THE ROLE OF COORDINATES IN GEOGRAPHIC INFORMATION SYSTEMS

Charles A. Chamberlain
A/Chief, Systems Development Section

J. Paul Henderson
A/Assistant Director, Data Management Group

Geodetic Survey Division
Canada Centre for Surveying
Energy Mines & Resources Canada
615 Booth Street
Ottawa Ont., K1A 0E9
613-995-4504

ABSTRACT

Coordinates provide the basis upon which all Geographic Information System (GIS) data may be spatially related. The judicious choice of a coordinate base (and its associated coordinate system) is therefore a critical step in the development of all GIS projects. As communication networks to support the increasing need for data sharing and distributed systems become more widely available, a knowledge of the coordinate systems becomes even more critical.

For about sixty years, NAD27 (North American Datum, 1927) has been the only "official" national datum for coordinates in Canada. A new national datum, NAD83 (North American Datum, 1983), is about to be introduced. The Maritime provinces use the ATS77 (Average Terrestrial System, 1977) datum. In addition many network adjustments related to these datums, such as May 76 (an adjustment of the Canadian primary framework based on Doppler control), MSA (Maritime Scientific Adjustment), and Southern Ontario 1974, are being used.

The paper outlines the differences between datums and adjustments, and identifies problems encountered in determining the relationships between some datums and adjustments. The need for a standard coordinate system for GIS development in Canada is stressed.

INTRODUCTION

Geographic Information Systems (GIS) contain two types of data, primary or positional data, and secondary or thematic data. The primary data consists of geodetic control, mapping control, and land survey systems. The secondary data includes environmental, natural resource, socio-economic, and infrastructure data (Lodwick, 1987). In this paper, we will concentrate on the primary data, in particular, the geodetic control. The geodetic control provides the spatial framework to support the secondary data. This framework must be reliable, consistent and accurate if it is to fulfill its intended role.

In order to make the correct choice in the selection of the geodetic control network, one must be cognizant of the differences between a geodetic datum and the various types of coordinates and network adjustments. This is becoming increasingly important as "networks" of GIS's are being created and data sharing among these systems is becoming a reality. This has been recognized by several authors, for example, (Walker & Usher,1988) who state: "*It was recognized that many land-related information systems were being developed on the premise that the coordinates provided under the base mapping programs would be the common interface which would allow for the sharing of information. It was further realized that if users were going to share information, they would need a consistent and reliable means of correlating the information.*"

The need for consistent geodetic control was identified by (Lachapelle, et el,1987) when they wrote: "*If incorrect assumptions are made, for example about the respective accuracies of the data being captured and the base to which this data is being added, or about the reference meridian for the plane coordinate projection zone being used, problems will ensue. The danger in not recognizing and addressing this issue is that resulting inconsistencies or errors in the data will erode user confidence in the GIS and its management will be called into question.*"

In the remainder of the paper we will define some of the above geodetic terms, describe the different datums and some adjustment systems in use within Canada, identify specific problems in the geodetic network and stress the importance of the adoption of a standard coordinate system for all GIS developments.

COORDINATE SYSTEMS AND DATUMS

The following paragraphs are taken, virtually unchanged, from (Junkins, 1988). They describe the concept of a geodetic reference system, a datum, various types of coordinates and control survey adjustment systems.

Concept of a Geodetic Reference System

A geodetic reference system (GRS) provides the accurate foundation for the organization of all information pertaining to land. When all spatial information is based on a standard GRS, it may be readily combined in any way, and by anyone, due to the universal compatibility imparted by the GRS.

There, is however, a duality in the interpretation of the term Geodetic Reference System. A research study done for the Federal Geodetic Control Committee (U.S.) states that "*A geodetic reference system can be characterized as a set of land positions whose spatial relationships are known.*" (Epstein & Duchesneau, 1984). This conflicts with the definition of the Geodetic Reference System 1980 (GRS80) adopted by the International Union of Geodesy and Geophysics (Moritz, 1980), which defines a comprehensive coordinate system for all geodetic applications. These common uses of the terminology need to be resolved in order to properly select a basis for GIS.

Geodetic Datum

In geodesy, it is traditional to represent the approximate shape of the Earth by an ellipsoid of revolution. By selecting an appropriate ellipsoid and its position as the coordinate surface for a large area, such as a country or a continent, a standard coordinate system, or datum, is created. The coordinate surface is specified ("given" as translated from the Latin "datum") by the location of the origin of the Cartesian coordinate system concentric with the ellipsoid, the directions of its axes, and the size and shape of the ellipsoid.

