SATELLITE ALTIMETRY APPLICATIONS FOR MARINE GRAVITY

Petr Vaníček
David Wells
Eugene Derenyi
Alfred Kleusberg
Rostam Yazdani
Terry Arsenault
Nick Christou
John Mantha
Spiros Pagiatakis

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

This is the final report for DSS Research Contract No. OSC84-00472 entitled "Satellite Altimetry Applications for Marine Gravity". Under this contract we were supposed to do the following:

- Derive the geoid from satellite altimetry data.
- Derive the geoid from marine gravity data.
- Compare satellite altimetry data with marine gravity data.
- Combine satellite altimetry data with marine gravimetric data.
- Compare in detail for two test areas (Hudson Bay and Labrador Sea) the gravity fields obtained from satellite altimetry and marine gravimetry.

The overall (unstated) goal has been to investigate how satellite altimetry can help in offshore geophysical studies conducted by means of gravity. Since (marine) gravity anomalies have been used for this purpose almost exclusively in the past, it was envisaged that the sea level surface obtained from satellite altimetry should be used to check the observed gravity anomalies or even predict anomalies where none exist.

Leaving aside the difference between altimetrically derived sea level and the geoid, i.e., neglecting for the moment the sea surface topography (see section 2.4), the conversion of the geoid into gravity anomalies poses some formidable problems. Determination of point gravity anomalies from the geoid represents a classical ill-posed problem [Molodenskij et al., 1960; Coleman and Mather, 1976]—cf. section 4.3. Only some areal averages of gravity anomalies are determinable, and the smaller the averaged areas the less stable the inversion becomes. The conversion of satellite altimetry, neglecting the sea surface topography, has been done already by several people, most notably by Rapp [1983].

The idea occurred to us then that at least checking the observed marine gravity anomalies could be done in the "geoid space" rather than the "gravity space." It is clearly much more simple to compute the geoidal height N from Δg than the other way around. The only problem with the classical Stokes's convolution used for this transformation is that Δg should be available all over the world thus ruling out the use of pure marine gravity data.

To remove this problem, one may carry out the Stokes convolution of Δg with the Stokes

kernel only over a spherical cap of a limited radius ψ—this operation we call here the *truncated Stokes integration*. The price one pays for doing this is that the truncated Stokes integration does not yield the full geoidal height N (above a selected reference ellipsoid) but only a partial contribution $\delta N(\psi)$—termed here, in the absence of a better term, the *truncated geoid*. All that remains to be done is to derive such an integration kernel which, used in a convolution with (the altimetrically derived) geoidal height N, also gives $\delta N(\psi)$.

This has turned out to be rather a tricky mathematical problem which would not have been solvable within the means of this contract. Fortunately, we have had at our disposal other means, namely the NSERC Strategic Grant "Applications of Marine Geodesy", under the auspicies of which we could carry out the requisite theoretical investigation. At the end we have succeeded: The result is shown in section 5.3.

Intuitively, the above described approach should not only allow one to test the goodness of the observed marine gravity anomalies within a spherical cap of radius ψ but also provide an alternative tool for geophysical studies. The *truncated geoid* $\delta N(\psi)$ by itself should give a unique view of the density distribution within the earth. Additionally, the changes of $\delta N(\psi)$ in response to varying ψ should offer yet another perspective. The uses of the geoid (instead of or in addition to gravity anomalies) for density distribution studies have increased in the recent past, particularly after it was realized that the different wavelengths of the geoid may have special meaning [Zlotnicki, 1984]. It is hoped that the technique we are here suggesting will prove equally fruitful, thus providing additional benefit unforseen in the contract.

One additional problem we have encountered while working on this project concerns the observed marine gravity data. The marine gravity data file has been improved by its custodian, the Earth Physics Branch of Energy Mines and Resources Canada in Ottawa, several times during the life of this project. The most up-to-date file was made available to us at the beginning of this year. It necessitated a recomputation of the already compiled gravimetric geoid for the eastern part of Canada (see section 3.3).

Contours and spot heights are the generally accepted means for the visual representation of spatial data sets, covering large areas. Although contour maps can represent three-dimensional figures with a high degree of fidelity, the surface portrayed is not easily perceivable by the viewer. Geographers have long recognized this problem and introduced colour, shading, and hachuring to

enhance the topography. There is substantial evidence that such techniques can also be applied successfully to geoscience data (see, for example, Haxby et al. [1983]). Therefore, provision of expertise and software for digital image analysis and for display techniques using colour graphics forms part of this contract.

2. BASIC DEFINITIONS

In this report we deal with three classes of earth surfaces:

- various surfaces related to sea level;
- reference surfaces within the ocean (isobaric, isopycnal, or others); and
- the geoid, a particular equipotential surface of the earth's gravity field.

In this chapter we define these surfaces and describe how they are related. We also introduce the two main data sources used in this report to determine these surfaces: satellite altimetry for sea level, and gravity anomalies for the geoid. These data will be discussed in detail in Chapter 3.

2.1 Sea Level

Sea level and its variations is a very complex phenomenon. A discussion of sea level proceeds most conveniently by first considering a ficticious scenario.

Consider the ocean as an undisturbed **homogeneous** fluid. With this statement we have obliterated the moon, sun, and planets; demanded that the earth's atmosphere also be a static fluid devoid of changing weather features; and mandated the elimination of other perturbing influences as well. We have left the earth's crust as it is, and the earth is still rotating. We have constructed an *idealized sea level* which is static, and spatially influenced only by variations in the earth's gravity field [Stewart, 1985; p. 261].

The actual sea level will be affected both temporally and spatially by many influences. The sea level surface with none of these perturbations removed is the *instantaneous sea level surface*.

The "long term" average of the temporal variations in sea level at a point is a *mean sea level* for that point. "Long term" may mean one month, or one year, or longer, and the corresponding monthly mean sea levels, annual mean sea levels, and so on, will in general be different. Monthly means at the same point may vary typically by 0.5 m, and annual means by 0.1 m.

Some perturbations may have static components, so that temporal averaging at a point will not necessarily recover our idealized sea level for that point.

To obtain a *mean sea level surface*, we must establish mean sea level at a network of points.

4

2.2 The geoid

The geoid is one of the equipotential surfaces of the earth's gravity field. As such, it is defined mathematically by the following prescription:

$$W(r) = const., \qquad\qquad (2.1)$$

where W stands for the earth's gravity potential. The value of the constant is selected so that this equipotential surface represents the mean sea level (surface) as well as possible. The idea is due to J.K.F. Gauss (1777-1855) who was the first to recognize the geoid as a mathematical figure of the earth.

To select the "right" equipotential surface, clearly, is not a simple task, particularly because the mean sea level changes with both the time (epoch) to which it is referred and the period of time over which the sea level has been averaged. But even if the mean sea level were not varying with time, how would one select the appropriate constant in eqn. (2.1) to make sure that the equipotential surface really approximates the mean sea level? With all the modern solutions, where satellite determined long wavelength features make the backbone of the solution, it is a matter of a proper selection of scale. The scale is calibrated against terrestrially determined heights and becomes part of the solution.

2.3 Sea Surface Topography

Sea surface topography is the departure of the actual sea surface from the geoid, or from the idealized sea surface defined in section 2.1.

We can use several "actual sea surfaces", each of which will have a different sea surface topography associated with it. If we include all wavelengths of the spatial variations, then at any instant we have an *instantaneous sea surface topography*.

Instantaneous sea surface topography includes features with amplitudes of up to (see for example, Chelton and Enfield [1985]):

- several tens of metres from waves,
- several metres from tides,
- one metre from strong geostrophic currents,
- several decimetres from atmospheric pressure variations,
- several decimetres from wind-driven piling up effects along coastlines,
- several decimetres from other ocean dynamic effects.

The features with the strongest temporal variations are those due to waves, tides, and atmospheric variations. If we use mean sea level at a network of points as our "actual sea surface," then most of these temporal variations (over the period used to compute the mean) will be eliminated, and we will have a *quasi-static sea surface topography*, which contains only semi-permanent perturbations, mainly due to ocean dynamics effects. Two other effects must also be taken into account:

- Significant local and regional secular variations in sea level due to crustal motion may exist.
- Eustatic changes (the advance or retreat of ice sheets) involve a global change in mean sea level of a few centimetres per century.

We will see that satellite altimetry involves sampling over a "footprint" of a few kilometres in diameter. Thus the altimetrically-measured "actual sea surface" suppresses short wavelength features (mainly those due to waves) resulting in an associated *medium and long wavelength instantaneous sea surface topography*. Since temporal variations in sea surface topography remain, and altimetry does not sample the entire ocean surface simultaneously, we obtain only topographic point values at different instants, rather than a coherent sea surface topography.

The temporal variations due to tides and atmospheric pressure variations can be modelled using a global tidal model, (usually that of Schwiderski [1980]), and meteorological records, respectively. An altimetrically-measured "actual sea surface" which has been corrected for tides and atmospheric pressure variations, results in a coherent sea surface topography which approximates the quasi-static sea surface topography.

To reduce this further to obtain the idealized sea surface which we seek is a more difficult task. We must model the topography due to ocean dynamic effects. This is discussed in the next section.

2.4 Dynamic Topography

Sea surface topography features due to ocean dynamics effects are closely linked to seawater density variations.

Seawater density varies with temperature (t), pressure (p), and salinity (s). *Specific volume* (α), is defined as the reciprocal of density, in units of volume per unit mass. The most general expression for specific volume is

$$\alpha_{(s,t,p)} = \alpha_{(35\,\%_o,\,0^\circ C,\,p)} + \delta_s + \delta_t + \delta_{s,t} + \delta_{s,p} + \delta_{t,p} + \delta_{s,t,p}$$

where

$\alpha_{(s,t,p)}$ is the specific volume of a sample of water of salinity (s), temperature (t) and pressure (p),

$\alpha_{(35 ‰, 0°C, p)}$ is the specific volume of an arbitrary "standard" sea water of salinity s = 35 ‰, temperature t = 0°C, and pressure p at the depth of the sample.

and the δ terms represent the contributions of various combinations of s, t, and p.

Sea water density variations are expressed in terms of the *anomaly of specific volume* ($\delta\alpha$)

$$\delta\alpha \equiv \alpha_{(s,t,p)} - \alpha_{(35 ‰, 0°C, p)} = \Delta_{s,t} + \delta_{s,p} + \delta_{t,p} + \delta_{s,t,p}$$

where

$\Delta_{s,t} = \delta_s + \delta_t + \delta_{s,t}$ is called the *thermosteric anomaly*.

The term $\alpha_{(35 ‰, 0°C, p)}$ expresses mainly the effect of pressure on specific volume. The anomaly of specific volume ($\delta\alpha$) represents the excess of the actual specific volume of the sea water at any point in the ocean, over the specific volume of the "standard" sea water, and is the sum of six anomaly terms. The thermosteric anomaly ($\Delta_{s,t}$) accounts for most of the effect of salinity and temperature, disregarding pressure. In water depths less than 1000 m, $\Delta_{s,t}$ is the major component of $\delta\alpha$. The term $\delta_{s,p}$ accounts for most of the combined effect of salinity and pressure; the term $\delta_{t,p}$ for the combined effect of temperature and pressure; and $\delta_{s,t,p}$ is usually so small that it can usually be neglected.

The *dynamic topography* (also called *geopotential thickness anomaly* and *dynamic height anomaly*) between two constant pressure surfaces p = p_1 and p = p_2 in the ocean is defined as the integral of the specific volume anomaly over the pressure surface difference [Levitus, 1982]

$$h_d = \int_{p_1}^{p_2} \delta\alpha \, dp \ .$$

A fundamental assumption in dynamic topography is that ocean dynamics effects disappear on a "zero" surface at which there is no motion of the water. The integral is taken from this "zero" surface to some other surface. Identifying this **surface of no motion** is the major problem of dynamic topography. Since this is critical to our application, it is worth reviewing in some detail.

Several methods have been used to determine the surface of no motion, differing in the characteristic that is assumed to be associated with no motion [Fomin, 1964]:

- the layer of minimum dissolved oxygen (Dietrich's method),
- the isopycnal (constant density) layer showing minimum distortion (Parr's method),
- computation of the vertical distribution of current velocity field based on salinity distribution (Hidaka's method), and
- the isobaric (constant pressure) surface for which the gradient of dynamic depth is a minimum (Defant's method).

It is clear that selecting one method to use in preference over the others is merely to select the assumption which appears the least subjective. Defant's method is the most widely used, and is used here. It should be emphasized that whatever method is used, the dynamic topography obtained is relative to the assumed level of no motion, and is not "absolute."

The assumption that a surface of no motion exists is based on the belief that in the deep ocean, sea water velocities are negligible. Recent work indicates this may not be a valid assumption [Olbers and Willebrand, 1984; Wunsch, 1978]. On the other hand, averaged over several months and tens or hundreds of kilometres, it may be that any reference level below 1000 m may yield adequate dynamic topography results for many applications, particularly over short distances [Pond and Pickard, 1983]. Over longer distances it is likely that a surface possessing small or zero velocity will have a complex structure which is not consistent with selecting a single isobaric surface as the level of no motion [Fomin, 1964]. In the absence of compelling evidence to the contrary, however, we have assumed that the level of no motion is represented by a single isobaric surface.

2.5 Gravity Anomaly

Gravity anomaly is the difference between actual gravity on the geoid and normal gravity at the "corresponding" point on the reference ellipsoid. Depending on how the actual gravity on the

geoid is obtained from the observed gravity, i.e., which formula is used for the vertical gradient of gravity, we get different kinds of gravity anomalies. Here we will use only the free-air anomaly typified by the use of free-air gravity gradient:

$$\partial g / \partial h = - 0.3086 \text{ mGal/m} . \tag{2.2}$$

If the gravity is observed on the geoid, then all kinds of anomalies are equivalent.

Depending on what gravity network the observed values of g are referred to, we have the derived gravity anomalies referred to the same network. The anomalies used here are referred to the International Gravity Standardization Network of 1971 (IGSN-71).

Lastly, the gravity anomaly can be related to different reference ellipsoids through the different existing normal gravity formulae. In our case, the Geodetic Reference System 1980 (GRS-80) is used: the normal gravity on the reference ellipsoid, γ_0, is computed from the International Gravity Formula 1980:

$$\gamma_0 = 978.0327(1 + 0.005\ 302\ 4 \sin^2\phi - 0.000\ 005\ 8 \sin^2 2\phi) \text{ Gal}. \tag{2.3}$$

2.6 Satellite Altimetry

Satellite altimetry techniques for observing the ocean surface topography have been one of the fundamental tools of the NASA Ocean Processes Program for more than 15 years [Tapley et al., 1982]. Currently, two major satellite altimetry data sets exist. These are from the GEOS-3 mission (1975-1978), and from the SEASAT mission which was operational for three and a half months in 1978. SEASAT and GEOS-3 raw data are available to researchers from NOAA/NESDIS [U.S. NRC, 1985].

The principle behind satellite altimetry measurements of the ocean surface lies in the fact that the satellite serves as a stable platform from which the radar altimeter measures the distance of the spacecraft above the instantaneous ocean surface, based on the travel time of short-pulse microwave signals. This distance measurement is an average height over the area covered by the radar footprint, the size of which depends upon the sea state.

The geometry of the altimeter measurement is shown in Figure 2.1 and it relates:

  h   the geodetic height of the spacecraft;
  a   the measured altitude above the ocean surface;

SEASAT

instrument corrections   altimeter   Orbit

atmospheric corrections α

                                                 h

geophysical corrections   Land

ocean surface
                                    ζ_s
Geoid ζ

bottom topography N

                                                          Reference
                                                          Ellipsoid
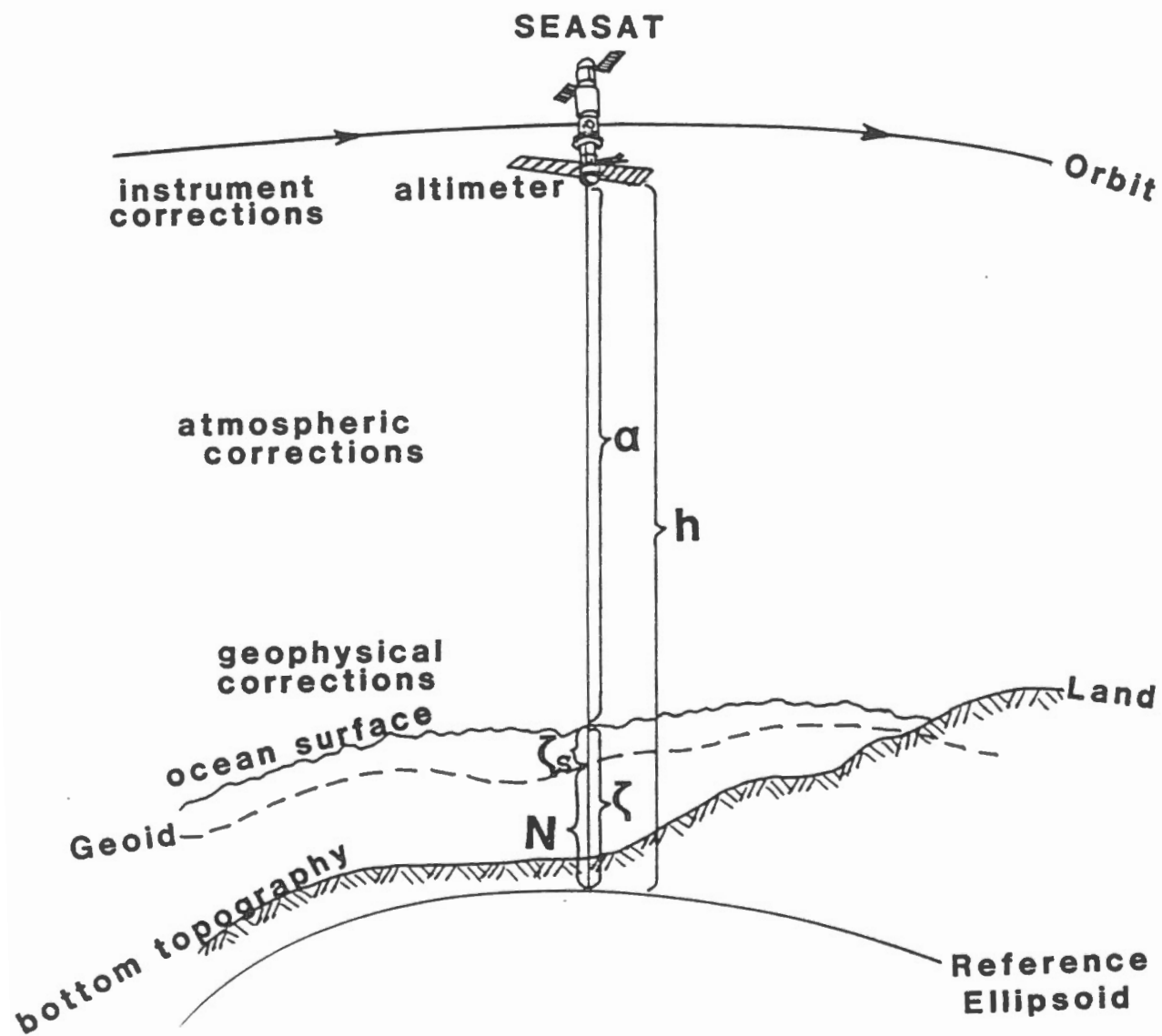
Fig. 2.1: Geometry of altimeter measurements.

10

N   the geoidal height;

ζ the height of the sea surface above the reference ellipsoid; and

ζs the dynamic sea surface height;

through the expression

$$\zeta s = \zeta - N = (h - a) - N .$$

In this equation, it is assumed that the altimeter measurement (a) has already been corrected for a number of effects, such as (see Lorell et al. [1980]):

(a) instrument corrections (instrument delays, altimeter off-nadir pointing errors, antenna centre of mass correction, residual biases, etc.);

(b) propagation medium effects (atmospheric path length corrections due to tropospheric dry and wet component delay and ionospheric delay);

(c) geophysical reductions (corrections due to solid earth and ocean tides, inverse barometer response of the ocean, sea state, wind pile up, etc.).

Orbital errors (such as those due to model gravity field, atmospheric drag, solar radiation pressure) are assumed to have already been accounted for in the geodetic height (h) of the spacecraft.

