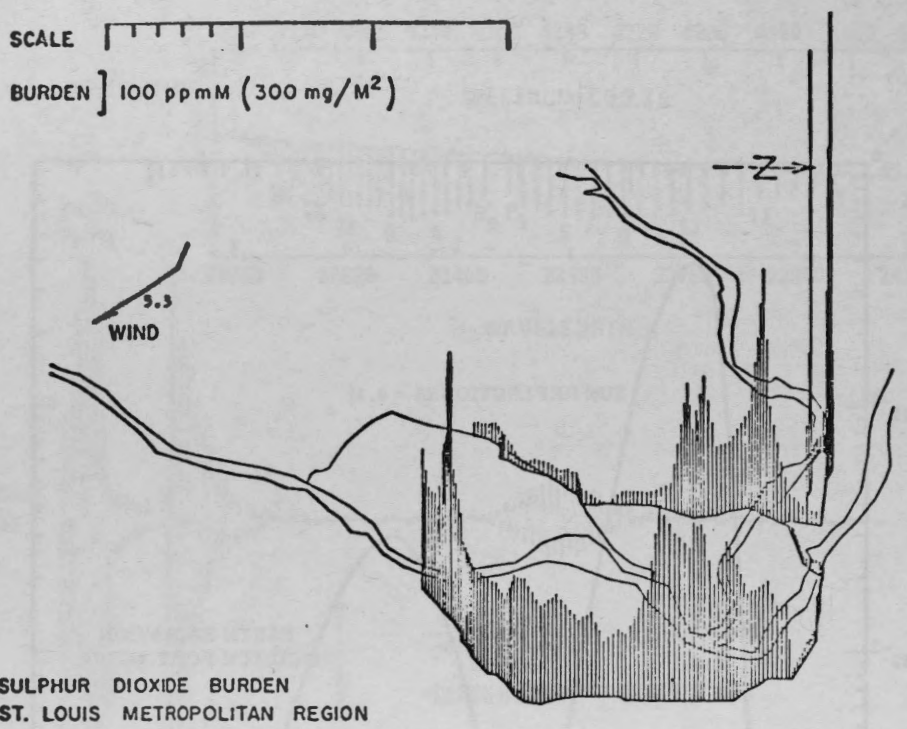
FIGURE 8

INTERFEROMETER SYSTEM BLOCK DIAGRAM

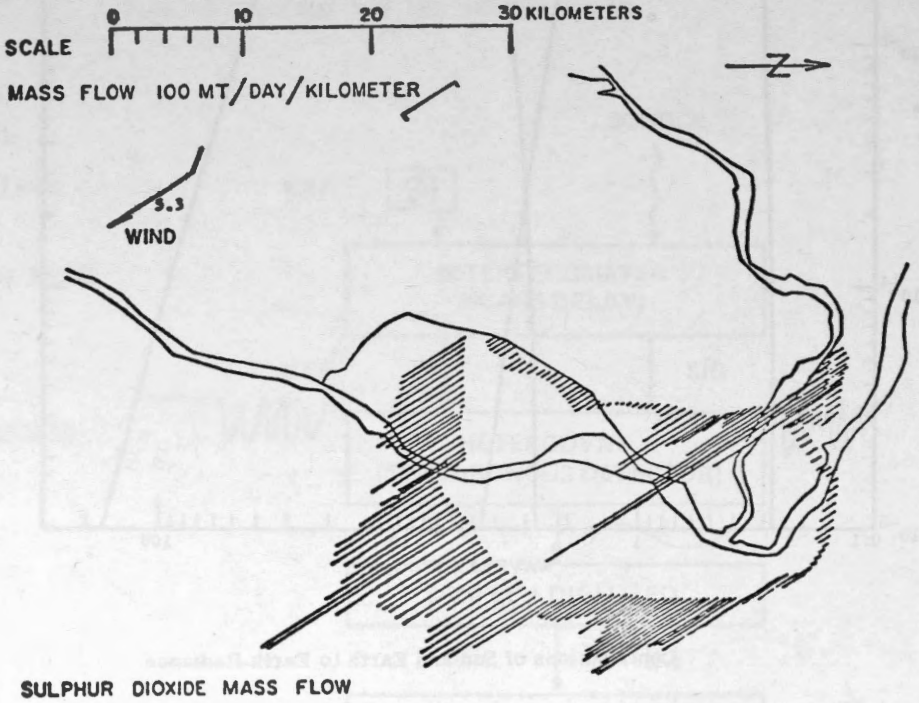


Contributions of Sun and Earth to Earth Radiance

FIGURE 9



SULPHUR DIOXIDE BURDEN
ST. LOUIS METROPOLITAN REGION
4 TH. MARCH 1970
ENVIRONMENTAL MEASUREMENTS INC.



SULPHUR DIOXIDE MASS FLOW
ST. LOUIS METROPOLITAN REGION
4 TH. MARCH 1970
ENVIRONMENTAL MEASUREMENTS INC.

FIGURE 10

DECISIONS ON COMBINING DATA FROM SEVERAL SENSORS

A.H. Aldred,
Research Scientist,
Forest Management Institute,
Canadian Forestry Service,
Ottawa, Ontario, K1A 0H3

ABSTRACT

The purpose of the paper is to set up a cost effectiveness model for evaluating the efficiency of combining information collected from satellite imagery, large-scale photos and ground measurements. A forest inventory problem is used to illustrate the approach and to indicate the cost advantages of using more than one imaging medium for certain problems. Some criteria are given for deciding when to use multiple sources of image data.

INTRODUCTION

Remote sensing platforms are being developed to operate over a wide range of distances from the earth. Airborne systems operate from a few hundred feet to 8 or 10 miles away; space vehicles provide coverage from distances of several hundreds of miles. The platforms accommodate many types of photographic sensors, scanners and other means of recording data. Consequently a wide selection of systems can be put together for various applications. Most research has concentrated on developing the potential of one system at a time, supported by a certain amount of ground data. Relatively little attention has been given to assessing how space and airborne systems can be combined to provide required data more efficiently.

Combined use of several data sources is the subject of this paper. It raises two questions which make up the main issue: does a method which uses data from more than one source perform better than one relying on a single source? If so, how do we decide what proportion of the work should be allocated to each source of data — i.e., how are the data sources best combined?

To answer the above questions, a model can be developed to evaluate the performance of the alternative methods of acquiring and combining the data. The evaluation serves to measure or rate how well the information requirements are met in terms of accuracy and cost. The rating thus sets up a criterion for selecting the best method and comparing it with the other methods. Further, the rating offers a means of optimizing the relative effort or expense of joint activities such as the combination of two or more sources of data. The model also provides a means of assessing the effect of other factors on the problem, such as changes in the budget, photogrammetric parameters, and other factors affecting the accuracy and costs of the work.

The purpose of the paper is to describe such a model, indicate its advantages and illustrate how it can be applied to a specific forest resource information problem. The example also illustrates the potential benefit of using inputs of data from more than one imaging medium.

THE COST EFFECTIVENESS MODEL

The development of the model consists of three main steps. The first defines precisely the information required. Included is a statement of how accurate the information should be and an indication of how much can be spent on the mission. The second step lists the alternative means at hand to collect the information. The third step develops a quantitative measure of how well the information requirements are met. The measure, used as a cost effectiveness criterion, permits the alternative methods to be rated and the best one to be selected for the job.

The criterion used is based on the trade-off relationship between the cost and accuracy of making an estimate. It can take one of two simple forms: choose the alternative that 1) minimizes the costs of achieving a given level of accuracy (traditional concept of efficiency) or 2) maximizes accuracy (minimizes error of estimate) given a limited budget. The following example illustrates how the trade-off relationship is used in the model and how the model is applied to choosing the best method of combining remote sensing data.

AN EXAMPLE: A FOREST INVENTORY PROBLEM

For illustrative purposes, suppose the manager of a forest property of known geographic location and size has to decide whether his area should be logged for pulpwood or withheld from harvesting for recreation purposes. In order to decide on the above, the manager wants to know how many coniferous trees above a certain size are on the property. He realizes that the estimate is subject to error and accordingly, that he must specify an acceptable margin of error. He also knows that it costs more to reduce the error of estimate and thus, that a ceiling must be placed on the budget.

The Information Requirement

The manager wishes to estimate the number of coniferous trees 30 feet or taller on his property and requires the estimate to be reliable within ± 4 per cent of the true value with a 67 per cent probability. The manager wants to have the above estimate for least cost but would be willing to pay up to $5,000 for the job. If the accuracy requirement cannot be achieved for $5,000, the manager settles for the best that can be done for this amount.

Alternative Means for Collecting the Information

The following is one of several approaches that could be taken to the problem of estimating the above quantity (*see*, for instance, multistage sampling described by Langley, 1969). Obtain ERTS or other small-scale photo coverage of the property and separate forest from non-forest (NF) areas such as water, natural or man-made clearings, rock, brush, and swamp. For the forested area, try to separate softwood (S), mixed hardwood and softwood (M), hardwood (H) forest from each other using image data such as shown in Figure 1[1]. Transfer the delineations in Figure 1 to a cover-type map shown in Figure 2, and planimeter the areas of the types. Use large-scale photo and/or ground samples to estimate the average number of coniferous trees 30 feet and taller per acre for the area, from which the total number of coniferous trees on the area can be compiled. The estimate thus obtained is the information the manager requires for his decision.

The above approach involves several possible methods of obtaining the required information:

- Method A. Input of ground data alone — Within the defined property, establish a simple random sample of ground plots of a specified shape and size. Estimate the characteristic of interest, called μ, which in the present case is the average number of coniferous trees per acre, 30 feet and taller:

$$\hat{\mu}_1 = \Sigma Y_i/n \quad \text{- - - - - - - - - - - - - - -} \quad (1)$$

where $\hat{\mu}_1$ is the simple random sampling estimator, Y_i is the number of conifers per acre 30 feet and taller obtained from the plot measurement on the ground and n is the number of such plots.

- Method B. Input of small-scale photo (or ERTS) and ground data — Within each j^{th} stratum defined by the description S, M, H, and NF (*see* Figures 1 and 2), establish n_j ground plots by simple random sampling. The total sample size n is allocated to the strata on the basis of stratum acreage, i.e., by proportional allocation. Estimate the required average number of trees per acre:

$$\hat{\mu}_2 = \Sigma W_j \overline{Y}_j \quad \text{- - - - - - - - - - - - - - -} \quad (2)$$

where $\hat{\mu}_2$ is the stratified random sampling estimator, W_j is the j^{th} stratum weight based on relative acreage size and $\overline{Y}_j$ is the result of a simple random ground plot sample in the j^{th} stratum.

- Method C. Input of large-scale photo and ground data — Establish an unrestricted random sample of m large-scale photo plots on the property. These random locations are photographed (scale 1:500 to 1:3000) and the coniferous trees 30 feet and taller are counted on the photo plots (Figure 3). Select from the m plots a random subsample of n plots which are measured on the ground. Form a regression relationship between the ground plots (Y) and their photo-plot

---

[1] It should be noted that the area shown in Figures 1 and 2 is used only to illustrate the approach to the problem. The acreage is far too small for effective application of ERTS-type coverage. Existing conventional aerial photographs would serve better. The benefits of space imagery would be realized on much larger areas — perhaps over half a million acres.

counterpart (X) as illustrated in Figure 4. Estimate the regression coefficient (b) and the simple correlation coefficient (r) between X and Y. Estimate the required average number of trees per acre:

$$\hat{\mu}_3 = \bar{Y} - b(\bar{X} - \bar{X}') \quad \text{- - - - - - - - - - - (3)}$$

where $\hat{\mu}_3$ is a double sampling for regression estimator, $\bar{Y}$ is the mean of n ground measurements, $\bar{X}$ is the corresponding mean of n photo-plot values, $\bar{X}'$ is the mean of m photo-plot measurements, and b is a least squares estimate of the regression coefficient (Raj, 1968, p. 150).

- Method D. Input of small-scale photo, large-scale photo and ground data — Use the small-scale photos as in method B to define the strata S, M, H and NF. Employ the regression relationship established in method C and illustrated in Figure 4 to estimate the stratum means and form the following estimator:

$$\hat{\mu}_4 = \Sigma W_j \hat{\mu}_{3j} \quad \text{- - - - - - - - - - - - - - (4)}$$

where $\hat{\mu}_4$ is the stratification/double sampling for regression estimator, W_j is as previously defined and $\hat{\mu}_{3j}$ is the stratum mean based on double sampling for regression.

The following lists the linear cost functions used to express how the budget (D) is expended by each method:

- Method A —

$$D = C_o + C_1 n \quad \text{- - - - - - - - - - - - - (5)}$$

where C_o is the sum of all fixed costs associated with the project and C_1 is the unit cost of establishing a plot on the ground.

The fixed costs include office overhead, equipment, compilation, and the cost of preparing a report.

The unit costs include all variable costs which depend on the number of units (plots) measured, such as manpower, accommodation, transportation, and certain supply costs.

- Method B —

$$D = C_o + C_p + C_1 n \quad \text{- - - - - - - - - - - (6)}$$

where the symbols are as above and C_p represents the additional fixed costs associated with procuring and using small-scale photos or ERTS imagery to delineate types (strata), transfer them to maps and planimeter acreages in order to calculate the stratum weights, W_js.

- Method C —

$$D = C_o + C_q + C_1 n + C_2 m \quad \text{- - - - - - - - (7)}$$

where C_q is the sum of fixed costs associated with the use of large-scale photos and C_2 is the unit cost (independent of fixed costs) of measuring a large-scale photo plot. C_q includes the cost of aircraft rental and installation of cameras and other necessary equipment. Film processing and printing, measurement instruments, and extra compilation costs are also included.

- Method D —

$$D = C_o + C_p + C_q + C_1 n + C_2 m \quad \text{- - - - (8)}$$

where all variables are as defined previously.

The Cost Effectiveness Criterion

The measure of effectiveness for evaluating the four methods is based on the variance of the estimators defined in equations 1 to 4. The variance equation applicable to each method is defined next followed by the use of the equations and the cost functions to draw up the effectiveness criterion for rating the methods.

- Method A —

$$\hat{V}(\hat{\mu}_1) = S^2/n \quad \text{- - - - - - - - - - - - - (9)}$$

where $\hat{V}(\hat{\mu}_1)$ is an estimate of the variance of the simple random sampling estimator and S^2 is an estimate of the population variance, $S^2 = \Sigma_i (Y_i - \bar{Y})^2/(n-1)$.

- Method B —

$$\hat{V}(\hat{\mu}_2) = (\Sigma W_j S_j^2)/n \quad \text{- - - - - - - - - (10)}$$

where $\hat{V}(\hat{\mu}_2)$ is an estimate of the variance of the stratified random sampling estimator for proportional allocation (Cochran, 1963)

and S_j^2 is an estimate of variance in the j^{th} stratum.

- Method C —

$$\hat{V}(\hat{\mu}_3) = (1-r^2(1 - n/m))(S^2)/n \quad - - - \quad (11)$$

where $\hat{V}(\hat{\mu}_3)$ is an estimate of the variance of the double sampling for regression estimator (adapted from Cochran, 1963, p. 339) and r is an estimate of the correlation coefficient between large-scale photo and ground-plot measurements.

The allocation of work to the photo- and ground-measurement activities is decided by the following optimization equation (adapted from Raj, 1968, p. 151).

$$m/n = \sqrt{(r^2/(1-r^2))(C_1/C_2)} \quad - - - - - \quad (12)$$

which, in conjunction with equation (7) permits m and n to be determined.

- Method D —

$$\hat{V}(\hat{\mu}_4) = \Sigma W_j^2(1-r^2(1-n/m))S_j^2/n_j \quad - - - - \quad (13)$$

where $\hat{V}(\hat{\mu}_4)$ is an estimate of the variance of the stratification/double sampling for regression estimator.

The allocation rules applied in methods B and C are used to determine the sizes of n_j and m_j.

The effectiveness criterion used to measure the efficiency (E) of the methods is derived from the variance equation:

$$E_k = (100)\sqrt{\hat{V}(\hat{\mu}_k)}/\hat{\mu}_k \quad - - - - - - - - - \quad (14)$$

where E_k is the standard deviation of the k^{th} estimator expressed as a per cent of the mean. The best method is the one that minimizes E subject to a budgetary constraint imposed by the cost functions defined in equations (5) to (8). The rating says that a method should provide an estimate of μ to within $\pm E$, 67 per cent of the time (or within about $\pm 2E$, 96 per cent of the time).

APPLICATION OF THE MODEL

To select the best method from the four given, it is necessary to have data on the variables defined earlier, such as the fixed and variable costs, the stratum weights, correlation coefficients between the large-scale photos and ground measurements, and variance data. Note that in trying to decide on the best method, we have not actually started the survey; we are merely planning it. Thus, we must depend on experience, tests and existing data to draw up reasonable estimates of values to use in the equations.

Table 1 gives the values required to evaluate the efficiency of the four methods. The estimates S^2, S_j^2 and r were calculated from existing data collected on or near the sample area, W_j was calculated from planimetering the delineated strata, and C_1 and C_2 were derived from speed trials and experience with the cost of carrying out photo- and ground-plot measurements. The fixed costs were also set on the basis of past experience.

RESULTS

The methods are evaluated in Table 2 assuming that a set budget of $5,000 was available. Method D was rated the best and would be well within the accuracy requirement of ±4 per cent. The other three methods would not. The budget would have to be increased to about $6,500 to permit the next best method to reach the 4 per cent accuracy level and to about $9,080 to meet the accuracy of method D.

Effect of the Budget

Since the four methods involve different amounts of fixed cost, it is helpful to know how the relative efficiencies (E_k) vary with changes in the budget (D). This can be done simply by substituting the cost function into the variance equation for each method. For instance, in method A, equation (5) is rewritten as:

$$n = (D-C_o)/C_1$$

and substituted into equation (13) yields

$$E_1 = 100\sqrt{S^2/n}/\hat{\mu}_1 = 100\sqrt{(S^2)(C_1)/(D-C_o)}/\hat{\mu}_1$$

where S^2, C_1, C_o and $\hat{\mu}_1$ are known values and D is a variable. The efficiency of the other methods may similarly be expressed in terms of dollars budgeted (D). The four resulting equations are plotted in Figure 5.

The figure shows that for budgets larger than $3,400, method D should be used. For budgets

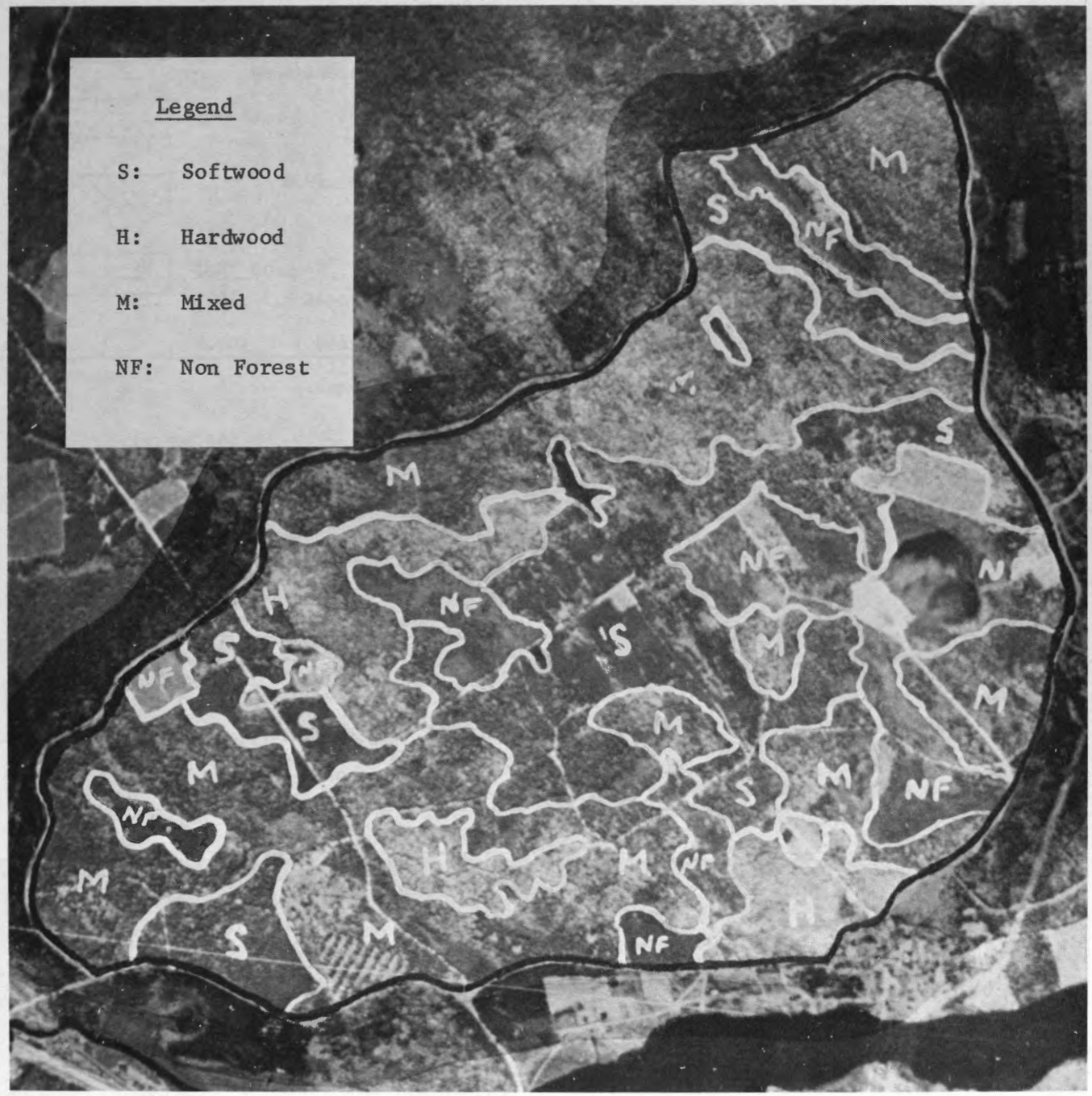


Figure 1. A six-times enlargement from a 70-mm color infrared photo showing forest-type delineations (contact scale about 1:160,000).

TABLE 1. Data used to evaluate the four survey methods

Estimated mean of population	$\hat{\mu}$ = 99.04 trees/acre			
Estimated variance of population	S^2 = 5550.96			
Regression relationship between large-scale photo (X) and ground measurement (Y)	Y = 17.78 + .993X			
Correlation coefficient of regression	r = .970			
	S	M	H	NF
Stratum weights	.255	.446	.103	.196
Estimated stratum variance	3240.3	2698.7	406.3	0
Plot size and shape	1/5-acre, square			
Variable costs	Ground: $50.00*/plot			
	Large-scale photos: $3.00			
Fixed costs	Method			
	A $ 500.00*			
	B 650.00			
	C 2,600.00			
	D 2,750.00			

*A finer breakdown in costs is available from the author.

TABLE 2. Efficiency rating of four methods: 1 standard deviation expressed as a per cent of the mean achieved with a $5,000 budget

Method	Rating (E)
A	7.85%
B	4.93
C	5.17
D	3.35

TABLE 3. Budget required by each method to achieve an accuracy of E = 4.0 per cent

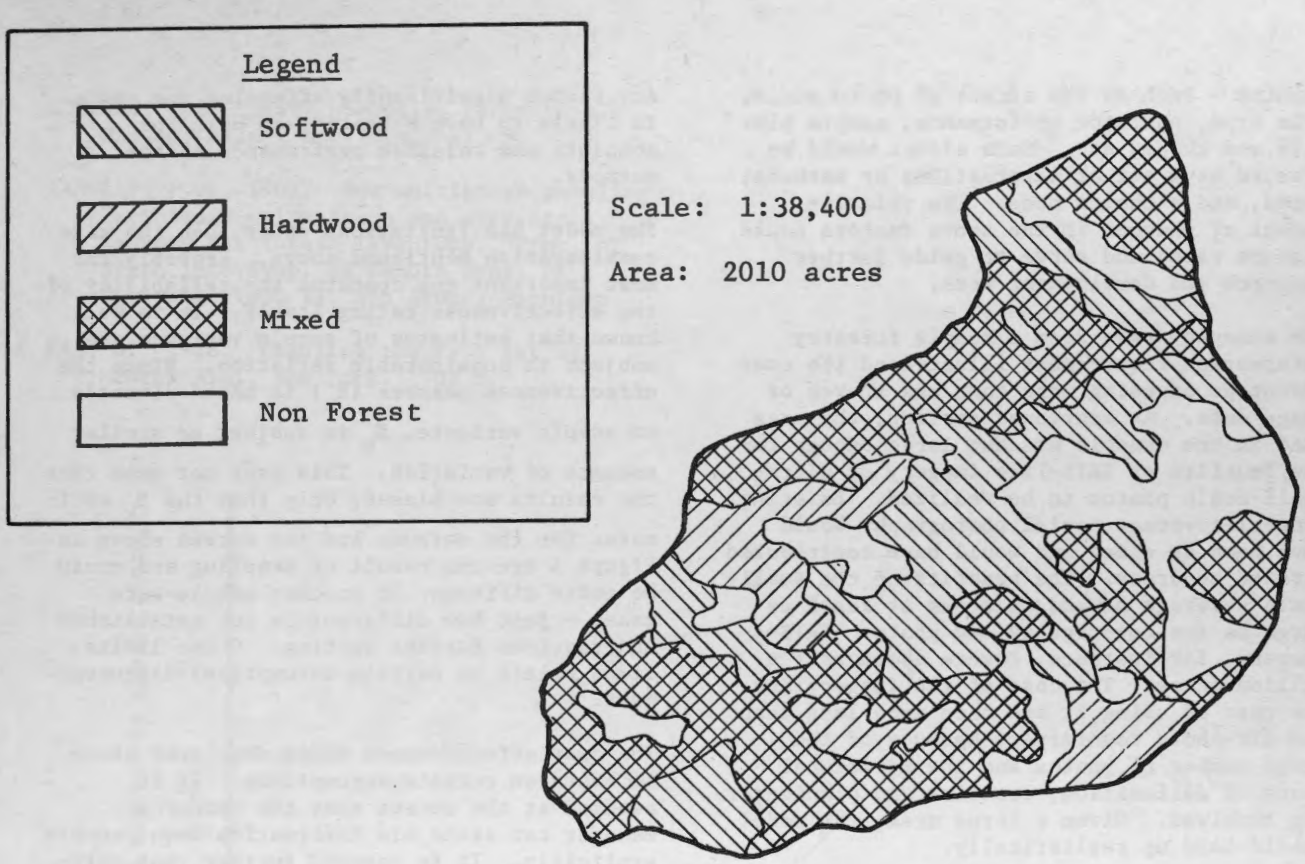
Method	Budget
A	$17,810
B	7,120
C	6,590
D	4,245

smaller than $3,400, method C should be preferred until a similar cross-over between A and C (not shown on the graph) takes place at $735. Method A should be used for budgets less than $735.

Table 3 shows the budget needed to reach the required accuracy level of E = 4.0 per cent. The difference between A and B of $10,690 indicates the amount to be saved from using small-scale photographs or perhaps ERTS imagery. The difference between A and C of $11,220 is an indication of the benefit of the large-scale photography alone. The difference between A and D of $13,565 indicates the advantage of combining both types of imagery.

CONCLUSIONS

The decision model described and illustrated above sets up the machinery for evaluating alternative methods of combining data. The evaluation or rating permitted the performance of the methods to be compared and the best one to be chosen. The methods could also be compared in relation to the effect of other factors, e.g. changes in the budget as illustrated. The influence of additional factors could be examined in a like



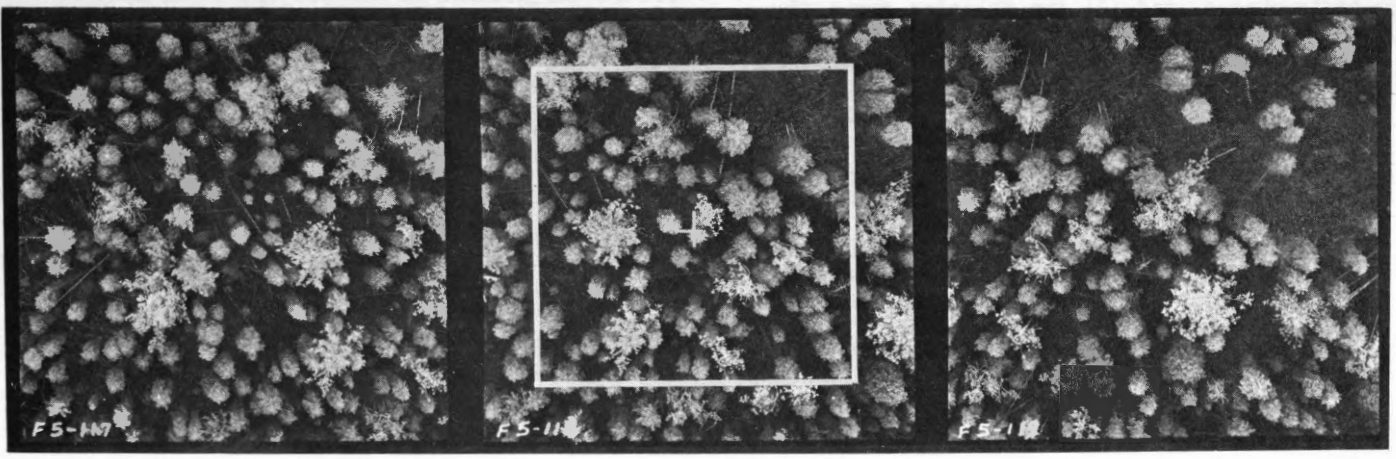


Figure 3. A stereogram of a 1/5-acre, square sample plot established on a large-scale photo (scale about 1:700).

fashion — such as the effect of photo scale, film type, operator performance, sample plot size and shape, etc. Each effect would be treated as a set of alternatives or methods, rated, and a choice made. The relative impact of changes in the above factors could also be rated and serve to guide further research and development work.