The realization of the datum is made by computing coordinates of points on the ground from a network of survey observations (NGS,1986). These coordinates refer to the datum, but do not define it. A datum is a coordinate system, not a series of estimated coordinates for survey points. New estimates resulting from adjustments or other computations may be computed without changing the datum.

Types of Coordinates

Ellipsoidal coordinates are only one way of expressing the position of a point. The concentric Cartesian system coincident with the principal axes of the ellipsoid is another obvious one. A third way that makes reference to the ellipsoidal surface is the great variety of map projections in common use. For a specific datum, the coordinates of a point may be expressed in any of these ways. Thus we may have

Geodetic Ellipsoidal	φ,λ NAD83
Geodetic Cartesian	X,Y,Z NAD83
Universal Transverse Mercator	N,E(UTM) NAD83

as equivalent expressions, since the indication of the datum automatically implies the relationships amongst the types of coordinates. These relationships are geometrical, and are well specified in the literature, such as Heiskanen & Moritz (1967, pg. 204) or Vanicek & Krakiwsky (1982, pp. 331-332) for ellipsoidal to Cartesian, and Thomas (1952) or Dept. of the Army (1958) for ellipsoidal to UTM.

Control Survey Adjustment Systems

The process of estimating the coordinates of physical points with respect to a specified coordinate system or datum is accomplished by computing a network of survey observations that relate the survey points to the datum and to each other. The result is a set of coordinate values for the points that constitute an adjustment system. Like the datum, it serves as a standard over a large area, such as a country or a continent. Various adjustments of the control survey network observations may be carried out on the same datum with different constraints, or the same data may be adjusted with respect to various datums.

Often, the adjustment system is given the same name as the datum to which it refers, thus leading to confusion between the two aspects. However, if only one set of adjustment values is published with respect to a particular datum, it is common and useful to imply both the datum and the adjustment by a single reference. Both NAD27 and NAD83 have this connotation. More accurately, the description of the source of coordinates for a control survey point should refer to the adjustment system, and published details of the adjustment should indicate the datum to which it refers.

CANADIAN DATUMS AND ADJUSTMENTS

The NAD27 (North American Datum, 1927) has been the "official" datum in Canada for some sixty years. It is a non-geocentric datum that uses the Clarke 1866 reference ellipsoid. It has been "realized" by the establishment of a control survey network that covers Canada. During the time that the network was being established (surveyed), various adjustments of the network (observations) have taken place. These adjustments were performed when new networks were added to the existing network, new observations were added to strengthen the existing network, or when significant improvement in coordinate accuracy could be realized by combining networks in an area. This piecemeal approach to the establishment of the network, however, has caused significant relative distortions throughout the network. Attempts to distribute, or minimize, distortions have given rise to area readjustments such as the 1974 Southern Ontario adjustment and the Maritime Scientific Adjustment.

The readjustment of the Southern Ontario network was carried out to provide the most accurate framework possible prior to planned network densification. Coordinate changes were not large (generally less than 3 metres). The Maritime Scientific Adjustment was used as the basis for an evaluation of the distortion existing in the (then) published coordinates for the geodetic networks in that part of the country.

Two specific examples of coordinate changes resulting from attempts to produce a consistent system are shown in Figures 1 and 2. Figure 1 shows displacements up to 30 metres in mapping control in the Arctic

Islands. The obvious inconsistency in the vectors clearly indicates the relative distortion in the pre-1975 coordinates. Figure 2 shows displacements in the geodetic control in Northwestern Ontario. Prior to 1971, the networks were constrained to previously determined coordinates near Thunder Bay from the original 1927 computations. Following network strengthening and readjustment (several attempts over 3 years) coordinate changes of up to 20 metres were necessary.

During the 1970's, it became apparent that the "patchwork readjustment" approach was not satisfactory. In 1979 the Maritime provinces completed a recomputation of their networks on the geocentric ATS77 datum. In 1978 the Canadian Geodetic Survey agreed to carry out an overall continental recomputation in cooperation with the United States National Geodetic Survey. This recomputation is now being finalized. The resulting NAD83 (North American Datum, 1983), will be introduced in Canada over the next few years. The NAD83 is a geocentric datum (GRS80 reference ellipsoid) that overcomes two crucial shortcomings in the NAD27; firstly, it removes the major distortions in the network coordinates and secondly, it is compatible with modern satellite positioning systems such as the Global Positioning System (GPS). Coordinate differences between NAD27 and NAD83 result from the removal of distortion and from the adoption of a geocentric reference ellipsoid. The change from a non-geocentric to a geocentric reference ellipsoid will have a major impact on the users. This change ranges from 70 metres easterly in Nfld to 120 metres westerly in the Yukon and UTM northings are changed by more than 200 metres in most of Canada.