3. DATA

In this chapter we describe the characteristics of the data sets used for the work reported in this report. There are four types of data:

- the marine gravity data which is to be improved and extrapolated using
- the SEASAT satellite altimetry data, and
- a recently-computed gravimetric geoid based on land and marine data;
- sea surface topography data which is used to reduce the altimetry to an approximation of the idealized sea surface defined in section 2.1.

3.1 Marine Gravity Data

We used 318 579 marine gravity data points supplied by the (then) Earth Physics Branch (EPB) of Energy, Mines and Resources Canada [Hearty, 1986]. Figure 3.1 shows the distribution of this data.

Gravity mapping of waters off eastern Canada began in 1964, and has been carried out mainly by the Atlantic Geoscience Centre (AGC), a division of the Geological Survey of Canada, with participation by EPB [EPB, 1986; pp 1-3].

Sea surface gravity measurements are subject to errors from the following sources [Talwani, 1970; Dehlinger, 1978]:

- navigational inaccuracies
- Eötvös effect errors
- mislevelling of the gyro-stabilized platform
- datum errors
- system calibration errors
- scale factor inaccuracies
- vibration and thermal stress noise.

The quality of marine gravity data depends on the position and velocity determination techniques used, the weather, and other factors. A conservative quality factor should represent the
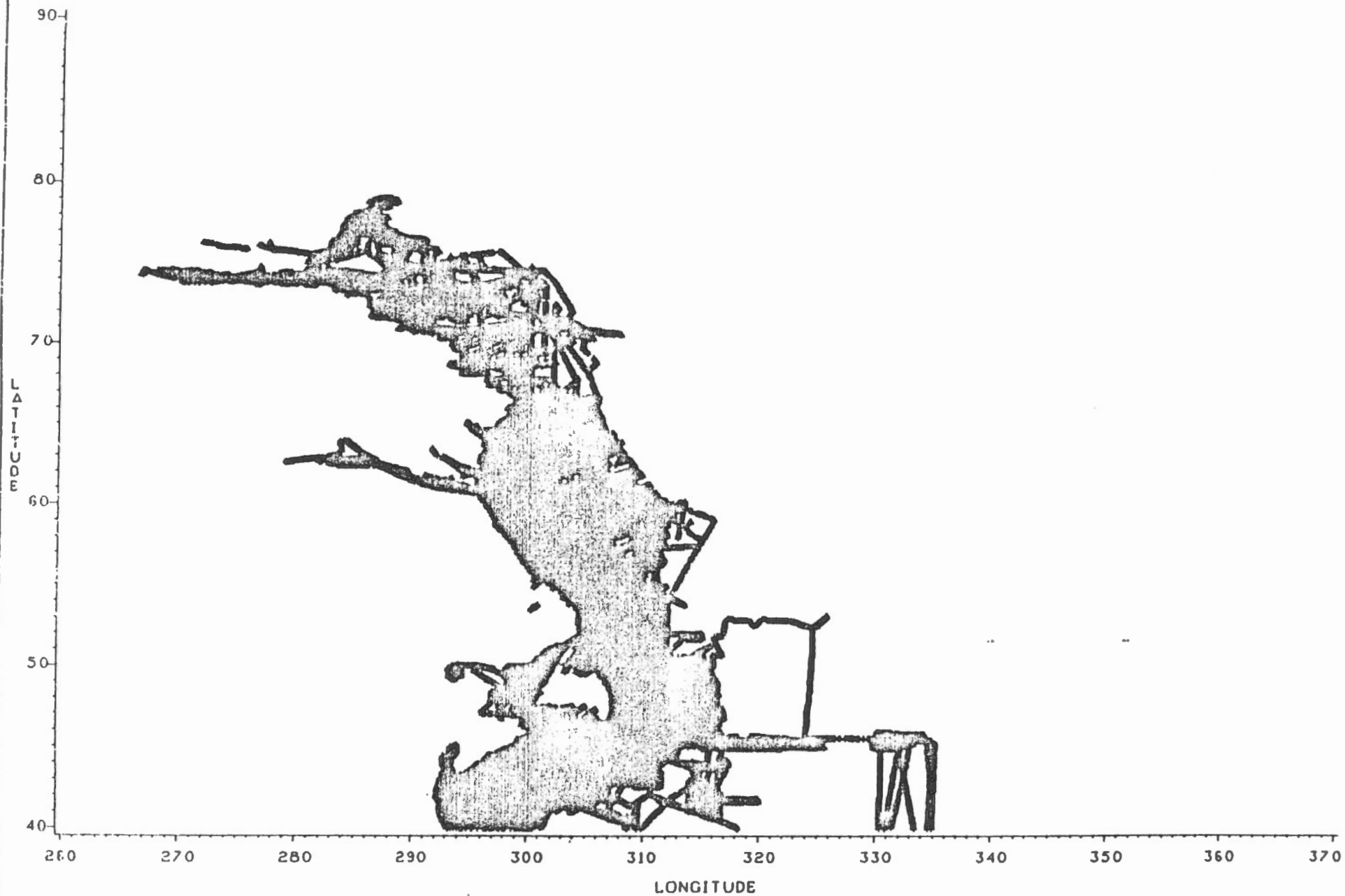
Fig. 3.1: Distribution of Marine Gravity

13

worst case variability within the section of data to which it refers. The quality factor codes used for the marine gravity data available to us are:

Quality Factor	Gravity Variability
0	unknown
1	± 0.01 mGal
2	± 0.05 mGal
3	± 0.1 mGal
4	± 0.5 mGal
5	± 1 mGal
6	± 3 mGal
7	± 5 mGal
8	± 10 mGal
9	>± 10 mGal

The eastern Canada offshore gravity data we have used has an overall accuracy of 2.5 mGals. The internal consistency varies from ±1 to ±5 mGals [EPB, 1986; p.47].

Homogeneous gravity data are produced by adjustment of the gravity field data. Field data may have one or more of the following problems, which cause problems in adjustment:

- Positions of the main survey lines and checks lines are not identical.
- Gravimeter drift calibration is difficult to model due to instrument failure, mechanical shock, or other discontinuities between port calibration visits.
- Due to the conservative approach taken, a lot of good data is rejected along with the bad.
- The quality factor is only a very rough indicator of data quality.

3.2   Satellite Altimetry

Three different sets of SEASAT satellite altimetry data have been acquired. Two of them are in the form of adjusted altimetry tracks, the third one being the "raw" SEASAT altimetry data set. The first of the adjusted altimetry data sets was obtained from the Department of Geodetic Science of The Ohio State University (OSU). The second became available from NASA's Goddard Space Flight Center (GSFC). The SEASAT "raw" observation records were obtained from the U.S. National Geodetic Survey (NGS). All data sets cover the geographical area between 35°N to 72°N and 260°E to 350°E.

The reason that we have acquired altimetry data only from the SEASAT mission is because the coverage, adjustment procedures, and accuracy results are superior to those obtained through the GEOS-3 mission. Moreover, we have chosen to acquire two adjusted SEASAT altimetry data sets because the adjustment procedures followed at OSU and GSFC are somewhat different.

In the following subsections, a more detailed description of these data sets is given.

It should be mentioned here that the techniques we have developed under the terms of this contract will be equally applicable to altimetry data collected by other missions. It may be a very interesting exercise to acquire some GEOSAT data from the U.S. military and perform similar analyses with these data. The TOPEX mission, now likely to take place in the 1990s, will be another excellent source of data to be exploited.

3.2.1 OSU altimetry data set

The OSU adjusted SEASAT altimetry data were provided on tape containing the information described in Appendix A of Rapp's [1982b] analysis of SEASAT altimeter data. The adjusted sea surface heights contained on the tape are referenced to the Geodetic Reference System 1980 (GRS80) with a = 6 378 137 m and f = 1/298.257.

The basic discussion of the OSU SEASAT altimeter data processing can be found in Rowlands [1981]. In his study, Rowlands carried out a primary adjustment (primary arcs) using a global distribution of altimetry tracks and fixing one long arc. The resultant RMS cross-over discrepancy after this adjustment was ±28 cm [Rapp, 1982a]. Next, these primary arcs were held fixed and four regional, as well as several special purpose, adjustments (local arcs) were carried out, the details of which can be found in Rowlands [1981] and Rapp [1982b]. The final adjusted data records form a data base, an edited subset of which was obtained for our study area. This subset consists of primary and local arcs in the North Atlantic and Hudson Bay areas and has approximately 142 000 data records in all. Figure 3.2 is a plot of the subset received and shows the distribution of the adjusted SEASAT satellite tracks.

Although the data set we received was supposed to be clean of "blunders" in the data records, we discovered a few "bad" records and have further edited the data set. The bad records corresponded either to out-of-range values of adjusted sea surface heights (ASSH), or to data points where positional information was wrong (outside the area of interest), or to a few data points where the tide information was unavailable even though the points were lying in the middle of the Atlantic Ocean. Thus several screening procedures were implemented before the data set was ready
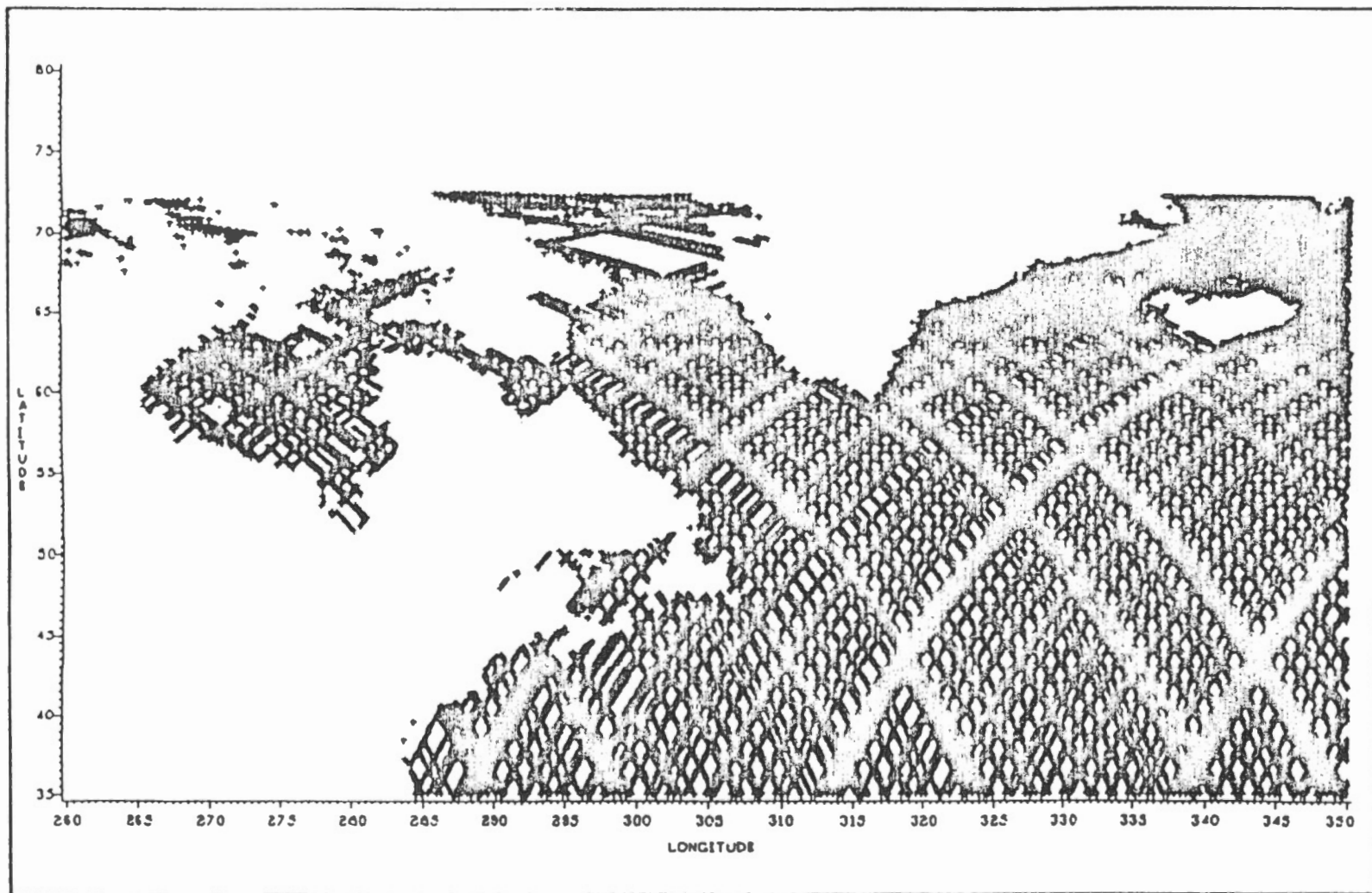
Fig. 3.2: SEASAT Subsatellite tracks in N. Atlantic; OSU Adjusted Altimetry Data Files

to be used.

At this point it should be mentioned that the OSU tape also contained information about the standard deviations of the individual ASSHs. The standard deviations in the file range from 1 cm to a maximum of 25 cm, which is an artificial cutoff value.

3.2.2 GSFC altimetry data set

The GSFC adjusted SEASAT altimetry data set obtained contains the following information: revolution number, modified Julian day, fraction of day, latitude, longitude, and adjusted sea surface height. The data set is in the form of several files, each file covering a diamond-shaped geographical area. These "diamonds" are shown in Figure 3.3 One of the diamonds pertinent to our study area was not included in the received tape (i.e., diamond #66).

The basic discussion of the SEASAT altimeter data processing performed at GSFC can be found in Marsh et al. [1984] and in Marsh and Martin [1982]. The global SEASAT data set was divided into 65 diamond-shaped regions, the maximum length of an altimeter ground track within a region being approximately 2000 km. For each "diamond," three control passes were selected from the three-day repeat era: one ascending and two descending passes which crossed the ascending one. The ascending control pass was held fixed and only the bias terms for the two descending tracks were adjusted.

A cross-over adjustment was carried out for each "diamond"; within each region the orbit error could be represented by a linear trend. The "diamonds" were overlapped by two to five degrees of arc in order to provide continuity. After the adjustment process, the RMS difference of the cross-over points was found to be below 20 cm [Marsh et al., 1984].

It should be pointed out, however, that even though all of the ASSHs are referred to the same ellipsoid (i.e., GRS80 ellipsoid), the common ASSHs in the overlapping "diamond" areas experience discrepancies sometimes of the order of 0.5 m. The reason for this is that, although the regional ASSHs have been computed with a high degree of precision, the absolute orientation of each diamond-shaped area with respect to the centre of mass of the earth is different. This orientation problem led us to treat the whole data set with caution. More details regarding this point will be found in section 3.4.4.

The approximate total number of records contained in the merged (i.e., all "diamonds" received) data set is 250 000. As in the case of the OSU data set, further screening procedures were employed to eliminate the few bad records from the received GSFC data set. These bad
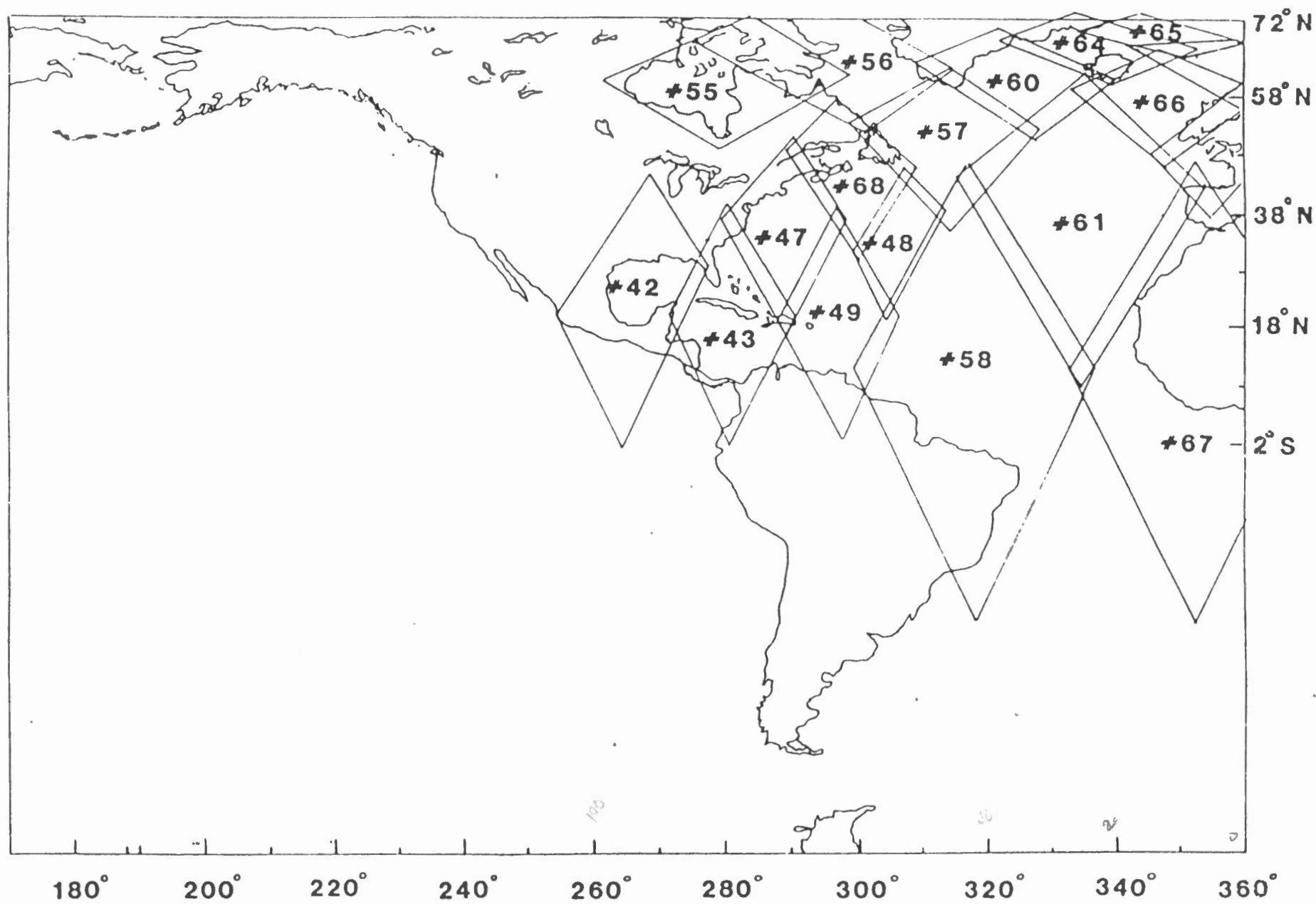
Fig. 3.3: Locations of diamond-shaped areas in N. Atlantic; GSFC-NASA Adjusted Altimetry Data Files

records were essentially erroneous ASSH values.

3.2.3 NGS altimeter data set

Three tapes were received from the U.S. National Geodetic Survey, containing the original SEASAT altimeter data as described in the Geophysical Data Record (GDR) in Lorell et al. [1980], covering the complete lifetime of SEASAT. The precision of this data set is described in Tapley et al. [1982].

Since we had two adjusted altimetry data sets (OSU and GSFC), it was superfluous to perform the same tasks as the two above institutes had done. In addition, time constraints prevented us from any further consideration of processing this "raw" data set to obtain ASSHs.

3.2.4 Comparisons and discussion on the altimetry data sets

From the preceding discussion it should be apparent that the GSFC and the OSU data sets are bound to have differences. These differences arise as a result of the different geographical split of the global SEASAT data set and the subsequent regional cross-over adjustments performed. It is obvious that different lengths of satellite orbital arcs have been used to minimize the cross-over discrepancies in the two data sets, thus leading to different linear trend removal and different accuracy assessments in terms of the post-adjustment RMS of these discrepancies. Intuitively we can expect that the GSFC data set has a higher, but regional, accuracy than the OSU data sets.

Another difference between the two adjusted data sets is that different atmospheric corrections have been applied. In addition, the modelling needed to be performed in order to attain a common orientation for all diamond-shaped GSFC subsets will affect the ASSHs directly. Some criteria for this task should be developed and assessed. While at NASA the atmospheric corrections supplied in the GDR have been used, OSU derived corrections from other sources [Rapp, 1982b]. However, this difference is less significant than the orbital bias modelling for the cross-over adjustment.