The example drawn from a simple forestry information requirement illustrated the cost advantage of using more than one source of image data. As mentioned earlier, the area used in the example was not large enough for the benefits of ERTS-like imagery or ultra-small-scale photos to be realized. Existing, general-coverage aerial photography would have been as cheap and would have contributed more to accuracy. The benefits of the small-scale coverage depend on areas at least as large as the coverage by one photo. ERTS-A imagery, for instance, covers about half a million acres. The cost of the imagery and the cost of using it are then much less than the air-photo counterpart because of the large number of photos and the additional costs of delineation, transfer and planimetering involved. Given a large area, the model should hold up realistically.

Subject to the above conditions, similar gains could be expected in estimating other parameters such as timber volume or basal area per acre. The success depends on (1) the correlation between the variable of interest and the image signatures used to stratify the forest on the small-scale photos or ERTS imagery, (2) the correlation between plot measurements on the ground and on the large-scale photos, and (3) the relative costs of photo and ground work. Other things equal,

- the higher the correlation between the image signatures (used to stratify the population) and the attribute of interest, the higher will be the performance of methods based on the stratification,
- the higher the correlation between large-scale photo and ground measurements, the higher will be the performance of methods based on double sampling,
- the lower the costs of either kind of photo work, the greater the reliance on the photos and the higher the performance,
- the higher the fixed costs occurring at any stage, the larger the budget (or the size of the project) required to absorb such costs before the efficiencies can be realized.

Any factor significantly affecting the above is likely to have a bearing on both the absolute and relative performance of the methods.

The model has limitations other than the size consideration mentioned above. Probably the most important one concerns the reliability of the effectiveness rating itself. It is well known that estimates of sample variance are subject to considerable variation. Since the effectiveness measure (E_k) is based directly on sample variance, E_k is subject to similar amounts of variation. This does not mean that the results are biased; only that the E_k estimates for the methods and the curves shown in Figure 5 are the result of sampling and could be quite different if another sample were taken — just how different is not established and requires further testing. Other limitations relate to certain assumptions discussed next.

The cost effectiveness model developed above is based on certain assumptions. It is assumed at the outset that the resource manager can state his information requirements explicitly. It is assumed further that reliable data are available on the costs, variances and correlation coefficients involved in the measurement and sampling problem. Pilot studies can provide most of the data but such studies are costly relative to data that should be available from similar past projects and related experience. Thus it is recommended that efforts be made to publish and otherwise organize the flow of such information to the planner of new projects. It must be further assumed that the various cost and variance functions are a realistic expression of what takes place in reality. Only validation tests, a trial run of the results from the model and experience can pin down the degree to which the model produces valid results.

The above limitations and assumptions must be carefully studied in relation to each application. If so, the model should provide an objective means of measuring the effectiveness of acquiring and combining several sources of data. Useful guidelines can be generated for deciding how to best assemble the multiple-input information system.

REFERENCES

ALDRED, A.H. and F.W. KIPPEN. 1967. Plot volumes from large-scale 70 mm air photographs. For. Serv. 13:419-426.

COCHRAN, W.G. 1963. Sampling techniques. 2nd ed. Wiley, New York. 413 p.

LANGLEY, P.G. 1969. New multistage sampling techniques using space and aircraft imagery for forest inventory. Proc. of Sixth Int. Symp. on Remote Sens. of Environ., Vol. II, Ann Arbor, Michigan.

RAJ, D. 1968. Sampling theory. 1st ed. McGraw-Hill, New York. 302 p.

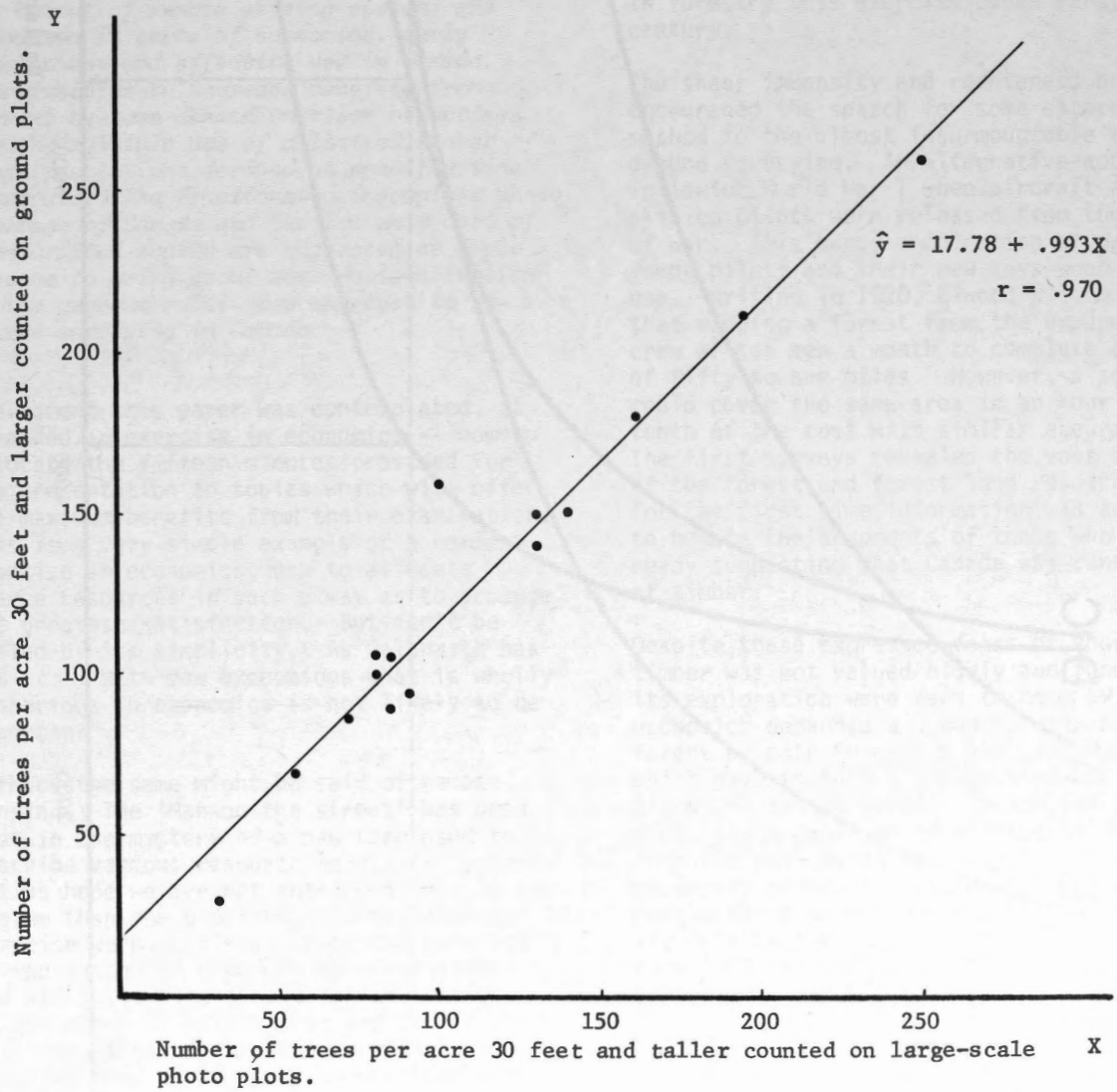


Figure 4. Regression relationship between the number of coniferous trees per acre 30 feet and taller counted on a photo plot and on a corresponding ground plot. Data are based on 15, 1/5-acre plots established on the test area by Aldred and Kippen (1967).

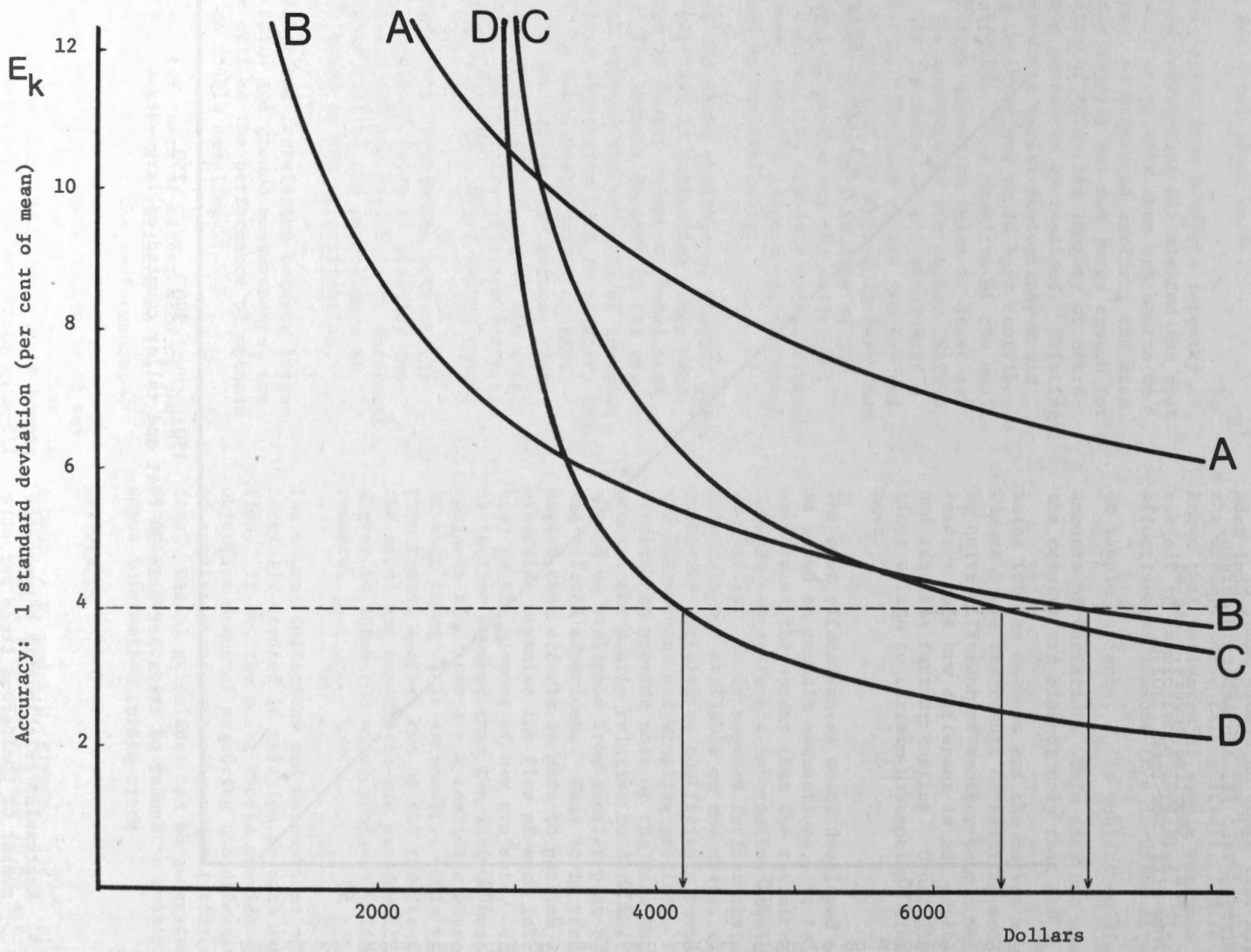


Figure 5. The relationship between the accuracy and the costs of the four methods.

THE ECONOMICS OF REMOTE SENSING OF FOREST LAND

H. Rae Grinnell
Michael S. Conway
Forest Economics Research Institute
Canadian Forestry Service
Ottawa, Ontario.

ABSTRACT

The output of remote sensing systems are discussed in terms of economics, early development and effective use in Canada. Opportunities to increase benefits from current systems depend on clear objectives for the multiple use of a limited number of image resolutions derived at specific time intervals. The fractionated incomplete photo coverage of Canada and the low unit cost of standardized scales are suggested as ample reasons to bring about some rationalization of the current multi-view approach to resource surveying in Canada.

The moment this paper was contemplated, it demanded an exercise in economics -- how to allocate the fifteen minutes provided for its presentation to topics which will offer the maximum benefits from their examination. This is a very simple example of a common practice in economics, how to allocate scarce resources in such a way as to produce the greatest satisfaction. But don't be misled by its simplicity. As Galbraith has put it: "With few exceptions what is wholly mysterious in economics is not likely to be important".

Perhaps the same might be said of remote sensing. The 'Man on the street' has been lost in the mystery of a new term used to describe various resource measuring systems. Let us hope we are not intrigued more by the system than the benefits. In the economic exercise we must clarify both the benefits or the output of a remote sensing system and its role. The output is information -- an inventory of conditions; and the information must be useful or generated by a specific need if it is to be considered as a benefit of the system. The role is merely that of collecting this information on command. There is nothing mysterious about either the output or the role - any mystery lies in the minds of those unable to interpret the output of the system. With these few points in mind let us explore the question of how best to derive the optimum benefits from our expenditures on sensing. In forestry this exercise began early in the century.

The sheer immensity and remoteness of Canada encouraged the search for some alternative method to the almost insurmountable task of ground surveying. An alternative appeared following World War 1 when aircraft and skilled pilots were released from the conduct of war. This happy combination of daring young pilots and their new toys soon found a use. Writing in 1920, Elwood Wilson noted that mapping a forest from the ground took a crew of ten men a month to complete an area of fifty square miles. However, a seaplane could cover the same area in an hour at a tenth of the cost with similar accuracy. The first surveys revealed the vast dimensions of the forest and forest land resources, and for the first time information was available to negate the arguments of those who were already suggesting that Canada was running out of timber.

Despite these expressed fears of shortage, timber was not valued highly and funds for its exploration were hard to come by. Thus economics demanded a synoptic view of the forest by both Foresters and Land Managers which gave to them a perspective not found elsewhere in the world. To achieve this view, the separation of extraneous detail from the real needs was essential to gain the necessary coverage at a realistic cost. Recently there have been suggestions that we are falling behind in our methods of collecting land resource data and losing this perspective. Whether this is true or not can only be determined by the users of the information output working in close harmony with the developers of sensing systems. Objectives must be real and evaluated carefully for costs and benefits. Above all they must be compared with alternatives which vie for the limited funds available within a renewable forest resource development and use program.

There are examples of proposed objectives for

sensing which lack this perspective. The accurate location of the "Tree Line" in Canada has been cited as important. When we can adjust the area of forest land in Canada from 1.3 billion to 800 million acres with a simple change in definition, a few million acres here or there have no national significance. Of course, if Prince Edward Island were to lose those acres, the response might be somewhat different.

A more subtle example lies in the field of mensuration where it is suggested that stratification and resulting statistical accuracy of national or regional forest inventories is one of the important objectives which justifies space platforms for very small scale imagery. When the whole forest industry is faced with serious economic problems, such ventures must be examined carefully since economic resource exploitation is only marginally affected by the statistical accuracy of resource data. It can be argued that data on spacial distribution of the anomolies of the resource are of much greater value to the planner and manager. When the resource is in surplus, or when only minor attempts are being made to fully use its capabilities, accuracy is of even less importance. We shall return to this question of objectives later.

Let us look at one particular aspect of the output of data from a photographic system and the user. A review of the literature uncovers repeated suggestions that scale must be tailored to the individual needs of the user. This may be true for some particular needs, but image resolution and signature identification are only partially the output of scale. The concept of a one-to-one relationship between benefit and scale introduces a level of complexity which, if eliminated, and a few standard scales accepted, could certainly increase benefit-cost ratio.

Such a standardization for forest land use does seem possible! It is suggested that failure to do so in the past is more an indication of insufficient utilization of highly qualified photo interpretation personnel rather than a need for greater resolution of images. Hodges said as much to the 1964 Seminar on Air Photo Interpretation, when he noted that skill not scale was the important problem. Standardization will be nearer if we accept the fact that present equipment, at a few scales, can provide all the detail heretofore found in a number of customized scales. Systematic analysis and improved interpretation skills are the other necessary ingredients for success. Thus it is patently obvious that in forest land use the important economic consideration is not the occasional use derived from a scale or image, but the extraction of the maximum amount of basic information from a few selected images.

In general, what renewable resource information do users need, and how can their needs be combined? At one time forest land exploitation in its narrow sense, dominated and determined the information requirements, but now social and economic values have changed! All components of the forest land resource are carefully analysed for their relationships and their ability to serve people. These objectives and the specific information needs may be derived by looking at the different administrative levels, their responsibilities to people, and the values of the forest resource.

This hierarchy is developed in Figure 1, to which has been added a suggested image resolution to satisfy each administrative level. The figure refers to black and white photography only. Whatever the size of the image, it must in aggregate, or in composition, present a physiognomy which exhibits the essential resource characteristics associated with planning and programs conducted by the appropriate administrative level.

As can be seen, there are four administrative levels of responsibility: National, Regional, District and Operational. Each level is a cohesive or unified expression of the data from the various components contained in the level below it, while at the same time providing a framework or outline of what is required from that level.

Looking at the type of data more closely, it will be appreciated that usage requirements may be periodic, infrequent or once in a century. Figure 2 suggests that the frequency of such coverage should reflect the rate of change of resource features. Exceptions may be anticipated, but it is felt that the stated requirements are sufficient and more frequent coverage would yield only marginal benefits. The custom coverage noted in the figure refers specifically to time, but may, of course, include sensing systems other than black and white where their economic efficacy has been demonstrated.

One further technical aspect of the information system should be clarified. It involves both the resource data framework and the timing. The resource data has two relatively static components, geological and climatic,

which determine the highly dynamic biotic resultant. The long term efficiency of any natural resource analysis system requires that the two static components be mapped first to establish meaningful natural boundaries for the associated ecosystems. Since the biotic resultant is under various levels of stress from man, priorities for detailed data must recognize high stress areas first, while reconnaissance type data should be sufficient for the low stress areas.

At this point, some mention should be made of images other than black and white for forest land use activities. Undoubtedly colour infra red, radar imagery and other techniques and systems have their place, but they too must be scrutinized for their benefits and costs. One sizeable constraint to their use is that the images received are often very difficult to interpret and the skills demanded are even greater than for black and white. Marshall clearly expressed the problem of images back in 1968 when he said, "in reference to infra red colour photographs of vegetation: a number of external factors which may affect colour reproduction of foliage ... include time of year, time of day, orientation of leaves to the sun, the moisture and mineral content of the soil and also the sensitometric characteristics of the film employed". In short, black and white offers greater potential for extensive use in forestry, but where special techniques appear to have value, the marginal benefits must be carefully measured.

Now let us move to a few practical examples of problems and possibilities. We have stressed the standardization of data which has as its objective the multiple use of a single coverage. An examination of Bowen's recent publication on "Air Photo Coverage for Canadian Forestry", and the earlier photo coverage maps published by the National Air Photo Library, uncovers an apparent lack of coordination of coverage, particularly between federal and provincial governments and between resource personnel. We are informed that some improvement in this situation is underway, but is it in fact? In the period 1968 to 1970 a sizeable portion of Eastern Ontario in the vicinity of Ottawa was photographed nine times for almost as many different images. If we can overcome two hurdles, beaurocratic isolation and personal taste for a particular scale of photograph, we might make some real economic headway in this field.

Another important economic question is whether the capabilities of our currently proven sensing systems are even close to being fully exploited. The National Forest Inventory of 1968 reports only 65% of the forest land of Canada as inventoried which implies comparable coverage by photography for that purpose. Some provinces have complete coverage of their forest land, while others have not photographed some districts for 24 or more years.

Although, without exhaustive analysis to derive accurate figures, there appears to be an increase in the annual photo coverage over the past few years. Bowen's report shows annual provincial coverage, at scales suitable for management planning, now exceeds 50 million acres per year which, if distributed nationally, would provide complete national coverage every ten years. Since the emphasis has been more on B.C., Alberta, Ontario, Quebec and New Brunswick, the coverage for other provinces falls short. They appear to be more dependent on federal coverage where we find various scales, lenses and films suggesting considerable experimentation. Nothing specific has been published on the economics of these activities.

Let us take a brief look at the specifics of black and white photography at the intermediate scale referred to in our figures and currently used more often than any other image today for forestry. It provides us with images of ground features with dimensions of approximately 12" x 12". Such photography, to be fully exploited in the regional planning field, should be placed in the broad framework of small photography in the range of 1/50,000 to 1/100,000, but for actual management and operations it is satisfactory without such a framework. It provides the basis for management maps and "on the ground" activation of programs. Its enhancement is derived from sampling and here we find that our large scale photography serves our purposes most economically. What costs can we anticipate?

For sizeable blocks of black and white photography at a scale of 1/15,840, and in a province with land cover typical of Ontario, the cost is about $3.50 per square mile or a little more than 1/2 cent per acre. To purchase such coverage from the National Film Library the cost would be approximately 3 cents per acre. Careful mapping, ground sampling and compilation of data for areas in excess of 500 square miles, would range between 5 and 10 cents per acre. Zilinsky and Cardwell of the Ontario Department of Lands and Forests report that photography using 35 mm cameras mounted on conventional Beaver air-

craft, has reduced their annual sampling to 1/5th of the cost when field work only was used. The simplicity and versitility of this technique and equipment certainly recommends its use to the forest land manager.

How do these costs compare with the actual management and operational costs? A wildlife clearing cut may cost $40.00 per acre. Reforestation will average about the same. Harvesting expenditures will approach or exceed $200.00 per acre. A one year delay in regeneration might cost $1.00 per acre or more. Against these expenditures, the cost of introducing recent, carefully interpreted photography, is insignificant. It doesn't require a Galbraith to come to that conclusion!

In summary, today's economic analysis is not the comparison of tested sensing systems with untested systems, but rather how well any system might be used to obtain the highest benefit cost ratio. Such an objective will be realized only upon the clear identification of useful objectives and the efficient and full exploitation of a minimum number of images obtained at the lowest cost. A corollary to this are two requirements: the use of qualified resource specialists for interpretation, and multi-sensing from a single platform. One final point, a sophisticated solution for a simple problem is a luxury Canada can ill afford.

Bibliography

Aldred, A.H. Potential of Satellite Photography in Forestry. Forest Management Institute, Information Report, FMR-X-11. Ottawa, 1968.

Aldrick, R.C. Space Photos for land-use and Forestry. Photogrammetric Engineering, April, 1971.

American Society of Photogrammetry. Selected Papers on Remote Sensing of Environment. July, 1966.

Bowen, M.G. Air Photo Coverage for Canadian Forestry. Forest Management Institute, Information Report, FMR-X-38. Ottawa, 1971.

Canada Program Planning Office for Resource Satellites and Remote Airborne Sensing. Resource Satellites and Remote Airborne Sensing for Canada. Report No. 5, Forestry and Wildlands. Ottawa, 1971.

_____. The Interdepartmental Committee on Air Surveys. Proceedings of a Seminar on Air Photo Interpretation in the Development of Canada. Ottawa, 1963.

_____. Proceedings, Second Seminar on Air Photo Interpretation in the Development of Canada. Ottawa, 1967.

Colwell, R.N. Remote Sensing of Natural Resources. Scientific American, January, 1968.

MacDowall, J. and B.H. Nodwell. Study of Sensors for Earth Resources Satellites and Remote Airborne Sensing. Centre for Applied Research and Engineering Design, McMaster University. Hamilton, 1971.

Marshall, A. Infrared Colour Photography Science Journal. January, 1968.

Sayn-Wittgenstein, L. and A.H. Aldred. A Forest Inventory by Large-scale Aerial Photography. Pulp and Paper Magazine of Canada, September 5, 1969.

University of Michigan, Willow Run Laboratory of The Institute of Science and Technology. Proceedings of the Fifth Symposium on Remote Sensing of the Environment. April, 1968.

Wilson, E. The Use of Aircraft in Forestry. American Forestry. June, 1920.

Zsilinszky, V.G. Supplementary Aerial Photography with Miniature Cameras. Paper Presented at International Congress of Photogrammetry. Lausanne, Switzerland, 1968. Photogrammetria 25(1969/70).

Figure I DATA FOR DECISIONS AND ACTION - FOREST LAND RESOURCES

Administrative Level	Responsibilities	Type of Data	Image to be Interpreted and Data Flow
National	- Fiscal policy - International markets - Regional economic needs - Basic research - Trans provincial research - Statistics - international and national - National good - National information framework	- Inventory of national and regional resource and their stresses - Derived from an aggregation of regional data	
Regional	- Comprehensive planning and development policies	- Regional land systems with dimensions of biotic state - Ecological time in terms of social & economic values - Macro stresses - Communications - Spacial distribution - Aggregate of qualitative & quantitative value with reference to utilization centres - Assessments	10' to 300'
District	- Integrated management planning	- The current state & capability of each resource value for management planning within regional plan. - Assessment	1'
Operational Unit	- Efficient Utilization	- Detailed operational data - Assessment	Detailed sample or complete survey

FIGURE 2 LAND PLANNING AND MANAGEMENT NEEDS FOR SENSING

Frequency of Aerial Surveys

Type of Data	Photo Image Small	Medium	Large
Regional Planning & Policies			
Land Systems	Once	-	-
Infrastructures	Infrequent	-	-
Capability	-	Once	-
Present Land Use	Infrequent	-	-
District Planning & Management			
Current Vegetative Cover, Wildlife	-	Once every 10-15 years	-
Ecological Capability and Range	-	Once	-
Depletion of Exploitation (Assessment)	Custom	Custom	Custom & Sample
Operations & Utilization			
Communications	-	Once every 10-15 years	Sample
Specific Detail Planning & Action Programmes	-	-	Custom
Depletion & Exploitation (Assessment)	Custom	Custom	Sample & Custom

Assessment includes: Cut-over, fire, insects, disease, drainage, wildlife populations, permafrost, flooding.

Assessment may require frequent coverage, but is usually custom work.

DATA HANDLING FACILITY OF THE CANADA CENTRE FOR REMOTE SENSING

W. Murray Strome,
Consultant,
Canada Centre for Remote Sensing,
Ottawa, Ontario.

ABSTRACT

The three major functions of the data handling facility are: a) to process the data received from the Earth Resources Technology Satellites (ERTS) to produce corrected photographic images and computer-compatible digital representations thereof; b) to process data remotely sensed by airborne sensors in a similar fashion; and c) to provide the tools required for research and development effort in automatic interpretation techniques. Because of the launch date of the ERTS-A satellite (May, 1972), the requirements for processing the ERTS data have received the most attention. Hence, the ERTS production system represents the dominant feature of the facility.

This paper describes the equipment and operating procedures of the facility as it will appear shortly after the launch of ERTS-A. The ERTS requirements are well defined and impose a heavy operational load upon the system. Most of the requirements for interpretation and the processing of airborne data have not yet been defined in detail. In fact, these requirements will likely be in a constant state of flux as new sensors and interpretation methods are developed. In order to cater to the changing needs of the latter functions the system must have a high degree of flexibility.

INTRODUCTION

The data handling facility of the Canada Centre for Remote Sensing must meet all the specialized data processing needs of the Centre. These needs have been divided into three categories: satellite, airborne and interpretation. (1)

The most urgent requirement is to process the data which will be obtained with the aid of the Earth Resources Technology Satellites (ERTS), the first of which, ERTS-A, is to be launched by N.A.S.A. during May, 1972. The data from these satellites will be used to produce photographic imagery and computer-compatible digital magnetic tape representations of these. The most complete description of the ERTS experiment is the ERTS Data Users' Handbook (2). The computational load and special equipment requirements for the processing of ERTS data are quite heavy, but well defined. Thus, a large part of the resources is devoted to this function on a semi-dedicated, stable basis.

The airborne requirements are much less rigidly defined, and are expected to change as new sensors and recording methods are developed under the direction of the Data Acquisition Division of the Centre. In order to meet the needs of the various airborne programmes, a relatively high degree of flexibility is essential. This is especially true with respect to the interfacing of special and varied playback equipment.

Similarly, the research into automated interpretation techniques requires special imaging and display equipment to be interfaced to the system. In addition, a flexible and powerful computing facility must be coupled to this special hardware.

The following sections outline the equipment and approaches which are being implemented in an attempt to meet the varied, and in some ways conflicting requirements of the data handling system.

GENERAL SYSTEM DESCRIPTION

Figure 1 is a block diagram of the data handling system. At the heart of this is a dual PDP-10 digital computer, manufactured by Digital Equipment Corporation. The dual processor configuration was chosen primarily to handle the generally conflicting requirements of high system stability and reliability for the ERTS production function, and the high degree of flexibility in both hardware and software for the airborne and interpretation needs. As a bonus, this configuration provides a high degree of redundancy for the ERTS production system, since most components can be quickly and efficiently switched back and forth.

The equipment is normally operated as if it consisted of two entirely separate and independent systems. The first is the dedicated, on-line ERTS production facility, which is used to control the processing of the video data. The second is the utility system. It is used for all general software and hardware development in support of the airborne and interpretation requirements, as a research tool for the centre scientists, and for the off-line data processing tasks which are necessary to prepare for the ERTS production runs. As experimental systems reach an operational status, they can be moved over to the production system.

ERTS PRODUCTION SYSTEM

Products

The primary output product of the ERTS production system is a geometrically and radiometrically corrected photographic latent image. The National Air Photo Library has the responsibility for processing the photographic film, duplicating it and distributing it to the users.

Two sensor packages will be flown in the ERTS satellites, a three-camera Return Beam Vidicon system (RBV) and a Multi-Spectral Scanner (MSS). The RBV covers the three spectral bands: .475 - .575, .580 - .680 and .698 - .830 micrometer. In ERTS-A, the MSS covers four spectral bands: 0.5 - 0.6, 0.6 - 0.7, 0.7-0.8 and 0.8 - 1.1 micrometers. In ERTS-B, a fifth band will be added operating in the region from 10.4 to 12.6 micrometers.

The distributed photographic imagery will be in a 9 1/2 inch format at a scale of 1:1,000,000. A nominal scene covers an area of 185 x 185 kilometers (100 x 100 nautical miles).

The RBV imagery will be corrected for geometric and radiometric distortion within the camera system. A set of three black-and-white images will be available for each scene. These may be in the form of prints, transparancies, or negatives. They will be annotated with information identifying position, spectral band, time and gray-scale. The spatial resolution of the RBV system is approximately fifty meters, and that of the MSS system, one hundred meters.

The MSS imagery will be corrected for radiometric distortion and for geometric distortion due to the scanning operation. In addition, where scenes are sufficiently cloud free and contain a sufficient number of identifiable ground control points, the images will be corrected to better than three resolution elements. Images will be framed to nominally match the RBV scenes. A set of four black-and-white images as well as two "false colour" images will be available for each scene from ERTS-A.

On request, the MSS data for a limited number of scenes will be made available on computer compatible, 9-track digital tape at either 800 or 1600 characters per inch density.