EFFECT ON GEOGRAPHIC INFORMATION SYSTEMS

The effect of these various adjustments and datums may or may not be significant to any specific GIS. If a GIS user creates a data base with coordinates from one datum where existing coordinate distortions are inconsequential for his purposes, these various datums and adjustments are not significant. However, if a GIS user wants to combine his data with that from another GIS, it will only be possible if the datum and adjustment are the same, or insignificantly different. It is up to each GIS user to know what significantly different means to him. It is also imperative that he be aware of the basis for the coordinates that are on his GIS. The addition of new data, such as remotely sensed data from space systems (LANDSAT, SPOT, ERS - 1) will again require that one be aware of the datum to which their coordinates refer.

Before combining data from different GIS's or adding data from a new source, a user must therefore verify that the coordinate differences are insignificant. This is usually done by comparing coordinates at a number of common points and looking for shifts, scale changes and rotations. If the differences are significant, one set of coordinates must be transformed to be compatible with the other.

Coordinate Transformations

The coordinate differences may be due to a change in datum or due to coordinate distortions. If a change in datum is the cause of the coordinate differences, then a datum transformation will remove the difference. Usually, a seven parameter transformation (eg. Vanicek & Krakiwsky, 1982) is used. This is a straight forward mathematical process if the parameters defining the change in datum are known. For the NAD27 these parameters are not known and they must be estimated. The Geodetic Survey of Canada is presently estimating the datum change parameters between the NAD27 and NAD83 datums. Once the change in datum is removed, there may also be coordinate distortions with which to contend.

Transformations due to coordinate distortions may be performed by one of several methods depending upon the size of the area, the magnitude of the distortions and the precision required. Junkins (1988) lists several possible solutions, including simple coordinate shifts, Helmert transformations, affine transformations and more elaborate distortion modeling.

The Geodetic Survey of Canada is creating a Canadian transformation between NAD27 and NAD83 based on the coordinates of the primary framework. This transformation includes both a datum change (as identified above) and a distortion model. There will however, be no attempt to model the various adjustments that have taken place over the years. The interested reader is again referred to (Junkins, 1988) for details.

CONCLUSION

Coordinates play a significant role in geographic information systems. To maintain compatibility the GIS user must be intimately aware of the source and accuracy of the coordinates used in his GIS. It is only with this knowledge that he can combine data from other GIS's and data from outside sources. When there are significant differences he must be aware of the different methods and tools available to him to transform his coordinates into one system.

The only consistent and accurate national datum available for use with GIS in Canada is the NAD83 datum. It is a modern system that is compatible with new positioning systems and remote sensing systems. It is also relatively free from coordinate distortion. However, as the use of GIS's expand, there will undoubtedly be users who will require a higher relative accuracy. In the short term, transforming existing data to NAD83 will cause extra cost and inconvenience, however, in the long term it will prove to be a valuable investment and provide universal compatibility to all GIS users and suppliers.

REFERENCES

Department of the Army, 1958, *Universal Transverse Mercator Grid*, Technical Manual TM5-241-8

Epstein, E. and T. Duchesneau, 1984. The Use and Value of a Geodetic Reference System. University of Maine at Orono.

Heiskanen, W. and H. Moritz,1967, *Physical Geodesy*, W. H. Freeman & Company, San Francisco.

Junkins, D. 1988. Transforming to NAD83, *Papers for the CISM Seminars on the NAD'83 Redefinition in Canada and the Impact on Users,* Ottawa, Ontario.

Lachapelle,G., E. Krakiwsky and E. Kennedy, 1987. Impact of the Navstar / GPS on Geographic Information Systems, *Proceedings of Geomatics Applied to Municipal Management, CISM, Montreal Branch,* Montreal, Quebec.

Lodwick, G. and M. Feuchtwanger. 1987. *Land-Related Information Systems*, University of Calgary Surveying Engineering Reports, Number 10010, Calgary, Alberta.

Lodwick, G. and M. Feuchtwanger. 1987. *Land-Related Information Systems*, University of Calgary Surveying Engineering Reports, Number 10010, Calgary, Alberta.

Moritz, H. 1980, Geodetic Reference System 1980, The Geodesist's Handbook, *Bulletin Geodesique*, Vol. 54, No. 3.