It also seems to us that different earth gravity models have been used for the processing of the data. Earth gravity model PGS-S3 has been used by OSU, while the PGS-S4 model was used for the processing at GSFC. If our understanding is correct, this might have an effect on the radial orbital error analysis, which will be primarily reflected in the cross-over discrepancies.

Another source of possible differences between these two data sets is the selection of the cross-over points utilised in the linear trend removal, and the selection of the fixed arc for the adjustment. This fact points to the possible existence of a bias term between the two data sets.

At present, it seems to us that no other apparent differences exist in terms of the set of the applied corrections. We have not performed a quantitative analysis of the differences between the two data sets, but it is in our future plans to look at this problem in more detail.

3.2.5 Gridding through bilinear surface

A grid of mean sea surface heights was calculated for both sets of adjusted SEASAT data. The function of the grid was to present the ASSH values at regularly spaced intervals of latitude and longitude, thus allowing for easy and fast graphical displays using either conventional surface representations (contour maps, three-dimensional plots) or colour graphics on a digital image analysis system, such as the ARIES II.

To determine a value representative of the ASSH at each grid intersection, the data in the surrounding area was fit using a weighted least-squares process to a surface-modelling function, and a local Cartesian coordinate system. The zero-order coefficient estimated from the process was then used as the evaluated ASSH at the grid location.

We have tried several surface-modelling functions starting with a bi-quadratic function in latitude and longitude. We found that in areas of poor data coverage and distribution, this choice often led to a singular normal equation matrix for the surface fit. We tried to lower the degree of the fitted surface, but still there were grid locations that presented the same problem as above. The final choice, after several trials that were carried out specifically in the Hudson Bay area, was a simple bilinear surface modelling function [Christou and Yazdani, 1986]

$$P_i(x_i, y_i) = \alpha_0 + \alpha_1 x_i + \alpha_2 y_i + \alpha_3 x_i y_i , \quad . \tag{3.1}$$

where

$$x_i = k_1(\phi_i - \phi_0) \tag{3.2}$$

$$y_i = k_2(\lambda_i - \lambda_0) \quad . \tag{3.3}$$

k_1, k_2: being numerical constants to convert from angular to length units;

$\alpha_0, \alpha_1, \alpha_2, \alpha_3$: the unknown surface coefficients to be estimated;

ϕ_o, λ_o: latitude and longitude of the jk grid point;

ϕ_i, λ_i: latitude and longitude of the altimetry data points in the surrounding area of the jk grid point;

x_i, y_i: the local Cartesian coordinates of the data points; and

$P_i(x_i, y_i)$: the associated ASSH values of the data points.

We can then write the "observation equations" [Vaníček et al., 1972]:

$$w_i \, P_i(x_i, y_i) = w_i \, \alpha_j^T \, \Phi_j(x_i, y_i) = w_i \, h_i, \qquad i=1,2,...,n \ . \tag{3.4}$$

Denoting $\Phi_j(x_i, y_i)$ by Φ, we obtain the "normal equations" in the form [Vaníček et al., 1972]:

$$\Phi^T \, w \, \Phi \, \alpha = \Phi^T \, w \, h \ , \tag{3.5}$$

where

$$w = \text{diagonal } (w_i)$$

and the solution is

$$\alpha = (\Phi^T \, w \, \Phi)^{-1} \, \Phi^T \, w \, h \ . \tag{3.6}$$

Then, α_o is the sought interpolated value:

$$h_o \equiv \alpha_o = ASSH_o \ . \tag{3.7}$$

Since in our approach a minimum of eight data points is required, there is a redundancy of at least 4, and we are able to evaluate the confidence limits of h_o as well.

The grid spacing chosen was 1/6 of a degree in both latitude and longitude. A rectangular window was selected about each grid point having a half-width of 0.5 degrees (i.e., all data within the window were used to calculate the predicted value). There is the option of varying the window size (i.e., using a larger window) to ensure that grid values could be calculated in the regions of sparse data coverage. As a weighting scheme, the associated ASSH standard deviations contained in the input file were used with the OSU data set. It was decided to use a ±15 cm standard deviation [Marsh, 1985] for all ASSH in the GSFC data set. It should be noted here that this gridding procedure leads to a smooth surface.

Due to the reasons explained in the previous sections with regard to the GSFC data set (i.e., no common orientation of all diamond-shaped areas), the merged and subsequently gridded data set is presenting problems. Thus this data set was excluded from all subsequent calculations and

displays. Figure 3.4 shows the computed grid of mean sea surface on a 10' x 10' grid using the OSU adjusted SEASAT altimetry data. In the sequel, this grid will be labelled the "OSU-grid" for reference purposes.

The "OSU-grid" and its asociated standard deviation were subsequently converted in a suitable raster-format and transferred to a digital image analysis system ARIES II for graphical colour displays and further image analysis processing technique.

Finally, it should be mentioned that the gridding algorithm provided gridded altimetry values over land in waste regions. To alleviate this defect, a special masking procedure was elaborated. The World Data Bank II (WDBII) coastline [Gorny, 1977] was utilised to mask the gridding artifacts over land. This coastline data set was transferred to the ARIES II system where it was generalized to the same pixel size as the altimetry (10' x 10' grid cells) and then transferred back to screen the altimetric grid shown in Figure 3.4.

3.3 UNB Gravimetric Geoid

Recently, we compiled a fairly detailed Canadian geoid [Vaníček et al., 1986] for the Geodetic Survey of Canada. It covers the region of $\phi \in \langle 42°, 70°\rangle$ and $\lambda \in \langle 220°, 317°\rangle$ with the exception of areas where there is an insufficient gravity data coverage. The eastern part of this gravimetric geoid is shown in Figure 3.5. The technique we have used for the compilation relies on two kinds of data:

(a) Low-order harmonic coefficients obtained from satellite orbit analysis. A (20, 20) field produced by Goddard Space Flight Center (GEM9) was chosen.

(b) Terrestrial gravity data, consisting of (i) point gravity anomalies; (ii) 5' x 5' mean gravity anomalies; and (iii) 1° x 1° mean gravity anomalies.

(i) **Point gravity anomalies.**

The point gravity anomalies were obtained from the Earth Physics Branch, Ottawa, and the Atlantic Geoscience Centre, Dartmouth. This data file contained 558 565 records.

(ii) **5' x 5' mean gravity anomalies.**

The original 5' x 5' mean gravity anomalies were supplied by the Geodetic Survey of Canada. It was decided to use the updated marine point gravity anomalies to compute new 5' x 5' mean gravity anomalies for the east coast of Canada and replace the original data in the file. The

69.6667
69.0000
68.3333
67.6667
67.0000
66.3333
65.6667
65.0000
64.3333
63.6667
63.0000
62.3333
61.6667
61.0000
60.3333
59.6667
59.0000
58.3333
57.6667
57.0000
56.3333
55.6667
55.0000
54.3333
53.6667
53.0000
52.3333
51.6667
51.0000
50.3333
49.6667
49.0000
48.3333
47.6667
47.0000
46.3333
45.6667
45.0000
44.3333
43.6667
43.0000
42.3333
41.6667
41.0000
40.3333
39.6667
39.0000
38.3333
37.6667
37.0000
36.3333
35.6667
35.0000

260   270   280   290   300   310   320   330   340   350

23

Fig. 3.4:   Mean Sea Surface from SEASAT Adjusted Altimetry in N. Atlantic.   Grid Interval 10' x 10'.
Contour Interval 2m.   OSU DATA FILES.   Sea Surface Heights are refered to GRS 80 Ellipsoid
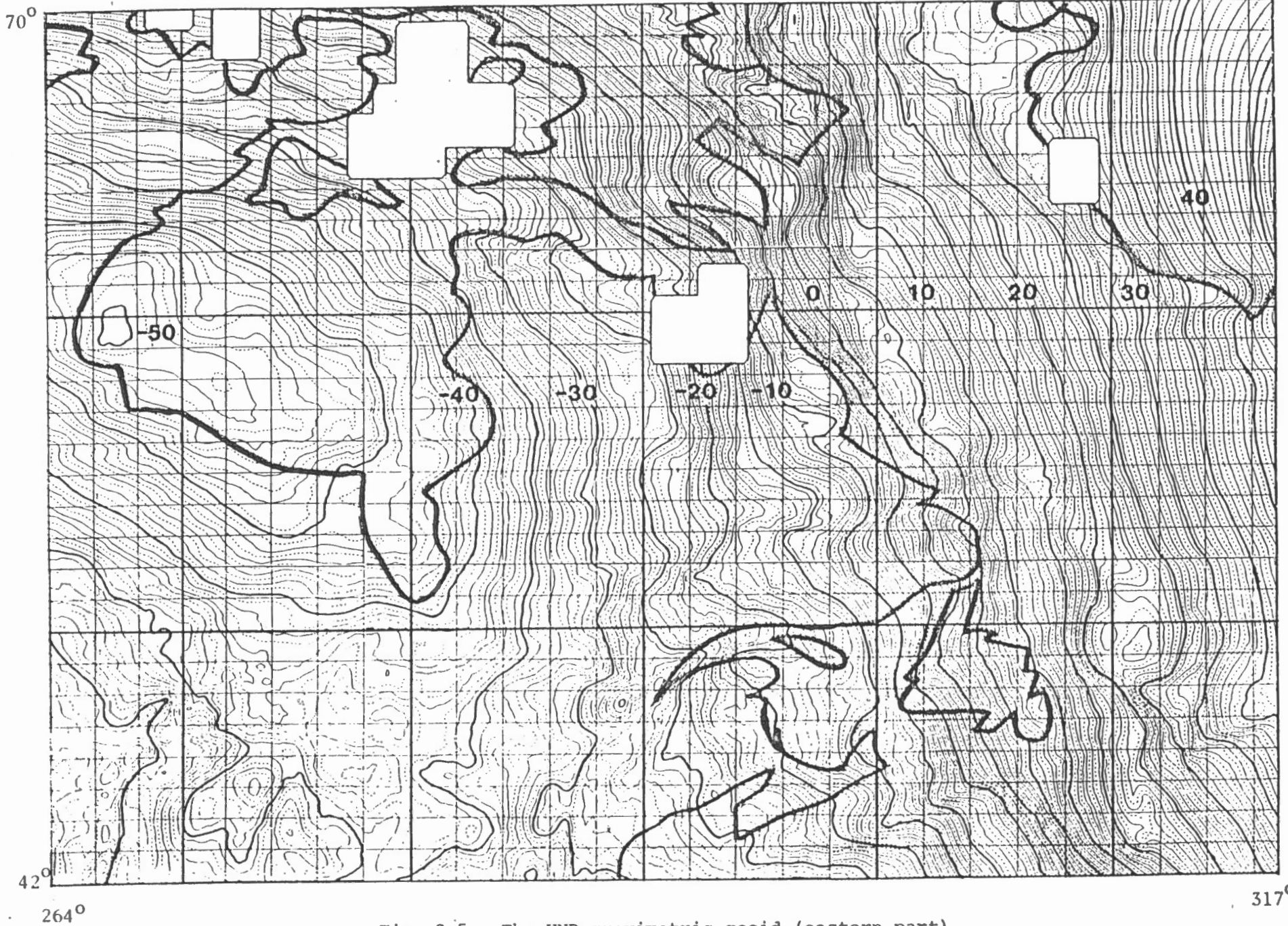
Fig. 3.5. The UNB gravimetric geoid (eastern part).

computation formulae are:

$$\overline{\Delta g} = (\sum_{i=1}^{n} \Delta g_i)/n , \qquad (3.8)$$

$$\overline{h} = (\sum_{i=1}^{n} h_i)/n , \qquad (3.9)$$

where Δg_i is the ith gravity anomaly in the 5' x 5' cell, and h_i is the ith height in the cell. The standard deviation of the new mean anomalies are obtained as geometrical averages:

$$\sigma_{\overline{\Delta g}} = \sqrt{(\sum_{i=1}^{n} \sigma_{\Delta gi}^2)/n} . \qquad (3.10)$$

If the 5' x 5' mean gravity anomaly could not be computed directly because of sparse data coverage, a weighted mean recommended by Kassim [1980] was utilized:

$$\overline{\Delta g_p} = (\sum_{i=1}^{n} W_i \Delta g_i) / (\sum_{i=1}^{n} W_i), \qquad (3.11)$$

$$\overline{h_p} = (\sum_{i=1}^{n} W_i h_i) / (\sum_{i=1}^{n} W_i), \qquad (3.12)$$

where $\overline{\Delta g_p}$ is the predicted mean gravity anomaly, Δg_i are gravity anomalies of the immediate surrounding cells, W_i is a weight equal to $1/s_i$ where s_i is the distance of the ith cell from the prediction cell, n is the number of cells used in the prediction, and $\overline{h_p}$ is the predicted mean height. The standard deviation of the $\overline{\Delta g_p}$ is:

$$\sigma_{\overline{\Delta gp}} = \sqrt{(\sum_{i=1}^{n} W_i^2 \sigma_{\Delta g}^2) / (\sum_{i=1}^{n} W_i^2)} \qquad (3.13)$$

where $\sigma_{\bar{\Delta}_{gp}}$ is the standard deviation of the predicted gravity anomaly. If of the eight immediately surrounding cells more than two are empty, then the area utilised for the prediction is expanded to include a second layer of surrounding cells. The prediction fails if there are more than four empty cells in the area utilised for prediction. In this case, the missing 5' x 5' mean gravity anomaly has been *replaced during the geoid computation* by the corresponding 1° x 1° mean gravity anomaly.

(iii) **1° x 1° mean gravity anomalies.**

These data were supplied by OSU, also named "January 1983 1° x 1° Mean Gravity Anomaly".

The low-order harmonics are taken as defining a (higher order) reference spheroid. A modified Stokes's approach is then used to compute geoidal heights above this spheroid from terrestrial gravity. The resulting relatively short radius of integration (6 degrees) allows us to compute the geoidal height on a rather dense grid; in this case a 10' x 10' grid.

It is difficult to assess objectively the accuracy of our gravimetric solution. GEM 9 is deemed to be accurate to about 1.75 m (1 sigma), the accuracy of the high frequency terrestrial/marine gravity contribution is estimated internally to be within ⟨5, 50⟩ cm. Comparisons with "independent" solutions show the following statistics (in the sense of "standard minus UNB"):

Rapp 180° x 180° [Rapp, 1983]	$\mu = 101$ cm,	$\sigma = 106$ cm
Wenzel 180° x 180° [Wenzel, 1985]	$\mu = 62$ cm,	$\sigma = 95$ cm
212 Doppler points in Canada	$\mu = 86$ cm,	$\sigma = 170$ cm

σ denotes the RMS with respect to μ.

Colour plots of the differences between the UNB geoid and the Rapp/Wenzel geoids are provided as external appendices.

It seems to us that the low frequency part of the UNB geoid agrees quite well with Wenzel's and Rapp's solutions as well as with the SEASAT (see section 3.4.2). This should not be surprising since Wenzel's solution relies heavily on GEM9 and so does the OSU adjusted SEASAT altimetry—at least in the final analysis.

It should be noted that the total UNB gravimetric geoid is referred to the reference ellipsoid defined by GRS-80. Since it is computed on the 10' x 10' grid, the UNB gravimetric geoid can be merged with the gridded altimetry, if such a merger is deemed desirable.

3.4 Sea Surface Topography

3.4.1 Levitus's dynamic topography data

Since a definite answer to the problem of "absolute" dynamic topography could not be given, we resorted to the estimation of relative dynamic topography using Levitus's specific volume anomalies [Levitus, 1982]. The data we have acquired is in the form of dynamic topography and it was supplied to us by J. Marsh [1985] of the GSFC.

This data set is based on the Levitus computed anomaly of specific volume for 33 reference levels of the global ocean. The product of the processing performed at GSFC is a tape that contains the dynamic topography for 33 reference levels (i.e., the integrated anomaly of specific volume at 33 different reference depths) on a 1° x 1° grid covering the whole globe.

We have selected reference level #25 (depth 1875.0 m) as the level of no motion, or better, as the pressure surface (1875 dbars) with respect to which we will obtain the relative dynamic topography. The reasons for that selection are underlying the following arguments:

(a) It may be claimed that an assumed level (or layer) of no motion generally lies in deep water.

(b) There is evidence of a strong circulation in the upper North Atlantic waters (down to approximately 1000 m) as well as in the deep North Atlantic waters (below 2000 m, approximately) [Pond and Pickard, 1983].

(c) Montgomery [1969] suggests as an optimum reference pressure for a world steric sea level map the 2000 db isobar.

(d) We have computed difference maps of dynamic topography between several reference levels below 1000 m depth and no significant dynamic topography differences were found (see Figures 3.6 to 3.8).

(e) We wanted to have as large a geographical area as possible covered, where dynamic topography estimates are available to reduce our SEASAT altimetry.

Going for deeper reference levels, it would disclose the possibility of having adequate area of overlap between SEASAT altimetry and Levitus's dynamic topography as close to the coastline as possible.

This dynamic topography data set referenced to the 1875 dbar surface has been processed to become available on the same grid spacing as the altimetry gridded data, i.e., on a 10' by 10' grid. For that purpose, linear interpolation in two directions, of the 1° by 1° grid of Levitus's dynamic topography, was applied. The 10' by 10' interpolated grid is shown in Figure 3.9. A simple