Figures 2 and 3 illustrate the proposed formats for the annotated MSS imagery.

Data Flow

Figure 4 shows the planned time schedule for the production of ERTS imagery. It will probably require modification in the light of experience obtained as the system commences operations, and is representative of the "best-case" conditions. As shown, it is hoped that under normal circumstances, imagery ordered in advance of a satellite pass will be shipped to the user

within one week of the pass.

Data is received and recorded on magnetic tape at the Prince Albert Satellite Station (PASS) and is to be shipped to Ottawa by air on the day it is recorded (day 1). It should arrive at the ground data handling facility early the following morning (day 2). Housekeeping information (giving satellite attitude and sensor status) will be extracted and one band of the MSS data will be processed in an uncorrected mode. Orbital predict information supplied by NASA will be used to supply annotation. The film produced by either the Laser Beam Image Recorder (LBIR) or the Electron Beam Image Recorder (EBIR) will then be delivered to the NAPL for processing. EBIR images are recorded on a 70 mm. format and will be enlarged to a 9 1/2 inch format.

The processed film will be returned to the data handling facility on the following morning (day 3). Preselected ground control points will then be manually identified and their positions accurately measured within the photographs with the aid of graphic digitizers which are connected to the utility system. These points are used by the off-line computer to produce an annotation/correction tape which will be used by the production system to control the generation of the corrected images. This operation is completed by the end of day 3. On day 4, the corrected production runs occur where the precision corrected MSS and accurately annotated RBV images are produced. This film is then delivered to NAPL where it is processed (and enlarged in the case of EBIR film) to produce the precision masters on day 5. On day 6, the preordered distribution copies are generated for shipment on day 7.

If there is a demand for them, uncorrected RBV images may be produced on day 2 for distribution by day 4.

Equipment

Refer again to Figure 1, the block diagram of the total system. One PDP-10 central processor with 48K 36-bit words, an operator's console, two nine-track 200 inch/second 1600 character/inch magnetic tape drives, paper tape reader/punch and two *DEC-tape drives forms the standard computer portion of the system. Operation is under the control of a specially developed software monitor.

Data received from PASS is on two magnetic tapes. The RBV data is played back on an ERTS unique video tape recorder similar to that used in the television industry. The MSS data is in digital form and is played back on a modified 28-track instrumentation recorder. Twenty-five of the channels are equipped with digital encoding/decoding electronics which allow a digital recording density of 10,000 bits per inch per track. Satellite attitude information is derived from the housekeeping data recorded on both tapes with the aid of the PCM telemetry demultiplexer.

Correction of the RBV data for geometric and radiometric distortions due to the camera systems is performed by the Radiometric/Geometric Correction Unit (RGCU). Annotation is added, under computer control, by the EBIR Controller. The latent photographic images are produced by the EBIR. The RBV images produced are expected to be comparable to that produced by the NASA bulk processing facillity.

MSS data is deskewed, reformatted and radiometrically corrected by the Multi-spectral Scanner Data Processor (MDP) under the control of the computer system. The data from the MDP may be directed to either the Electron Beam or Laser Beam recording systems for the production of photographic images, or directly to the computer for the production of digital computer compatible magnetic tapes.

Colour photographs may be produced directly by the LBIR or from the black-and-white negatives produced by the EBIRs using a colour composite printer/enlarger within the NAPL facility.

The LBIR is described in detail in another paper presented at this symposium.

---

* DEC-tape is a registered trade-mark of the Digital Equipment Corporation.

UTILITY SYSTEM

The remainder of the system shown in Figure 1 is intended primarily to service the airborne and interpretation requirements. Some off-line ERTS processing must also be done with this part of the system. The basic computer configuration is identical to that dedicated to ERTS production. In addition, there are standard seven and nine track digital tape drives, movable head disc drives with removable disc packs, card reader, line printer, high speed printer/plotter, and line-scan equipment to support multiple slow speed terminals. The graphic digitizers used for the preparation of ERTS annotation/correction tapes are connected to the system as terminals.

The utility portion of the facility is operated under a standard time-sharing monitor system which allows a number of users simultaneous access to the system. Standard software available to the time-sharing user includes Assembler, Extended Fortran IV, Algol, Basic, Cobol, Interactive Text Editor, and a large number of special utilities. It is relatively easy to add non-standard devices to this system. This is the case with respect to both the hardware interface required and to the software modifications required in the monitor.

As a first step to providing support to the airborne programmes, a high speed analogue-to-digital converter will be added to the system. This will allow computer processing of data recorded on analogue magnetic tapes from airborne sensors, such as infrared line-scanners. Eventually, it is expected that other playback devices will be interfaced to the computer.

The first non-standard equipment to be provided to support the research and development effort into automatic interpretation techniques will be a high resolution, full colour video display system. This system will be controlled by the computer and will operate in a time-sharing mode with the other activities.

Many other general functions will be performed by the facility. A few examples of these are: reducing photographic images (colour or black-and-white transparancies) to digital form with the aid of the LBIR in scanning mode; maintaining a library of references to remote sensing literature which may be accessed through the Technical Information Retrieval system; and maintaining an inventory of all remotely sensed data which is processed by the facility.

REFERENCES

(1) L. W. Morley et al. Organization for a National Program on Remote Sensing of the Surface Environment. (A Working Paper) October, 1970

(2) National Aeronautics and Space Administration. Data Users Handbook, NASA Earth Resources Technology Satellite. Document No. 71SD4249 September 15, 1971

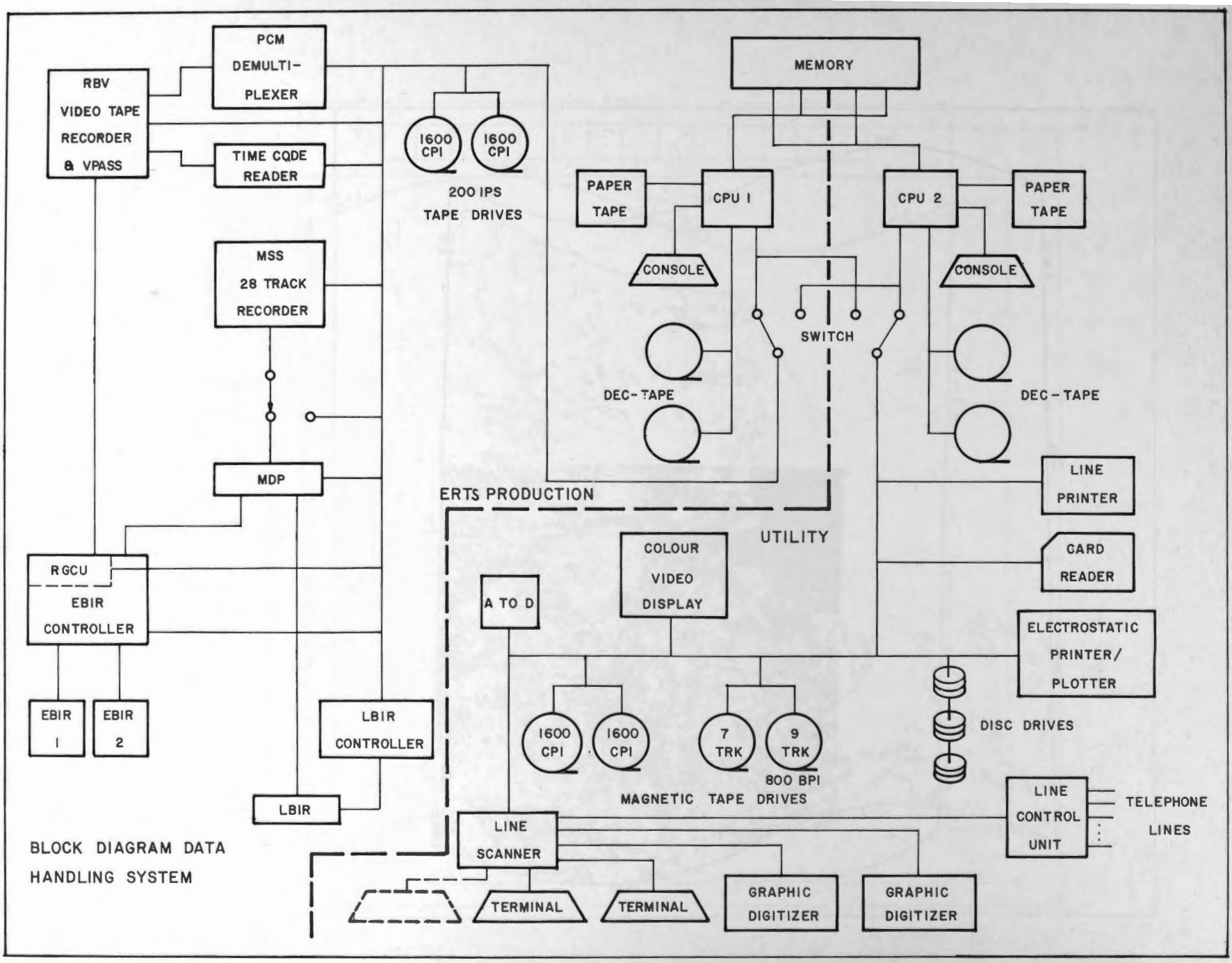


Figure 1. Block diagram data handling system

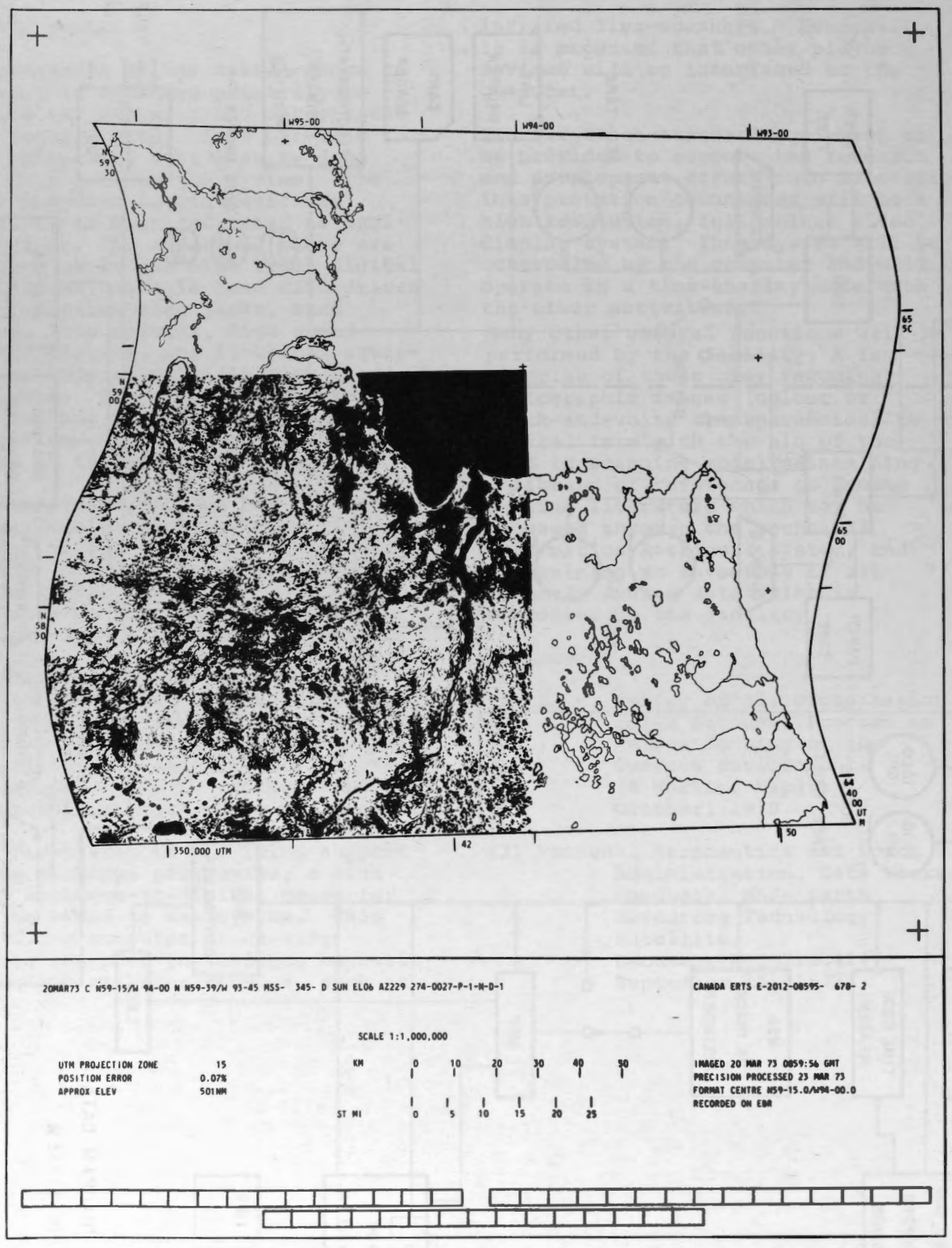


Figure 2. Sample MSS data produced by EBIR

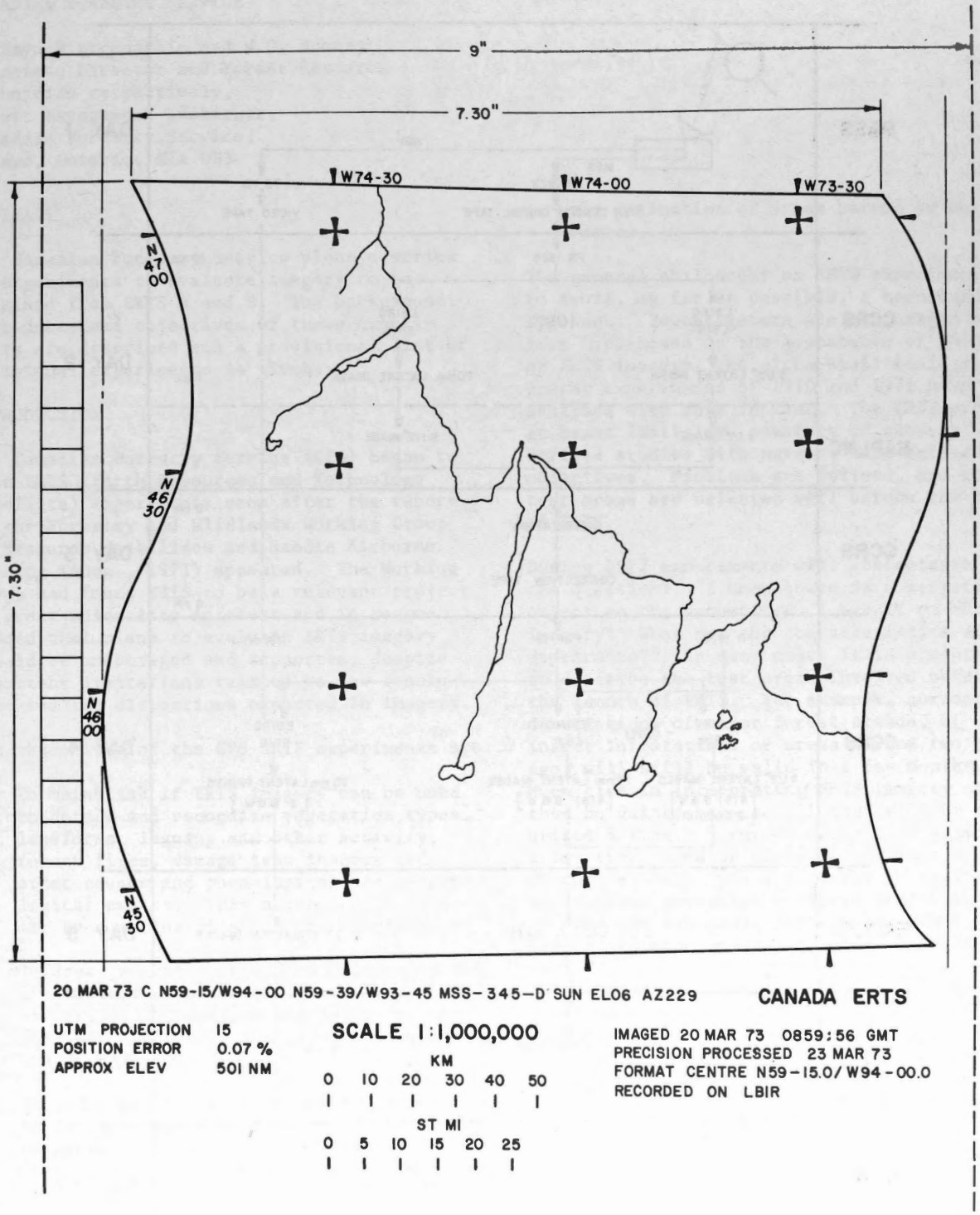


Figure 3. Sample MSS data produced by LBIR

ERTS PRODUCTION SCHEDULE

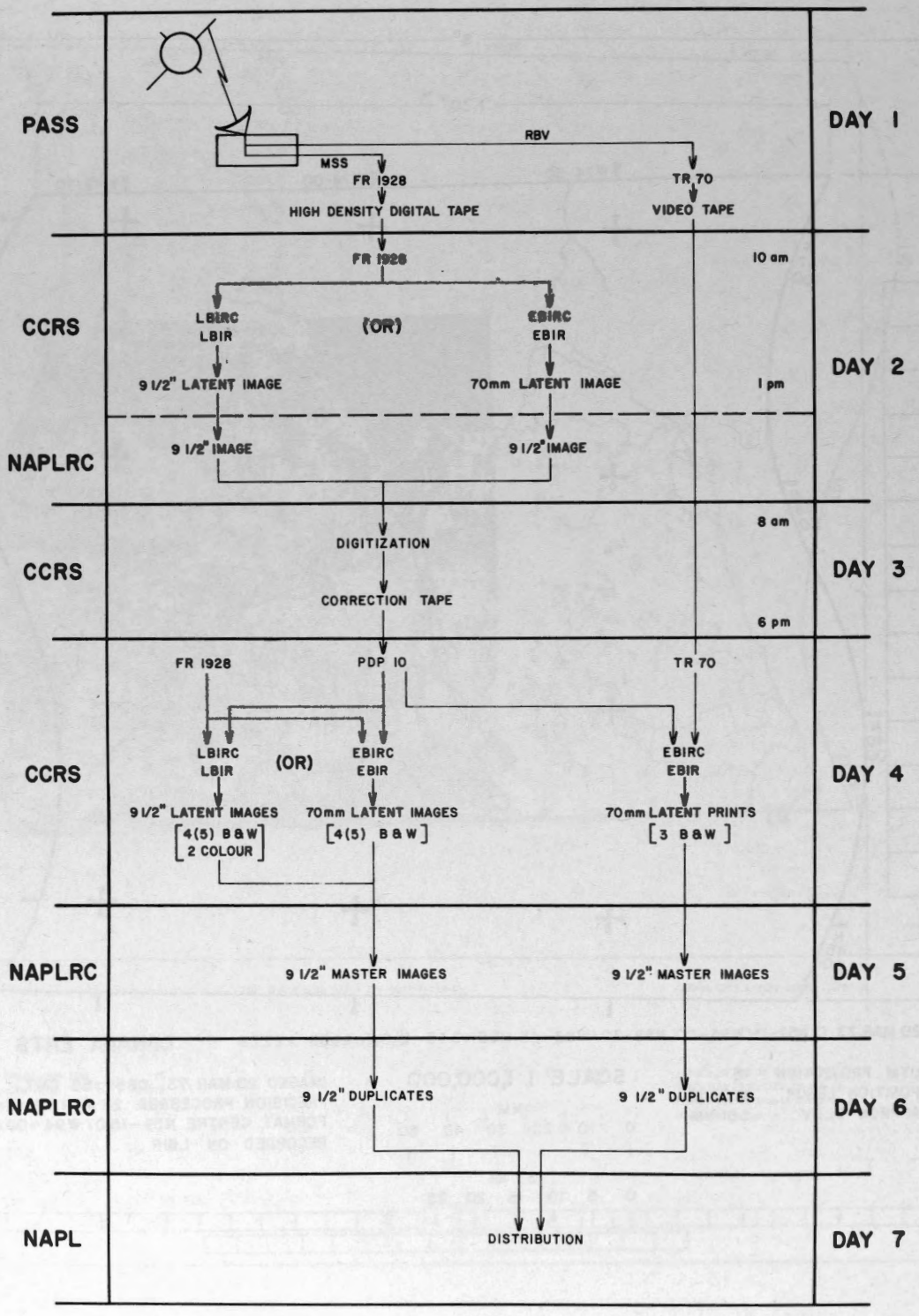


Figure 4. ERTS production schedule

THE ERTS EXPERIMENTS OF THE CANADIAN FORESTRY SERVICE

L. Sayn-Wittgenstein and W.C. Moore,
Associate Director and Forest Research Technician respectively,
Forest Management Institute,
Canadian Forestry Service,
Ottawa, Ontario, K1A 0H3.

ABSTRACT

The Canadian Forestry Service plans a series of experiments to evaluate imagery to be obtained from ERTS A and B. The background, procedures and objectives of these experiments are described and a provisional list of individual experiments is given.

INTRODUCTION

The Canadian Forestry Service (CFS) began to plan ERTS (Earth Resources and Technology Satellite) experiments soon after the report of the Forestry and Wildlands Working Group on Resource Satellites and Remote Airborne Sensing (Anon., 1971) appeared. The Working Group had found ERTS to be a relevant project of great scientific interest and it recommended that plans to evaluate ERTS imagery should be encouraged and supported, despite important limitations related to low resolution and the distortions expected in imagery.

The objectives of the CFS ERTS experiments are:

1. To establish if ERTS imagery can be used to detect and recognize vegetation types, landforms, logging and other activity, forest fires, damage from insects and other causes and phenological and meteorological events. This objective is to be met by a series of specific experiments.

2. To draw general conclusions concerning the limitations (e.g., resolution and position errors) of ERTS and on its value for mapping and monitoring the forest resources and environment.

3. To undertake research to improve interpretation techniques. This includes image enhancement, pattern recognition, the development of interpretation keys.

4. To make operational use of ERTS imagery. Plans will be made after review of preliminary results from (1) and (2) above. Immediate applications may lie in the mapping of major vegetation zones and in the estimation of areas burned or harvested.

The general philosophy on ERTS experiments is to avoid, as far as possible, a haphazard approach. Investigators are encouraged to form hypotheses on the appearance of phenomena on ERTS imagery; the ultra-small-scale photography experiments of 1970 and 1971 have been analysed with this in mind. The ERTS program, at least initially, consists of clearly defined studies with narrow, well-defined objectives. Problems are defined, and most test areas are selected well before the launch of ERTS.

During 1972 experiments will concentrate on the question: "I know there is a certain object on the ground; can I see it on ERTS imagery? What are the characteristics of its appearance?" In many cases it is acceptable to describe the test areas involved before the launch of ERTS. For example, current descriptions of major forest stands, of old insect infestations or areas burned two years ago, will still be valid in a few months. Much time in interpreting ERTS imagery can thus be gained because all that will be required during the summer of 1972 is a confirmation that no major disturbances have occurred. Of course, this approach cannot be applied to short-lived phenomena or where significant changes are expected, for example, in active insect infestations or phenological events.

ERTS experiments for which plans and preparations have begun are listed (Table 1) and the areas involved are shown (Figure 1). The experiments are numbered and these numbers will be used for cross-references throughout the text below. We expect the list of proposed experiments to grow substantially next summer if initial interpretation shows that ERTS was a success as far as forestry is concerned.

The proposed experiments are associated with current programs of the CFS; many deal with objectives which are now met by other remote sensing methods. Also they involve data,

techniques and approaches developed for other projects. The following sections indicate the close relationship between current programs, current methods and the ERTS experiments.

MAPPING FOREST LAND AND VEGETATION

One of the biggest requirements for remote sensing in forestry stems from the continuing need to map and describe forest land, forest types and other vegetation.

Aerial photographs have for decades been used in the preparation of "forest-type maps". These maps typically show the species composition, height and density of forest stands. They are a good summary of the essential characteristics of a forest, and find use in forest management, planning of logging operations, road construction and fire protection. One of their main rôles is in estimating timber volume, as will be discussed later.

Few expect ERTS to yield imagery which can be used for the mapping of forest types with the precision of the usual forest inventories, but imagery may be useful for reconnaissance inventories, for mapping major vegetation zones and to monitor major changes in the forest environment.

For example the CFS is involved in a crash program to map the vegetation within the 1600 mile long corridor considered for transportation and pipeline routes along the Mackenzie River. It will be established whether satellite imagery could be used to accomplish such a task more efficiently in the future. The interest is not only in vegetation mapping but also in the recognition of landforms and in observing the interrelationship between human activity, land and vegetation. Experiments FMI-6 and FMI-10 apply directly; recent work (Gimbarzevsky 1972) in using high-altitude photography for terrain analysis will be used in bridging the gap between satellite imagery and ground observation.

A significant forest classification for Canada (Rowe 1959) recognizes forest regions based upon characteristics of tree-species distribution and topography. With the help of ultra-small-scale photography obtained under the 1970 aircraft program, the Forest Management Institute has already established the value of a synoptic view for such tasks (Nielsen and Wightman, 1971). Several ERTS experiments will deal with the recognition of major forest and vegetation types (AL-2, BC-3, FMI-8, 10, NF-1). These experiments have a high probability of success because ERTS imagery will have a quality between that of ultra-small-scale aerial photos, which contain far more information than necessary, and weather-satellite imagery, which has been shown to have some possibilities for recognizing major vegetation zones, although there are severe limitations due to low resolution (Aldred 1968).

One chronic problem, for which there is not yet a satisfactory solution is in the updating of vegetation maps and statistics. The value of forest resources changes with time. These changes could be beneficial or ruinous, short-term or long-term, trivial or significant — but combinations of such changes are universally present. Up-to-date information on the distribution, quantity and quality of the forest resource is required for forest management and research.

But forest-cover maps are often outdated before they are completed, for example, timber may have been harvested or burned; in addition there are the more gradual changes through growth. Subsequent sections will discuss ERTS experiments dealing with the detection of forest damage, logging and other human activities; these studies will give an indication of the value of ERTS for keeping maps and statistics up-to-date.

TIMBER VOLUME ESTIMATION

Forest inventories involve the preparation of forest-type maps and the estimation of timber volume by species, size and quality. The two are almost inseparable because timber volume estimates are usually derived through ground- or photo-sample plots established within strata defined by the forest-type maps and area estimates based on these maps. Estimates may be required in detail for small areas or as broad summaries for several thousand square miles.

Timber volume estimation is so important that there inevitably will be tests to establish whether variables observed on ERTS imagery are correlated with timber volume. One traditional approach is based on "stand volume tables" which give timber volume from photo-estimated tree species composition, stand height and canopy density. No one expects a similarly close relationship for ERTS imagery, but as long as there is a significant relationship, there will be attempts to incorporate such imagery into forest inventory procedures, as has already been done with Apollo 9 photographs (Aldrich 1970, Langley 1969).

An important contribution of the CFS forest

inventory capability will be in documenting ERTS test sites. Conventional aerial photography and the skill of interpreters trained in inventory tasks will be used in describing vegetation, disturbances and other conditions on test sites. Large-scale aerial photography and radar altimetry methods developed by the Forest Management Institute during the last ten years will play an important role and the statistical methods developed for timber volume estimation will be used in the collection of data necessary for the interpretation of ERTS imagery.

FOREST DAMAGE

The CFS includes the Forest Insect and Disease Survey, one of the World's largest and best organized operations for monitoring insect infestations and the distribution and changes in insect populations; this information is summarized in comprehensive annual reports. Several major active insect outbreaks are under observation, and reliable information on the location of old outbreaks is available. This information will be used to establish if ERTS imagery might be used to detect new outbreaks and to map old ones. One of the experimental areas now being documented concerns a major spruce budworm outbreak in Ontario (ON-2); another concerns several insect outbreaks in British Columbia (BC-1) and builds directly upon small-scale aerial photography experiments carried out in 1971 (Harris 1972).

On several occasions the CFS has been involved in the mapping and assessment of pollution damage to vegetation. In one study small-scale photography, aerial observations and ground checks were used to map levels of damage from SO_2 emissions near Wawa, Ontario (Murtha, 1972). This study is also the basis of an ERTS experiment.

Similarly descriptions and maps resulting from studies of burned areas, of wind damage, floods and landslides will be used in experiments to test the value of ERTS for detecting such phenomena. For example, one study (Al-3) involves a large burn in Alberta; in another an attempt will be made to detect and describe an area in northern Ontario where a strong twister caused extensive damage a few years ago (FMI-2). Experiments to monitor forest fires of known size and intensity are planned for ERTS B (ON-1).

LOGGING AND OTHER HUMAN ACTIVITY

There are many reasons why efficient methods of detecting and describing human activities in the forest are desirable. One, mentioned earlier, is simply to keep maps and inventory statistics up-to-date, to have at any time an account of the size of the forest resource. Other reasons are the need to measure the progress of logging and construction operations and in some cases to observe the impact of human activity on the environment.

In theory, satellite imagery and even high-altitude aerial photography would appear to be the best methods for monitoring changes because they provide an opportunity for the rapid and repeated coverage of large areas. In practice the opposite approach has often been taken; if activities such as logging or road construction were monitored from the air, one used large scales of photography. The reason for this was probably not so much the requirement for great detail, as the fact that the areas involved were small and that their locations were known precisely; thus it was practical to use light aircraft and relatively simple cameras.

However, logging activities can be detected and even precisely described on photographs at very small scales, such as 1:160,000; this has been shown in a paper presented at this Symposium (Wightman 1972). The CFS ERTS experiments will provide several opportunities to examine the appearance of logging operations (e.g., BC-4, FMI-8). Attempts will also be made to record seismic lines in the Mackenzie valley (FMI-6) and to follow the progress of a power line under construction in British Columbia (BC-4); we will also establish if the large log booms in the Fraser River delta appear on ERTS imagery.

PHENOLOGICAL AND METEOROLOGICAL PHENOMENA

ERTS may well be at its best when recording changing phenomena. The big advantage will lie in sequential coverage and in the possibility of observing phenomena simultaneously at many locations distributed over a large area.