National Geodetic Survey, 1986. *Geodetic Glossary*, Charting and Geodetic Services, National Ocean Service, National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

Vanicek, P. and E. Krakiwsky, 1982. *Geodesy - The Concepts*, North - Holland Publishing Company

Walker, W.J. and W.D. Usher. 1988. The Impact of the Redefinition and Readjustment of the North American Datum on Land-Related Information Systems in Alberta, *Papers for the CISM Seminars on the NAD'83 Redefinition in Canada and the Impact on Users*, Ottawa, Ontario.

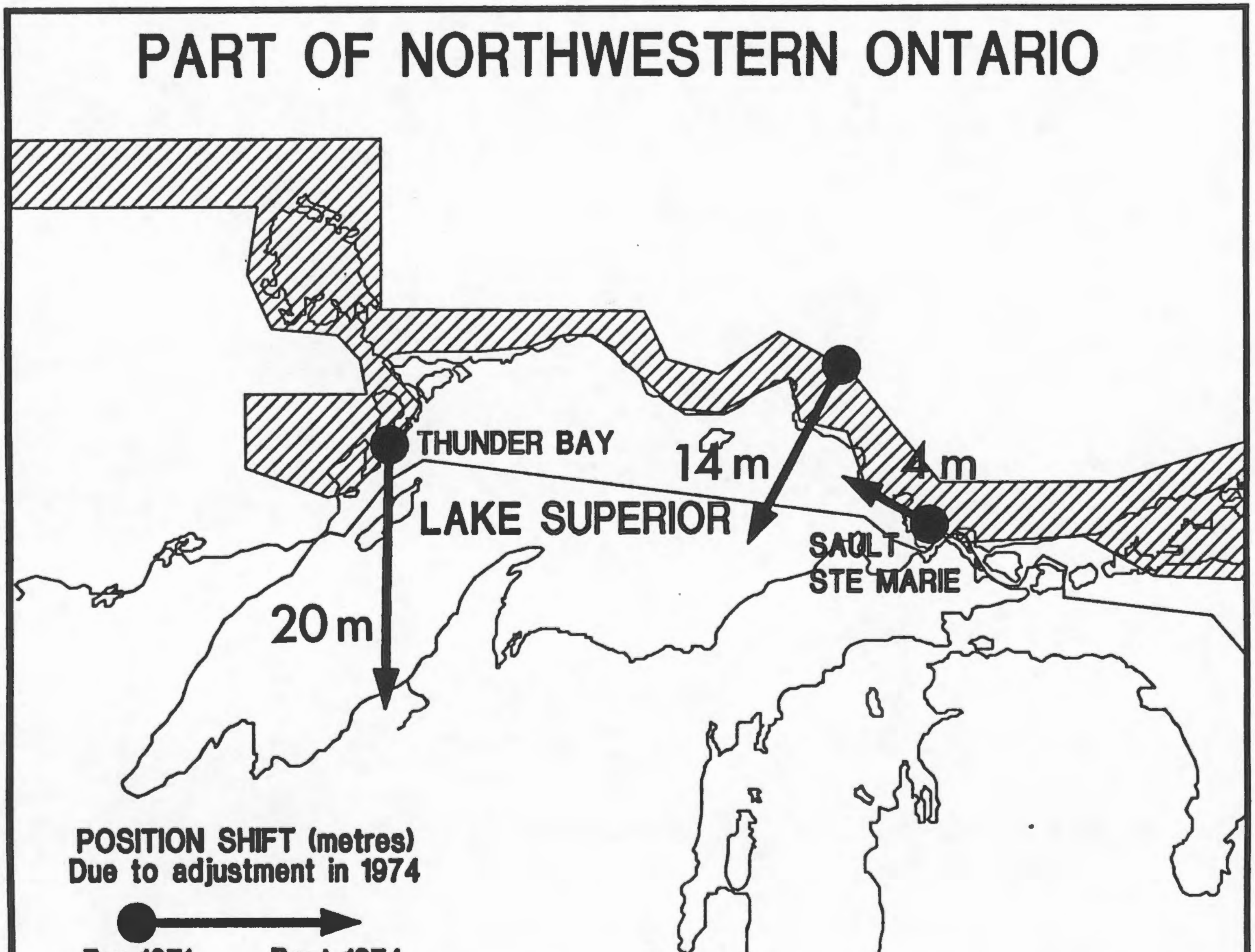
PART OF NORTHWESTERN ONTARIO
THUNDER BAY
LAKE SUPERIOR
20 m
14 m
4 m
SAULT
STE MARIE
POSITION SHIFT (metres)
Due to adjustment in 1974
Pre 1971
Post 1974

FIGURE 2

FIGURE 1: ARCTIC ISLANDS ADJUSTMENT