Fig. 3.6: Difference of Dynamic Topography between levels #18 and #17. (Difference = level 18 - level 17). LEVITUS DATA FILES. Contour Interval is 2.5 dynamic centimetres. (Level 18 : 950 dbars, level 17 : 850 d bars)
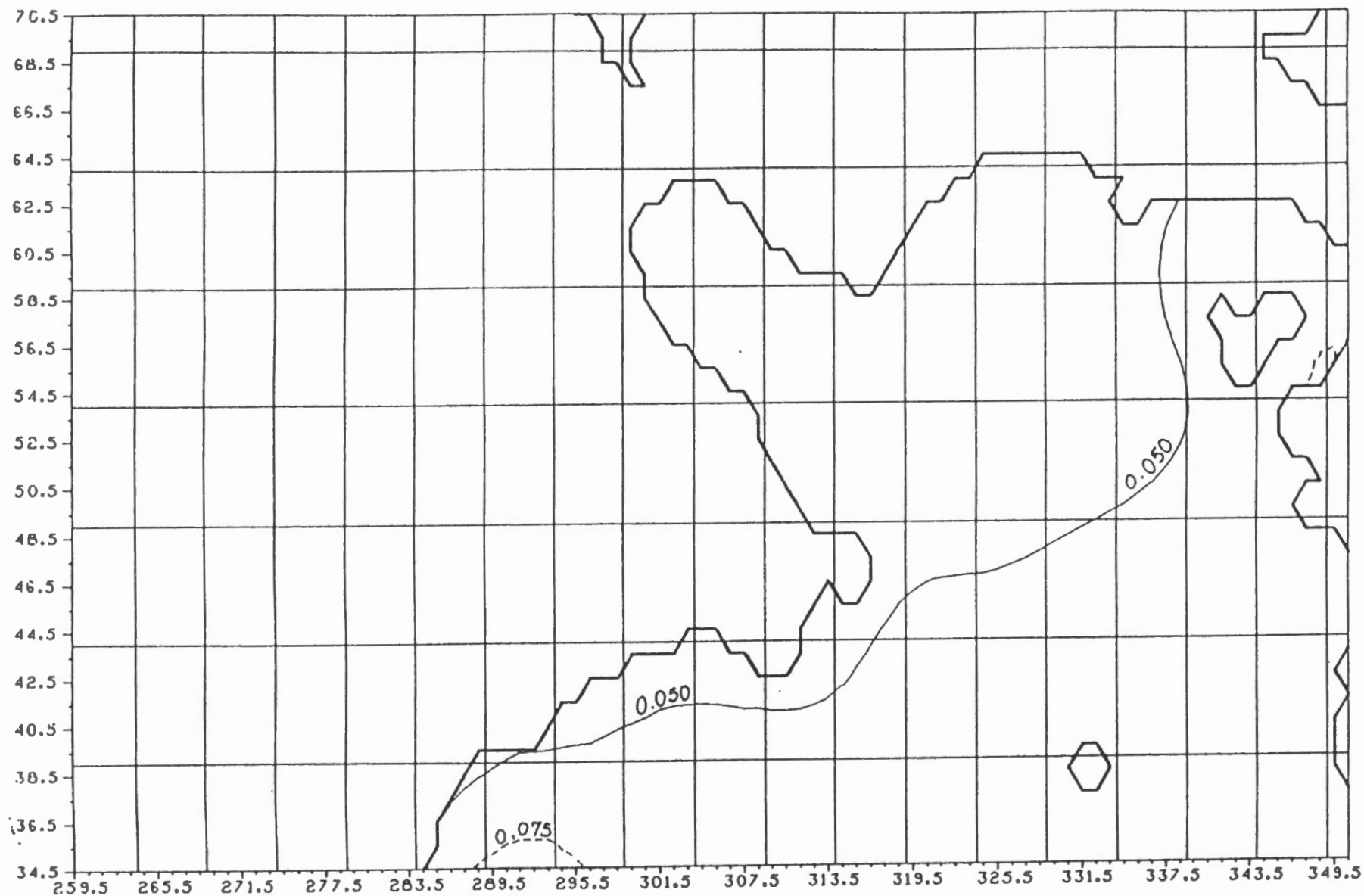
Fig. 3.7: Difference of Dynamic Topography between levels #19 and #18. (Difference = level 19 - level 18) LEVITUS DATA FILES. Contour Interval is 2.5 dynamic centimetres. (level 19 : 1050 dbars, level 18 : 950 dbars)
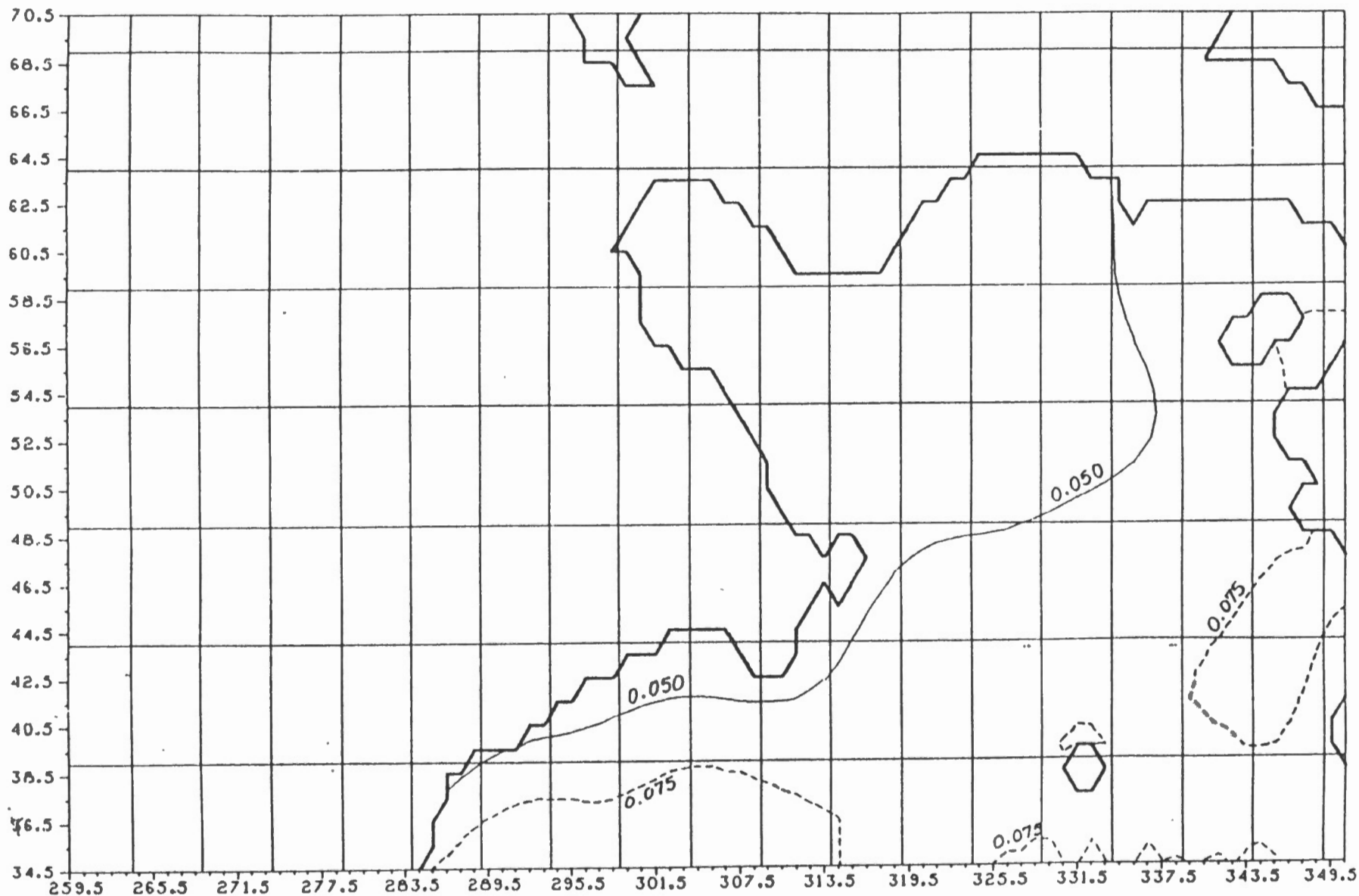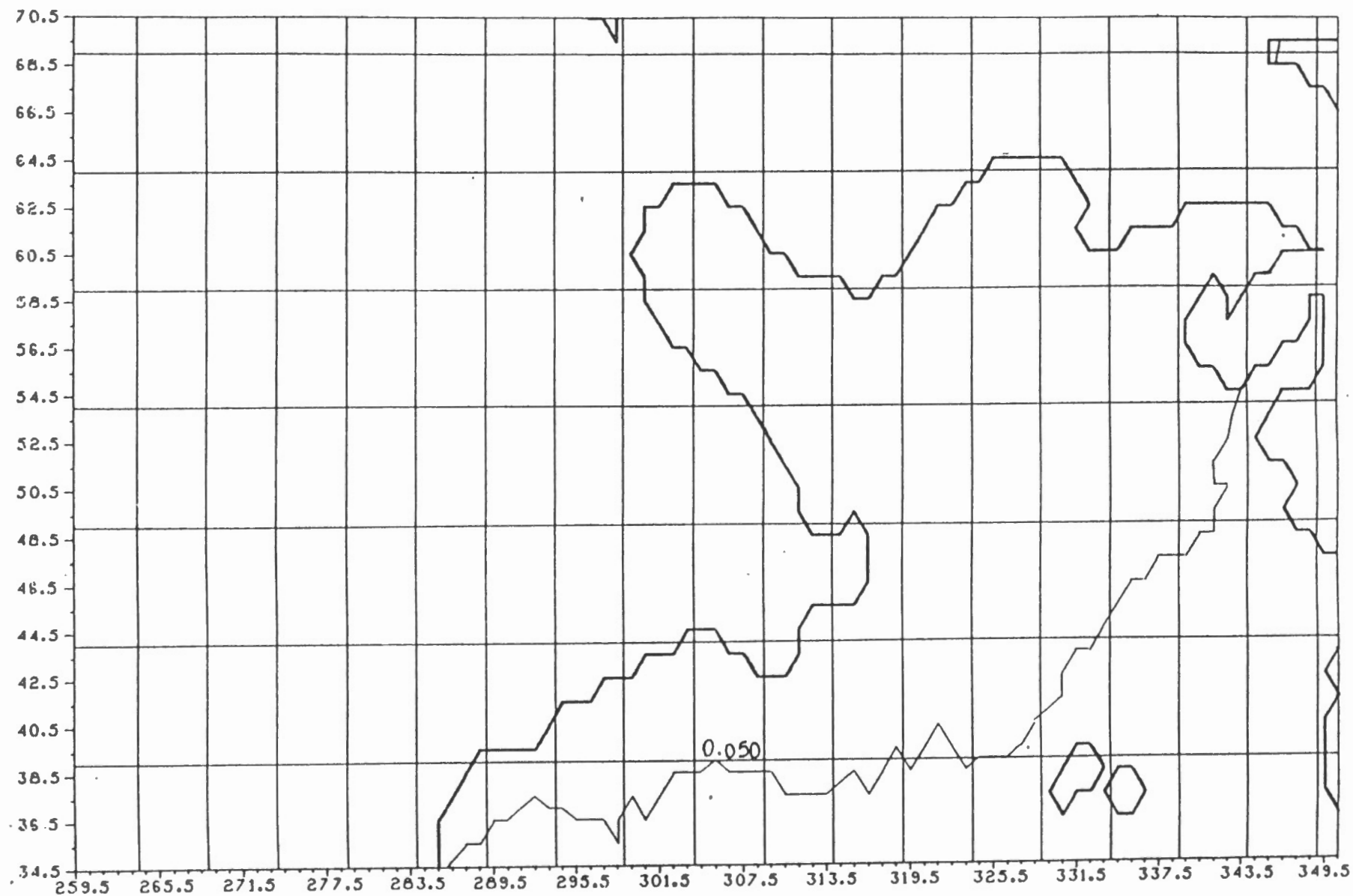
Fig. 3.8: Difference fo Dynamic Topography between levels #22 and #21. (Difference = level 22 - level 21) LEVITUS DATA FILES. Contour Interval is 2.5 dynamic centimetres (level 22 : 1350 dbars, level 21 : 1250 dbars)
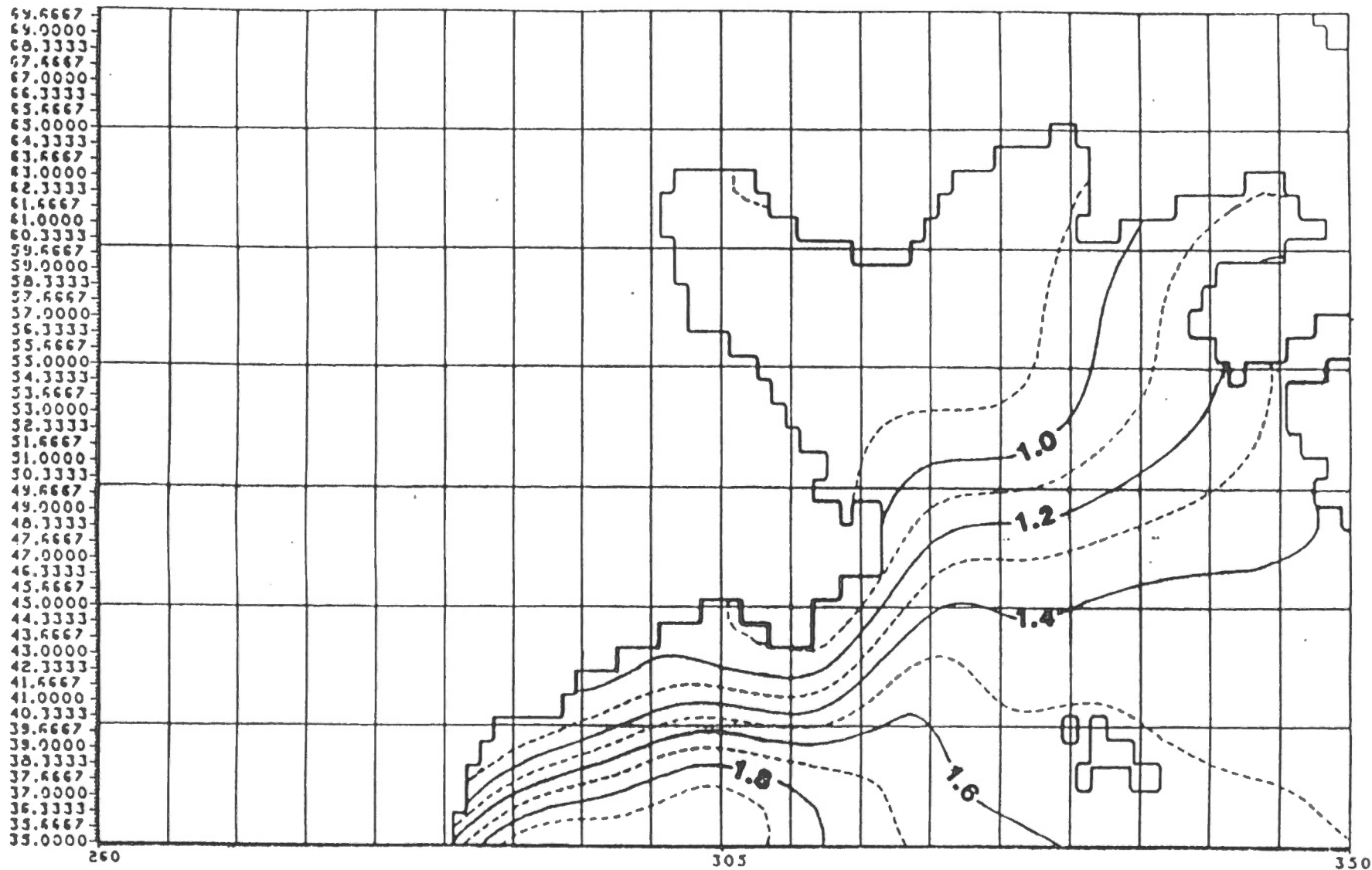
Fig. 3.9: Levitus Ocean Dynamics Topography in N. Atlantic relative to the 1875.0 dbars pressure surface. Grid interval 10' x 10'. Contour Interval 0.1 dyn. metres.

FORTRAN 77 program was written for that purpose. First, the latitudinal grid lines were interpolated on 10' by 10' spacing. The resultant data set was then passed through the same program and the longitudinal grid lines were interpolated on 10' by 10' intervals. Fast sorting routines available on the IBM 370 computer facilitated the cost effectiveness and simplicity of the developed program. The above choice of reference level #25 is providing us with dynamic topography estimates in most of the North Atlantic Ocean. With regard to the Hudson Bay area, no dynamic topography estimates will be considered.

3.4.2   Altimetry minus UNB geoid

Another way of determining the SST is to subtract the geoid from the sea surface determined by the altimeter. This difference is shown on Figures 3.10 through 3.18 for the OSU SEASAT data and the UNB gravimetric geoid. A colour plot of these differences is provided as an external appendix to this report. The mean difference for the area covered is -0.114 m. The rms with respect to this mean is 0.827 m. Both the mean and the rms are strongly influenced by areas of unreliable altimetry data in the northern part of Hudson Bay, the Hudson Strait, around Baffin Island, and along the coast of Greenland. Figure 3.19 depicts those areas where the estimated standard deviation of the gridded altimetry exceeds a value of 0.25 m. Figure 3.20 shows the standard deviations of the difference between the OSU SEASAT altimetry and UNB gravimetric geoid estimated from the individual standard deviations of altimetry and geoidal heights according to the law of covariance propagation.

A second statistical analysis of the differences between the altimetry and UNB geoid was done by excluding the areas above $\phi=62°$ thus virtually eliminating all areas with high SEASAT standard deviations. The mean difference of this reduced area of coverage was computed as +0.036 m. The rms with respect to this mean was 0.534 m. This agreement in the mean certainly is somewhat unexpected. Based on the error budget for the geoidal computation (see section 3.3), a mean difference of up to one metre could have been produced by the errors in the long wavelength features of the GEM9 spheroid alone. This, on the other hand, shows a weak point in this procedure: any long wavelength SST signal in the difference between the altimetry and gravimetric geoid may be distorted or overpowered by long wavelength computation errors.

We recognize a long wavelength feature in the area of Davis Strait to Labrador Sea and further south. It can be roughly described by a half wavelength of 10 degrees and an amplitude of 0.5

Figure 3.10: Differences between SEASAT altimetry and UNB gravimetric geoid

SAS

34
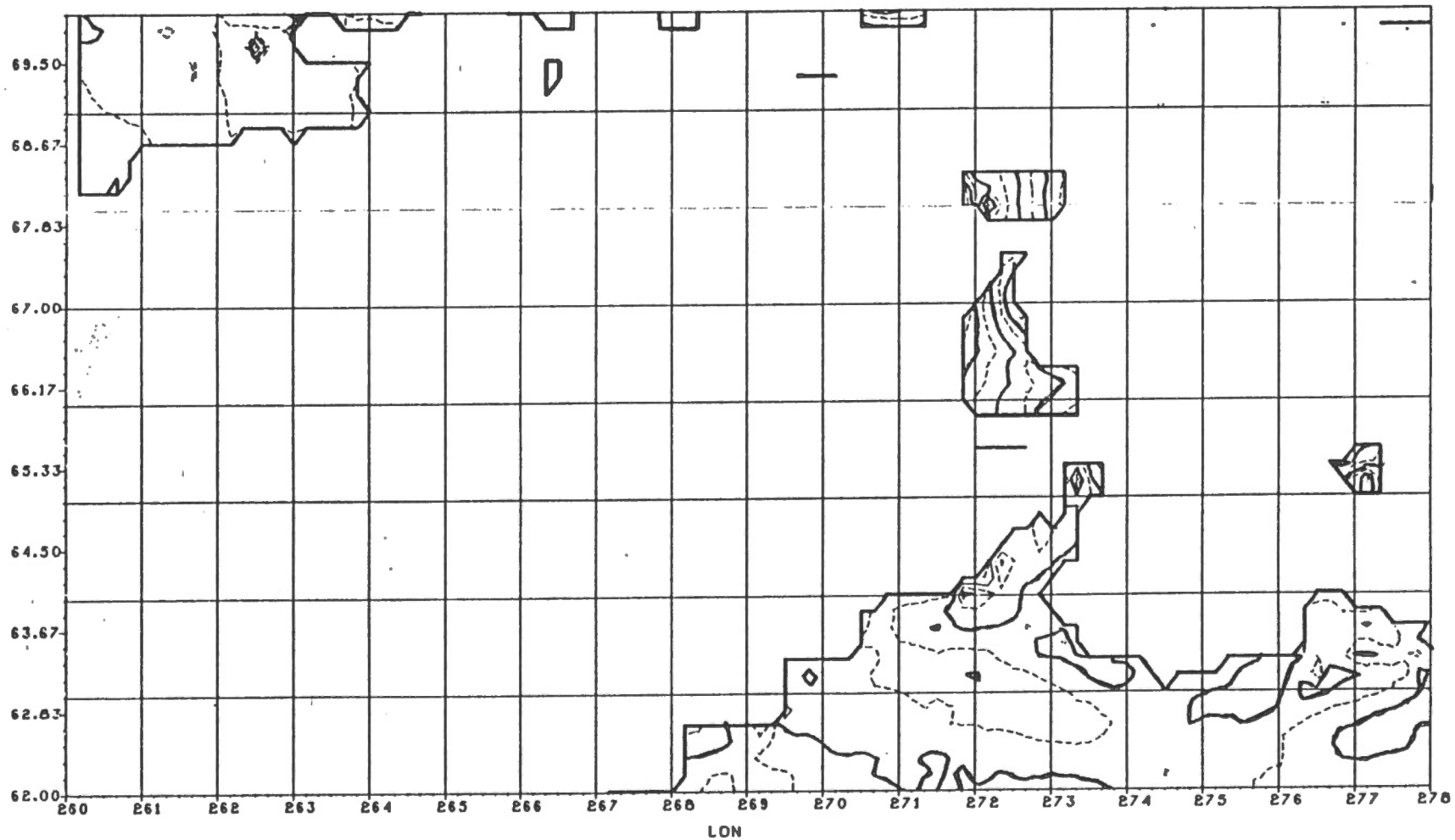
Fig. 3.11: SEASAT altimetry minus UNB geoid.
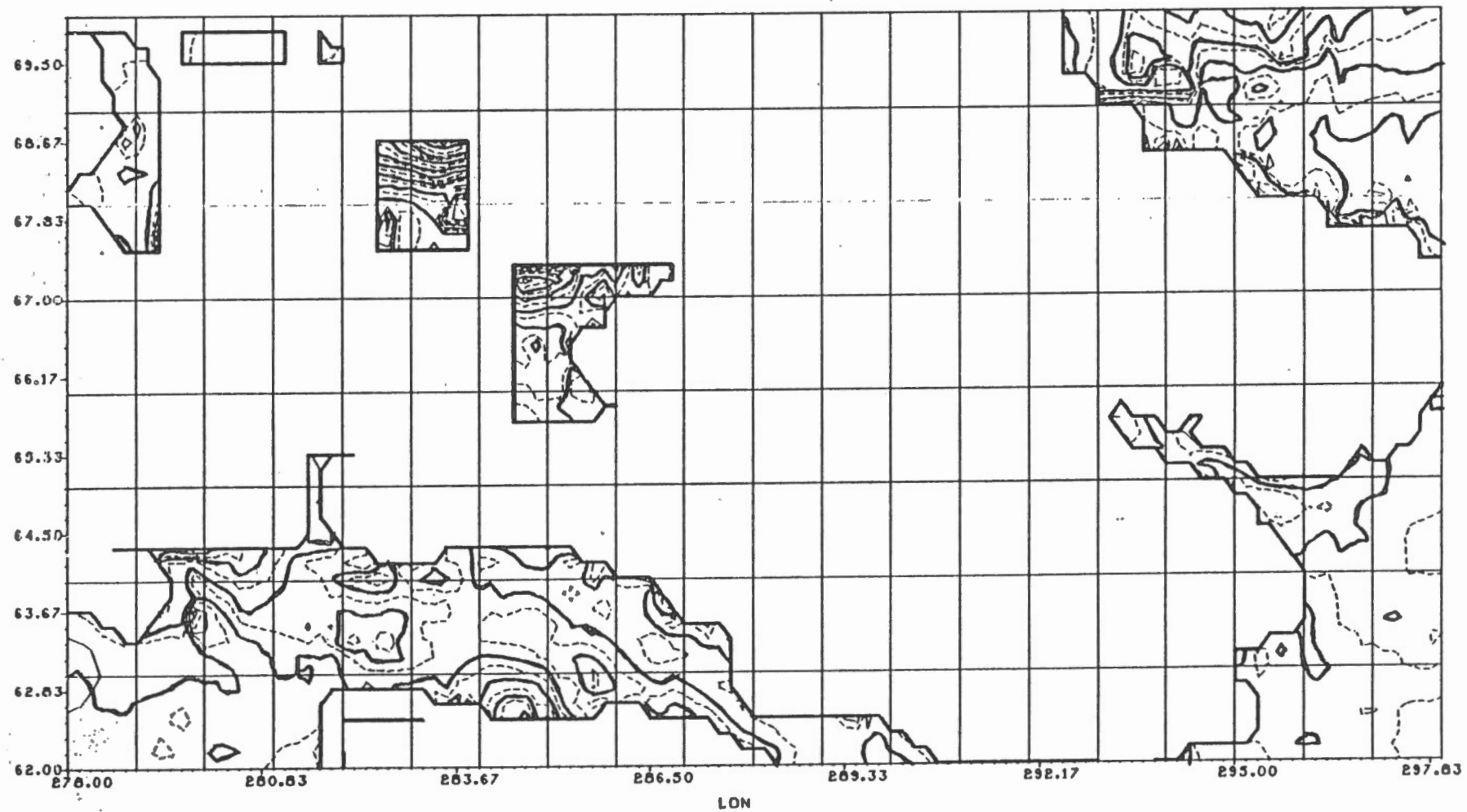
Fig. 3.12:   SEASAT altimetry minus UNB geoid.