There is for example meager information on the dates and sequence of phenological events in a forest. For example, when and at what rate does deciduous foliage appear; when, and in what order do trees of different species lose their leaves or change to fall colours? How are these events related to meteorological factors? What is the combined effect of weather and phenological characteristics of trees on insect populations, on the success of plantations and on the rate of water run-off?

Two experiments are planned to examine the leafing out and leaf fall of deciduous forests and to follow other seasonal changes. One study (FMI-7) will concentrate upon a transect from Richmond to Moosonee, Ont. which already is the site of other remote sensing experiments (Nielsen and Wightman 1971, Wightman 1972). The other (FMI-9) is to compare ERTS imagery with a map showing lines joining points of equal phenological development for the Maritime provinces.

Three experiments deal specifically with meteorological factors. One (AL-1) is a study of snow packs during chinooks, involving in particular, the relationship between the rate of disappearance of snow packs and landforms and elevation. Experiment AL-4 involves an analysis of the patterns of snow melting on a large area. In British Columbia there is a study of the distribution of snow, fog and clouds; a comparison of ERTS and weather satellite data is also involved (BC-2). The CFS will follow with interest the results of other ERTS experiments in meteorology.

THE ADMINISTRATIVE AND ORGANIZATIONAL CONSIDERATIONS

The Canadian Forestry Service is a decentralized organization. Its programs are essentially divided among two types of organizations: (1) regional establishments, such as the Pacific Forest Research Centre in Victoria or the Great Lakes Forest Research Centre in Sault Ste. Marie, which deal with problems particularly relevant to a geographic area, and (2) specialized institutes such as the Forest Pathology Institute, the Forest Fire Research Institute and the Forest Management Institute which deal with matters on a discipline basis. ERTS experiments are planned in the regional establishments and in the institutes, but the Forest Management Institute, centre of CFS remote sensing activities, coordinates the ERTS program.

The initial functions of the Forest Management Institute included dissemination of ERTS information to regional establishments and the discussion of the potential of ERTS with interested scientists in these establishments. The Forest Management Institute also is the link between the CFS and the Canadian Centre for Remote Sensing.

The ERTS experiments within regional establishments evolved more or less spontaneously and on a voluntary basis. No pressure was exerted on individual scientists to make proposals and commitments. As a result, the distribution of experiments is somewhat uneven; in part the Forest Management Institute has filled the remaining gaps.

Many of the ERTS investigators (Table 1) are not specialists in remote sensing, but scientists in such fields as entomology, forest and land inventory, hydrology and forest fire research, or they may be persons who know the ERTS test area involved particularly well. With few exceptions it is proposed that the investigators will make the first interpretation of imagery but this evaluation will be followed by discussions, comparisons and cooperation with others involved within related experiments, or with specialists at the Forest Management Institute who have access to interpretation equipment and methods not available at regional establishments.

Investigators have been issued with guidelines for the documentation of test areas. These guidelines for example discuss the size of areas to be chosen. They also give specific information to be followed in describing the location of test areas; for example the area should be shown on a map or aerial photograph, map sheet or photo number should be given, and the UTM grid coordinates should be shown. Each investigator should comment on the reliability of the ground observations; he should if possible include black and white photographs illustrating the area on the ground, and he should cite any important literature references that might be used to obtain further background information on the area. Investigators were supplied with a sample description of a hypothetical test area.

The purpose of these rather detailed instructions was to ensure that no significant information would be omitted and to standardize procedures. This will facilitate the comparison of results of different investigators. Also the Canadian Forestry Service plans to issue the results of its initial experiments in a joint report to which all investigators will contribute; standardization will be a help in completing this report.

CONCLUSIONS

The Canadian Forestry Service faces a few problems in its program of ERTS experiments. One, naturally, is the lack of experience with ERTS imagery. This problem will surface as soon as investigators place orders for imagery. Few guidelines are available in deciding what type of imagery to request and when to insist upon corrected imagery.

Another consideration is that the ties between CFS investigators and the suppliers of ERTS

TABLE 1

PRELIMINARY LIST OF CFS ERTS EXPERIMENTS

Experiment No.	Subject (Location)	Investigator	Remarks
FMI-1	SO_2 damage (N. Ontario)	P.A. Murtha, FMI*	Extension of current remote sensing study.
FMI-2	Wind damage (N. Ontario)	W.C. Moore, FMI	Ont. Lands & Forests and Federal meteorological data available for Sudbury 1970 tornado.
FMI-3	Landslide damage detection (Que., Ont., B.C.)	W.C. Moore, FMI	Landslides at St. Jean Vianney, Que.; South Nation, Ont.; and Frank, Alta.
FMI-4	Drainage basin study (S. Ontario)	W.C. Moore, FMI	Rideau Valley Conservation Authority Report, 1968, provides excellent ground truth.
FMI-5	Delineation of forest flooding (Ont., Man.)	W.C. Moore, FMI	Ottawa, Rideau and Red Rivers offer three different sets of circumstances.
FMI-6	Seismic line detection (Mackenzie Valley)	W.L. Wallace, FMI	Extensive activity in the Arctic may have important environmental implications.
FMI-7	Leaf development variations (S. to N. Ontario)	U. Nielsen and J.M. Wightman, FMI	Richmond-Moosonee strip was flown in 1970 and re-scheduled for 1972.
FMI-8	Mapping of forest regions (Canada)	U. Nielsen and J.M. Wightman, FMI	1970/71 high-altitude experiments relevant.
FMI-9	Maritime forest phenology (New Brunswick)	L. Sayn-Wittgenstein, FMI	Ground truth based on a study by Forbes and Webb, MFRC.
FMI-10	Arctic landforms and ecology (Mackenzie Valley)	L. Sayn-Wittgenstein, P. Gimbarzevsky, FMI	Involves Mackenzie Transportation Corridor.
BC-1	Forest insect attacks (S. British Columbia)	J.W.E. Harris and R.F. Shepherd, PFRC	A dynamic phenomenon; sequential coverage required.
BC-2	Snow and fog distribution (Vancouver Island)	E.T. Oswald, PFRC	Microclimates in the forest regime.
BC-3	Tundra, subalpine, montane and grassland vegetation (B.C. Interior)	E.T. Oswald, PFRC	Canada Land Inventory survey in 1971 should provide accurate ground truth.
BC-4	Logging, regeneration, and human activities (Vancouver Island and B.C. Interior)	Y. Lee, PFRC	Dynamic phenomena that might be monitored by ERTS.

Table 1 (continued)

Experiment No.	Subject (Location)	Investigator	Remarks
AL-1	Chinook snowpack sublimation (S.W. Alberta)	D.L. Golding, NFRC	Relevant to forest and watershed management.
AL-2	Boreal vegetation (N.W. Alberta)	P. van Eck, NFRC	Involves the most important forest region of Canada.
AL-3	Forest fire incidence patterns (Central Alberta)	J. Niederleitner, NFRC	Vegetation succession after fire included in study.
AL-4	Snow-melt patterns (Alberta)	J.M. Powell, NFRC	Some ground studies have been completed; significant for forest regeneration studies.
AL-5	Forest management (Rocky Mts. East Slope)	C.L. Kirby and K. Froning, NFRC	Well documented test area with many species.
ON-1	Forest fire sizes and intensities (E. Ontario)	C.E. Van Wagner, PFES	ERTS B experiment
ON-2	Spruce budworm infestations (N. Ontario)	W.L. Sippell, GLFRC	Major outbreaks might be delineated.
NF-1	Lichen forest recognition (Labrador)	W.C. Wilton and R. Wells, NFRC	An important northern forest type.
NF-2	Strip cutting of black spruce (Newfoundland)	R.C. van Nostrand, NFRS	A major logging area with well-documented silvicultural experiments.

*FMI — Forest Management Institute, Ottawa; PFRC — Pacific Forest Research Centre, Victoria; NFRC — Northern Forest Research Centre, Edmonton; PFES — Petawawa Forest Experiment Station, Chalk River; GLFRC — Great Lakes Forest Research Centre, Sault Ste. Marie; NFRC — Newfoundland Forest Research Centre, St. John's; MFRC — Maritimes Forest Research Centre, Fredericton.

imagery (the Canadian Centre for Remote Sensing and ultimately NASA) are far looser than the essentially contractual relationships that bind NASA and the individual U.S. ERTS experimenters. Thus, for example, how certain can one be of obtaining the supporting aircraft data during 1972? No doubt this will be clarified during the next few weeks, but plans had to be completed simply on the assumption that data would be available.

Then there is the risk of negative results, and this does not merely refer to the possibility of failure of the satellite or readout system. Rather, the CFS has incurred considerable risks by selecting and documenting test sites before imagery is available. It is to be expected that no acceptable imagery will become available for some of these sites, because of cloud cover or technical problems, but we hope that the percentage will be small.

This early selection of test sites was made for two reasons. First, to speed up the process of ERTS interpretation and secondly to cut down on what Robert N. Colwell calls "ghee-whizz research". "Ghee-whizz research" arises when someone picks up a piece of ERTS imagery and exclaims: "I have this image and

ghee-whizz, I can see something on it; ghee-whizz, it's a lake, or a forest!" and off we rush to the printers. One substitute for this haphazard and arbitrary way of arriving at conclusions is through the selection of test sites after *a priori* consideration of significant problems and important areas.

In analysing the results of the interpretation of ERTS imagery one will have to distinguish exactly between (a) observing an object on ERTS imagery, only because it was known that the object had to be there, (b) detecting an anomaly or change, but being unable to identify it and (c) correctly identifying an unexpected object. If the second of these alternatives is very common, as it well may be, there will be great emphasis upon sampling procedures using ground work and aircraft.

Satellite imagery will probably be another step forward in expanding and improving the techniques available for surveying the forest environment. It is a part of the progression linking field-work, large-scale photography, conventional medium- and small-scale photography and ultra-small-scale aerial photography.

The major problem in making efficient use of satellite imagery will be in extracting useful information from the mass of available data.

We do not think the answer to this problem lies in low resolution systems, which achieve "automatic data reduction". The requirements for data are too varied to justify such arbitrary truncation at any stage of data collection. The answer will lie in the development of efficient statistical methods for retrieving desired information as Aldred (1972) has discussed at this Symposium, and it will lie in clear ideas about the purpose, benefit, cost and efficiency of any proposed task.

REFERENCES

ALDRED, A.H. 1968. Potential of satellite photography in forestry. Forest. Manage. Inst., Dep. of Forest. and Rural Develop., Inform. Rep. FMR-X-11. 18 p.

ALDRED, A.H. 1972. Economics of combining remote sensing data collected at several scales. First Can. Symp. Remote Sensing, Ottawa, Can. Centre Remote Sensing, Feb. 7, 8, 9.

ANON. 1971. Report No. 5, Forestry and Wildlands. Resource Satellites and Remote Airborne Sensing for Canada, Inform. Can., Cat. No. M75-2/5. 17 p.

GIMBARZEVSKY, P. 1972. Terrain analysis from high-altitude aerial photographs. First Can. Symp. Remote Sensing, Ottawa, Can. Centre Remote Sensing, Feb. 7, 8, 9.

HARRIS, J.W.E. 1972. High-level photography for forest insect damage surveillance in British Columbia. First Can. Symp. Remote Sensing, Ottawa, Can. Centre Remote Sensing, Feb. 7, 8, 9.

LANGLEY, Philip G. 1969. New multi-stage sampling techniques using space and aircraft imagery for forest inventory. Sixth Int. Symp. Remote Sensing Environ., Univ. Mich., Oct. 13, 14, 15, 16. II: 1179-1192.

MOORE, W.C. 1972. Remote airborne sensing of water pollution: Rideau River drainage basin. First Can. Symp. Remote Sensing, Ottawa, Can. Centre Remote Sensing, Feb. 7, 8, 9.

MURTHA, P.A. 1972. Sulfur dioxide damage delineation on high-altitude photographs. First Can. Symp. Remote Sensing, Ottawa, Can. Centre Remote Sensing, Feb. 7, 8, 9.

NIELSEN, U., and J.M. WIGHTMAN. 1971. A new approach to the description of the forest regions of Canada using 1:160,000 color infrared aerial photography. Forest Manage. Inst., Can. Forest. Serv., Dep. Environ., Inform. Rep. FMR-X-35. 25 p.

ROWE, J.S. 1959. Forest regions of Canada. Can. Dep. North. Aff. Nat. Resour., Forest. Br., Bull. 123. 71 p.

WIGHTMAN, J.M. 1972. High-altitude photography records and monitors logging operations. First Can. Symp. Remote Sensing, Ottawa, Can. Centre Remote Sensing, Feb. 7, 8, 9.

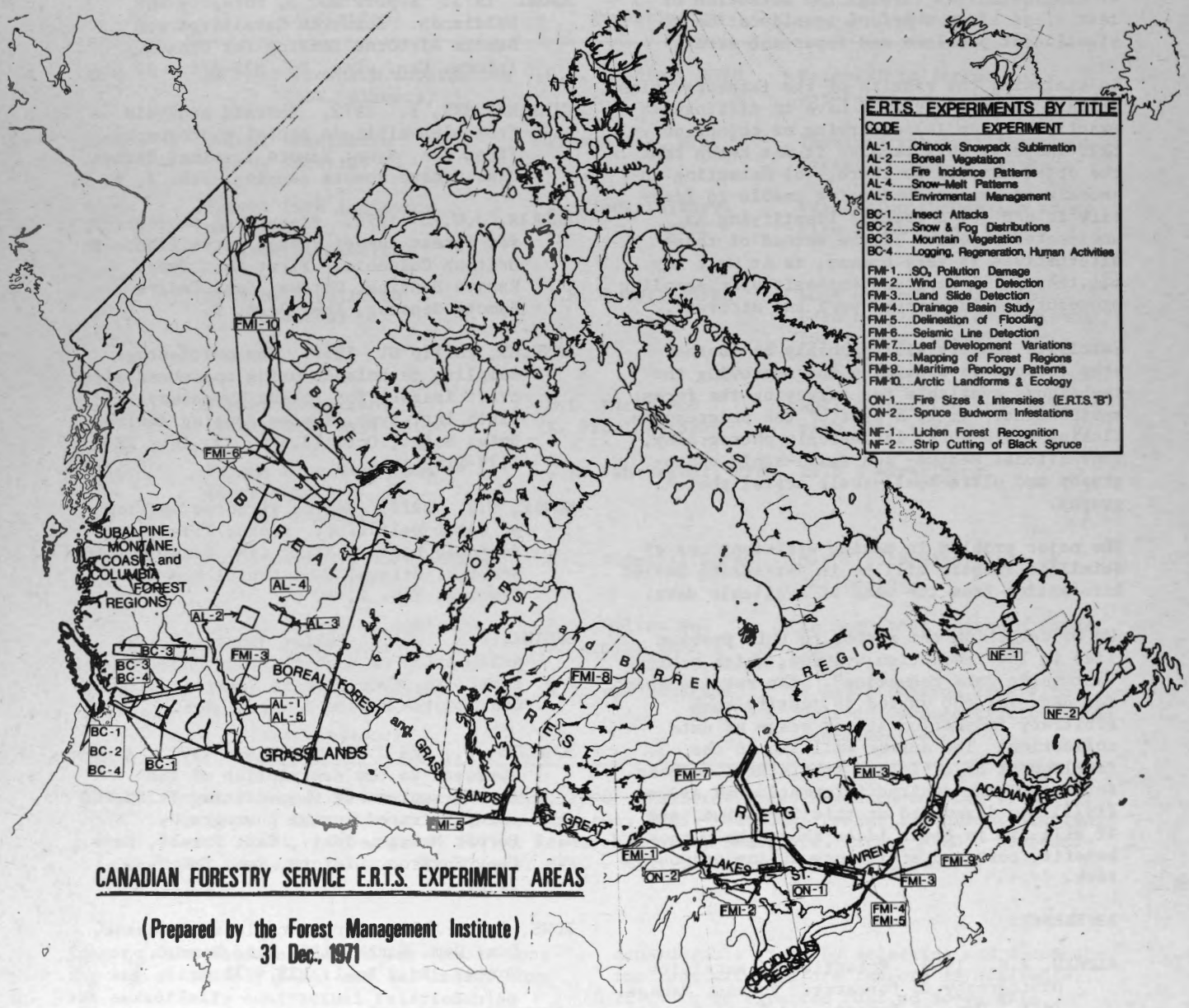

E.R.T.S. EXPERIMENTS BY TITLE
CODE EXPERIMENT
AL-1.....Chinook Snowpack Sublimation
AL-2.....Boreal Vegetation
AL-3.....Fire Incidence Patterns
AL-4.....Snow-Melt Patterns
AL-5.....Enviromental Management
BC-1.....Insect Attacks
BC-2.....Snow & Fog Distributions
BC-3.....Mountain Vegetation
BC-4.....Logging, Regeneration, Human Activities
FMI-1....SO2 Pollution Damage
FMI-2....Wind Damage Detection
FMI-3....Land Slide Detection
FMI-4....Drainage Basin Study
FMI-5....Delineation of Flooding
FMI-6....Seismic Line Detection
FMI-7....Leaf Development Variations
FMI-8....Mapping of Forest Regions
FMI-9....Maritime Penology Patterns
FMI-10...Arctic Landforms & Ecology
ON-1.....Fire Sizes & Intensities (E.R.T.S."B")
ON-2.....Spruce Budworm Infestations
NF-1.....Lichen Forest Recognition
NF-2.....Strip Cutting of Black Spruce
SUBALPINE, MONTANE, COAST, and COLUMBIA FOREST REGIONS
BOREAL FOREST
BOREAL FOREST and GRASSLANDS
GRASSLANDS
BOREAL FOREST
BARREN REGION
GREAT LAKES ST. LAWRENCE REGION
ACADIAN REGION
CANADIAN FORESTRY SERVICE E.R.T.S. EXPERIMENT AREAS
(Prepared by the Forest Management Institute)
31 Dec., 1971

Figure 1

EXPERIMENTS IN AERIAL REMOTE SENSING FOR HIGHWAY ENGINEERING

Cal D. Bricker
Department of Highways and Transport
Edmonton, Alberta

SUMMARY

The Alberta Department of Highways and Transport is researching various remote sensing methods to assess their possible application to highway engineering. The object of the program is to assess photographic and non-photographic sensors as they become available, by evaluating each type and combinations of types to currently used panchromatic air photography.

Ground test sites were selected that contained a diversity of cultural and natural features encountered in highway route location in Alberta. These sites, previously well covered by existing panchromatic photography at various scales, were field investigated prior to and immediately after each sortie to establish ground truths.

The remote sensing sorties were flown to flight-planned specifications in varying combinations of photographic and non-photographic sensors. In addition to the contracted coverage, remote sensing imagery flown by the Remote Sensing Centre of Canada was incorporated into the program as it became available.

Interpretation experiments were carried out in each type of imagery, in combinations of types and in photographically reproduced variations of the original, to define the type (s) that may at present produce the maximum usable data for highway location studies.

This paper describes generally and not specifically the first of these experiments in a continuing program of analyzing types of remote sensing imagery as they become available and comparatively assessing the cost/value of each type with that of panchromatic air photography.

INTRODUCTION

The Alberta Department of Highways and Transport at present exclusively utilizes panchromatic air photography for gathering data applicable to highway engineering. Although the photography is used for planning, construction, maintenance and many secondary uses, this paper deals only with its major application - highway route location.

The majority of air photography is flown by aerial survey companies under contract to the Department and the remainder procured from provincial and federal air photo libraries. All of the contracted air photography is available to the public by purchase through the Provincial Government's Air Photo Library. The planning branch has over the past five years purchased by contract a yearly average of 1860 flight line miles of new photography in the form of the film and two sets of photographs and an additional 9839 contact prints from other sources. The photography used primarily for air photo interpretation in route location studies is flown at scales of 1 in. =500 ft, 1 in. =100 ft and 1 in. = 2000 ft with specific areas requiring more detailed study at larger photo scales. Library reprints usually at a scale of 1 in. =2640 ft are used mostly for regional terrain studies and are supplemented in a few areas of the Province by older 1 in. =3333 ft coverage.

With air photo interpretation an established part of route location studies and indications of yearly increases in the amount of photography required, the recent availability of new aerial sensors and camera systems suggests an even more efficient means of obtaining data. As our experience with remote sensing, other than panchromatic air photography, was limited, an evaluation program designed to our specific requirements was established. The objective being to assess the potential value to cost of each type of imagery and combinations of types and to evaluate them against a known factor - the procurement and interpretation of panchromatic air photography. As the cost of procuring aerial imagery continuously varies and the value of the interpretation fluctuates with the skill and experience of the interpreter, the cost/value estimations are certainly more probable than actual. Basically, the program is to assess the benefits derived from each

remote sensing method and evaluate these benefits with the present cost.

METHOD

A series of remote sensing flights were made of three test sites in different areas of the Province. The criteria in selecting the sites was that they contain classic examples of different cultural and natural features encountered in route selection studies, that they were accessible for ground investigation and that they had been previously well covered by panchromatic air photography at various photo scales. Each type of imagery was then individually and comparatively interpreted utilizing all the background studies and ground truth information available.

Airborne coverage of test site one - Edmonton-East, was co-ordinated with the faculty of agriculture of the University of Alberta and Grumman Ecosystems on a cost-sharing basis. Although separate interpretation studies were made in Edmonton with different, but related, objectives, an exchange of information was made throughout at the working level.

The test site was flown in August 1970 in color and infrared false color exposed simultaneously on a time-coincident and field of view-coincident frame by frame basis at a scale of 1 in. =1760 ft. Thermal imagery in 4.5 to 5 microns was obtained on the same sortie and repeated that night in 0.5 to 5 microns.

Imagery from flights sponsored by the Remote Sensing Centre of Canada in color, infrared false color, infrared black and white, black and white (red and green) and thermal scanning at 3 to 5 microns was also incorporated into the program as it became available.

Test site two - Devon, Alberta, was flown in August 1970 at a scale of 1 in. =1000 ft with panchromatic, color and infrared false color. Test site three - Edson-East, was flown in September 1971 at a scale of 1 in. =1000 ft in color and infrared false color.

Coverage of sites two and three flown specifically for the Department by Alberta organizations was semi-simultaneous, that is, each site was taken with one camera on the same sortie as fast as film and filters could be changed.

GROUND INVESTIGATION

The sites were thoroughly investigated on the ground utilizing available studies, maps and reports. Ground truths were established and key features photographed in some instances in panchromatic, color and infrared false color for comparison with similar airborne films. Later ground investigation checked key features against their signatures on the imagery.

An experiment to assess the readability of ground targets for use as control points in mapping was attempted. Ground targets in various sizes, shapes, colors and reflective surfaces were laid out on different terrain backgrounds adjacent to one of the sites. The concept and construction of the targets was more successful than the results obtained - the aircraft missed the target area.

MOSAICS

Semi-controlled mosaics were constructed of each area for background study purposes. Those for regional studies were compiled from small- and medium-scale existing photography and of each test site from the program's larger scale photography.

An uncontrolled mosaic was also constructed from the night time thermal imagery to provide a regional 'picture' during detailed study of each line. The centre portion of each line-strip was photographically enlarged and printed on paper with a glossy finish and normal mosaic construction techniques used with adjustments to compensate for the distortion inherent in line scanning. Lateral oblique distortion and improperly aligned flight lines necessitated variations in the amount of lateral overlap used to butt together line-strips, resulting in varying amounts of distortion throughout the mosaic. Generally, the mosaic was very acceptable and of great assistance in relating an object under study to the overall view.

All of the mosaics were photographically copied, and by different techniques photographic and blue-print paper copies produced. The blue-print paper prints, although lacking the clarity of the original, were an economical method of annotating preliminary study data.

INTERPRETATION

Each test site was quantitatively interpreted and selected terrain features such as granular deposits, vegetation, sand dunes and muskeg were subject to qualitative study.

The interpretation techniques used in the program to date are:

- Separate interpretation of the panchromatic and color prints and some of the infrared false color in print form.

- Separate interpretation of the panchromatic and color negatives and the infrared false color transparencies.

- Interpretation of the panchromatic, color and infrared false color in various combinations to utilize the complementary attributes of each type.

- Interpretation by cross-pairing successive frames of different imagery.

- Separate and combined interpretation of the day and night thermal imagery.

- Comparative analysis of the day thermal imagery with the simultaneously acquired color and infrared false color.

- Comparative evaluation of small-scale 70 mm imagery to medium-scale 9 inch format photography for regional terrain studies.

- Selections of different types of imagery covering the same terrain area were distributed among engineers and technical personnel presently using black and white air photography for opinions as to the value of each type to their particular field.

The first phase experiments have indicated that a more comprehensive extraction of data is possible by employing other methods of remote sensing than black and white photography. We are now assigning specific sensors to specific areas of study for a more thorough evaluation.

Time allows only a brief summary of some of the results obtained to date.

Color photography is preferred by the majority of personnel for interpretation as it appeared more realistic than the varied tone black and white photography. It also supplied more data than any one of the other types of imagery.

Combinations of black and white, color and infrared false color, although time consuming to interpret, supplied the largest amount of information.

Infrared false color more clearly defining wet and dry areas and tree types has an advantage over other imagery for mapping drainage, patterns and the study of muskeg. Indications are that this film has a possible application in interpreting areas of potential slumping.

The identification of ground signatures indicating possible sources of sand and gravel is more clearly defined on small-scale color, and in certain instances infrared false color, than on black and white.

The application of infrared black and white, color, and infrared false color to mapping drainage patterns, muskeg and granular deposits is to be the subject of further and more exacting experiments.

CONCLUSIONS

In mediums other than panchromatic photography our limited experience with other systems led to many varied productive and non-productive interpretation techniques, some of which, without further controlled tests, may have led to incorrect assumptions and attributed an importance or lack of importance to types of imagery not in keeping with their true value.

Although our experiments are imcomplete, indications are that all of the photographic sensors have an application to highway engineering. A cost/value assessment based on today's cost of procuring aerial imagery and our interpretation studies to date, indicate - (1) panchromatic air photography is the best dollar value for programs involving large quantities of imagery (2) combinations of panchromatic, color and infrared false color supply more information than a single type and are the best cost to value if employed in small terrain areas for specific requirements.

Non-photographic thermal imagery requires more expertise in its procurement and interpretation before an assessment can be made of its true capabilities. Our experiments have indicated that it has a potentially valuable application in specific areas of highway engineering.

A PROPOSED ORGANIZATION FOR THE EFFICIENT INTERPRETATION OF REMOTE SENSING DATA

H. W. Thiessen, Director,
Interdepartmental Planning Division,
Alberta Department of the Environment,
Edmonton, Alberta.

SUMMARY

In Alberta, we have now had one season's experience of remote sensing service with the Remote Sensing Center. This experience included several sources of photography in a variety of locations throughout Alberta requested by numerous provincial government agencies under the auspices of the interdepartmental Conservation and Utilization Committee. In addition, several departments of the federal government, and the University of Alberta, have participated in remote sensing photography within Alberta during the same period.

Prior to this time, we have had other experiences with private contractors utilizing remote sensing in several interdisciplinary studies. These experiences have convinced us that the existing organization and use of remote sensing is not meeting the challenge of the need for this capability, nor is it anticipating the full potential of the technology.

We would propose the development of an organization involving interdisciplinary analysis comprised by a variety of disciplines and supported by several levels of government and research institutions. In this manner, we may realize the full potential of this science with the greatest efficiency to our taxpayers.

INTRODUCTION

In the closing hours of the first remote sensing symposium to be held in Canada, it is apparent that much time and effort has been expended in the preparation of technical papers describing the capability and potential of this rapidly developing technology. Far less, however, has been said on how this sophisticated technology might be interpreted and applied to the multitude of problems facing our nations resource managers and administrators; and still less has been said on the administrative organization which might translate the imagery, data and jargon of the specialist into the language and facts of the decision maker.

It is my personal belief that the gap separating remote sensing technology and the capability to interpret it for efficient action is progressively widening and that some very positive measures are required to bridge it. It should be emphasized at this point that the views put forth in this paper are the sole responsibility of the author and do not necessarily reflect those of his employer, the Government of Alberta, although they may reflect the partial opinions of numerous co-employees with whom the author has discussed this situation over the past several years, especially Mr. B. Patterson and Dr. G. L. Nielsen of the Alberta Department of the Environment.

The proposal is intended to describe at a conceptual level an organizational relationship which could correct the current administrative hiatus within Alberta, and possibly in other provinces. For the benefit of those not familiar with the Alberta Government structure, organizational details will be omitted, although occasionally specific clarification will be provided.

ESTABLISHMENT OF A REMOTE SENSING INTERPRETATION INSTITUTE

It is proposed that a Remote Sensing Interpretation Institute should be created whose primary purpose would be to service remote sensing data users in any given region. This regional Institute could be synonomous with a province or with a geographic region of Canada although the writer prefers the former in order that interprovincial matters would not complicate its establishment.

PRINCIPAL ELEMENTS OF THE INSTITUTE

There are several critical elements which would be required in the structuring of the proposed Institute. Their neglect in the establishment would in all likelihood create numerous weaknesses or even failure; on the other hand however their successful inclusion will be extremely difficult and could prevent the Institute from becoming a reality.

Governmental and Non-Governmental Participation

In addition to governments, the organization must include universities as well as the aero-survey industry. Not only must their presence be physical, but it must be fiscal as well.

Intergovernmental Participation

Municipal governments as represented by metropolitan areas and regional planning commissions must be represented as well as the federal and provincial governments. As both rural and urban planning accelerates, the need to involve that level of government with the greatest degree of responsibility for immediate short term planning becomes more evident.

Inter-Disciplinary Participation

The imagery and data results of remote sensing does not recognize the artificial and sometimes arbitrary barriers created by universities and perpetuated by professional associations. Although the resources and the profession traditionally associated with their management are discrete and defineable the environment is not. It is holistic and its complex interrelationships can best be identified and interpreted by teams of multi-disciplinary professionals.

Inter-Departmental Participation

Since many of our government departments can be identified by their professional staff and since each department has a specific function to perform, it is imperative that each department has the opportunity to participate in advising and also staffing the Institute. Not only will these contributions result in better developed solutions but also in a better appreciation of the complex interrelationships of environmental systems.

Core Interpreters

The core staff, in addition to the necessary clerical and administrative staff, should consist of qualified professionals seconded from a variety of agencies and disciplines with a common interest and capability of remote sensing interpretation. For example, they could include a regional planner specializing in transportation; a forester specializing in forest protection research; a hydrologist specializing in moisture regime; an agrologist specializing in crop production and a geographer specializing in land use planning. Each of these would interpret the same data uniquely on the basis of his professional training and related work experiences, however, when the five interpretations were collated, additional conclusions would be reached based on the inter action of the five uni-disciplinary views resulting in a far more comprehensive interpretation than any single one discipline could have made.

It is proposed that this core staff be seconded from their regular employment for one or two years at which time they might withdraw and return to their former employment. In this system, they would remain in touch with the practical application of their interpretations and not become isolated from the administrative and resource management responsibilities throughcomfortable ensconcement within research laboratories. The seconding could take place over irregularly staggered periods, consequently new disciplines and fresh techniques would be the order of the day. A beneficial side effect would result from the seconded interpreter returning to his former employment, practising some of his newly acquired interdisciplinary interpretative skills as well as passing them on to his less fortunate colleagues.

Administration of the Institute

The Institute should be under the direct administration of the provincial government and managed by a full time permanent administrator. He would be guided by a board of directors comprised of representatives of the various users including the federal government, the provincial government, municipal government, universities and the aero-survey industry.

Within the provincial government a permanent

interdepartmental advisory committee should be structured in order that adequate consideration would be given to the various provincial needs and transmitted to the administrator through the provincial director on the board. A similar federal committee would also be recommended.

The physical plant of the Institute could be attached to the provincial natural resource or environment research component. There are several examples of this relationship in Alberta already which appear to be reasonably functional.

FUNCTIONAL RESPONSIBILITY OF THE INSTITUTE

The Institute would serve as the provincial counterpart of the Canadian Center of Remote Sensing and therefore have similar responsibilities of a provincial nature. Principal functions would include the following:

- Service and Co-ordinate all Applications for Remote Sensing Data in the Province

  The Institute would co-ordinate, arrange and schedule requests for data, whether obtained directly from the Canadian Center of Remote Sensing or required to be flown on request. Under this system the various provincial users would be kept informed of what data was available within Alberta, what data would become available and how they might obtain additional data on areas to be flown.

  During the past several years, we have experienced some degree of confusion in this general area within Alberta. This was partly as a result of unco-ordinated governme nt action at both levels and partly as a result of over zealous aero-survey operators.

  This function would also apply to the necessary arrangements for adequate ground truthing as well as the actual planning required prior to specifying the sensing criteria.

- Undertake Primary Interpretation of all Alberta Data

  All Alberta data would be interpreted to a primary level by the Institute and the resultant information made available to the co-operating users.

  More specific or secondary interpretation could be made upon request, although a user fee charge would be levied on the co-operator. The distinction between primary interpretation and secondary interpretation is subject to further study and clarification. It could be arbitrarily decided that primary interpretation included the collation and elaboration of the multidisciplinary team.

- Maintenance of Equipment Inventory for Use by Co-operating Agencies

  In many instances the secondary interpretation could possibly be done more effectively directly by a co-operator providing that he had access to some of the equipment. For this reason, it would appear more efficient that the various agencies pooled their equipment inventory with the Institute resulting in one extremely well equipped center available to all rather than each agency independently maintaining an under-equipped semblance of an interpretation unit.

- Maintenance of Data Library and Training Information

  This is a natural adjunct growing out of the above three functions which must be provided to maintain continuity of research and development.

CONTRIBUTING RESPONSIBILITIES OF THE CO-OPERATING AGENCIES

In addition to a great deal of effort and goodwill, very specific contributions would be required from the various co-operating agencies.

- Provincial Government

  The provincial governm ent should supply the space and the administrative and clerical staff as well as the secondment of qualified personnel in those areas of expertise common to the provincial jurisdiction, plus a portion of the annual operational budget.

- Federal Government

  The federal government should provide re-

search grants, a portion of the annual operational budget in the form of a grant plus secondment of qualified personnel in those areas of expertise common to federal jurisdiction.

- Municipal Government

  The municipal government should provide the secondment of qualified personnel in those areas of expertise common to its jurisdiction or provide a salary grant in lieu thereof in order that such staff could be hired directly by the provincial government.

- Universities

  The universities and possibly technical institutions should provide interpretative assistants in the form of graduate or technical students and also research expertise where available. They should also provide teaching capability where possible.

- Aero-Survey Industry

  The industry should provide the major equipment required to equip the Institute. In the past the industry has been actively encouraging the use of sensing and photography but appears to have taken the view that its responsibility has ended with the processing of the data. A far greater market for its capability would develop if the industry also accepted a responsible role in the interpretation of the data to providing the necessary equipment. The industry has an opportunity to improve its image through co-operation rather than divisive competition.

NON CO-OPERATING USERS

There would be numerous users in the private sector who would not be co-operators. These could have interpretation of their data done on a cost plus basis with the primary interpretation information becoming available to all of the co-operating users.

CONCLUSION

Our nations resource managers can gain inestimable benefits from this technology at a saving to the taxpayers and consumers. We must, however, devise a more efficient organizational system of interpreting the data and disseminating the information to prospective users. The proposed Institute can accomplish this.

THE CANADIAN FORCES AIRBORNE SENSORS FAMILIARIZATION PROGRAM.

Lieutenant Colonel M. Sugimoto
Directorate of Aerospace Combat Systems
Canadian Forces Headquarters
Ottawa, Ontario.

INTRODUCTION

Gentlemen, as already mentioned, that (slide 1) is the subject of my paper. Although that appellation implies coverage of a wide range of airborne sensors, the program is in fact limited to a consideration of remote sensors for reconnaissance and surveillance.

The outline of my paper is as shown on the slide (slide 2). Time permitting, I would also like to show you some slides of infrared linescan imagery obtained with our equipment.

BACKGROUND AND AIM

I should like to start then with a few words on the background to our program. It is evident that remote airborne sensors have a part to play in almost every conceivable role to which the Canadian Forces could be assigned. Yet with respect to some of the most promising types of airborne sensors available today for surveillance and reconnaissance, we lack both the operational and the technical expertise to select and to employ operationally those airborne sensors which will enable us to accomplish our assigned roles and missions.

We don't know enough about them to define specific requirements; even if we could do that much, at this point in time we would not be able to specify the most effective equipment to meet those requirements; and finally, we lack the operational expertise to put those sensors to their optimum use. Compounding all of this is the fact that with respect to our role of arctic surveillance, the means of carrying out that role are still subject to review. We do believe that the arctic is an area where sideways looking airborne radar and infrared linescan, ie, SLAR and IRLS systems, have great potential yet our lack of knowledge is greatest on these two types of sensors, especially in the northern environment.

This serious lack of knowledge is recognized at the highest levels within the Canadian Forces and the Defence Research Board. Therefore, the Canadian Forces, in cooperation with the DRB, have embarked on the program which is the subject of this paper. The broad aim of the program is to acquire the knowledge and skills necessary to specify requirements for, to acquire and to operate airborne surveillance and reconnaissance sensors which will enable us to accomplish our assigned roles and missions.

APPROACH

That, briefly, is the background and aim of our program. I shall now outline the approach which we have taken. In general, it has been our intent to procure limited quantities, by that I mean one or two of each, of appropriate state-of-the-art equipments and to conduct a flight program aimed at gaining first-hand experience on the sensors now available for our surveillance and reconnaissance tasks, and determining their full capabilities and limitations.

I would like to emphasize what I have just said because I think therein lies an important difference between our current activities and the activities of many of you in the audience today. Most of you have rather specific tasks you wish to perform with remote airborne sensors. In many cases it matters not if you do the job yourself or if you have others do it for you. On the other hand, we in the Department of National Defence must be able to do the job ourselves. We must be able (a) to formulate requirements, (b) to procure the most effective equipment to meet those requirements, and (c) to put those sensors to their optimum use over a wide range of surveillance and reconnaissance tasks.

To this end, we have procured the equipment shown on this slide (slide 3) two sets of IRLS and one set of SLAR equipment. Since SLAR and IRLS are the most complex types among the several most promising types of sensors for our purposes, we have concentrated our program initially on these two types.

METHOD

That being our approach, what then is our method? ie, how do we intend to carry out our program. For guidance of staffs involved in planning or conducting the program, we have produced a detailed statement of our objectives for each of the two types of sensors currently programmed. The general program objectives are as shown on this slide (slide 4). These have been amplified in considerable detail to ensure that no important considerations are overlooked, and conversely, that no unnecessary investigations are undertaken. This slide (slide 5) shows operational objectives for example. Each of these have been further amplified, and as an example we can look at sea going targets as shown on the next slide (slide 6). The early stages of the program have been well defined, but it is impossible at this time to state categorically when, or after what studies, the program should be terminated. Nevertheless, the program for each sensor has been divided into six phases as shown on this slide (slide 7). These are not mutually exclusive in order or time. They have been established to provide a logical order for the familiarization of each sensor, and they help to identify and to assign responsibilities and tasks to specific organizations. Keeping these phases in mind and using our detailed objectives as a guide, we have assigned responsibilities and tasks to the organizations shown on this slide (slide 8). The coordination and management of the program remains in the Canadian Forces Headquarters here in Ottawa.

CURRENT STATUS AND FUTURE OUTLOOK

Now I should like to summarize our current status and future outlook. I would like to deal with the IRLS system first. I mentioned earlier that we had procured two sets of IRLS equipment, one set has been installed in a CF100 aircraft, and a program involving some 200 hours of flying is currently underway, basically for the technical evaluation. The second set has been installed in an Argus long range patrol aircraft, and a program involving some 400 hours of flying is currently underway, basically for the operational evaluation. Since the Argus aircraft provides us with a platform to perform inflight adjustments and experiments, some of the technical and scientific evaluations are being done using the Argus. Now I shall turn to the SLAR system. The one set which we have procured has been installed in an Argus aircraft. This particular aircraft is the one that has been on static display at Upland throughout the day. The installation and checkout flights have just been completed and we are now ready to commence a program of some 400 hours of flying.

The SLAR and IRLS program will extend over at least the next eighteen months. By that time we hope we have acquired sufficient knowledge and experience on these two sensors to answer many questions. For example: what is the effectiveness of these sensors: how should they be employed; under what conditions and against which targets; what coverage can be provided and how often. These are but a few of the many questions which we hope to be able to answer. In addition we hope to have a better understanding of the human factors involved in the effective use of this equipment by aircraft crews as well as ground interpreters. We hope to have a better feel for the quantities, types of equipment, and the level of support required.

As the program evolves, areas where cooperation would be mutually beneficial will no doubt become apparent. I speak of cooperation with other departments and agencies with whom some areas of cooperation have already been or are being exploited. Other areas will be exploited as they become apparent.

Our program currently deals with only SLAR and IRLS, but we are considering other sensors, for example, low light level TV and forward looking infrared. However, nothing definitive has been planned to date with the exception of some work, which the Communications Research Centre has undertaken on our behalf, to explore the feasibility of adapting a state-of-the-art coherent airborne search radar to function as a SLAR.

Now before I conclude my paper, I believe I have enough time to show you the slides of IRLS imagery. All of the slides which you are about to see were obtained with our Reconofax XIIIA mounted in the Argus aircraft. The first slide (slide 9) shows the ferry boat between New Brunswick and PEI - the JOHN HAMILTON GREY - taken at night from an

altitude of 500 feet. The wake is noticeable for close to 5000 ft aft of the ferry. I would ask you to note the length of the vessel and compare it with the vessel in the next slide (slide 10). This is the same ferry seen from the same altitude. In the previous slide, the V/H ratio was incorrect as it is in this one but in the opposite direction. This next slide (slide 11) is a shot of the naval yards in Halifax Harbour taken at dusk from an altitude of 1000 feet. The next slide (slide 12) shows Halifax Harbour again from an altitude of 2000 feet at dusk. The thermal effluent from the Nova Scotia Light and Power Company plant is very noticeable. The next slide (slide 13) shows the airfield at Chatham, New Brunswick taken at dusk from 500 feet. The aircraft in various stages of run up or shut down can be seen. The roll stabilization was inoperative during this shot and the effects are visible especially on the edges of the hangars. The next slide (slide 14) shows a freighter taken at dusk from 500 feet. This shot was taken using manual gain control in lieu of automatic gain control. Consequently the cargo is recognizable on the deck. The next slide (slide 15) is a standard daytime photo of that same freighter. The next slide (slide 16) is a shot of an offshore drill rig taken during daylight from 1000 feet. The next slide (slide 17) is a shot of the oil rig on Sable Island taken during daylight from 1000 feet. Buildings and equipment are noticeable but not recognizable. This next and last slide (slide 18) is a shot of the tip of Sable Island taken during daylight from 1000 feet showing, would you believe, a herd of seals.

Gentlemen, that concludes my presentation. I hope I have given you a clear understanding of the background and aim of our program, and given you an appreciation of the approach and method we have used, as well as a feel for the current status and future outlook of our program. I thank you.

TITLE

THE CANADIAN FORCES AIRBORNE SENSORS FAMILIARIZATION PROGRAM

SLIDE NO I

OUTLINE

1. Background and Aim
2. Approach and Method
3. Current Status and Future Outlook

SLIDE NO 2

EQUIPMENT PROCURED

IRLS

Reconofax XIIIA - H.R.B. Singer

SLAR

APS 94D - Motorola

SLIDE NO 3

GENERAL PROGRAM OBJECTIVES

- Identify and Procure Sensors
- Conduct Comprehensive Flight Trials
- Establish System Capabilities and Limitations
- Determine Technical and Operational Factors
- Define Operational, Training and Maintenance Procedures.

SLIDE NO 4

OPERATIONAL OBJECTIVES

To detect, classify and identify Surface Targets such as:

1. Seagoing Targets
2. Deployed land Forces
3. Settlements and Installations
4. Natural Surface Features
5. Pollution
6. Lake, River and Coastal Targets.

SLIDE NO 5

SEA GOING TARGETS

Determine effects of weather, sea state, water temperature, etc. on typical IRLS imagery of -

a. Surface Vessels
b. Submarine, Periscopes or Snorkels
c. Surface Lakes
d. Hazards to Shipping
e. Life Saving and SAR Signalling Equipment

SLIDE NO 6

PROGRAM PHASES

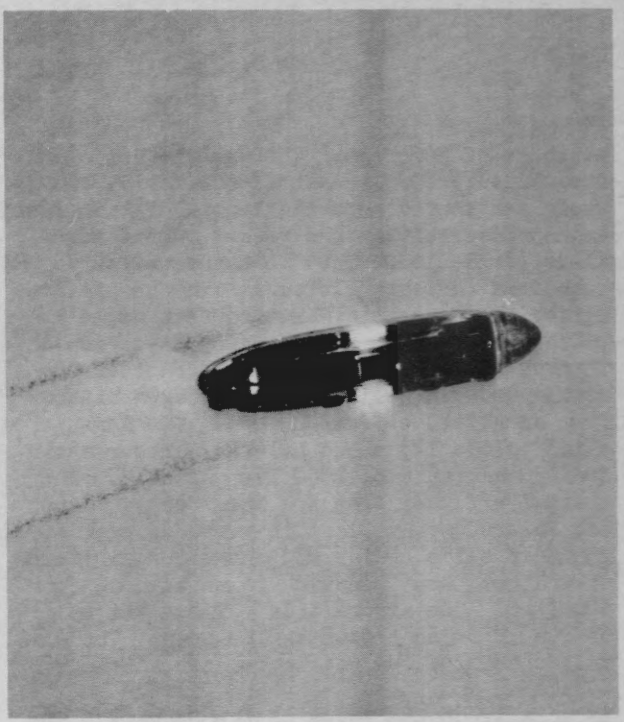
I. Procurement and Installation
II. Operational Demonstration
III. Technical Evaluation
IV. Operational Evaluation
V. Scientific Evaluation
VI. Continuing Applied Research Efforts

SLIDE NO 7

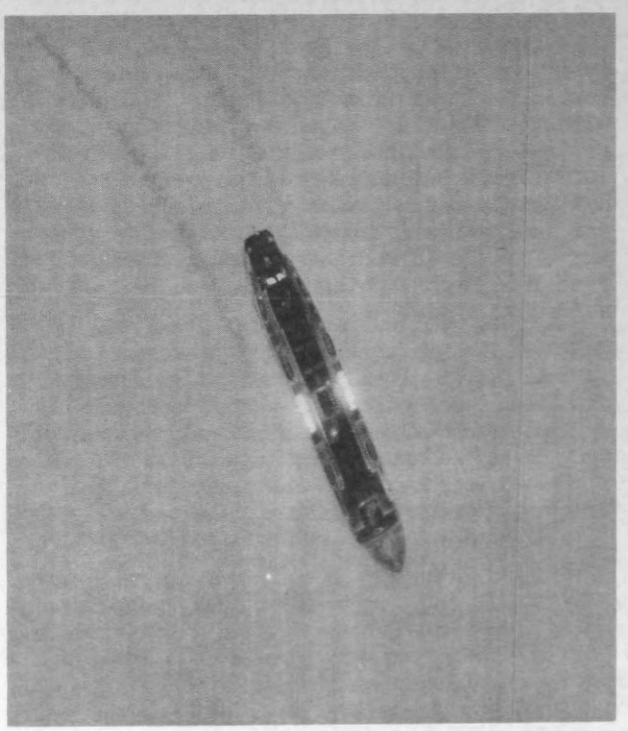
ORGANIZATIONS INVOLVED

Technical Aspects	-	Aerospace Engineering Test Establishment- Cold Lake
Operational Aspects	-	Maritime Proving and Evaluation Unit - Summerside
Interpretation	-	Defence Photo Interpretation Centre - Rockcliffe
Scientific	-	DRB/CRC

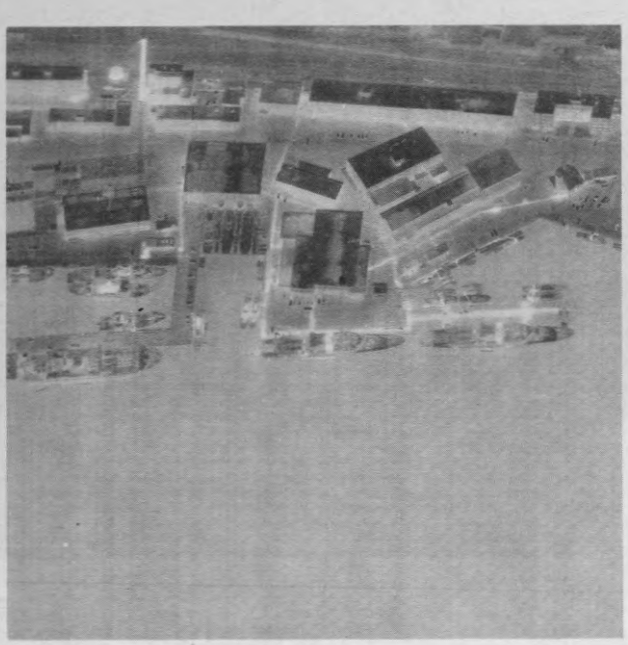
SLIDE NO 8



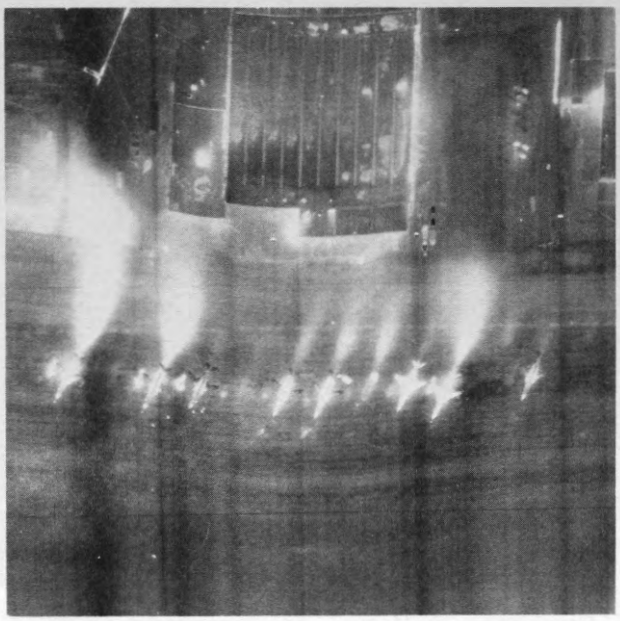
SLIDE NO 9



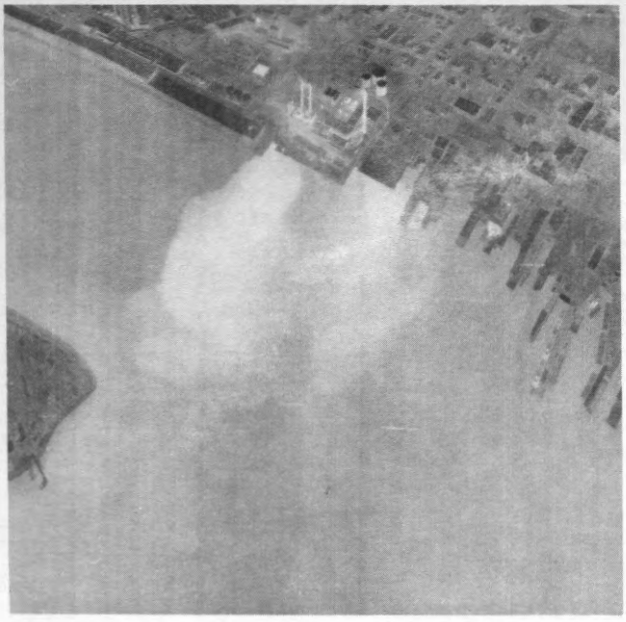
SLIDE NO 10



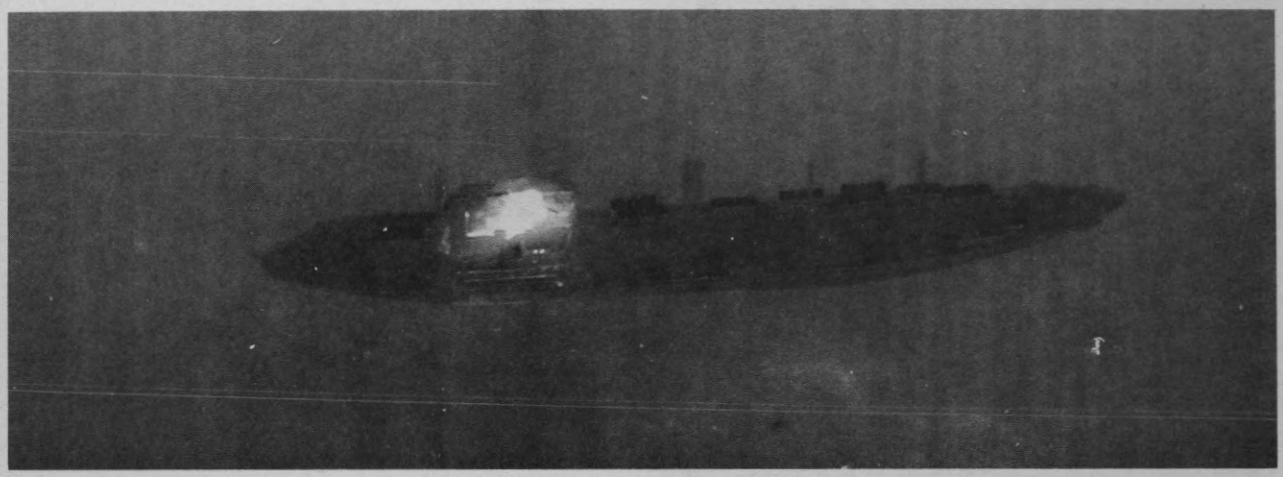
SLIDE NO 11



SLIDE NO 13

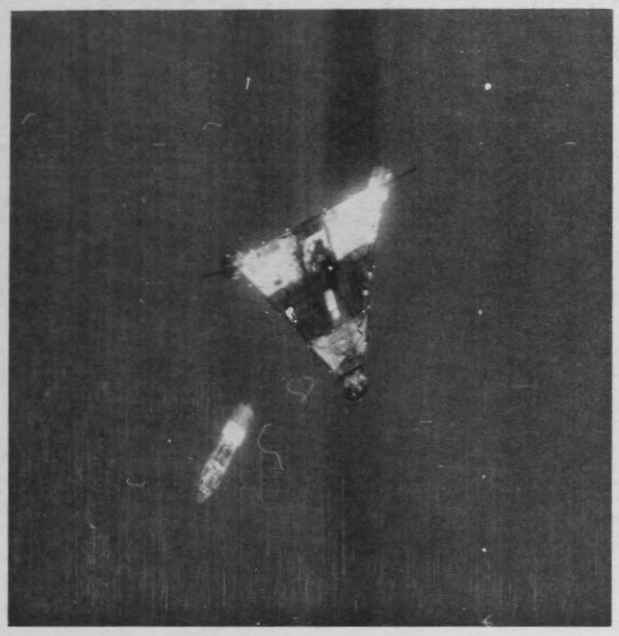


SLIDE NO 12

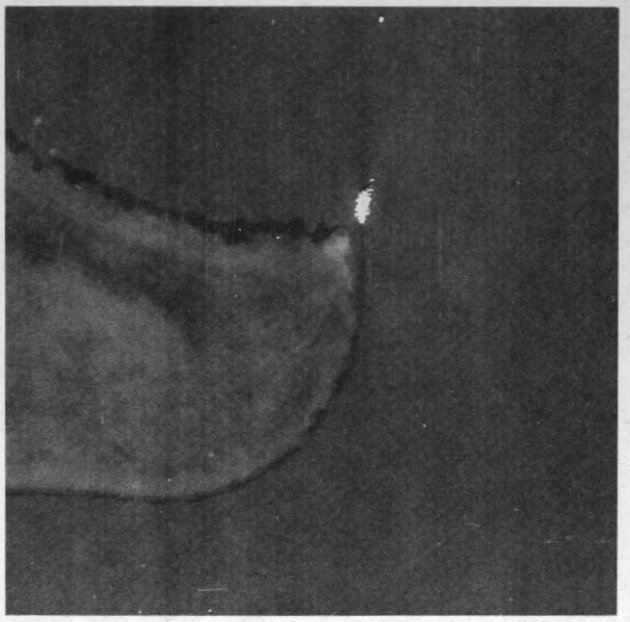


SLIDE NO 14

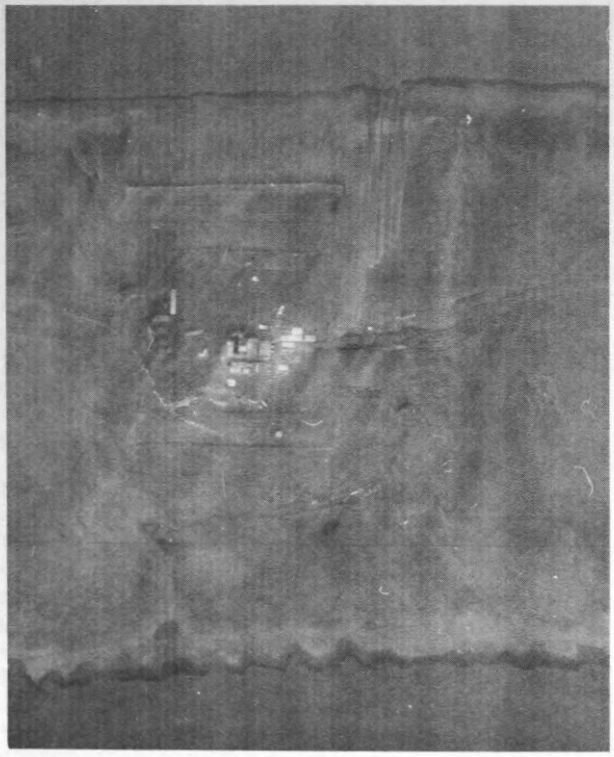
SLIDE NO 15



SLIDE NO 16



SLIDE NO 18



SLIDE NO 17

THE APPLICATION OF MULTISPECTRAL REMOTE SENSING TO THE STUDY OF SOIL PROPERTIES AFFECTING EROSION

G.F. Mills,
Manitoba Department of Agriculture,
Pedology Section,
Winnipeg, Manitoba.

ABSTRACT

Medium and high altitude multispectral photography and low altitude thermal infrared scanning was obtained for the Wellwood area of Manitoba during the summer of 1971. An evaluation of the various kinds of imagery was carried out to determine if these techniques could be applied successfully to identify the soil properties which affect erodibility and to determine the extent and severity of soil erosion in the study area.

Examples of the various kinds of imagery, the signatures obtained and an evaluation of the imagery in light of the ground truth collected in 1971 is given. These preliminary analyses indicate that no single film-filter combination or image type is best for identifying or mapping all of the soil properties which are significant to soil erosion. None of the image types were dependable for differentiation of soil texture. Variation in soil organic matter content could be identified by colour infrared, red band and thermal infrared imagery. Soil moisture properties could be differentiated on the near infrared black and white and colour imagery as well as with the thermal infrared data. An assessment of topographic pattern and slopes could be accomplished on all image types, but most easily with panchromatic black and white red band and colour infrared imagery. Study of soil moisture and organic matter contents could be made semi-quantitative with sufficient detailed ground truth collection.

INTRODUCTION

The field study of erosion and delineation of its extent and severity presents many difficulties. In general, erosion differences and their significance can be evaluated only in relation to each individual soil type. In the past, estimation of erosion losses has been accomplished by setting standards on each soil for reference. Arduous and time consuming ground-truth collection in the form of on-site soil inspection has been necessary in order to compare representative sites to the standards. The delineation of the areal extent of erosion usually has been dependent on the use of aerial photography. Again, certain limitations are inherent in the use of this technique. The effects of erosion may have been masked by crop or vegetative cover at the time of flight. Recent aerial photographs are not always available for the area in which erosion is to be assessed. Correlating the grey tones of black and white photography with the various colours found in soils is difficult. Variation in film exposure and development can also cause interpretive problems.

We have had the opportunity recently to evaluate the new techniques of multispectral remote sensing in Manitoba. There is very little research available pertaining to remote sensing and erosion studies. However, remote sensing has been shown useful for evaluation of certain soil properties, some of which are directly related to soil erosion. For example, wavelengths in the thermal infrared range are specific to soil temperature characteristics, which in turn relate to other soil properties such as moisture and organic matter content.