SAS
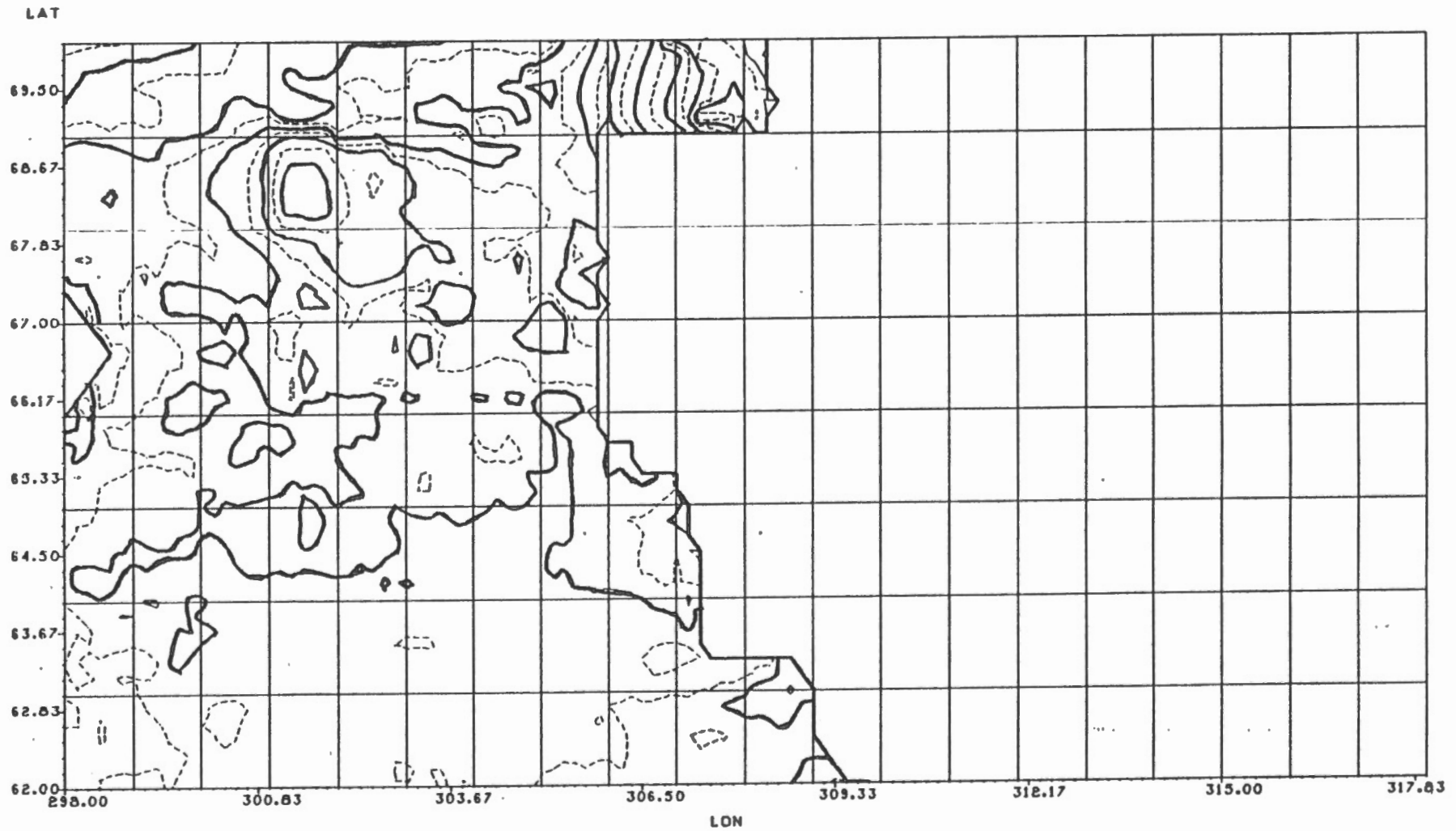
Fig. 3.13:   SEASAT altimetry minus UNB geoid.

Fig. 3.14:   SEASAT altimetry minus UNB geoid.
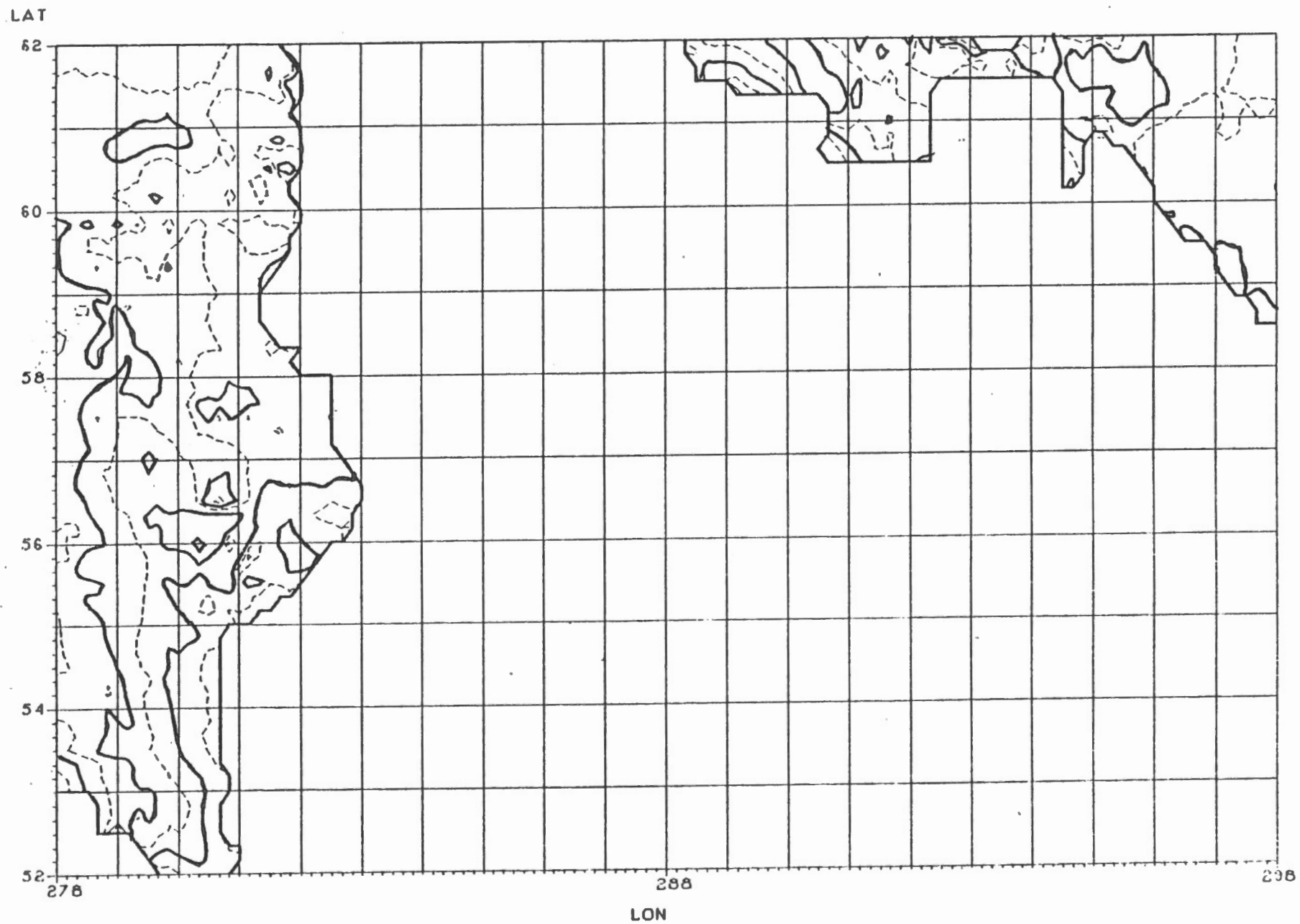
Fig. 3.15:   SEASAT altimetry minus UNB geoid.
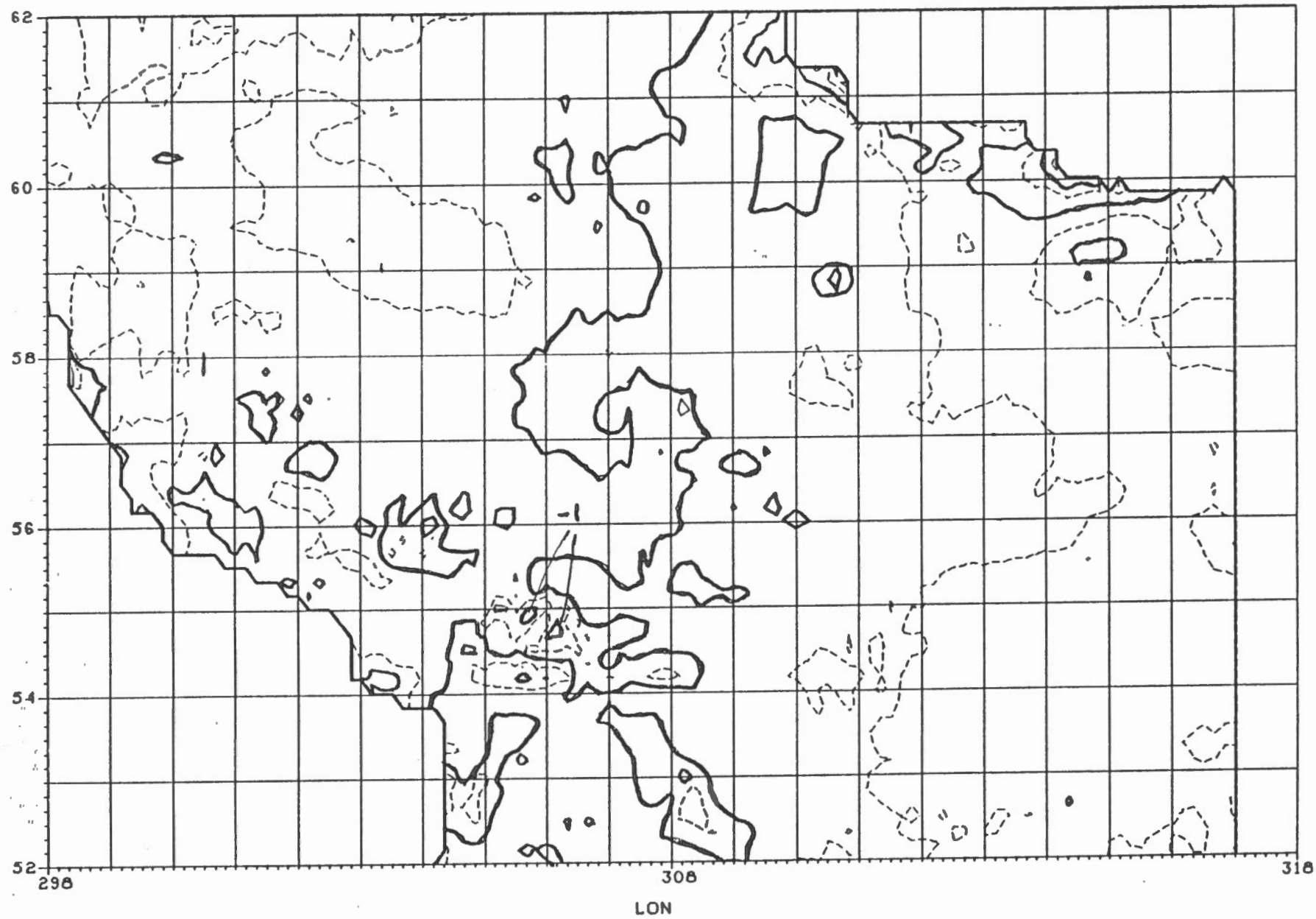
Fig. 3.16:   SEASAT altimetry minus UNB geoid.
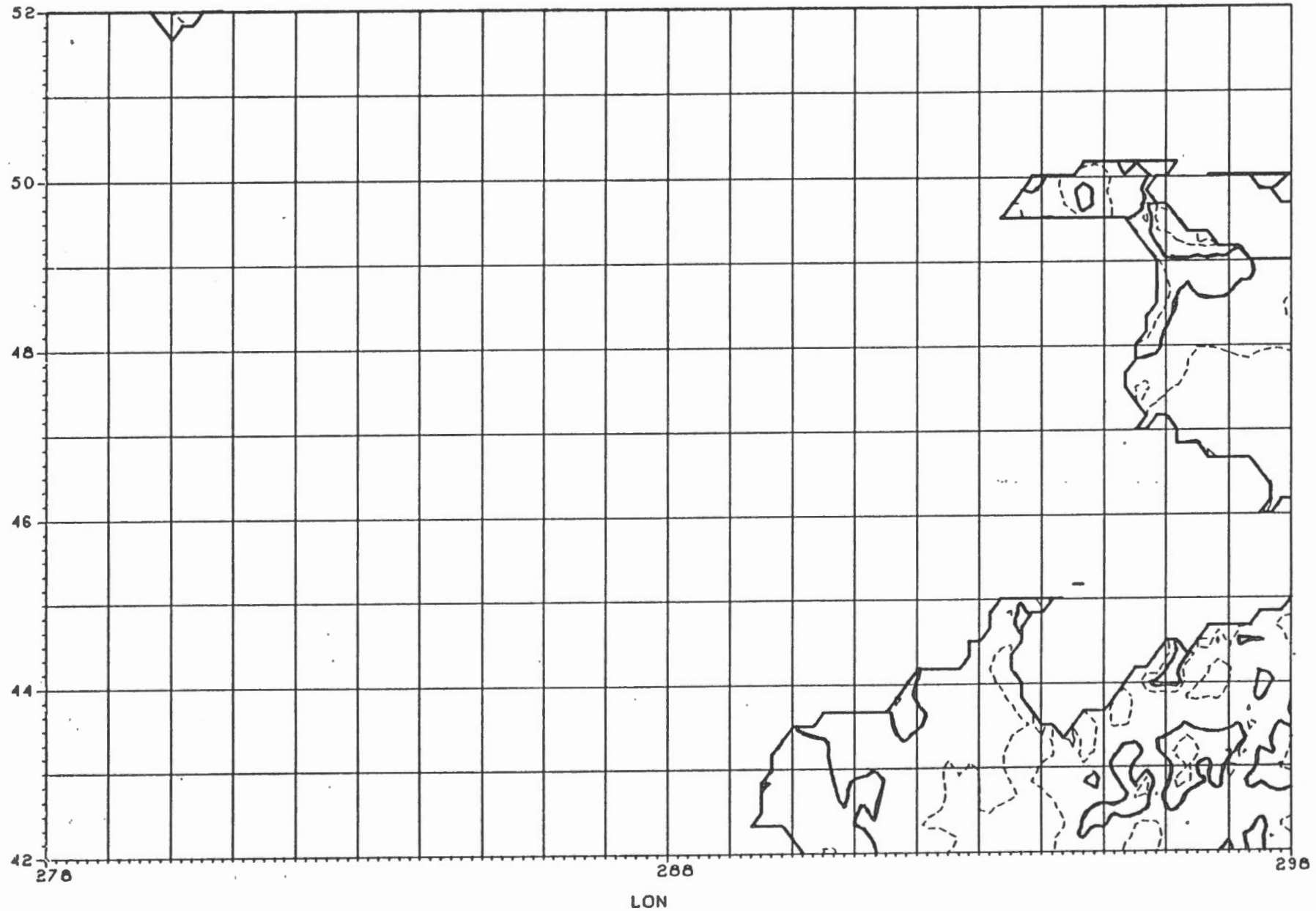
Fig. 3.17: SEASAT altimetry minus UNB geoid.

LAT

52

50

48

46

44

42

298

308

318

LON
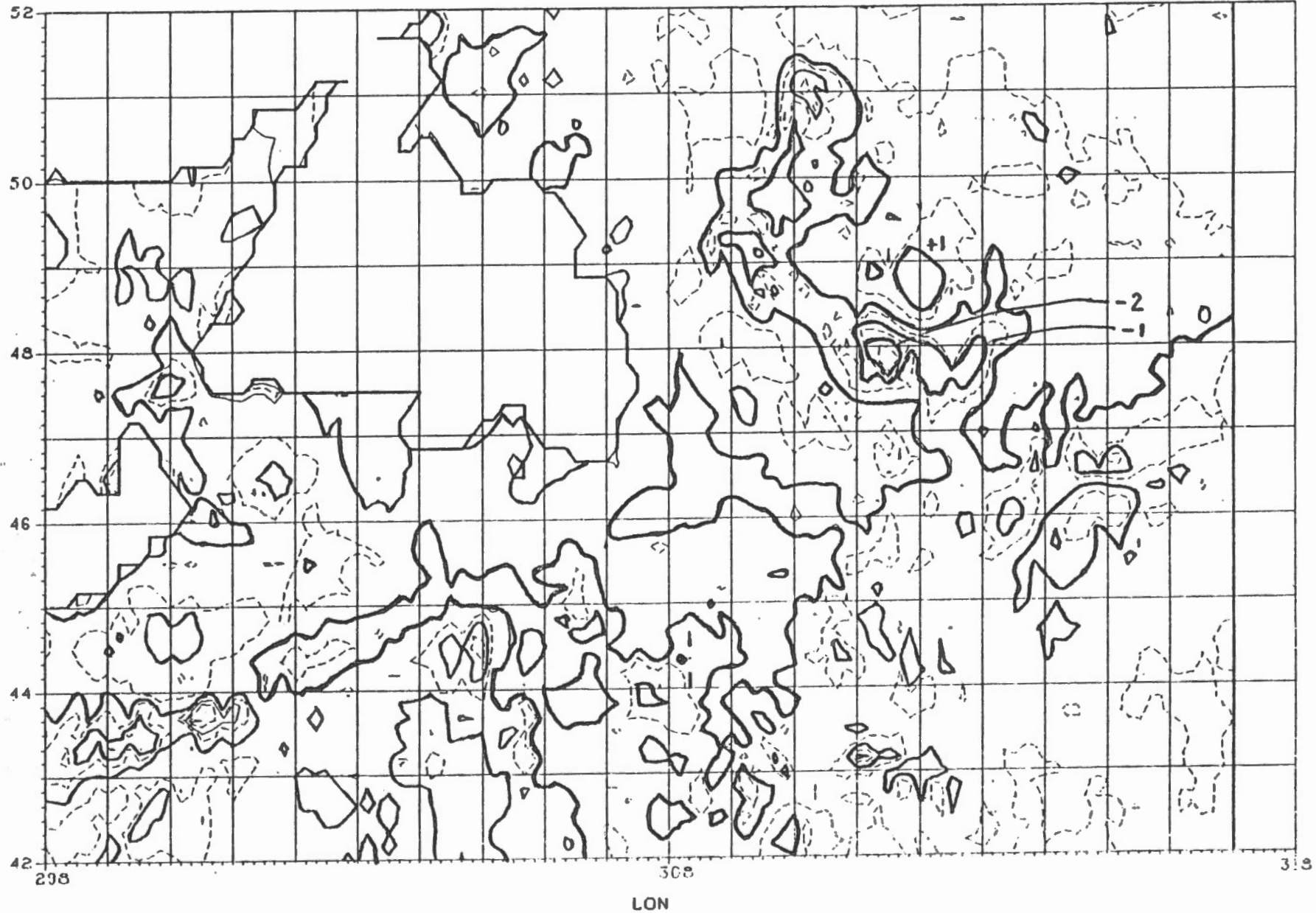
Fig. 3.18:   SEASAT altimetry minus UNB geoid.
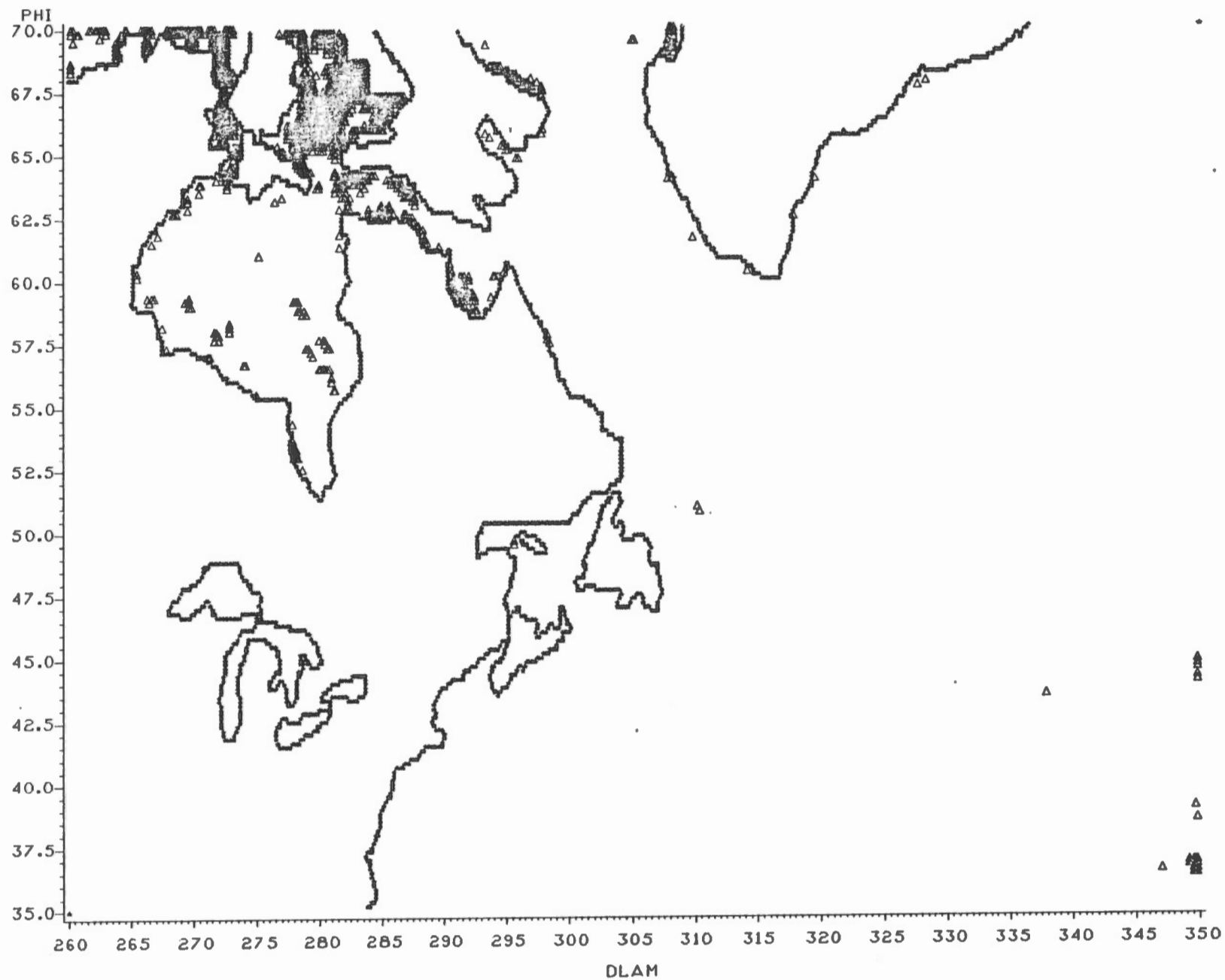
Fig. 3.19: Location of gridded OSU SEASAT altimetry data with
standard deviations exceeding 0.25 m.

42

shaded areas: Standard Deviation larger than 50 cm

dotted line: Standard Deviation equals 25 cm

70°

42°
265°

318°

43
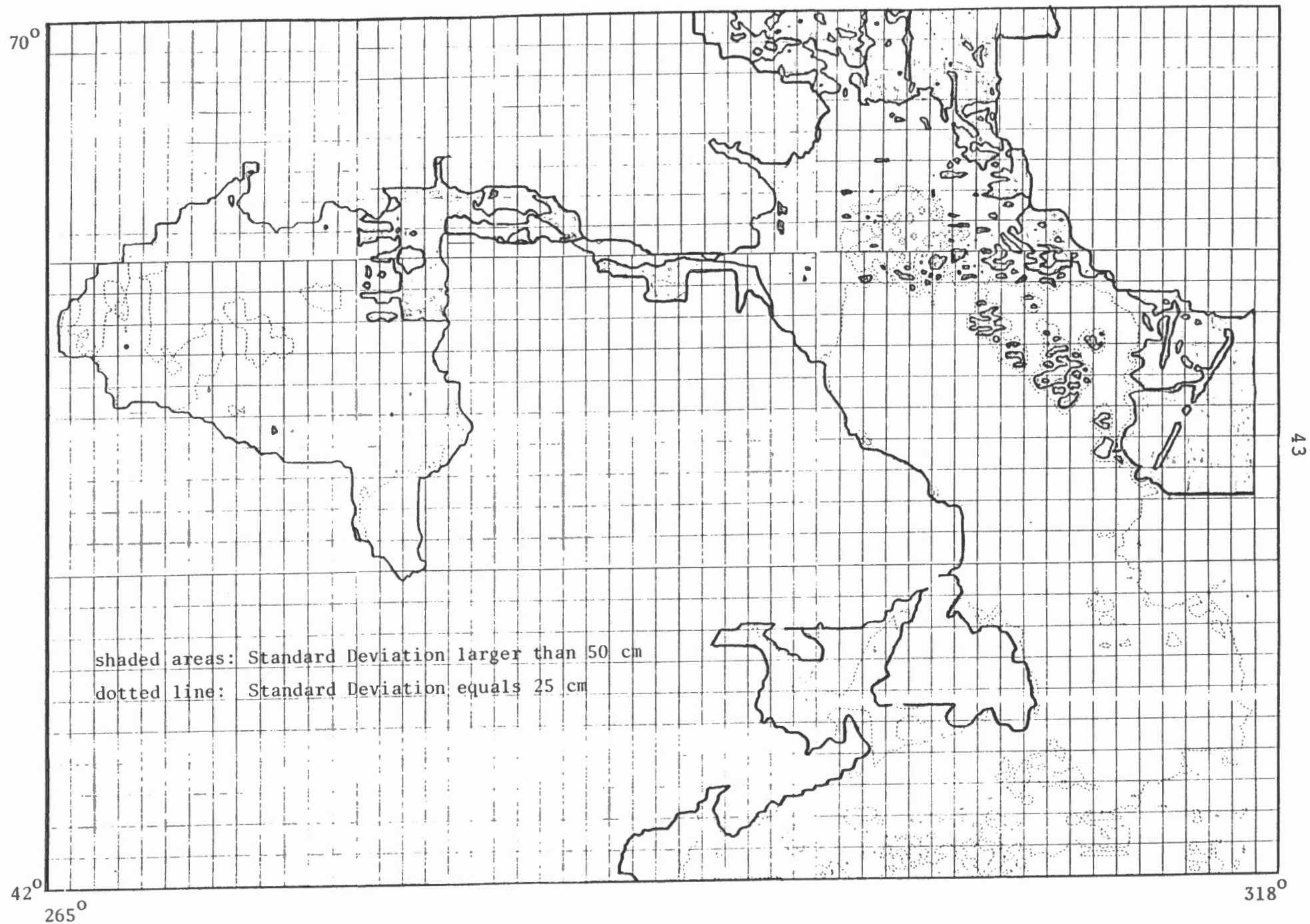
Fig. 3.20: Standard deviation of SEASAT altimetry minus UNB geoid.

metre. Another long wavelength feature is visible in Hudson Bay. There is a remote possibility that this is real SST. Much more probable is that it reflects the uncertainty in the tilt of descending SEASAT orbital arcs coming from the cross-over point adjustment [Rapp, 1983]. If this is the case, the tilt may be justifiably removed at some stage.

Superimposed on the long wavelength structures we see many very localized features in Figures 3.11 to 3.18. These local SST structures are very pronounced along the ridge of the continental shelf off the coasts of Labrador, Newfoundland, and Nova Scotia.

The local variations are of the order of several metres over typical distances of 1° (~100 km). Neither the gravimetric geoid nor the gridded SEASAT altimetry are expected to contain errors of this characteristic. Thus these local structures are believed to represent changes of true local SST. The correlation with ocean bottom topography is very obvious: the largest variation (more than 3.5 metres over a distance of 1°) is sitting right on Flemish Cap.

3.4.3 Zero frequency response technique

In 1982, Merry and Vaníček published results of their analysis of partial point SST differences for several ports in eastern Canada. For their analysis they used the zero frequency response technique [Merry and Vaníček, 1982] applied to a few available forcing phenomena. The results represent only a partial SST because not all forcing phenomena were available. Conspicuously missing among these phenomena are sea currents and salinity variations.

The ports for which the analyses were carried out are: Father's Point, Saint John, Yarmouth, Boutilier Point, Halifax, and Pictou (see Figure 3.21). Because of the non-availability of global averages for even the observed forcing phenomena (atmospheric pressure, air temperature, river discharge, and wind stress components), only differences of the part of the SST between pairs of ports could be evaluated. These are:

Father's Point - Saint John	- 68 mm
Father's Point - Yarmouth	- 206 mm
Father's Point - Halifax	- 92 mm
Father's Point - Pictou	- 228 mm

Boutilier Point is too close to Halifax to be mentioned separately.

■Tide gauge
● River gauge
★Meteorological station

SAGUENAY
RIVER

ST. MAURICE RIVER

ST. LAWRENCE RIVER

FATHER'S POINT

GULF OF ST. LAWRENCE

CHARLOTTETOWN

ST. JOHN RIVER

PICTOU

SAINT JOHN

HALIFAX
BOUTILIER POINT

U.S.A.

YARMOUTH

GULF OF MAINE
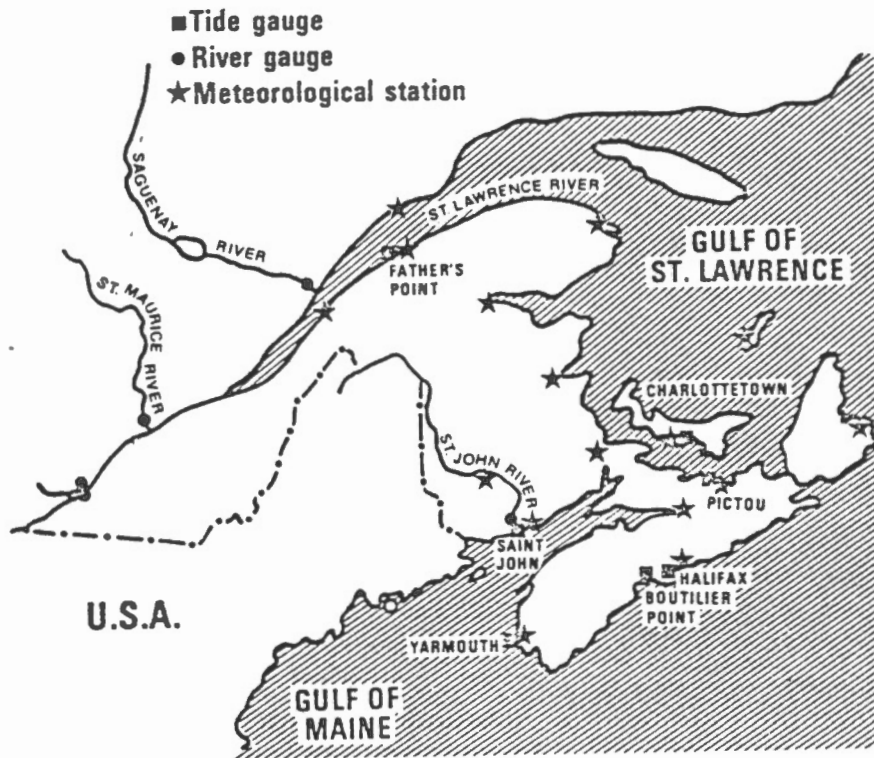
Fig. 3.21

Location of stations for zero frequency response technique.

3.4.4 Comparisons of results

When comparing SST as determined by Levitus (section 3.4.1), the SEASAT altimetry minus UNB geoid (section 3.4.2), and the zero frequency response technique (section 3.4.3), we have to keep in mind that these SSTs do not necessarily represent the same thing.

Levitus's dynamic topography models the long wavelength structure of the global SST. It is based on data collected through many years in different seasons. Levitus's SST can be computed for areas with a certain minimum depth. Since this depth is larger than the depth on the continental shelf, Levitus's SST is available only in the open ocean areas (cf. Figures 3.6 through 3.9). The dynamic topography is positive everywhere. Since the geoid is defined to be the mean global sea level, the Levitus dynamic topography differs from the SST by a positive constant.

The SST determined by the difference between the SEASAT altimetric sea surface and the UNB gravimetric geoid is obviously affected by errors in both surfaces. The characteristics of the altimetry errors are small long wavelength errors resulting from errors in the adjustment of the SEASAT orbital arcs and errors due to spurious reflections in coastal regions resulting in erratic behaviour of the altimetric sea surface in coastal regions. The altimetrically determined sea surface has a very high spatial resolution in areas of good coverage. The relative errors over distances of up to a few hundred kilometres are expected to be at or below the decimetre level in areas of good coverage.

The errors in the gravimetrically determined UNB geoid can be separated into long wavelength errors of the order of about 1 m due to inaccuracies of the underlying reference spheroid (GEM9) and more local errors resulting from approximation in the Stokes integration process and errors in the gravity data used. These local errors are believed to be at or below the decimetre level over distances of a few hundred kilometres. The spatial resolution of the gravimetric geoid in marine areas depends somewhat on the spatial resolution of the gravity data used in the integration. Off the eastern coast of Canada, rather dense point gravity data was available mainly on and somewhat beyond the continental shelf areas. Outside this area of point gravity data coverage, the gravimetric geoid is based mainly on 1° x 1° mean gravity anomalies. The point gravity data will lead to a very high spatial resolution of the gravimetric geoid, whereas 1° x 1° mean gravity anomalies will limit the spatial resolution to features larger than some 100 km. Therefore, for the SST derived from altimetry and the gravimetric geoid we may expect long wavelength errors in the metre range and local to medium range (several hundreds of kilometres) accuracies of the decimetre level. The best spatial resolution will be obtained on the continental shelf. SST close to the shoreline may be unreliable due to errors in the altimetry.

The zero frequency response technique results described in section 3.4.3 provided results for differences of SST between 5 stations located along the coast of the Maritime Provinces of Canada. Since Levitus's SST is available only for the open ocean areas, a comparison between these two techniques is ovbiously not possible. Comparing Figures 3.10 and 3.21, we see that the locations of Father's Point, Pictou and Saint John are not covered by the SST determined from altimetry and geoid. The remaining two stations are obviously located on the shore where altimetry is contaminated by errors. Whereas the altimetric SST is very smooth at a level of 50 to 60 cm in the Yarmouth region, it behaves rather erratically in the Halifax area (see Figure 3.17). Thus a comparison with the 11.4 cm SST difference determined by the zero fequency response technique seems inappropriate.

Figure 3.22 shows the difference between Levitus's dynamic topography and SST determined by the difference between the altimetric sea surface and the gravimetric geoid. Because of the limited extension of Levitus's dynamic topography close to the coast and limitations of the gravimetric geoid at a longitude of 317°, the area of overlap is rather small. A statistical analysis of the differences gives a mean of 82 cm and an rms with respect to the mean of 54 cm. The mean of 82 cm will contain the positive constant of Levitus's SST mentioned above and some of the long wavelength errors of the altimetric/gravimetric SST. The variations of the SST differences characterized by the rms of 54 cm are as large as the variations of the altimetric/gravimetric SST alone (see section 3.4.2). Thus we conclude that the two types of SST discussed here show no significant common features in the overlapping area.

Based on the discussion above, we further conclude that the three types of SST described in sections 3.4.1 through 3.4.3 complement each other by providing the best results for different wavelengths and different areas.
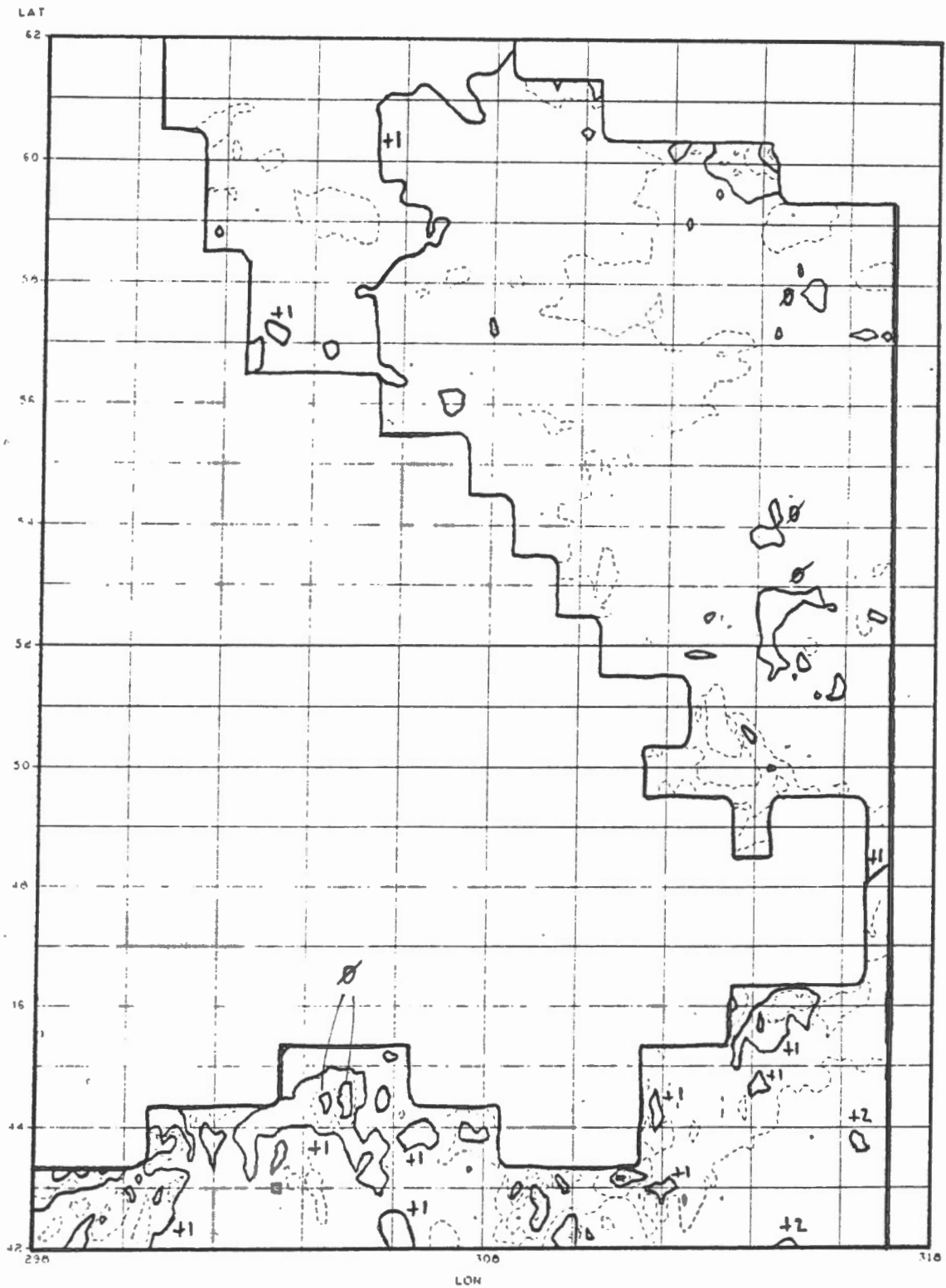
Fig. 3.22

Levitus SST minus (altimetry-UNB geoid).