These and other remotely sensed data can be evaluated either from the point of view of assessing how much new information can be gained from a particular wavelength, or what new information can be gained by interpretation of two or more image types in unison. In evaluations of imagery of various wavelengths, emphasis should be placed on what new information can be gained from that wavelength; not necessarily which imagery provides the most information.

The purpose of this paper is to determine how the recently developed techniques of remote sensing can be used as an aid in the study of soil erosion. Pursuant to this objective, one must determine the relationship between the response on the various kinds of imagery and the soil properties which affect erosion.

SOIL EROSION AND ERODIBILITY

The effects of soil erosion are many and varied. There are losses to the overall productive capacity of the agricultural industry dependent on the soil. The individual farmer-producer can suffer direct loss of income either in the form of higher production costs (chemical and fertilizer inputs) or in reduced yields. To this must be added the cost of damage to the environment resulting from air and water pollution. These agricultural and environmental effects from soil erosion further affect the overall economy through costs involved in preventative measures related to conservation and the maintenance of engineering works. The cost of protective measures to control soil erosion during the past forty years in the United States has been more than eleven billion dollars (Jacobson, 1969). In spite of these tremendous efforts towards ameliorating soil erosion affects, there appears to be no ready solution to the problem.

Erosion has been defined as the wearing away of the earth's surface by the forces of water and wind (Soil Survey Manual, 1951). This is so-called natural erosion occurring in the environment undisturbed by man. Erosion, as used in soil science, refers to "accelerated erosion", or erosion of soil resulting from exposure of the soil surface to runoff or wind action. Burning, overgrazing, forest cutting or tillage are activities which destroy or weaken the vegetation protecting the soil and so contribute to accelerated erosion.

Soil erosion processes can be divided into two classes, water erosion and wind erosion, according to the moving agent. Both kinds of erosion have taken, and are taking place, on Manitoba soils. In subhumid climatic areas such as Manitoba, soil erosion by wind has occurred to some degree on virtually all of the cultivated agricultural soils. The effects of water erosion become more important on the cultivated soils above a certain percent slope class.

The inherent erodibility of a soil is a complex property dependent both on its infiltration capacity and on its capacity to resist detachment and subsequent transport by rainfall, runoff and wind action (Wischmeier, W.H. and J.V. Mannering, 1969; Anderson, D.T. et al., 1966). Soil properties that contribute significantly to variability in soil loss include particle size distribution, organic matter content, pH, structure and bulk density of the plow layer and subsoil, pore space filled by air, degree of aggregation, type of parent material and various interactions of these variables. It has been shown that long time average soil losses may vary more than thirty-fold just due to some of these basic soil differences (Olson and Wischmeier, 1963). Wischmeier and Mannering (1969) have shown that there was a complex interrelationship among these soil properties which was used to obtain an empirical expression of the erodibility of a soil. More recently, Wischmeier et al (1971) have presented a nomographic solution for evaluation of the erodibility of a soil by water. This solution is a function of only five soil properties, namely percent silt, percent sand, organic matter content, structure and permeability.

MATERIALS AND METHODS

Description of Study Area

A portion of the upper Assiniboine Delta in southern Manitoba was chosen for this study (Wellwood Study Area). Detailed ground-truth collection to ascertain the relationship between soil properties at the time of flight and the image obtained was carried out at eleven sites on an east-west transect through the Delta about one and one-quarter miles south of Wellwood. An additional six ground truth sites to evaluate severity and extent of erosion were selected throughout the study area. The location of the sites for ground-truth data collection with respect to the various kinds of imagery and to the soils of the area are shown in Figure 1.

The detailed ground-truth information involved the assessment of the following factors:

- soil type and texture
- soil profile description
- soil drainage
- soil moisture content (surface and 20 cm)
- soil temperature (surface and 20 cm)
- soil colour
- surface condition (percent trash cover, direction of cultivation, degree of roughness, cloddiness and aggregation)
- slope class and aspect
- vegetative cover

Ground-truth data collection related to the degree of erosion involved assessment of the following factors:

- a soil profile description to assess the amount of soil removed. Three classes were defined:

- E1 - slight to no erosion
- E2 - moderate erosion - 25 to 50 percent of productive topsoil lost
- E3 - severe erosion - more than 50 percent of productive topsoil lost

- measurement of slope class and length of slope

The soils of the area are in the Stockton, Wellwood and Firdale Associations (Ehrlich et al., 1957). Soil textures vary from moderately coarse (Stockton loamy sand) to medium (Stockton fine sandy loam) to medium and moderately fine (Wellwood loam and Firdale clay loam). Soil drainage varies from well through imperfect to poor. In general, topography of the Stockton fine sandy loam and Wellwood loam is level; the Stockton loamy sands are undulating to level and Firdale loams and clay loams are level to rolling. Erosion varies from none to slight in the Wellwood loams and Stockton fine sandy loams and from moderate to severe in some areas of Firdale loams and clay loams and Stockton loamy sands.

Description of Imagery

The remote sensing data used in this study were of two kinds: (1) thermal infrared scanning obtained from low altitude flights and (2) photographic imagery obtained from medium altitude and high altitude CF-100 flights. The specifications of the imagery obtained from the various flights are presented in Table 1. In addition, panchromatic black and white photographs obtained in July, 1948, October, 1958, and September, 1964 at an altitude of 8,000 feet a.s.l. were used in this study.

RESULTS AND DISCUSSION

The results obtained from the remote sensing studies carried out in the Wellwood area of Manitoba must be evaluated in light of limited ground-truth collection aimed specifically at soil erosion. It must also be realized that the objective was not to determine which kind of imagery was better for studying soil erosion than another, but rather to make a multispectral interpretation with respect to a single or even several soil properties.

The properties of the soils as determined at the detailed ground-truth sites in the Wellwood study area are summarized in Table 2. Prints of the 3-5 μ wavelength thermal infrared imagery of selected ground-truth sites in the study area are shown in Figures 2 and 3. Ground-truth data with respect to degree or severity of erosion are presented in Table 3. Multispectral photography is presented from soil areas affected by wind and water erosion (Figure 4, Firdale Association) and from soil areas in which wind has been the chief cause of soil erosion (Figure 5, Stockton and Wellwood Associations).

The Identification of Soil Properties Affecting Erodibility

Examination of the photographic and thermal infrared imagery showed both direct and indirect relationships between the signatures obtained and soil properties affecting erodibility.

Soil Texture - Differences in texture of an order of magnitude from moderately fine to moderately coarse-textured did not provide consistent characteristic signatures on any of the imagery. Inferences can be made about the texture of a soil as it may be reflected in the evidence of past erosion. The landform pattern on Figures 2c and 5c indicate that the area of coarser-textured Stockton soils has undergone more erosion than adjacent areas of finer-textured soil.

Organic Matter Content - This soil property is generally reflected in the surface soil colour. The ability of multispectral photographic imagery to differentiate soil colour depends on moisture content, vegetative cover and various other conditions of the surface at the time of flight. Aerocolour (Figure 5c), colour infrared (Figure 4e) and red band black and white (Figures 4c and 5c) imagery show the surface soil colour quite well under both crop and summerfallow conditions. Panchromatic black and white imagery varied in its reliability for identifying lighter coloured soil areas (low organic matter content) and dark coloured soil areas (higher organic matter content), (Figures 4a, 5b and 5e). The field indicated by arrows (centre of Figure 4a) shows the masking effect of vegetation and crop cover on soil colour. The 1948 imagery (Figure 5b) of the Stockton soil shows the lighter coloured surface much better than does the 1964 imagery (Figure 5e) in which the soil colour has been masked out entirely by crop cover (marked by arrows). Black and white imagery in the green and near infrared bands did not add any new information to the assessment of this soil property.

Inferences of organic matter content from thermal infrared imagery depends on soil temperature as affected by reflectance and absorbance properties of the soil. The

Firdale soil at site 11 had a greyish brown surface colour and a soil temperature of 26.0° C (Figure 3b). The strong signature noted at site 11 is partly due to reflected energy. The Wellwood soils at sites 1 and 6 with similar level topography, had very dark grey to black surface colours and soil temperature values of 31.0° C and 27.0° C, respectively on the PM imagery (Figure 2a).

Soil Moisture - This soil property is reflected both by darker tones of moist soils compared to that of dry soils, and by the response of vegetation to varying moisture conditions. Colour photographs and black and white near infrared imagery show good tonal differences for soil areas with higher moisture contents. On the colour imagery (Figure 5f) the higher moisture content in the drainage channel (broken arrow) than in the surrounding area is indicated by a more luxuriant dark green vegetative growth. The more vigorous vegetative growth at moister sites also shows a stronger signature on the near infrared black and white imagery (Figure 4d and 5d, broken arrow). Similarly, luxuriant vegetative growth provides a stronger signature on the colour infrared imagery (Figure 4e) in the form of bright pink colours. Panchromatic black and white photography was less useful for depicting variation in soil moisture as the signature obtained is highly dependent on conditions at time of flight.

Soil moisture is one of the most important factors influencing the thermal properties of a soil. The kinds of soil properties which may be differentiated using thermal infrared imagery are those which affect soil temperature and moisture relationships. In evaluating the extent to which thermal infrared sensing can be used under Manitoba conditions, it was necessary to ascertain the magnitude of possible temperature and moisture differentiation. Three distinct signatures are evident at sites 1, 6 and 7 in Figure 2a. The three sites occur on well drained Wellwood soils with level topography. Although there was some variation in surface condition among the three sites, it is likely that the differences in signature are largely due to variations in the surface moisture contents. The values obtained for surface moisture content and temperature at sites 1, 6 and 7 are 2.8, 3.5 and 9.3 percent moisture and 31.0° C, 27.0° C and 22.0°C, respectively. A Stockton very fine sandy loam (site 3) and a Wellwood loam (site 4) occurred both within one field, and therefore had the same management history. Ground-truth data from these two sites at the time of the PM flight showed only slight differences in moisture content and temperature values (1.0 percent moisture, 26.5° C and 2.0 percent moisture, 25.0° C, respectively). Such slight variation did not provide visual differences in the imagery (Figure 2b). This does not preclude that with the use of more sophisticated electronic analysis that these slight differences could not be detected and even measured quantitatively. By establishing the degree of response to various temperature-moisture relationships such as these, it should be possible to infer which soils have a higher moisture content and therefore, not as highly erodible by wind.

Slope and Topographic Pattern - These properties are external to the soil but nevertheless are properties which affect the erodibility of a soil and which can be determined by remote sensing. The use of panchromatic black and white photography with stereoscopic coverage has long enabled a ready determination of overall topographic pattern. The information so obtained on amount of relief and length of slopes is valuable, as both of these factors have a direct bearing on the erodibility of a given parent material (Figure 4a). Black and white imagery in the red band and near infrared colour imagery (Figures 4c and 4e, respectively) show the overall topographic pattern and provide approximately as much topographic information as the panchromatic black and white photography. The stereoscopic image on the red band and near infrared colour imagery, although not as sharp as that found on good quality panchromatic photographs, did provide much more clearly defined stereo images with sharper boundaries than those found on the green band or on the near infrared black and white imagery.

The effect of slope and topographic pattern on the thermal infrared imagery is shown in an area of Firdale loam with moderately rolling topography (sites 9 and 10, Figure 3a). In the AM imagery the effect of temperature differences and low sun-angle shading combine to produce a marked three dimension picture of the relief. Ground-truth was collected at site 9 (south aspect) and site 10 (north aspect). Imagery of the north aspect had darker tones than the south aspect which showed light tones. This variation in imagery is related in part to the 3° C difference in surface soil temperature between the south aspect (site 9, soil temperature 10° C) and the north aspect (site 10, soil temperature 7° C). In the PM imagery the relief is washed out (due to levelling out of temperatures and higher

sun angle) so that the entire field shows as a uniform high energy source. Soil temperatures for sites 9 and 10 at the time of the PM flight increased to 29° C and 27.5° C, respectively.

If the thermal infrared scanning is carried out shortly after sunrise when temperature differences are at their maximum, the three dimensional effect noted in Figures 3a and 3b is produced. This technique could be used to delineate topographic pattern and length of slopes. It will be noted that this expression of relief is very similar to imagery produced by SLAR.

The Identification of the Degree and Extent of Soil Erosion

The reliability of remote sensing techniques for the recognition of the severity of soil erosion and delineation of its areal extent depends on how successful these techniques are in identifying some of the soil properties discussed in the foregoing section. Examination of photographic and thermal infrared imagery of both eroded and non-eroded soils in areas of level topography (eroded by wind) and in areas of rolling topography (both wind and water erosion) revealed the following results.

1. The reliability for identification of the degree of erosion depended more on vegetative cover and soil conditions at the time of flight than on the kind of imagery (Figure 4). In the panchromatic black and white imagery (Figure 4a), the eroded areas (arrows) are defined very well in the northern portions under summerfallow condition, but the pattern is not so evident under the crop cover (southern portion). It will be noted on the imagery from the red band black and white (Figure 4c) and near infrared colour (Figure 4e) that it is possible to delineate the more severely eroded areas under both crop cover (north field) and summerfallow (south field).

   An example of wind erosion in the Stockton soil area (arrow on Figure 5b) is clearly shown on the 1948 panchromatic black and white imagery but is almost completely masked on the 1964 panchromatic black and white photos (Figure 5e). The same eroded pattern however, could be defined fairly well on the red band black and white imagery (Figure 5c) and the colour photography (Figure 5f) under all kinds and densities of crop cover present at the time of sensing.

   Comparison of the 1948 panchromatic imagery to the 1971 multispectral imagery indicated no significant change in extent of eroded pattern in this area of Stockton soils.

   Thermal infrared imagery indicated the areal extent and degree of erosion only insofar as the thermal properties of the soil had been affected. Lower organic matter content (lighter soil colour and cooler soil temperature), and ablation of the finer soil materials (lower moisture holding capacity and warmer soil temperature) are two such changes in the soil brought about by erosion (Figure 2c). As none of the ground-truth sites suitable for evaluation of the degree and extent of erosion were covered by data from the infrared scanning it was not possible to ascertain whether such differences in tone of the imagery were indeed due to changes in the aforementioned soil properties caused by erosion, or whether the differing images are only reflective of the soil conditions occurring on non-eroded sites of coarser-textured soils.

2. A complete range in degree of erosion is found in the study area. Ground-truth studies carried out in 1971 indicate limitations in the use of multispectral photographic imagery for assessing accurately the severity of erosion. For example in the Firdale soil area(Figure 4a), ground-truth data (site 12a, Table 3) indicate a non-eroded soil and corresponding dark tones on the imagery. This is in contrast to the light tones found at site 12b where all of the A and B horizon has been removed. Similarly, in the Stockton soil area (Figure 5b), ground-truth data (site 15, Table 3) indicated slight to no erosion and corresponding dark grey image tones. Adjacent soil areas depicting light grey to white image tones are characterized by severe erosion. With access only to limited amount of ground-truth data in 1971, it was not possible to establish a correlation with the varying degrees of erosion between these two extremes.

   In general, the extent of severe erosion could be delineated as well on the colour, near infrared colour, and black and white red bands as on the panchromatic black and white imagery. All of these kinds of imagery show tonal differences in eroded areas but limitations for

identification of variation in degree of erosion between the two extremes were similar to those for the panchromatic black and white photography. The black and white green and near infrared bands did not differentiate by their tonal qualities, the boundaries between eroded and non-eroded soil areas.

CONCLUSIONS

1. The preliminary results obtained from these studies indicate that more detailed ground-truth collection under controlled conditions is needed in order to adequately assess the overall application of remote sensing to soil erosion studies.

2. The relationship between the image obtained at a particular wavelength and the physical factors producing it, can only be understood through measurement and quantification of the ground condition at time of sensing.

3. A single wavelength cannot provide all the answers to terrain and vegetation analysis. The greatest possible amount of information can undoubtedly be gleaned from a multispectral approach using thermal infrared sensing as a supplement to colour, and black and white photography in various wavelengths.

4. It remains for future research to decide how costs in terms of investigation time and imagery processing, balance off against the additional benefits of having other kinds of spectral imagery available. The use of multispectral remote sensing techniques will undoubtedly stand in better perspective when applied to the study and evaluation of the complete spectrum of resources to be found in an area.

ACKNOWLEDGEMENTS

Acknowledgement is extended to the Manitoba Department of Agriculture for support of this portion of the 1971 Manitoba Remote Sensing Program. Thanks are also due to Mr. C. Jenkins, Manitoba Department of Mines, Resources and Environmental Management for suggesting the topic of this paper. The author gratefully acknowledges the assistance of Mr. Grant Fraser, Soils and Crops Branch, Manitoba Department of Agriculture for providing additional ground-truth data and for his critical review of the manuscript.

REFERENCES

1. Anderson, D.T., et al. 1966. Soil erosion by wind, cause, damage and control. Publication 1260. Canada Dept. of Agriculture.

2. Ehrlich, W.A., et al. 1957. Report of Reconnaissance Soil Survey of Carberry Map Sheet Area. Soils Report No. 7, Manitoba Soil Survey.

3. Jacobson, Paul. 1969. Soil erosion control practices in perspective. Journal of Soil and Water Conservation, July-August, 1969. Vol. 24, No. 4, pp. 123-126.

4. Olson, T.C. and W.H. Wischmeier, 1963. Soil erodibility evaluations for soils on the runoff and erosion stations. Soil Sci. Soc. Amer. Proc. 27: 590-592.

5. Soil Survey Manual. 1951. Agr. Handbook No. 18. United States Dept. of Agriculture. Government Printing Office, Washington, D.C.

6. Wischmeier, W.H. and J.V. Mannering, 1969. Relation of soil properties to its erodibility. Soil Sci. Soc. Amer. Proc. 33: 131-137.

7. Wischmeier, W.H., et al. 1971. A soil erodibility monograph for farmland and construction sites. Jour. of Soil and Water Conservation. 26: 189-193.

Table 1

Specifications of Imagery Obtained for the Wellwood Study Area

Date	Altitude	Film	Filter	Wavelength (μ)	Band	N.A.P.L. Roll No.
July, 1948	8,000 ft. a.s.l.	Pan. B & W	Minus Blue	0.5-0.7	Visible	A 11553-63
October, 1958	8,000 ft. a.s.l.	Pan. B & W	Minus Blue	0.5-0.7	Visible	A 16398-173
September, 1964	8,000 ft. a.s.l.	Pan. B & W	Minus Blue	0.5-0.7	Visible	A 18595-63
August 8, 1971	11,000 ft. a.s.l.	Aerocolour	NAV	0.5-0.7	Visible	CN 1221-280
"	11,000 ft. a.s.l.	TRI X B & W	W-12 + 44	0.5-0.6	Green	BN 1220-280
"	11,000 ft. a.s.l.	TRI X B & W	25-A	0.6-0.7	Red	BN 1219-280
"	11,000 ft. a.s.l.	IR Aerographic	89-B	0.7-0.9	Near infrared - black & white	BN 1218 IR-280
August 5, 1971	32,000 ft. a.s.l.	TRI X B & W	W-12 + 44	0.5-0.6	Green	BN 1208 -396
"	32,000 ft. a.s.l.	TRI X B & W	25-A	0.6-0.7	Red	BN 1207 -396
"	32,000 ft. a.s.l.	IR Aerographic	89-B	0.7-0.9	Near infrared - black & white	BN 1206 -396
"	32,000 ft. a.s.l.	Aerochrome IR	W-12	0.6-0.9	Near infrared - colour	CP 1209 IR-396
May 20, 1971 07:50 13:00	2,000 ft. a.g.l.	Infrared scanner image	-	3.0-5.0	Intermediate infrared-thermal	-

Table 2

Soil Properties and Surface Condition at the Detailed Ground-Truth Sites in the Wellwood Area

Soil Type and Classif-ication*	Site No.	Soil Property							Surface Condition					
		Surface Texture	Colour	Depth (cm)	Soil Temp. °C		Soil Moisture %		Topography		Trash cover (%)	Direct. of Cult.	Roughness	Aggregation**
					AM	PM	AM	PM	Slope	Aspect				
Wellwood (O.Bl)	1	VFSL	10YR3.5/1	1 20	1.0 7.0	31.0 9.0	5.8 29.3	2.8 26.7	level		0	N-S	Crests 8" apart, 3" troughs	fgr, some cogr & vco sbky
Stockton (O.Bl)	3	LVFS	10YR3/1	1 20	- 2.0 6.0	26.5 10.0	2.6 15.7	1.0 15.4	0-½%	E-NE	0	N-S	Crests 6" apart, 3" troughs	vfgr - single grain, some cogr & vco sbky
Wellwood (O.Bl)	4	L-VFSL	10YR3.5/1	1 20	- 1.0 7.0	25.0 9.5	4.3 31.1	2.0 30.0	level		0	N-S	Crests 6" apart, 3" troughs	m & cogr, some vco sbky
Wellwood (O.Bl)	6	Loam	10YR3/1	1 20	0.0 7.0	27.0 10.0	13.6 30.1	3.5 31.1	level		30	NNE-SSW	Crests 10" apart, 1" troughs	f-mgr, some vco sbky
Wellwood (O.Bl)	7	Loam	10YR3/1 & 10YR3.5/1	1 20	3.0 7.0	22.0 10.0	18.5 30.2	9.3 29.9	level		90	N-S	Smooth	silt wash on surface of aggregates
Firdale (O.DG)	9	Loam	10YR5/2	1 20	10.0 11.0	29.0 12.0	4.3 20.2	1.9 21.7	12%	S-SE	0	N-S	Crests 6" apart, 3" troughs	m-cogr, some co-vco sbky
Firdale (O.DG)	10	Loam	10YR5/2	1 20	7.0 8.0	27.5 11.0	3.6 21.2	1.9 21.7	12%	N-NW	0	N-S	Crests 6" apart, 3" troughs	m-cogr, some co-vco sbky
Firdale (O.DG)	11	VFSL	10YR4/2-5/2	1 20	11.5 9.0	26.0 12.0	2.7 21.0	1.1 19.6	0-½%	E-SE	40	E-W	Crests 12" apart, 5" troughs	m-cogr, some co-vco sbky

* O - Orthic
Bl - Black
DG - Dark Grey

** f - fine
m - medium
co - coarse
vco - very coarse
gr - granular
sbky - subangular blocky

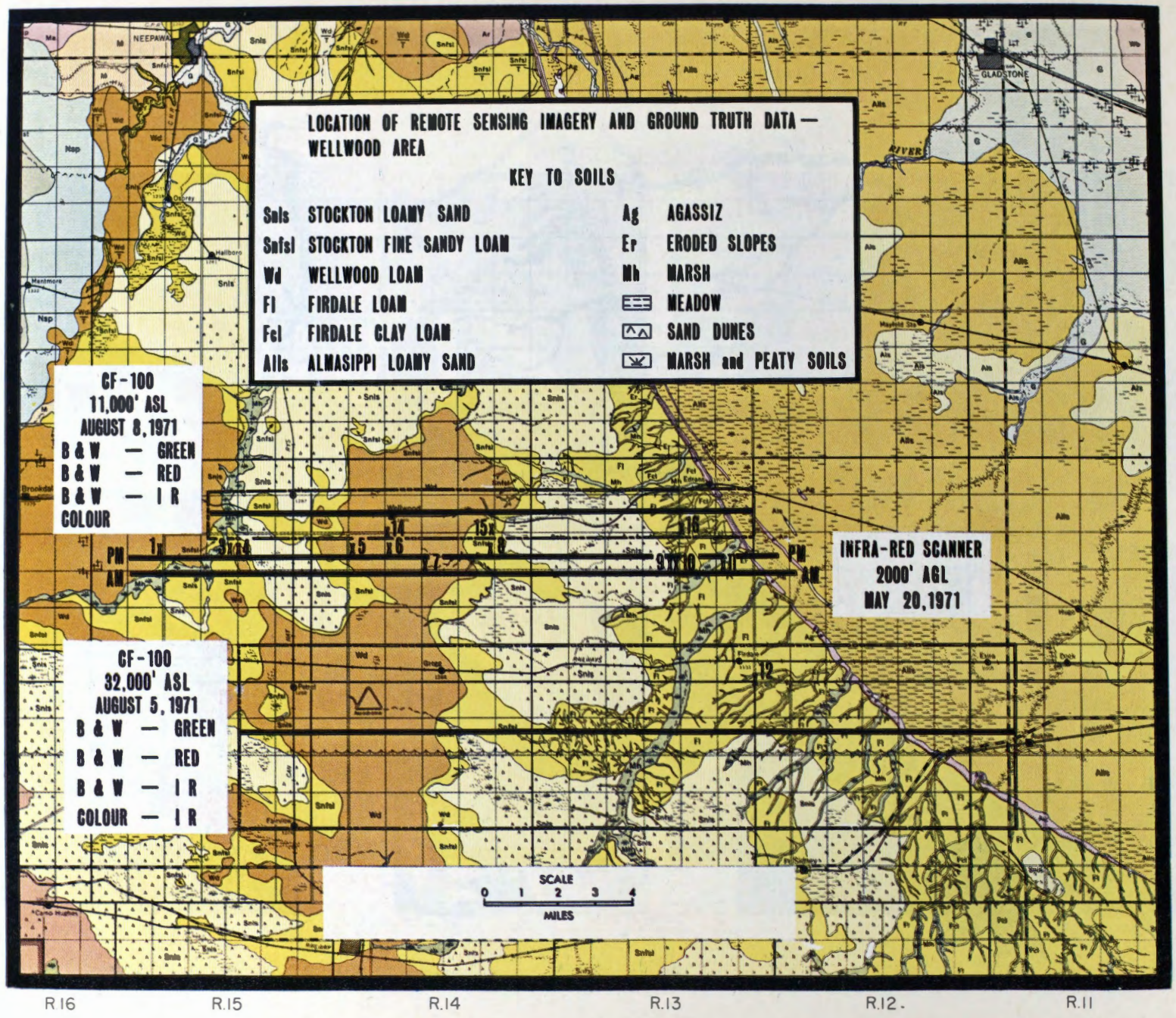


Figure 1. Location of remote sensing imagery and ground-truth data - Wellwood area, Manitoba.

Figure 2. Thermal infrared imagery (3-5 μ, 13:00 hours) of ground-truth sites on the Wellwood and Stockton soils. Ground-truth sites indicated by x.

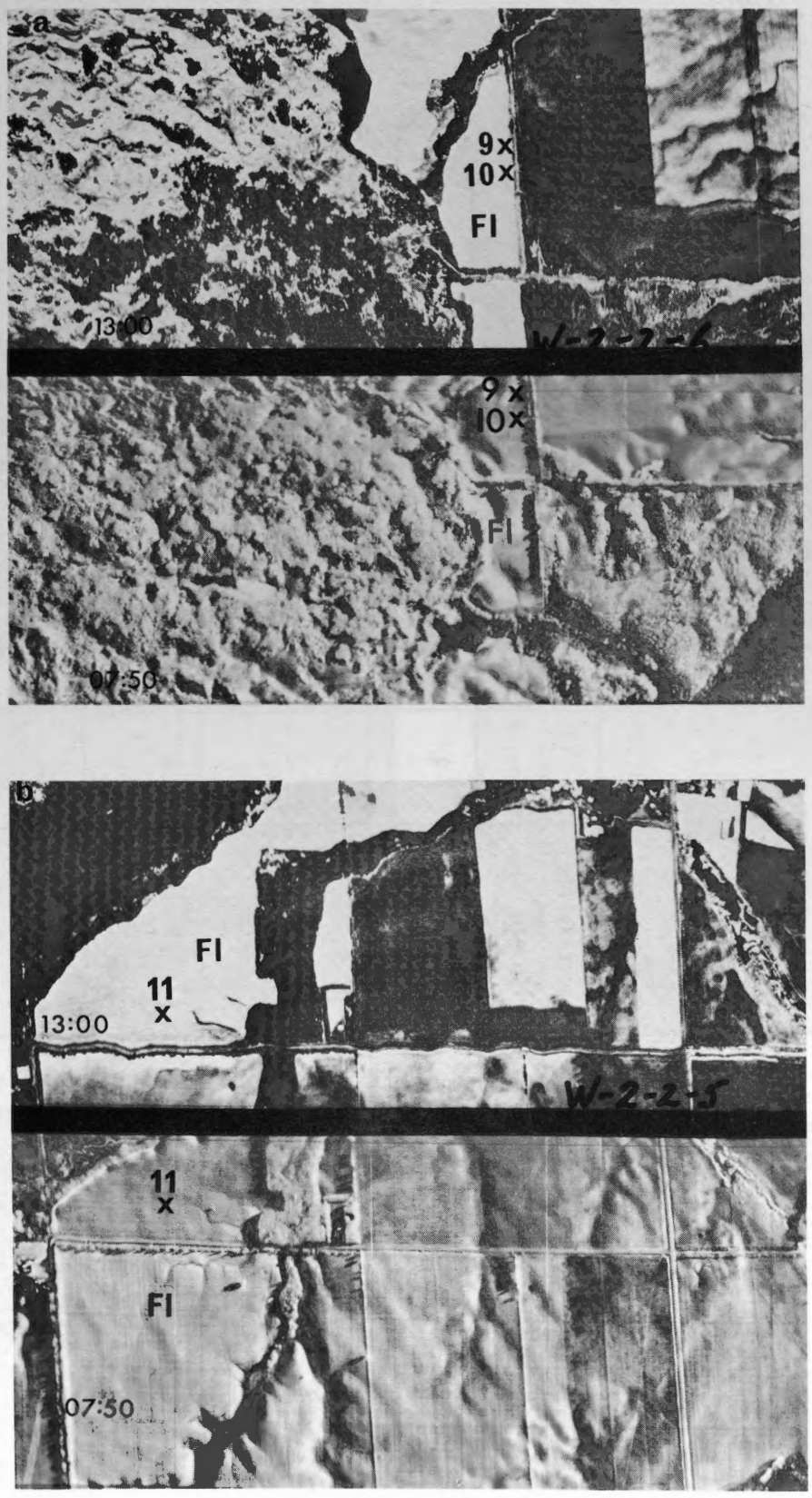


Figure 3. Thermal infrared imagery (3-5 μ, 07:50 and 13:00 hours) of ground-truth sites on the Firdale soils. Ground-truth sites indicated by x.