4. EXISTING TRANSFORMATIONS

This section describes existing techniques for the transformation between different functionals of the gravitational field of the earth. For a more detailed coverage of the subject, see Vaníček and Krakiwsky [1982].

4.1 Description of the Gravitational Field

The *gravitational potential* W at a point P with coordinates $X = (x, y, z)^T$ is

$$W_P = G \iiint_{\text{Earth}} \| X - X' \|^{-1} \rho(X') \, dX' . \tag{4.1}$$

G is the gravitational constant and ρ denotes the mass density. Subtracting from W a *normal potential* U, defined by a simple analytical expression, we obtain the *disturbing potential* T:

$$T = W - U . \tag{4.2}$$

Included in the definition of the normal potential is the assumption that it is produced by an earth model having the same total mass M and the same centre of mass as the actual earth. A particular equipotential surface of the normal potential, $U = U_o$, coincides with the mean earth ellipsoid. The equipotential surface of the gravitational potential having the same numerical value, $W = U_o$, is called the *geoid*. The relation between these surfaces is depicted in Figure 4.1.

By definition, the separation between the geoid and the ellipsoid is the *geoidal undulation*, positive if the geoid is outside the ellipsoid. All three potentials W, U, and T are harmonic outside the gravitating masses, i.e.,

$$\nabla^2 U = \nabla^2 W = \nabla^2 T = 0 . \tag{4.3}$$

49
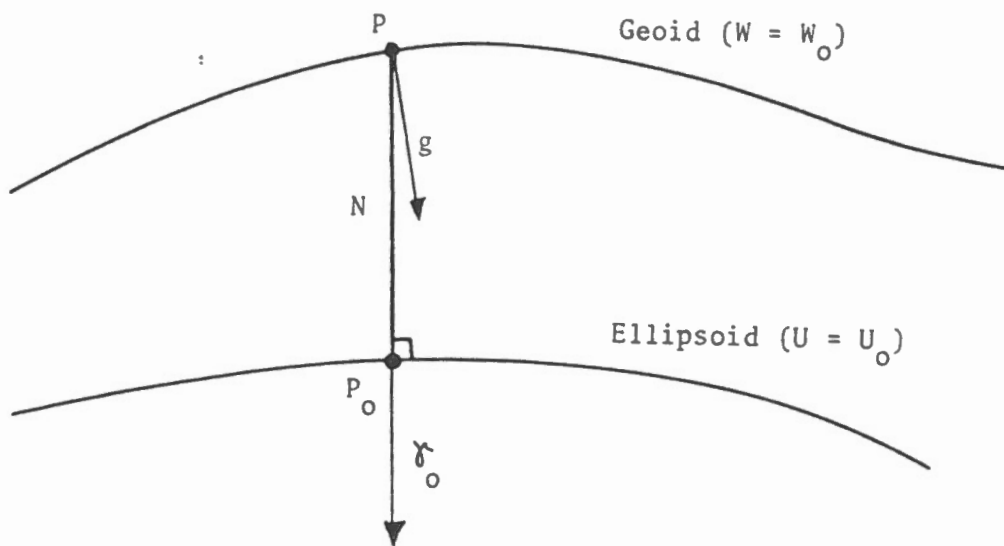
Figure 4.1: Geoid - Ellipsoid geometry

Because of this property, the potentials can be represented in a series expansion of solid spherical harmonics (outside the gravitating masses), e.g.,

$$W(r, \lambda, \phi) = GM/r \sum_{n=0}^{\infty} (R/r)^n \sum_{m=0}^{n} (J_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + K_{nm} \overline{Y}_{nm}{}^s (\phi, \lambda)), \qquad (4.4)$$

where R is a mean earth radius, M is the total mass of the earth, J_{nm} and K_{nm} are constants depending on the mass distribution of the earth, and

$$\left. \begin{array}{c} \overline{Y}_{nm}{}^c(\phi, \lambda) \\ \\ \overline{Y}_{nm}{}^s(\phi, \lambda) \end{array} \right\} = \overline{P}_{nm}(\sin \phi) \left\{ \begin{array}{c} \cos m\lambda \\ \\ \sin m\lambda \end{array} \right\} \qquad (4.5)$$

are orthogonal functions on the sphere. $\overline{P}_{nm}(\sin \phi)$ is a normalized associated Legendre function (e.g., Abramowitz and Stegun [1964]). The first subscript, n, in the expansion (4.4) denotes the **degree** of a particular gravity field constituent, the second subscript, m, denotes the **order**.

Since by definition the potentials U and W are produced by mass distribution having the same centre and the same total mass, the disturbing potential T does not contain any constituents of degree 0 and 1.

$$T(r, \lambda, \phi) = GM/r \sum_{n=2}^{\infty} (R/r)^n \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda)) \qquad (4.6)$$

The disturbing potential T and the geoid undulation N are related by Bruns's formula [Vaníček and Krakiwsky, 1982]

$$N = T/\gamma_0 , \qquad (4.7)$$

where γ_0 is the normal gravity at the ellipsoid (cf. Figure 4.1)

$$\gamma_0 = \| \nabla U \big|_{P_0} \| . \qquad (4.8)$$

The gravity anomaly Δg is defined by

$$\Delta g = g - \gamma_0 , \tag{4.9}$$

where g is the gravity at the geoid

$$g = \| \left. \nabla W \right|_P \| . \tag{4.10}$$

The relation between disturbing potential and gravity anomaly is given by the **fundamental gravimetrical equation** of geodesy

$$\Delta g = - \partial T/\partial H + 1/\gamma_0 \, \partial g/\partial H \, T . \tag{4.11}$$

In spherical approximation (approximation error of the order 1/300), we obtain from eqns. (4.6), (4.7), and (4.11)

$$N(\phi, \lambda) = R \sum_{n=2}^{\infty} \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda)) \tag{4.12}$$

$$\Delta g(\phi, \lambda) = \overline{\gamma} \sum_{n=2}^{\infty} (n-1) \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda)) \tag{4.13}$$

with a mean gravity $\overline{\gamma}$.

Equations (4.12) and (4.13) show that the transformation between geoidal undulations and gravity anomalies is very simple if one of the fields is known in terms of a spherical harmonic expansion. However, in general, we will not know the expansion coefficients and will have to attempt the transformation in a different way.

4.2 Stokes's Transformation

The Stokes transformation relates the geoidal undulation at a particular point P to the gravity anomalies all over the earth.

$$N_p = R/4\pi\gamma \int_{\psi=0}^{\pi} S(\psi) \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha) \sin\psi \, d\alpha \, d\psi . \tag{4.14}$$

ψ is the spherical distance between the dummy point and the point of interest, and α denotes the azimuth. The integral kernel $S(\psi)$ is known as the *Stokes function*, having the respresentation (e.g., Vaníček and Krakiwsky [1982])

$$S(\psi) = \operatorname{cosec} \psi/2 - 6 \sin \psi/2 + 1 - 5 \cos\psi - 3 \cos\psi \ln(\sin \psi/2 + \sin 2\psi/2). \tag{4.15}$$

The Stokes function has a singularity of the order ψ^{-1} at $\psi=o$. The Stokes function can be expanded in an infinite series according to

$$S(\psi) = \sum_{n=2}^{\infty} (2n+1)/n-1 \, P_n(\cos\psi) \tag{4.16}$$

with Legendre polynomials P_n [Abramowitz and Stegun, 1964].

The representation (4.16) is of particular interest, if we want to determine certain frequencies of the geoid, i.e., certain degrees n at the expansion (4.12). Using eqns. (4.12), (4.13), (4.14), and (4.16), and employing the orthogonality relations of spherical harmonics on the sphere, it can be shown that a term of degree n in the expansion (4.16) extracts only geoidal undulation of this degree n from the complete spectrum of Δg (eqn. (4.13)); each term of eqn. (4.16) operates as a band pass filter for a particular gravity field constituent.

In practical applications, these spectral properties of $S(\psi)$ are utilized, if low degree constituents of N are known from other sources and the observed gravity anomalies Δg serve for the determination of high degree constituents of N only. A recent example of this technique can be found in Vaníček et al. [1986].

Denoting

$$
\begin{aligned}
S^l(\psi) &= \sum_{n=l+1}^{\infty} (2n+1)/(n-1) \, P_n(\cos\psi) \\
&= S(\psi) - \sum_{n=2}^{l} (2n+1)/(n-1) \, P_n(\cos\psi)
\end{aligned}
\tag{4.17}
$$

and

$$N = N^l + \delta N^l \tag{4.18}$$

with

$$N^l = R \sum_{n=2}^{l} \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda)) \tag{4.19}$$

and

$$\delta N^l = R \sum_{n=l+1}^{\infty} \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda)), \tag{4.20}$$

the above discussed filtering properties lead from eqn. (4.14) to

$$\delta N^l = R/4\pi\gamma \int_{\psi=0}^{\pi} S^l(\psi) \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha) \sin\psi \, d\psi \, d\alpha . \tag{4.21}$$

The primary advantage of using the integration kernel $S^l(\psi)$ instead of $S(\psi)$ can be seen in Figure 4.2. This figure depicts the graphs of $S(\psi)$ and $S^{20}(\psi)$. It is obvious that the latter decreases much faster with growing spherical distance ψ than the original Stokes function $S(\psi)$. A direct consequence of this is that we can truncate the integration (4.21) at some appropriately chosen ψ_0 without committing significant errors. This fact can be interpreted that the nearby regions contain enough information to determine the high degree part of the geoidal undulations and that for the low degree part, the necessary information is partially contained in the more distant zones.

Since the kernel function $S^l(\psi)$ in eqn. (4.21) blocks the low degree part of the gravity field, this part may be subtracted from Δg beforehand without changing the integration result. Denoting (cf. eqn. (4.13))
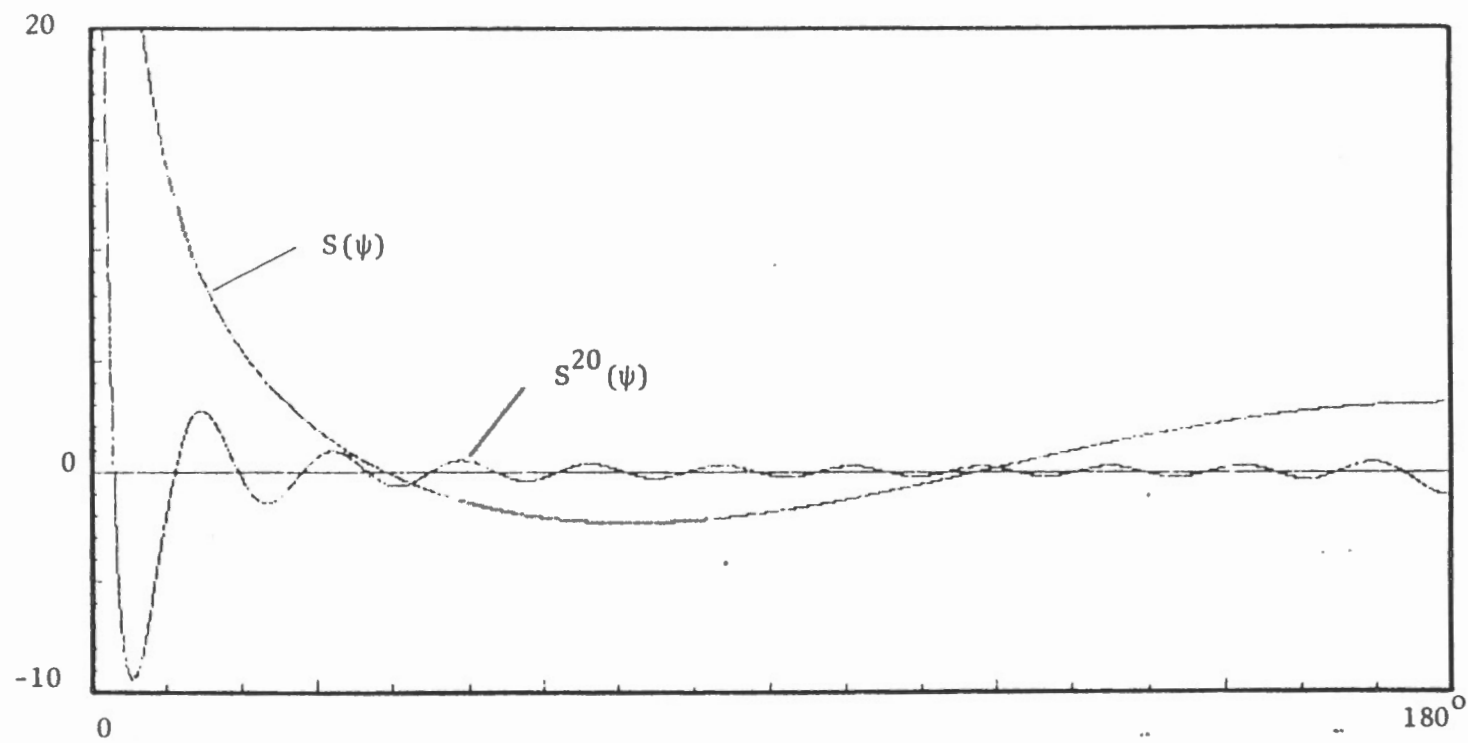
Figure 4.2: Stokes functions $S(\psi)$ and $S^{20}(\psi)$

$$\delta\Delta g^l(\phi, \lambda) \quad = \overline{\gamma} \sum_{n=l+1}^{\infty} (n-1) \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda))$$

$$= \Delta g(\phi, \lambda) - \overline{\gamma} \sum_{n=2}^{l} (n-1) \sum_{m=0}^{n} (\overline{J}_{nm} \overline{Y}_{nm}{}^c(\phi, \lambda) + \overline{K}_{nm} \overline{Y}_{nm}{}^s(\phi, \lambda)) \qquad (4.22)$$

we obtain from eqn. (4.21)

$$\delta N^l = R/4\pi\gamma \int_{\psi=0}^{\pi} S^l(\psi) \int_{\alpha=0}^{2\pi} \delta\Delta g^l(\psi, \alpha) \sin\psi \, d\psi \, d\alpha \ . \qquad (4.23)$$

The replacement of Δg by $\delta\Delta g^l$ becomes mandatory if we truncate the integration at some $\psi_0 < \pi$. In this case, the spherical harmonics are no longer orthogonal in the integration domain, and the filter properties of the integration kernel $S^l(\psi)$ discussed above are lost, at least partially.

When evaluating eqns. (4.14) or (4.23), we may choose to integrate first along the parallels ψ=const. Denoting the average gravity anomaly along ψ=constant by

$$\overline{\Delta g}(\psi) = 1/2\pi \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha) \, d\alpha \qquad (4.24)$$

$$\delta \overline{\Delta g^l}(\psi) = 1/2\pi \int_{\alpha=0}^{2\pi} \delta \overline{\Delta g^l}(\psi, \alpha) \, d\alpha \qquad (4.25)$$

we obtain from eqns. (4.14) and (4.23)

$$N_P = R/\gamma \int_{\psi=0}^{\pi} F(\psi) \, \overline{\Delta g}(\psi) \, d\psi \qquad (4.26)$$

$$\delta N_P{}^l = R/\gamma \int_{\psi=0}^{\pi} F^l(\psi) \, \delta \, \overline{\Delta g^l}(\psi) \, d\psi \qquad (4.27)$$

56

with

$$F(\psi) = 1/2 \, \sin\psi \, S(\psi) \qquad (4.28)$$

$$F^l(\psi) = 1/2 \, \sin\psi \, S^l(\psi) \qquad (4.29)$$

A graph of $F(\psi)$ is shown in Figure 4.3. It can be seen that the singularity of $S(\psi)$ for $\psi=0$ has been compensated for by $\sin\psi$ (eqn. (4.28)). This figure illustrates again that a truncation of the integration (4.26) leads to rather large errors in N, since $F(\psi)$ has considerable power in the long wavelength constituents.

A comparison of $F(\psi)$ with $F^{20}(\psi)$ and $F^{90}(\psi)$ in Figure 4.4 shows that the removal of the long wavelength constituents leads to a rapid decrease of the kernel for growing spherical distance ψ. Thus any truncation at $\psi_0 < \pi$ will give, for eqn. (4.27), smaller truncation errors than for eqn. (4.26).