Table 3

Description of Ground-Truth Sites for Soil Erosion Classification

Soil Type and Classification*	Site No.	Profile Characteristics		Topography		Degree of Erosion**	Image tone on Pan. B & W
		Horizon	Depth, ins.	Slope, %	Length, yds.		
Firdale loam (O.DG)	12a	A B C	3 2.5	2	100	E_1	dark grey
	12b	C A and B removed		12	75	E_3	light grey
	12c	A B C	5 5.5	12	65	Native Trees	dark grey - very dark grey
Stockton fine sandy loam (O.Bl)	15	A B C	5 7	2	200	E_1	dark grey

* O - Orthic
Bl - Black
DG - Dark Grey

** Soil erosion classes:
E_1 - none to slight erosion
E_2 - moderate erosion - 25-50% of the productive topsoil lost
E_3 - severe erosion - more than 50% of the productive topsoil lost

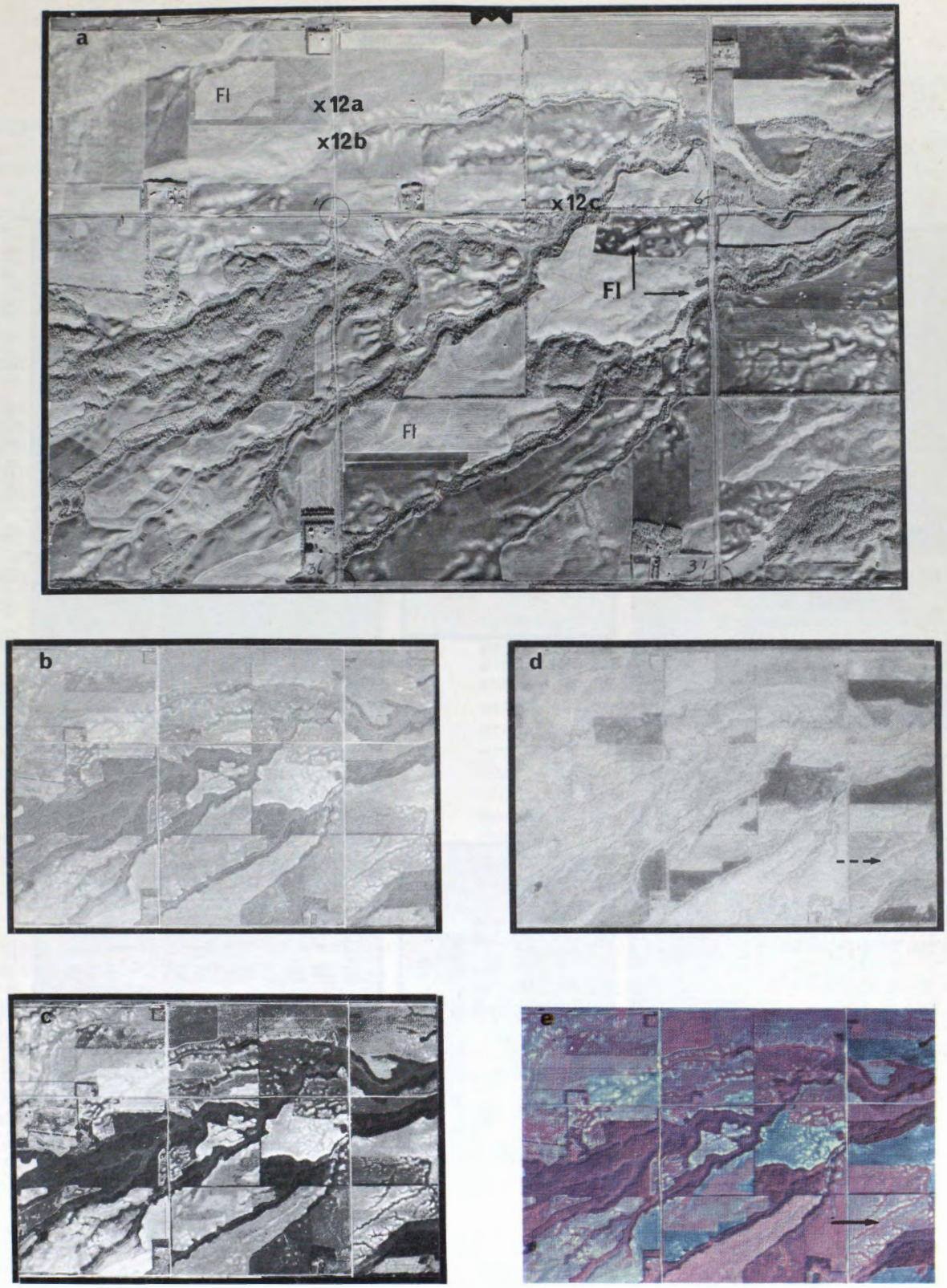


Figure 4. Multispectral photographic imagery of the Firdale soils
(a) Panchromatic black and white (October 29, 1958)
(b) Green band, black and white
(c) Red band, black and white
(d) Near infrared, black and white
(e) Near infrared, colour
Note: Imagery at b, c, d, and e taken August 5, 1971.

Figure 5. Multispectral photographic imagery of the Stockton and Wellwood soils (a) Green band, black and white (b) Panchromatic black and white, July 22, 1948 (c) Red band, black and white (d) Near infrared, black and white (e) Panchromatic black and white, September 15, 1964 (f) Colour Note: Imagery at a, c, d and f taken August 8, 1971.

HIGH FLYING HARDWARE AND

THE COMMON MAN

R.W. Higgs,
Traffic Engineer,
Federal Department of Public Works,
Ottawa, Ontario, Canada.

SUMMARY

An attempt is made in this paper to correlate the work of several independent researchers with reference to the tundra, treeline, muskeg, permafrost border as it occurs in the Northwest Territories.
The relationship of air mass types to this boundary and the location of cloud cover is discussed.
The usefulness of Essa-8 weather satellite "photography" in corroborating field information, and its possible extension to other disciplines is indicated.

INTRODUCTION

Precisely 200 years ago, Samuel Hearne completed his famous Coppermine Journey. He accomplished this feat without benefit of aerial photographs or satellite information and in the face of great hardship marked by severe climate, shortage of food, inadequate shelter and uncooperative Indians. The route which he chose is indicated on figure 1. It is with this specific location that the following paper is primarily concerned.

Hearne, of course, kept close to the barren lands for, I suspect, easier going and for the caribou for food; while remaining near the treeline for firewood and shelter. Unfortunately, from what we are told by Bryson *[3] this was also the storm path as you will see if you read Coppermine Journey.*[9]

No doubt, when he lost his compass Sam would have welcomed the Navy navigation satellite system*[14] and the portable receivers which the U.S. Army is presently experimenting with which work in conjunction with the satellites. Presumably, it will shortly be impossible to get lost no matter where you are on the earth's surface.

SATELLITE PHOTOGRAPHY

Last February (1971), I was fortunate enough to obtain from the Satellite Data Laboratory in Toronto several ESSA-8 weather satellite photos of the Yukon/NWT region. My intended use was, hopefully, to try and determine if some relationship existed between cloud cover and permafrost boundaries.

If any of you have ever studied these weather satellite photos and these were hand-picked cloud-free best examples, you will know that something less than 40% of Canada is cloud-free for any length of time. Not only is it difficult to draw any relationships but it is hard enough to recognize prominent surface features. Only Great Slave Lake seems to have a special working agreement with the Almighty.

CLOUD COVER AND AIR MASS

Brown*[7] states "Variations in cloud cover throughout the permafrost region may cause significant differences in the amount of Solar radiation received by the ground surface" and further "cloud cover is generally greater east of Hudson Bay than in Western Canada: this could influence the distribution of permafrost but no detailed information is available".

According to Bryson* "air mass and air stream patterns are profoundly affected by topography, winds being swept into the interior plains through gaps in the Cordillera. He further states that the arctic front seems to be topographically anchored year round at the northern end near Aklavik". Bryson implies that the dominance of a particular air mass over the tundra and a different air mass over the boreal forest may be of more than casual interest to the scholar concerned with possible relationships between climate and vegetation communities in the north".

By tracing trajectories of air masses to their sources, Bryson defined four general air masses as Arctic, Pacific, Atlantic and United States. (p.30 Bryson)

The latter two defined air masses have little bearing on western Canada since neither one penetrates to any great extent, being confined to the region south and east of Hudson Bay.

Favoured by a strong zonal westerly flow, Pacific air, however, penetrates 2,500 miles inland from its source so that it occurs with a frequency equal to Atlantic air along the east shore of Hudson Bay. Bryson divides Pacific Air into several sub-groups.

Yukon Pacific air crosses the cordillera through the Ward River gap; Alaskan air enters the MacKenzie Valley from Alaska, apparently through the Peel and Pelly Valleys." It would appear that these air masses are the

ones which influence the Boreal forest zone. "Maximum temperatures were taken as being fairly representative of an air mass". It is implicit in the air mass concept that different sources produce air of different mean characteristics.
"Evidently outbreaks of southwestern and Pacific air follow the same general pattern in summer as in winter". That they should is not very surprising in the light of the abundant evidence showing the profound effect of topography in the air mass and air stream patterns".

SOLAR RADIATION

In northern latitudes less than 50% of the sun's radiation reaches the earth's surface, some 35% being reflected or absorbed by cloud cover and another 20% being scattered or absorbed by the atmosphere. (Baier *15)

Cloud cover then plays a large part in controlling temperature and this, of course, is obvious to anyone who spends anytime outside on a mild cloudy day.

Since the cold arctic cloud mass hangs over the tundra zone, it is readily apparent that radiation reaching the earth surface here is very much less than in the cloud free areas of the boreal, regardless of other factors such as latitude, hours of daylight, etc.*2

Approximately half of the energy from the sun is contained in the infra-red portion of the spectrum. The ultraviolet rays are abruptly terminated by ozone and oxygen absorption. Absorption of the infra-red rays by water vapour and carbon dioxide also occurs.

PHOTOSYNTHESIS

"Under conditions of adequate light intensity and duration and available water supply, temperature becomes the dominant control over plant growth" - *17 (Met. mem. 17-1964 p.2) "The concept of" growing degree - days was first suggested by the French physicist Reaumar as early as 1735 but it apparently has not been widely used in Canada even today although a comprehensive survey of the subject was published by Holmes and Robertson of the Canada Department of Agriculture. (pub. 1042) It can be seen that a 100-day growing season approximates the tundra line whereas a 200-day growing season is required for production of fruit in Canada (Figure 4).

It is surprising indeed that only some 5% of the radiation absorbed in the daytime is used for the purpose of photosynthesis (see Rad. Energ, Its receipt and disposal.) Baier[15]

PERMAFROST (Figure 5)

The Division of Building Research at the National Research Council stated in May 1971 that substantial research is still required for the development of remote sensing techniques for detecting permafrost.*11

Brown*7 acknowledges that microclimatic factors are also important in influencing the distribution of permafrost, i.e., Net radiation, evaporation, condensation and conduction -convection. Wind speeds are lower in areas of dense growth, therefore, there is less evaporation.

According to Brown, the southern boundary of continuous permafrost zone is based on a Russian value of minus 5°c (23°F) isotherm of mean annual ground temperature, which corresponds to a 17°F mean annual air temperature, more or less. "More field tests are necessary to assess its validity in this country. since there is virtually no field information on the line separating the continuous and discontinuous permafrost zones. The 50% isopleth for Arctic Air passes between Fort Reliance and Artillery Lake where the forest border is quite clearly defined. According to Bryson, the agreement in position of the tree line, the storm track, and the air mass boundary is amazingly good. Apparently, the permafrost, muskeg and tundra boundaries are equally well-defined by this line.

BOREAL FOREST

Comparing Figure 3 (Bryson's Figure 19) with ESSA-8 photos taken in March 1969, it can be seen that the southern limit of the Arctic air mass coincides with the southern boundary of the heavy cloud cover on the photos. Bryson*3 has shown that this line is coincident with the tree line in the Northwest Territories. It also appears likely that the southern boundary of the Boreal is delineated by the northern limit of the mild Pacific air mass as can be seen when we compare Figure 3 with Figure 6 (Warkentin's*6 Figure 4.4), although Bryson passes this off as too coincidental. The frequency of the Alaska/Yukon air mass is highest in the Lake Corridor whereas the Arctic and Pacific air masses are lowest. Larsen*4 has indicated that Air Mass Frequency is directly related to vegetation in the Tundra. Warkentin states "Canadian forest are spatially

associated with climatic conditions which are warmer than those of the Tundra and moister than those of the grasslands. The transition from forest to grassland takes place over a moist to dry gradient and that from forest to tundra over a warm to cold grandient." (see Bryson)[3]

Bryson sought to prove that the southern edge of the boreal forest was also a meteorologically defined boundary and in this way to show that "there is a natural regionalization of climatic complexes which results ultimately in biotic regions."

Warkentin has indicated that there is a spatial correspondence between the tree line and the summer thermal conditions as expressed by the Thornthwaite index of thermal efficiency, an index of 12 being equivalent to the tree line.

The thermal conductivity of bare rock is many times that of silt; thus, the barren lands which have been nearly swept clean of all loose material by glaciation give up heat readily and add to the permafrost regime.

Radforth's[8] terrazoid airform pattern is found almost without exception in the northern part of the Northwest Territories and northern Quebec and is commonest where permafrost is continuous on lands where the underlying mineral terrain is almost flat.

The southern boundary of Radforth's terrazoid classification again approximates the tree line/tundra, air mass, cloud cover, and permafrost boundary already discussed. That the tundra is, in fact, a topographically-defined region seems evident from the direction of flow of all the rivers, north of the tree line. Retardation of uplift after glaciation may be a contributing reason to the maintenance of the permafrost.* Barr[16]

NORTHERN CONSTRUCTION

Turning to northern construction; even in surveying the shortest distance between two points is a straight line.

If we draw a line from Norman Wells to Winnipeg, we pass south of the Precambrian Shield and all the major lakes that lie therein. We, therefore, avoid all the hard-rock excavation which would be necessary. We also avoid all or most of the permafrost zone. (Permafrost at Fort Vermilion is absent whereas at Tundra Mines, east of Yellowknife, it reaches a depth of 900'.)

It can be seen then that the best place for a road, for example, is along the line* Figure 1 which not only avoids hard-rock and permafrost but involves a minimum of major water crossings as well. Even migrating birds prefer this route*[10]. Hours of bright sunshine per annum are higher west of this route than to the east in the tundra *[2]. Evidently, where the hours of sunshine exceed 2,000/annum, permafrost does not exist. Would satellite photos help in this choice? I doubt it. They do, however, seem to enforce the choice in that most of the heavy cloud cover hangs over the shield north of Brown's "southern limit of permafrost". In fact, the clear areas on these satellite photos, and therefore probably the warmest since they get the most sunshine, are along the line I have indicated (map).

The difficulties entailed in locating suitable borrow for highway construction in the Arctic together with the problem of excavating on frozen ground adds greatly to the cost of any project. Surely, then one would be well advised to avoid these permafrost areas as much as possible. Certain advantages accrue to the location proposed, for e.g.:

Most of the mining areas in Canada occur in the southern or outermost edges of the Canadian Shield at or near the contact with the sedimentary rocks.*[19]

Oil deposits too are to be found along this zone, particularly at Norman Wells and in the Tar Sand deposits near Fort McMurray. The latter are reported to exceed half of the world's known reserves.

In many places, this alignment is over existing roads and only short sections of new construction are required in order to complete the connections. In many cases, existing railway lines parallel the roads. As well, air routes follow the Mackenzie River. The timber, utilities and people are here also.
It seems most logical then to not only complete the road network along this alignment but to utilize it to facilitate construction of any other facilities This line is focussed on the resource potential of the north.[11]

The line which I am proposing largely follows MacKenzie's route as well as that of Pond.*[6] (Figure 2-8) It lies mainly in the areas of Devonian, Ordivician and Silurian rocks. Any new mineral or oil discoveries are most likely to be found along this line rather than in the tundra.

CONCLUSION

As Dr. Robert Harwood[5] reported recently, cloud photography is by no means the only meteorological information which can be obtained from satellites. Air and ground temperature, relative humidity, winds and ozone distribution can all be deduced in varying degrees by precise measurement of the electromagnetic radiation reaching the satellite from below.

Judging by the similarity of results obtained independently by different researchers as correlated in this paper, it should be possible to utilize our weather satellites not only to predict the weather but also to validate boundaries of air masses, permafrost, muskeg, plant regimes and all the myriad and wonderful things that we have been so laboriously determining from field measurements in the past, possibly, even property boundaries.*[21]

According to the National Research Council, the heat exchange mechanism at the ground surface is so complex that the occurrence of permafrost cannot be accurately predicted from air temperatures alone. Knowledge of the mean ground temperature is deemed essential for the prediction of the effect of temperature changes.

Meteorological satellite information which is normally taken during hours of darkness with infra-red scanners, provides the general pattern of long-wave radiation from the earth's surface. More detailed study of this information may be of tremendous value in predicting permafrost boundaries particularly in the discontinuous zone. The newly-developed Raman Laser*[20] should help in this regard.

It was reported recently that NOAA[14] intends to record all cloud information on microfilm. Gray-scale recording with a potential of 64 gray-scale levels is to be used. This will facilitate future studies.

REFERENCES

1. ESSA - 8 Satellite Photos 1969-70 Meteorological Branch, Toronto Weather Office.
2. Climatic Normals Vol. 3, 1968 - Met. Branch, DOT, sunshine cloud, etc.
3. Geographic Bulletin Vol.8, No.3, 1966, P.225 - R.A. Bryson - air masses streamlines and the Boreal forest.
4. Arctic, Vol.24, No.3, Sept. 1971, p.177 - J.A. Larsen - Vegetation Relationships with Air Mass frequencies: Boreal Forest and Tundra.
5. New Scientist and Science Journal, Sept.16, 1971 - Dr. Robert Harwood - Mapping the atmosphere from space.
6. Canada - A geographical Interpretation - John Warkentin - Editor - 1968 Cdn. Association of Geographers.
7. Permafrost in Canada - R.J.E. Brown 1970, Figure 3.
8. Muskeg Engineering Handbook - NRC 1969, Chap.3, p.53, Airform patterns, N. W. Radforth.
9. Coppermine Journey - Farley Mowat (Samuel Hearne's exploration)
10. Canadian Wildlife Cookery - Frances Maciliquham (Bird Migrations)
11. Canadian Geotechnical Journal, Vol.8 1971, p.236 - Note 1 (needed research) Construction on Permafrost. (p.248)
12. Journal of Remote Sensing, Vol.1, No.5, Nov./Dec. 1970 - Remote Sensing of the Earth's Surface - G.A. Rabchevsky
13. Photogrammetric Engineering, Vol.36, No.5, May 1970, p.491 - Fidelity of Space TV by Dr. K.W. Wong (Distortion characteristics of ESSA - 7 Photography)
14. Engineering Graphics, Nov. 1971, p.22
15. Radiant Energy, Its Receipt and Disposal - W. Baier, Canada Dept. Agriculture, 1971, Fig. 3.3
16. Arctic, Vol.24, No.4, Dec. 1971, p.249 - Postglacial Isostatic Movement - William Barr (also Northern Pipelines: The Canadian Position Wm. J. Fox).
17. Meteorological Memoir 17 - 1964, p.2
18. Montreal Gazette, Jan. 22, 1972, p.9

Figure 1. Samuel Hearne's route, 1772, ---------, proposed route ————

Figure 2. ESSA-8 satellite photograph

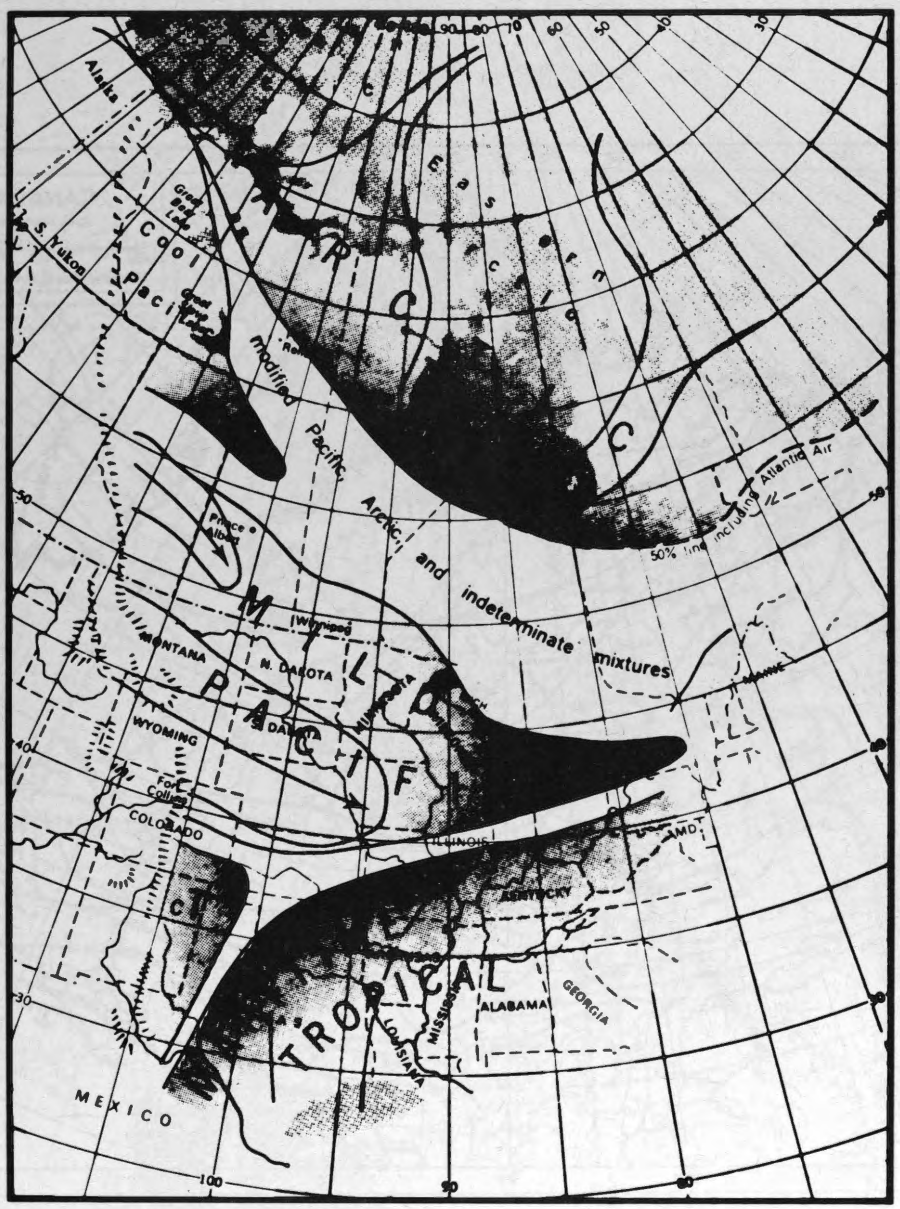


Figure 3. Air masses (Bryson's Figure 19)

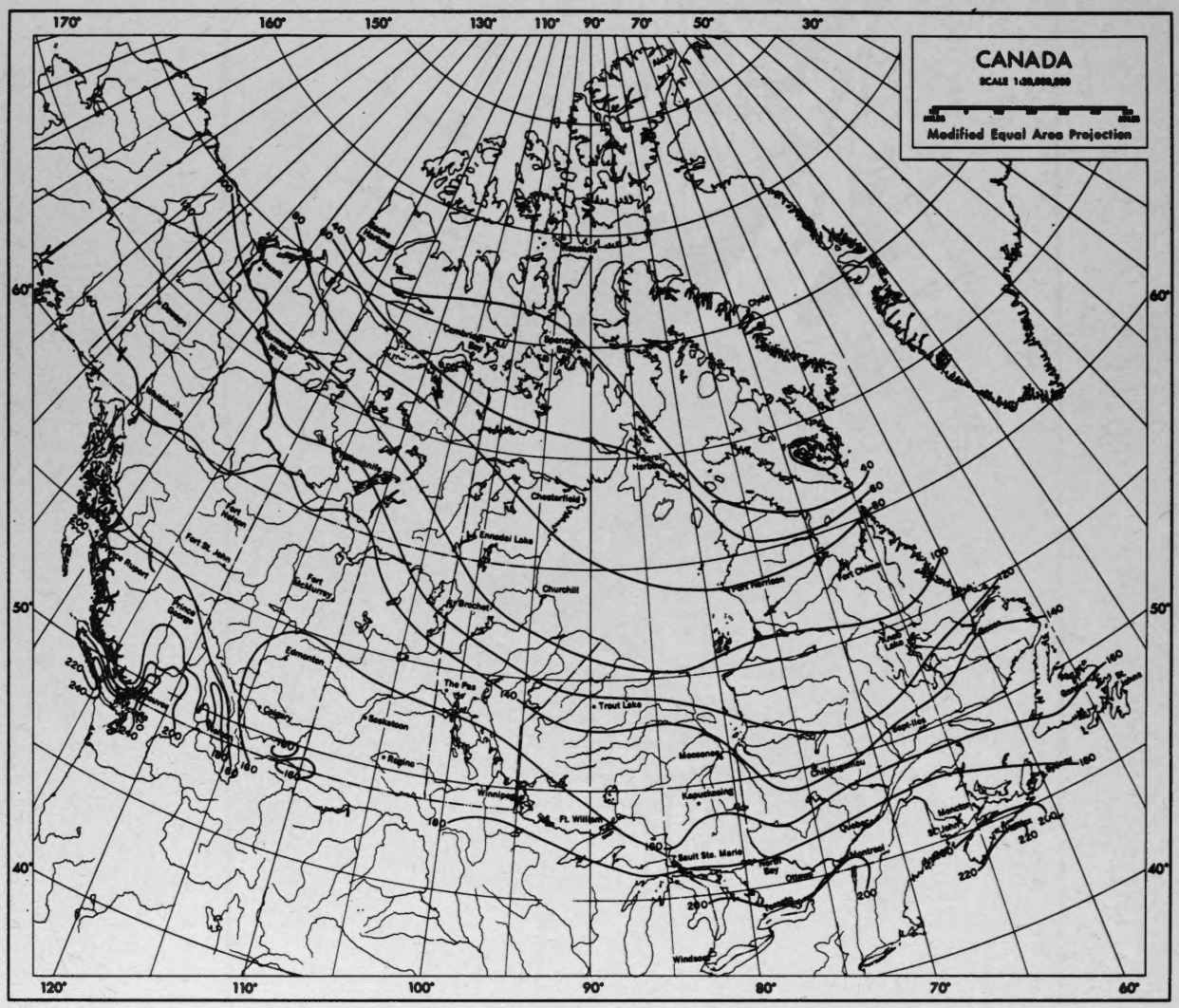


Figure 4. Total number of days in growing season

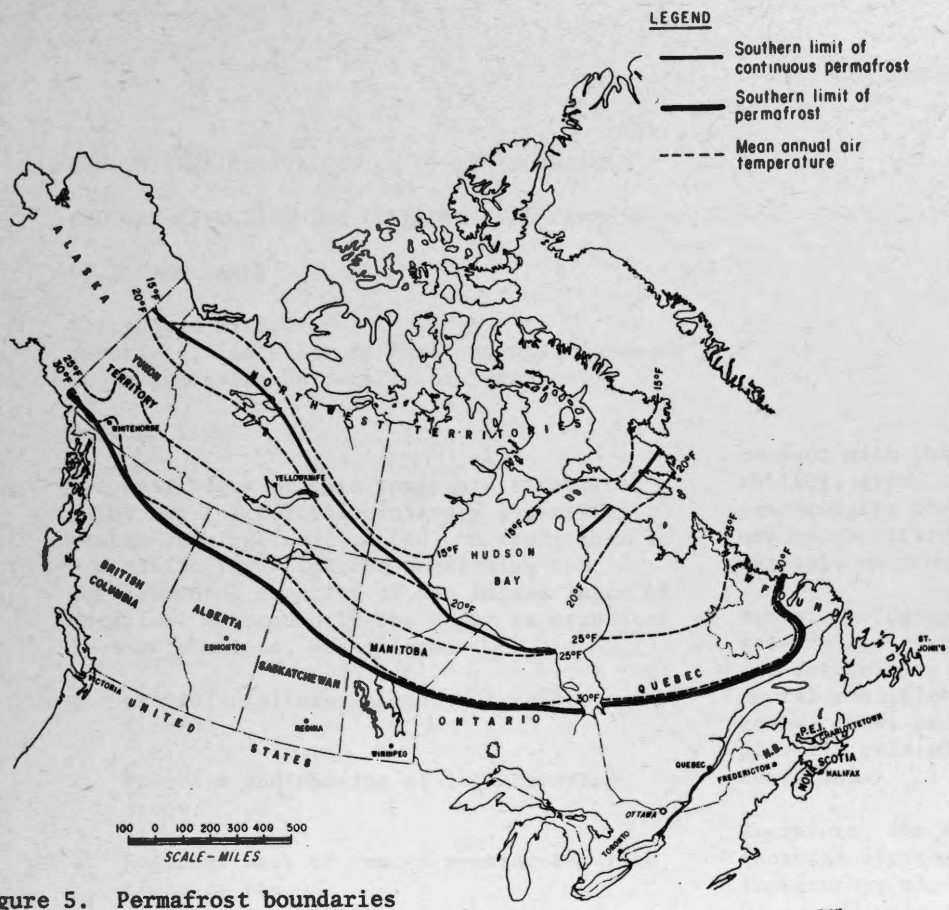


Figure 5. Permafrost boundaries

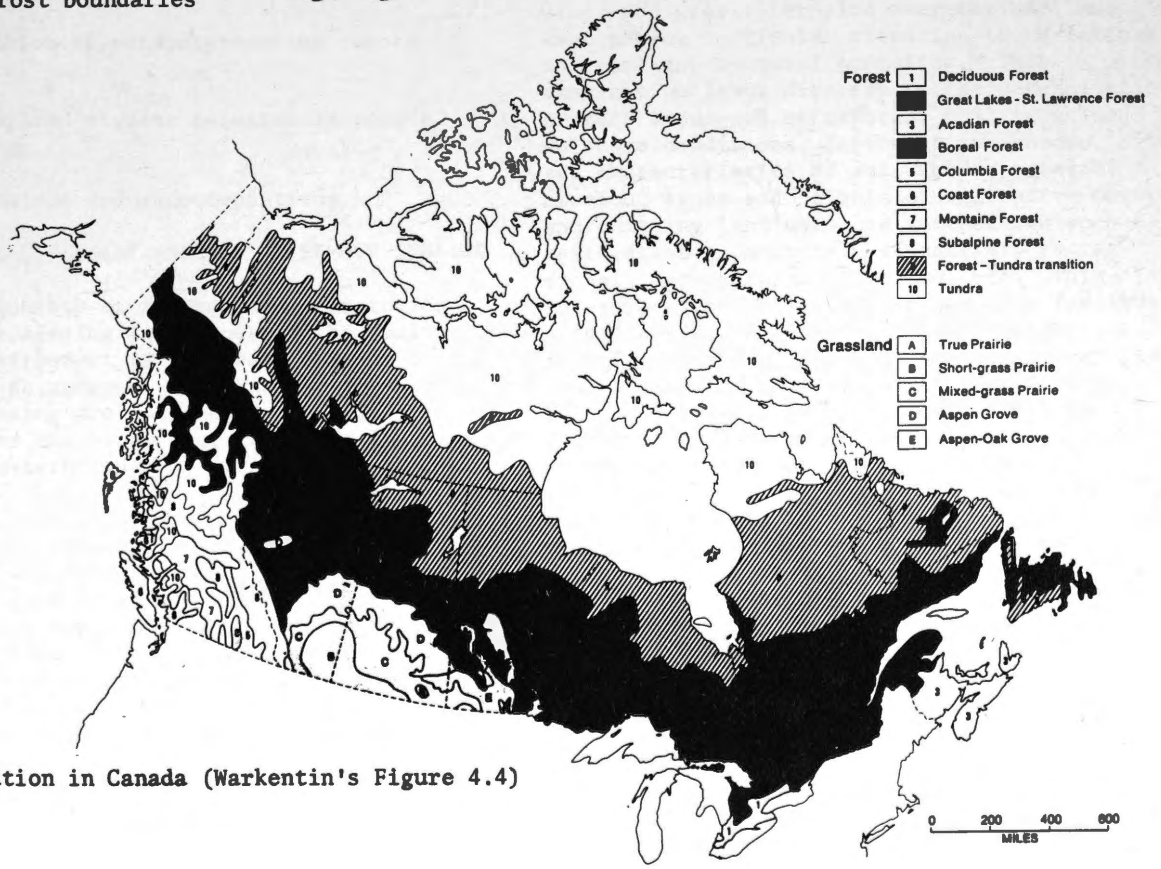


Figure 6. Vegetation in Canada (Warkentin's Figure 4.4)

REMOTE SENSING EVALUATION OF ENVIRONMENTAL FACTORS AFFECTING THE DEVELOPMENTAL CAPACITY OF INLAND LAKES

S.J. Glenn Bird,
Associate, Institute of Environmental Sciences and Engineering University of Toronto

INTRODUCTION

The objectives of this paper are first, to point out the multidisciplinary aspects of remote sensing, and, second, to apply them to a specific investigation concerning the developmental capacity of the inland lakes of Ontario. Consequently the paper is presented in five sections, as follows:

1. Multidisciplinary aspects of remote sensing.
2. Purposes and phasing of lake capacity study.
3. Regional uses of remote sensing for lake capacity study.
4. Evaluation of parameters using remote sensors.
5. Specialized studies required in remote sensing.
6. Conclusions and recommendations.

1. MULTIDISCIPLINARY ASPECTS OF REMOTE SENSING

The relationship of airphoto interpretation and remote sensing techniques to the evaluation of the environment is discussed briefly in the following paragraphs. The basic principles of remote sensing are outlined in the first section for the benefit of those not familiar with these techniques.

The aerial photograph accurately records all natural and cultural features on the earth's surface. The format usually consists of black and white (panchromatic photographs, taken on a continuous basis of all but the northernmost tips of Canada, by various federal and provincial governmental agencies. In addition, particular film-filter combinations, using colour and infrared colour photography, together with infrared scanning, have been used to remotely sense certain areas, in order to accentuate terrain characteristics for special studies, without necessitating actual contact with the land surface. The availability, area, coverage, and economy of the panchromatic photography has resulted in its use by qualified personnel to assess and evaluate environmental conditions.

Aerial photographic interpretation is the science of recognizing significant patterns and determining their relationship to environmental conditions. Within major climatic zones, these patterns are consistent, and directly related to particular terrain conditions.

Therefore, the airphoto interpreter performs a thorough stereoscopic examination of the topography; micro and macro drainage; water, wind, and gravity erosion; vegetation; land use; paying particular attention to variations in tonal and textural anomalies. This examination leads directly to the determination of rock types and structures, soil types and moisture conditions, depths of overburden, size and characteristics of watersheds, rates of erosion, types and heights of vegetative cover, and existing land uses and intensities such as residential, commercial, industrial, recreational, agricultural, forested, etc. This interpretation investigation normally involves a regional evaluation to determine major terrain divisions, with the subsequent detailed analysis of the specific areas of interest. This analysis may be highly detailed, the degree being defined by the aims and objectives of the investigation.

Unusual patterns, indicating particular natural or cultural conditions are always significant, as these represent the interaction of individual terrain characteristics. Selective field checking of predetermined points within these areas normally validates the causes of these anomalies as originally inferred from the airphoto evaluation.

Airphoto interpretation, performed by an experienced analyst, involves all terrain aspects of the environment, and reveals areas of ecological imbalance or high sensitivity.

Consequently, the terrain data defined using this technique is highly relevant to a large number of scientific disciplines.

From personal experience in both research and consulting, it has become evident that the total interrelationship of the terrain data interpreted from aerial photographs is seldom considered to the most beneficial extent. To cite a simple example, certain soil types and moisture conditions may be related to bearing capacity or slope stability by a civil engineer, to a geochemical test area by a geologist, to fertility for crops by an agriculturalist, to tree growth capabilities by a forester, to animal habitat suitability by a zoologist, to the over-all vegetative capability by a biologist or botanist, to the effect upon the natural balance by an ecologist, and so on. For example, an air-photo archaeological study of the Indian reserve at Parry Island for the Royal Ontario Museum related terrain patterns to cultural practices of over 100 years ago.

From the procedures and examples cited above, it is apparent that airphoto interpretation is capable of contributing valuable information to a variety of scientific disciplines. Consequently, the airphoto interpreter is most effective when functioning in a multidisciplinary atmosphere. The mutual interchange of ideas clarifies the significance of information obtainable from aerial photographs, and increases the sensitivity of the interpreter to the needs of the various disciplines.

Townsite selection within a large geographical region is a common Canadian problem, wherein every physical aspect of the environment must be considered, thus ensuring, for example, adequate water supply and relative freedom from both short and long term pollution. Ecological, social and economic factors are vital considerations. The aerial photographic mosaic, covering a relatively large area is easily constructed and provides a convenient and useful media of data presentation.

2. PURPOSES AND PHASING OF LAKE CAPACITY STUDY:

GENERAL OBJECTIVES:

The objective of this study is to provide the Department of Municipal Affairs with the simplest possible criteria for making decisions concerning land development throughout the entire province of Ontario and in particular those areas adjacent to the recreational lakes of Ontario. The study was divided into three phases with the fourth being the continual refinement of the systematic model.

The four phases and their interrelationships are illustrated in the Study Flow Chart, designated as Figure 1.

Expertise from a number of scientific disciplines must be involved in the investigation, due to the broad scope of the study and the variety of effects upon the environment. Therefore, as the Institute of Environmental Sciences and Engineering at the University of Toronto represents a large number of disciplines, it was logical that we were engaged by the Department of Municipal Affairs to conduct this study. The personnel involved, their university departments, and particular fields of specialization are listed in Table 1.

Phase I

This phase, which has now been completed, dealt with a review of the existing information and the preparation of a detailed study design for Phase II.

Phase II

Phase II is concerned with the physical properties of land and water as they relate to recreational development. The most influential properties would be determined and individually weighted depending upon their relative importance. This would lead to a system whereby the fewest and simplest measurements possible could be taken for a specific lake and be entered into a workable computerized mathematical model. The lake would then be classified in regard to its development potential, and specific constraints determined, thus providing the information necessary for logical decisions based on scientific principles.

It is expected that the study design outlined under Phase I will result in an approximate definition of the following:

(a) The major parameters affecting each specific recreational land use.

(b) The relative importance of each of these parameters and an approximate weighting factor for each, the summation of which would indicate the water and land capability of that lake and watershed area to withstand each recreational land use.

(c) The number of categories into which the lakes of Ontario could be placed, based on similarity of physical characteristics.

(d) A series of guidelines for logical constraints for each type of recreational land use and lake category.

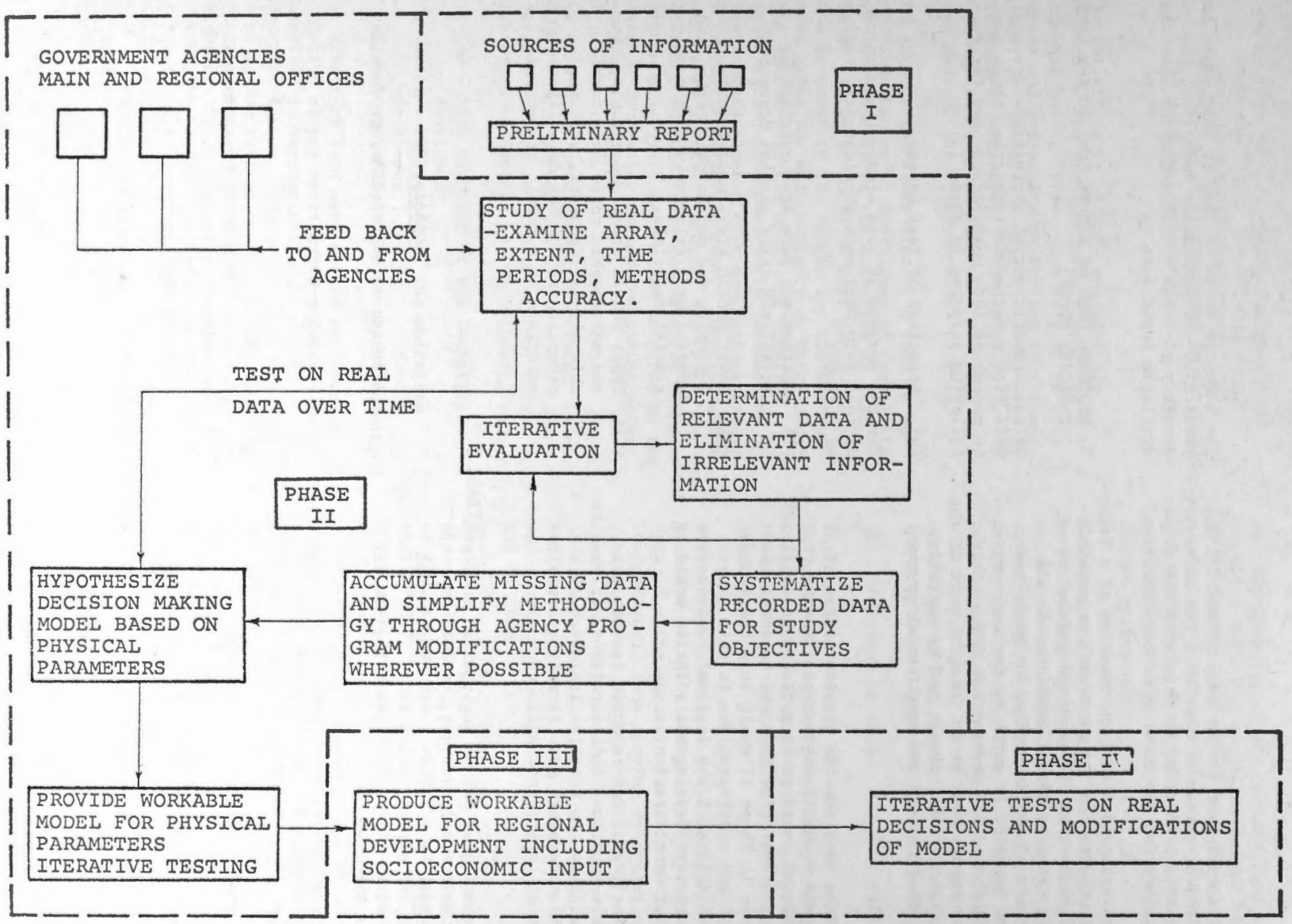


FIGURE 1. STUDY FLOW CHART

(e) The development of the most economical and efficient system of obtaining the necessary parameters in order to determine the proper category for the lake under consideration.

(f) The instigation and continuation of a training program for the personnel responsible for evaluating the required parameters of lake category classification. As all members of the investigation group have been involved in teaching for many years, this should not present any difficulties. This aspect must be an integral part of the program for this phase, and is essential for an efficient and operational system.

Phase III

This phase involves the considerations of social and economic parameters, leading to recreational development policies for regional planning areas. The physical constraints developed in Phase II would form the foundations of any considerations in this phase. The predictions of the regional consequences of alternative development policies would be the major consideration.

While the social and economic factors are extremely important considerations, it must be realized that the relative influence of the physical parameters must firstly be resolved because of their basic influence.

The computer would be used in this phase for predictive and simulation modelling, the accuracy of these models depending upon the available input data.

3. REGIONAL USES OF REMOTE SENSING FOR LAKE CAPACITY STUDY:

As panchromatic aerial photographs are available from governmental agencies, these would be assembled into mosaics and used for the following purposes in Phase II:

(a) Selection of Test Lakes:

For the purpose of this study, the province divided into four main physiographic regions as follows:

(i) Sedimentary rock area overlain by relatively thick glacial deposits.

(ii) Lower Shield area, underlain by igneous intrusive and metamorphic rocks, and covered by a relatively thin mantle of glacial materials.

(iii) Upper Shield area, underlain by igneous and metamorphic rocks, and overlain by thin glacial deposits, with the exception of thicker depositions of finer materials as a result of ponding.

TABLE 1

Name	Department	Speciality
S.J.G. Bird Project Director	Civil Engineering	Airphoto interpretation, groundwater
P.H. Jones Associate Project Director	Civil Engineering	Pollution Control
I.M. Burton	Geography	Resource Management, Environmental perception
B. Falls	Zoology	Wildlife ecology
F.E.J. Fry	Zoology	Fisheries ecology, limnology
J. Ganczarczyk	Civil Engineering	Water pollution
T.C. Hutchinson	Botany	Pollution ecology
G.K. Rodgers	Great Lakes Institute	Physical limnology, hydrology
A.K.F. Turner	Civil Engineering	Computer programming geology

(iv) Hudson Bay Clay Belt, the physical characteristics essentially being controlled by the relatively thick deposition of fine-grained soils resulting from glacial ponding.

The variations in climate are directly related to the physiographic regions on a regional basis. Also, such factors as lake depths and shapes are controlled to a large extent by the bedrock formations and glacial overburden. Therefore, the parameter of physiographic region exercises a strong control over many of the fixed physical factors.

The size of the water area of a lake is an important factor in developmental capability, and consequently in the selection of test areas. While it would be most desirable to be able to quantify the ranges for the surface water areas for each category at this time, this is not feasible until the detailed investigation more accurately defines the relationship between size and various types of recreational development. Therefore, the categories of small, medium, and large, for the selection of test lakes are relative within each physiographic region and selected on the basis of increasing the understanding within the lake ecosystem.

The density of development is related, in most cases, to the percentage of shoreline under recreational use and the terms low, medium and high could be defined by the following ranges:

Low - less than 20% of the shoreline developed
Medium - 20% to 60% of the shoreline developed
High - over 60% of the shoreline developed.

The effect of land uses within the watershed, exclusive of the water surface and adjacent shoreline on recreational land use is a very important factor. The benefits of studying lakes having various degrees of development, with the other parameters for test lake selection being relatively constant, would provide the opportunity to assess the influence of the types of land uses on the water quality and related water and land capabilities.

The location of a lake within a chain, or whether it is in confined drainage area, is another important criterion for selection, due to the influence of both surface and subsurface waters from related water and land areas. Wherever possible, lakes were selected in the same chain, thus providing more direct relationships.

As size, density of development, and location within a chain may be initially and approximately determined from topographic maps, the selection of the test lakes for this phase of the study would be finalized by a detailed examination of the existing black and white aerial photographs.

However, the use of high altitude photography would be extremely useful in establishing waterway connections, particularly infrared colour which so reliably registers moisture content variations, thus confirming inflow and outlet points. Therefore, for the purpose of selection of test lakes, this photography should be used as the second stage after topographic maps to verify chain position.

(b) Recording of Data:

The airphoto mosaics of each test area would be analyzed for the locations of optimum field test points, evaluated to best satisfy the investigative requirements of all disciplines involved.

Copies of the mosaics would be used for location of the field points for existing data, for the specific field investigations to be conducted, thus serving as the visual control format for the recording of all field locations. This procedure will ensure accuracy of location, ease of communication for field personnel, and increased efficiency and accuracy of filing, analysis, and processing of data.

Previous photographs of the test lakes would also be obtained, representing a time lapse of five to ten years with the most recent photography. This would provide the opportunity to analyse the development trends and environmental changes (i.e. offshore vegetation, erosion, accretion, water level changes, etc.) over these time periods. This comparison should become a matter of policy, being performed on a continuous basis as new photography becomes available.

This time lapse would be considerably reduced by the use of high altitude photography, both colour and infrared colour, which would indicate both vegetal and land-use changes on a regional basis.

4. EVALUATION OF PARAMETERS USING REMOTE SENSORS:

The investigative team listed the parameters affecting the types of recreational land usage and their effect upon each land usage was evaluated, in order to arrive at the study

priorities. The environmental parameters were divided into the following major and minor categories, indicated as follows:

Basin physiography
Water physical variables
Water biological variables
Water chemical variables
Existing land use within the watershed

A. Basin physiography:

* 1. Lake surface area
* 2. Lake shape (length of shoreline/unit area)
* 3. Lake depth (bathymetry)
* 4. Lake geological formation
* 5. Surface drainage characteristics
* 6. Subsurface drainage characteristics
* 7. Water flows (as flush factor)
* 8. Precipitation
* 9. Other meteorological parameters (thunder, electric storms, see A16, A18)
* 10. Solar radiation (number of clear days, amount of energy received)
* 11. Topography (rel. to A4 and also A1, A2, A3)
* 12. Soil type (rel. to A5)
* 13. Depth of overburden
* 14. Offshore vegetation
* 15. Volume
* 16. Wind direction (fetch)
17. Air temperature
18. Winds (rel. to A16)

B. Water physical variables:

1. Temperature when not frozen (see also B15)
* 2. Snow cover
* 3. Ice cover
* 4. Currents
* 5. Wave action
* 6. Intensity of water exchange (stratification and mixing)
7. Colour
8. Turbidity
9. Light penetration
10. Odour
11. Taste
12. Suspended solids
13. Settleable solids
* 14. Floating materials (oil, scum, foam, litter, etc.)
15. Water temperature (summer)
16. Evaporation

C. Water biological variables:

1. Coliform
2. Streptococcus
3. Salmonella
4. Virus (pathogenic)
5. Over-all algal counts
6. Presence of indicator groups of algae
* 7. Specific fish
8. Plant pests
9. Chironomids
10. Mayfly larvae
11. Sphaerotilus natano

D. Water chemical variables:

1. BOD
2. COD
3. TOC
4. D.O.
5. D.O. diurnal variations
6. Carbon dioxide
7. Carbon dioxide diurnal variations
8. Carbonates
9. pH
10. Alkalinity
11. Acidity
12. P contents
13. N contents
14. Ca contents
15. K contents
16. Mg contents
17. Fe contents
18. Mn contents
19. Heavy metals contents
20. SO_4 contents
21. Other toxic substances (biocides)
22. Irritating substances
23. Dissolved solids
24. Radionuclides

E. Socio-economic characteristics of the area:

* 1. Population density
2. Land values
* 3. Degree of development
* 4. Water treatment and wastewater treatment and disposal
* 5. Roads (access to shoreline assumed)
6. Railways
* 7. Powerlines
8. Pipelines - water
9. Pipelines - other than water
* 10. Sewerages
11. Agricultural land usage
* 12. Forested land usage
* 13. Urban land usage
* 14. Industrial land usage

The recreational activities were considered in three main categories, shown with the sub-categories as follows:

A. On the water:

1. Fisheries

2. Swimming and scuba diving
3. Canoeing
4. Sailing
5. Ice-boating
6. Skating
7. Snowmobiling
8. Rowing
9. Use of powered cruisers and house boats
10. Use of powered boats - inboard run-abouts
11. Use of powered boats - outboard run-abouts (muffled)
12. Use of powered boats - outboard run-abouts (unmuffled)
13. Use of aircraft (landing and takeoff)
14. Water skiing

B. On the shoreline:

1. Residential - permanent occupancy, year round
2. Residential - occasional occupancy, year round
3. Residential - occasional occupancy restricted seasonally
4. Marinas — Commercial Activity
5. Tourist resorts — " "
6. Retail outlets — " "
7. Hunting and fishing camps — " "
8. Aircraft docking facilities — " "
9. Camping grounds — Day Use
10. Picnicking — " "
11. Sunbathing — " "
12. Hiking (trails) — " "
13. Overnight docking facilities — " "
14. High density amusement facilities
15. Public beaches
16. Golf courses
17. Wildlife sanctuary

C. Within the watershed:

1. As B1 to B16 except B13 and B15
2. Ski resorts
3. Snowmobile trails
4. Arenas, racetracks, etc.

Therefore, the parameters were identified which were of critical importance to the largest number of recreational land uses, and also those land uses which were influenced by the most parameters. The ratings of importance were based upon the fact that the purpose of this study, in the first stage of the more detailed investigation, is the preparation of very approximate guidelines for recreational development constraints.

Those parameters having critical importance are identified by an asterisk, and their relation to remote sensing evaluations are discussed in the following section of this report.

Of a total of 93 parameters, the 32 marked with an asterisk were considered to be the most critical in the first stage of Phase II, and are discussed individually, regarding the potential of remote sensing evaluations, "mosaic" referring to existing panchromatic photography, under the assumption that the following specific terrain factors have been evaluated from the same media:

(a) Surface drainage efficiency and direction, both within the watershed and including inflow-outflow characteristics of the individual lakes.

(b) Topographic slope characteristics.

(c) Soil types and moisture conditions.

(d) Depths of overburden.

(e) Areas of rock outcrop and swamp, including offshore vegetation and stagnant bays, augmented by frequent infrared colour photography, from spring through to fall.

A. Basin physiography:

1. Lake surface area: easily measured from mosaic

2. Lake shape (length of shoreline per unit area): easily measured from mosaic

3. Lake depth (bathymetry): information available or requires field measurements

4. Lake geological formation: combination of factors (b), (c), (d) and (e)

5. Surface drainage characteristics: factor (a); augmented by infrared colour photography

6. Subsurface drainage characteristics: factor (c); augmented by infrared colour photography

7. Water flows (as flush factor): factor (a)

8. Precipitation: information available

9. Other meteorological parameters: information available

10. Solar radiation: information available

11. Topography: factor (b)

12. Soil type: factor (c)

13. Depth of overburden: factor (d)

14. Offshore vegetation: factor (e); augmented by infrared colour and high altitude photography

15. Volume: information available, or requires field measurements

16. Wind direction: information available
Fetch: easily measured from mosaic

B. Water physical variables:

2. Snow cover: information available, or requires field measurements

3. Ice cover: requires field measurements

4. Currents: indicated by pattern on airphoto mosaic also factor (e), may require field measurements, augmented by colour photography

5. Wave action: indicated by pattern on airphoto mosaic, also factor (e)

6. Intensity of water exchange: factors (a) and (e)

14. Floating materials: requires field measurements, augmented by high altitude photography

C. Water biological variables:

7. Specific fish: information available, or requires field measurements

E. Socio-economic characteristics of the area:

1. Population density: information available, or requires field survey

3. Degree of development: may be assessed from airphotos with selective field checking and updating from 1969 photography; augmented by high altitude photography

4. Water treatment and wastewater treatment: requires field measurements, although septic tile bed performance and sewage lagoon sites may be assessed and located respectively using airphotos

5. Roads: easily located on mosaics

7. Powerlines: main lines easily located on mosaics

10. Sewerages: requires field measurements; augmented by infrared colour photography

12. Forested land use: easily measured from mosaic

13. Urban land use: measurable from mosaic, requires specific field checking, augmented by colour photography

14. Industrial land use: measurable from mosaic, requires specific field checking.

Therefore, the interpretation of the physical factors, in combination with the information readily visible on the mosaics provides complete data for 14 of the above parameters, and partial data for 7, out of a total of 32. It is also pertinent that, of the other parameters, information is available either completely or in part for 10 of these.

5. SPECIALIZED STUDIES REQUIRED IN REMOTE SENSING:

The investigations required to accurately register the terrain and water characteristics to provide useful data for the physical disciplines are listed in Table 3. The type of photography required is shown in capital letters beside each investigation. It is assumed, unless otherwise indicated, that for colour photography only a haze filter is required, and that for infrared colour photography, either a Wratten 8 below a flying height of 5,000 feet, or a Wratten 12 above this height, is required.

It is pertinent to indicate the relationship of the remote sensing studies proposed to those investigations required for the other disciplines involved, as shown in Table 2.

Table 2 will provide the reader with an opportunity to assess the contribution of remote sensing to the other disciplines, although the individual studies for the other disciplines are not detailed in this paper due to the space limitations.

6. CONCLUSIONS AND RECOMMENDATIONS:

(a) Remote sensing has the greatest value and impact when contributing input for a variety of physically oriented disciplines.

TABLE 2

Remote Sensing Study No.	Biology (Total-3)	Fisheries (Total-4)	Limnology (Total-7)	Pollution (Total-4)	Wildlife (Total-3)
A1	1	1	5	3	3
A2	1	1	4	2	0
A3	3	1	6	4	2
A4	0	3	4	2	0
A5	0	0	3	2	2
A6	3	3	7	4	3

TABLE 3

SPECIFIC STUDIES IN REMOTE SENSING

Remote Sensing Study No.	Investigation	Application
A1	To evaluate existing photography (PANCHROMATIC), maps and reports on geology, soils, groundwater, etc.	To provide basic data on physical and cultural features within the watershed areas including environmental changes, and provide input as required for other study programs.
A2	To establish relationships between various types of agricultural usages, soil types, and groundwater flow. (INFRARED COLOUR PHOTOGRAPHY).	To attempt to reduce contribution of nutrients of water bodies by economical controls of surface water flow, thus assisting the department in regulating agricultural practices with the watershed.
A3	Monitoring seasonal contributions of septic tank and tile bed effluent to adjacent shoreline through groundwater movement, using spectral slicing in a time continuum study of environmental changes (INFRARED COLOUR PHOTOGRAPHY, 0.58 to 0.72 MICRONS IN 0.02 SLICES).	To assist in developing criteria for regulating the use of septic tank and tile bed sewage disposal systems.
A4	To establish the bathymetry of a lake from airphotos, to locate offshore bars, and to determine optimum film-filter combinations for water penetration (COLOUR PHOTOGRAPHY).	Techniques developed to aid in categorizing lakes.
A5	To relate microdrainage patterns airphotos to depth of overburden, (PANCHROMATIC AND COLOUR INFRARED PHOTOGRAPHY).	To establish the reliability of prediction of depth of overburden, and the effect of microdrainage characteristics on septic tile bed performance.
A6	Detection and mapping of stagnant water trapped in isolated bays of lakes (COLOUR AND INFRARED COLOUR PHOTOGRAPHY).	To establish land-use constraints along shorelines of stagnant areas.

(b) The existing panchromatic photography provides the basis for the acquisition of accurate and useful terrain data. The uncontrolled or semi-controlled mosaics are satisfactory for most interpretive studies, and serve as a useful media for the organized presentation of data.

(c) A multidisciplinary study for a regulatory agency on development must employ the most economical, and readily available imagery of the area involved - panchromatic photography at the present time. The training of personnel to properly evaluate and assess the input data for recreational development is simplified by the acceptance of the value of panchromatic photography for terrain evaluation.

(d) Regional evaluations, such as the selection of test lakes, and the subsequent categorization of other inland water bodies would significantly be enhanced by using high altitude infrared colour photography. This is particularly true in establishing lake inlets and outlets, one of the prime classification criteria.

(e) Time lapse photography, available over most areas only in panchromatic, is necessary to truly evaluate the changes in shoreline features (offshore vegetation, erosion, accretion, offshore bars, water level changes, etc.), land use within the watershed, and changes in drainage patterns caused by development. A program of infrared photography, taken along the shoreline on an annual basis, should be implemented to trace the encroachment of offshore vegetation, and thus the nutrient contributions to the water body. Consideration should also be given to taking this type of photography over the watershed areas on a regular time interval, determined by the rate of development.

(f) Basin physiographic features, such as surface and subsurface drainage and offshore vegetation, would be more completely assessed using infrared photography.

(g) Water physical variables, such as current effects and large quantities of floating materials would be more accurately evaluated by colour and high altitude photography respectively.

(h) Socio-economic characteristics of the region and the most appropriate auxiliary imagery are listed:

  (i) Degree of development: high altitude photography

  (ii) Sewerage outfalls: infrared colour photography

  (iii) Urban land use: colour photography

(i) The following studies should be conducted using the photography indicated to efficiently assess the full potential of remote sensing techniques in determining lake developmental capacities:

  (i) Relationships between agricultural usages, soil types and groundwater flow: infrared colour photography.

  (ii) Septic tile bed effluent contributions: infrared colour photography using spectral slicing.

  (iii) Water penetration for lake bathymetry: colour photography.

  (iv) Relationship of microdrainage patterns to depth of overburden: colour infrared photography.

  (v) Stagnant water detection: colour and infrared colour photography.

(j) The imagery obtained from the ERTS satellite program will serve as a continuous regional monitor of changing environmental conditions within inland water bodies, particularly in regard to turbidity during spring runoff or after high intensity rainfalls. The imagery to be obtained from participation in this program should be of considerable value for assessment of other significant parameters. However, only a controlled program over a large test area will establish the applications to recreational terrain evaluation.

REFERENCE:

Institute of Environmental Sciences and Engineering, University of Toronto. Environmental Analysis and Classification of Lakes in Ontario for Recreational Development. Annotated Bibliography. Part 2 of 2 parts.

PUBLISHED BY
THE DEPARTMENT OF ENERGY,
MINES AND RESOURCES,
OTTAWA.