4.3   Molodenskij's Transformation

The Molodenskij transformation [Molodenskij et al., 1960] can be viewed as the inverse of the Stokes transformation discussed in the previous section. It relates the gravity anomaly at a certain point P to the geoidal undulations all over the earth.

$$\Delta g_P = - \gamma/R \, N_P + \gamma/4\pi R \int_{\psi=0}^{\pi} M(\psi) \int_{\alpha=0}^{2\pi} (N(\psi,\alpha)-N_P) \, \sin\psi \, d\psi \, d\alpha . \qquad (4.30)$$

The integration kernel is given by

$$M(\psi) = - 1/4 \, \mathrm{cosec}^3 \, \psi/2 , \qquad (4.31)$$

and, therefore, has a singularity of the order of ψ^{-3} at $\psi=0$. It can be expanded in a series of Legendre polynomials (e.g., Mather et al. [1976]) according to
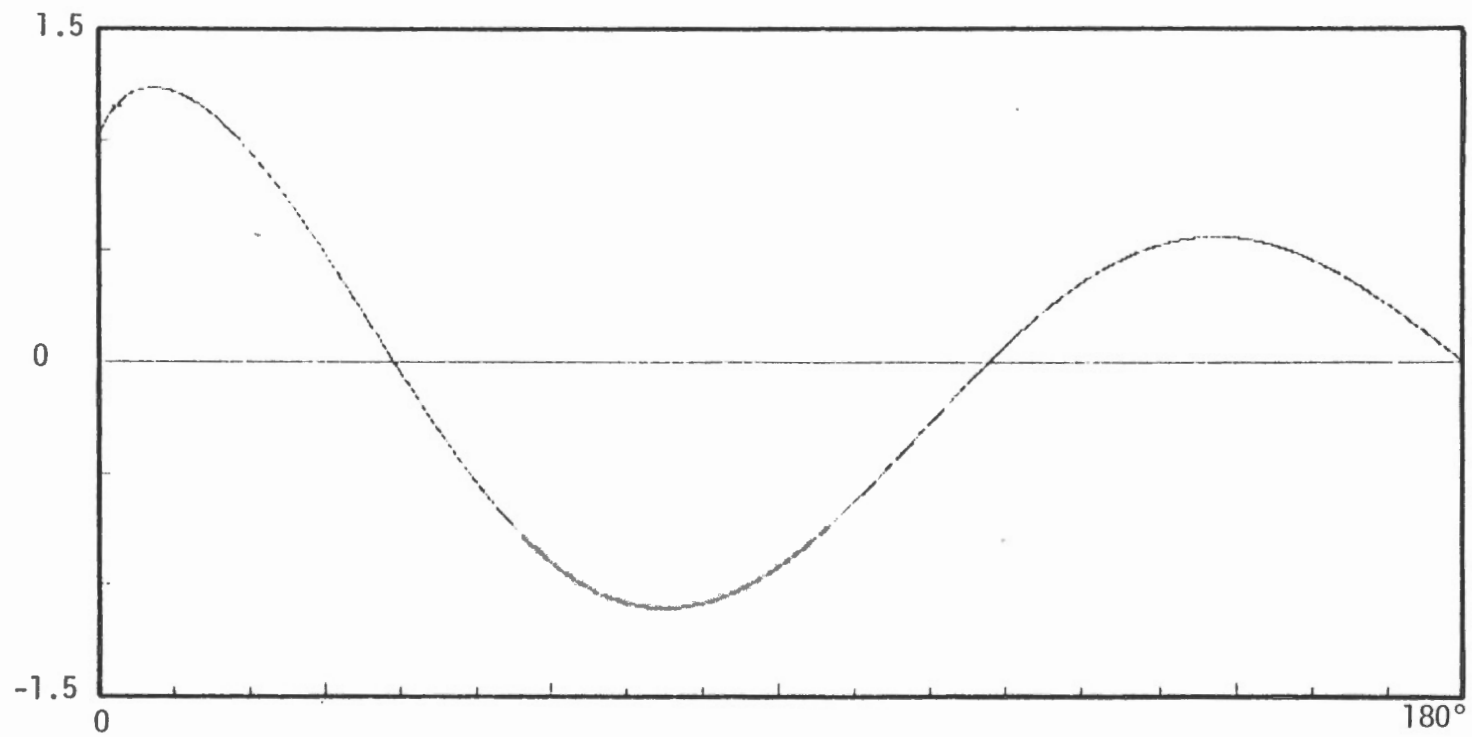
Figure 4.3: Graph of F(ψ).
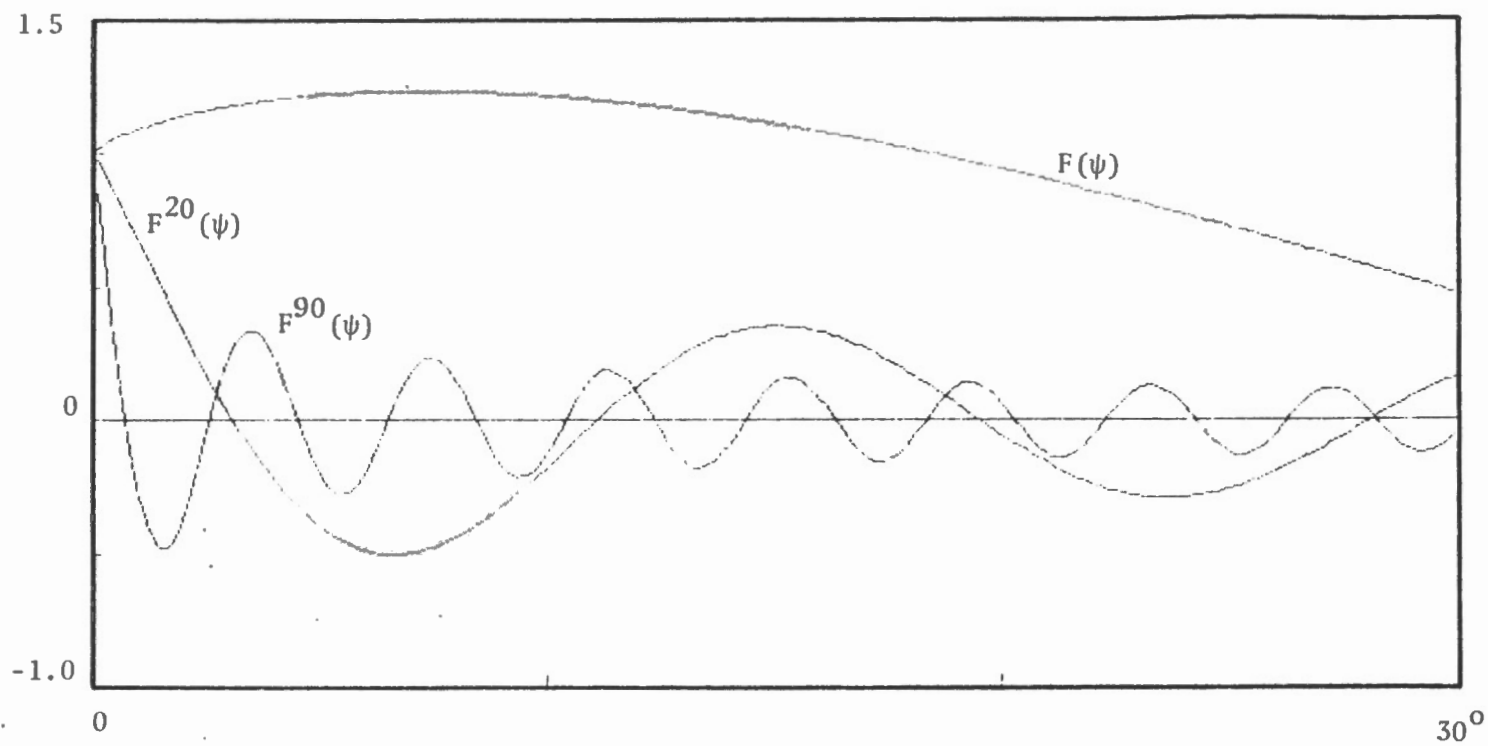
Figure 4.4: Comparison of $F(\psi)$, $F^{20}(\psi)$ and $F^{90}(\psi)$

$$M(\psi) = \sum_{n=0}^{\infty} n(2n+1) \, P_n(\cos\psi) . \qquad (4.32)$$

This series expansion is valid for $\psi > 0$. However, the series (4.32) is not convergent in the ordinary sense and cannot be used to replace eqn. (4.31) in numerical computations. Closely related to this property of the series expansion is the **instability** of the Molodenskij transformation. Changing the geoidal undulation $N(\psi', \alpha')$ at some point (ψ', α') by a small increment dN leads to

$$d\Delta g_P = (\partial \Delta g_P)/(\partial N(\psi', \alpha') \, dN \qquad (4.33)$$

and differentiating eqn. (4.30), we obtain

$$d\Delta g_P = -\gamma/16\pi R \, \sin\psi' \, \text{cosec}^3 \, \psi'/2 \, dN . \qquad (4.34)$$

Obviously, a small increment dN leads to large changes in Δg_P for small ψ'. This instability of Molodenskij's transformation has prevented its application for numerical computations. It will, however, be used in section 5.3 for analytical derivations leading to a different approach for the transformation from geoidal undulations to gravity anomalies.

4.4 Rapp's Application of Least-Squares Collocation

Rapp [1983], in his computation of 1° x 1° and 5° equal area mean gravity anomalies, has chosen to use the least-squares collocation. He evaluates the averages using sea surface heights referred to the GEM9 reference spheroid uncorrected for SST. These heights, δH, are selected to cover the block for which the average $\delta \overline{\Delta g}$ (referred to GEM9 spheroid) is sought plus a 2° wide strip around it. The formula used is

$$\overline{\delta \Delta g} = C_{\delta \Delta g, \delta N}(C_{\delta N, \delta N} + \text{diag}(\sigma_A{}^2))^{-1} \, \delta H , \qquad (4.35)$$

where

$C_{\delta \Delta g, \delta N}$ is the cross-covariance matrix between the anomalies and geoidal heights, both

60

referred to GEM9 reference ellipsoid,

$C_{\delta N, \delta N}$ is the auto-covariance matrix of the geoidal heights, and

σ_A are the standard deviations of the altimeter measurements.

The standard deviation $\sigma_{\overline{\Delta g}}$ of the mean anomaly is then evaluated from the following expression:

$$\sigma_{\delta \Delta g}^2 - \sigma_{\overline{\Delta g}}^2 = C_{\delta \Delta g, \delta N}(C_{\delta N, \delta N} + \mathrm{diag}(\sigma_A^2))^{-1} C_{\delta N, \delta \Delta g}, \tag{4.36}$$

where

$\sigma_{\delta \Delta g}^2$ is the expected mean square value referred to the GEM9 spheroid.

Using this approach, Rapp predicted 34 973, 1° x 1° mean anomalies with an average standard deviation of 5.1 mGal. The agreement of Rapp's predictions with terrestrial predictions is quite reasonable. Of the 34 973 values, 11 235 terrestrial values have standard deviations of 15 mGal or smaller. Out of this subset, only 38 disagreed with the SEASAT-derived values by more than 40 mGal. A total of 17 356 terrestrial values have standard deviations no larger than 99 mGal. From this subset, 379 disagreed with the SEASAT-derived values by more than 40 mGal.

5. THE TRUNCATION

This section deals with truncated Stokes's integration and integral forms of geoidal undulation. Its goal is the derivation of mathematical tools for the *local* comparison of gravity and satellite altimetry, in terms of both gravity anomalies and geoidal undulations.

In these derivations, we will also consider a known long wavelength gravity field in terms of a spherical harmonic expansion of the disturbing gravitational field (eqn. (4.6)) to be known up to some degree l.

Spherical harmonics gravitational field representations have been compiled up to degree 180 [Rapp, 1981] and 200 [Wenzel, 1985]. However, the estimated standard deviations of the potential coefficients $\overline{J}_{nm}$ and $\overline{K}_{nm}$ indicate that the coefficients are rather poorly determined for $n>100$. Therefore, we have decided to adopt only the coefficient up to degree $l=90$ as being known for our purposes.

The following section 5.1 describes the spheroidal Stokes integration within a local area. In this approach, only marine gravity data from the neighbourhood of the point of interest is used. Thus the resulting integral contains only local features. This integral should not be confused with the geoid. Firstly, all long wavelength features are missing since a corresponding spherical harmonic expansion of the gravity field has been removed from the input data. Secondly, the integration is truncated at a very small spherical distance and thus even the high frequent part of the gravity field is only partially represented in the integral. Stokes's integration requires continuous input data within the area of integration. Since the marine gravity data is available only in discrete points or tracks, a technique for quasi-analytical integration over sparse data is developed in section 5.2. The main contribution to the problem of comparing altimetry and marine gravity can be found in section 5.3. It describes the derivation of an integral kernel that extracts from a geoidal surface (as determined from altimetry) the same integral form as the above mentioned truncated spheroidal Stokes's integration. The comparison between truncated spheroidal Stokes's integration and the integration with the new kernel over the altimetric surface is expected to show how well altimetry and gravimetry agree locally. Since this is a new derivation, we describe in detail all the problems encountered in our work and their solutions. Section 5.4 describes the procedure implemented for the numerical integration over gridded altimetry data.

5.1 Partial Stokes's Integration

The transformation of gravity anomalies Δg into geoidal undulations (section 4.2) requires an integration all over the earth.

For practical applications, this integration usually has to be truncated at some radius ψ_0 because

- gravity data is not available all over the earth, and
- the computing burden is too large for complete integration.

Especially for the inverse transformations, described in section 5.3, we are interested in keeping the truncation radius ψ_0 as small as possible. As will be seen in section 5.3, the radius ψ_0 is related to the truncation radius ϕ_0 for integration on altimetry derived geoidal heights. To enable this altimetry integration to give results close to the coast line, we have to keep the radius ϕ_0 of integration as small as possible.

For a truncation radius ψ_0, we obtain, from eqn. (4.14)

$$\delta N^{l,\psi_0} = R/4\pi\gamma \int_{\psi=0}^{\psi_0} S(\psi) \int_{\alpha=0}^{2\pi} \delta\Delta g^l(\psi, \alpha)\, \sin\psi\, d\psi\, d\alpha \tag{5.1}$$

Integrating first with respect to the azimuth α, we obtain

$$\delta N^{l,\psi_0} = R/\gamma \int_{\psi=0}^{\psi_0} F(\psi)\, \delta\, \overline{\Delta g}^l(\psi)\, d\psi \tag{5.2}$$

with

$$F(\psi) = 1/2 \sin\psi\, S(\psi) \tag{5.3}$$

and

$$\delta\, \overline{\Delta g}^l = 1/2\pi \int_{\alpha=0}^{2\pi} \delta\Delta g^l(\psi, \alpha)\, d\alpha . \tag{5.4}$$

63

To perform the integration (5.1) and (5.2), continuous data $\delta \Delta g^l$ is required within $\psi \le \psi_0$.

5.2 Numerical Integration Over Sparse Data

In the marine environment, we will usually not find continuous coverage of $\delta \Delta g$'s in $\psi \le \psi_0$. We may, in extreme cases, end up with just one track of gravity observations within the integration area.

To facilitate an integration over this sparse data distribution, we may fit an analytical surface to the data and use the surface coefficients in the integration.

A particular surface representation in polar coordinates ψ, α is

$$\delta \Delta g^l(\psi, \alpha) = \sum_{i=0}^{I} a_i \psi^i \sum_{j=0}^{J} (b_j \cos j\alpha + d_j \sin j\alpha) \tag{5.5}$$

Performing the integration with respect to the parameter α according to eqn. (5.4), we obtain

$$\delta \overline{\Delta g}^l(\psi) = \sum_{i=0}^{I} b_0 a_i \psi^i = \sum_{i=0}^{I} C_i \psi^i. \tag{5.6}$$

The α dependence has disappeared due to the orthogonality of the trigonometrical functions in the interval $(0, 2\pi)$.

In order to keep the numerical effort as small as possible, we may directly fit a curve $\delta \overline{\Delta g}^l(\psi)$ to the $\delta \Delta g^l(\psi, \alpha)$ by disregarding their azimuth dependence.

Having determined a set of coefficients C_i representing the radial dependence of $\delta \Delta g^l$, we can perform the integration with respect to the variable ψ. Instead of evaluating the kernel $F(\psi)$ using eqns. (5.3) and (4.17), we approximate it in the area of interest by a simple algebraic polynomial

$$F(\psi) = \sum_{i=0}^{I} h_i \, \psi^i . \tag{5.7}$$

This approximation procedure avoids the repeated numerical evaluation of eqn. (4.17) during a numerical integration and, moreover, facilitates an analytical integration of eqn. (5.2).

Inserting eqns. (5.6) and (5.7) into eqn. (5.2), we obtain

$$\delta N^{l,\psi_0} = R/\gamma \int_{\psi=0}^{\psi_0} \left(\sum_{i=0}^{I} h_i \, \psi^i \right) \left(\sum_{j=0}^{J} c_j \, \psi^j \right) d\psi . \tag{5.8}$$

Integrating eqn. (5.8) gives finally the partial geoidal undulation from truncated Stokes's integration in terms of the curve coefficients C_i and the truncation radius ψ_0

$$\delta N^{l,\psi_0} = R/\gamma \sum_{i=0}^{I} \sum_{j=0}^{J} (h_i \, C_j)/(i+j+1) \, \psi_0^{i+j+1} . \tag{5.9}$$

The approximation (5.7) has to follow two objectives: the degree of the polynomial should be as small as possible to save computation time during the evaluation of eqn. (5.9), and the polynomial degree must be high enough to avoid any significant errors due to the approximation.

In the present context, the function $F(\psi)$ will be used for $\psi_0 < 2.5°$ only. The result of the integration, $\delta N^{l,\psi_0}$, will be of the order of several metres. Thus the polynomial approximation (5.7) should be accurate to 10^{-3} in the interval $0 \leq \psi \leq 2.5°$ to avoid integration errors larger then several millimetres. From Figure 4.4 we see that the $F(\psi)$ is close to $+1$ in this interval. Therefore, we require an approximation accuracy of $1 \cdot 10^{-3}$ in both the absolute and the RMS sense. We can evaluate the value $F(0)$ by performing some limit operations on eqn. (4.29) to obtain

$$F(0) = 1 . \tag{5.10}$$

To guarantee optimal accuracy for $\psi = 0$, we fix the coefficient h_0 in eqn. (5.7) a priori. The other coefficients h_i are determined from a least-squares curve fit to numerical values computed from

eqn. (4.29). A polynomial of degree six was found to fulfill the accuracy requirements stated above.

The coefficients for this polynomial are listed in Table 5.1. These coefficients require a polynomial evaluation with ψ in units of degrees.

TABLE 5.1

Coefficients of approximating polynomial.

i	h_i
0	1.000 00
1	0.390 33
2	-0.961 70
3	2.343 08
4	-3.372 95
5	2.485 83
6	-0.726 34

For these values, the approximation errors remain within $\pm 10^{-3}$. The RMS of the approximation errors is $0:24*10^{-3}$.

5.3 The "Truncation" Kernel

5.3.1 Truncated Stokes's integration contribution from altimeter data

Geoidal undulations N can be computed from (observed) gravity anomalies Δg by means of Stokes's integration (e.g., Vaníček and Krakiwsky [1982]):

$$N = R/(4\pi\gamma) \int_{\psi=0}^{\pi} S(\psi) \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha) \, d\alpha \, \sin\psi \, d\psi \; . \tag{5.11}$$

If for any reason the integration is restricted to a spherical cap of radius ψ_0, the resulting truncated Stokes's integration gives

$$N^{\psi_0} = R/(4\pi\gamma) \int\limits_{\psi=0}^{\psi_0} S(\psi) \int\limits_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha)\, d\alpha\, \sin\psi\, d\psi, \qquad (5.12)$$

with

R being the mean radius of the earth

γ the mean gravity at the earth's surface

ψ, α the spherical coordinates.

Any integration (eqn. (5.11) or (5.12)) is inherently restricted to an accuracy of $1/300$ because of spherical approximations. In a marine environment, we may have gravity anomaly observations Δg and "observed" geoidal undulations N from altimetry. If the Δg's are processed according to eqn. (5.12), we look for a way to compare the resulting N^{ψ_0} with the observed altimetric N. Especially we search for an operator F that gives N^{ψ_0} if applied to N:

$$N^{\psi_0} = F(N(\psi, \alpha)). \qquad (5.13)$$

To derive this operation, we recall several derivations and equations that can be found in modern textbooks on physical geodesy.

5.3.1.1   Molodenskij's truncation coefficients

If we want to extend the integration domain of eqn. (5.12) over the complete sphere without changing the result N^{ψ_0}, we replace the Stokes function $S(\psi)$ by $\overline{S}^{\psi_0}(\psi)$, defined by

$$\overline{S}^{\psi_0}(\psi) = \begin{cases} S(\psi), & \psi \le \psi_0 \\[2mm] 0, & \psi > \psi_0 \end{cases} \qquad (5.14)$$

and write formally

$$N^{\Psi_0} = R/(4\pi\gamma) \int\limits_{\psi=0}^{\pi} \overline{S}^{\Psi_0}(\psi) \int\limits_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha)\, d\alpha\, \sin\psi\, d\psi\ . \tag{5.15}$$

The Stokes function can be represented by

$$S(\psi) = \sum_{n=2}^{\infty} (2n+1)/(n-1)\, P_n(\cos\psi)\ , \tag{5.16}$$

with Legendre's polynomial P_n of degree n. $\overline{S}(\psi)$ can also be represented in a Legendre series expansion

$$\overline{S}^{\Psi_0}(\psi) = \sum_{n=0}^{\infty} (2n+1)/2\, \alpha_n(\psi_0)\, P_n(\cos\psi) \tag{5.17}$$

with

$$\alpha_n(\psi_0) = -\, Q_n(\psi_0)\ \text{if}\ n < 2$$

$$\alpha_n(\psi_0) = 2/(n-1) - Q_n(\psi_0)\ \text{if}\ n \geq 2 \tag{5.18}$$

and the Molodenskij truncation coefficients

$$Q_n(\psi_0) = \int\limits_{\psi=\psi_0}^{\pi} S(\psi)\, P_n(\cos\psi)\, \sin\psi\, d\psi\ . \tag{5.19}$$

The coefficients $Q_n(\psi_0)$ can be evaluated numerically using recurrence relations, the most efficient of which seems to be the algorithm of Paul [1973].

5.3.1.2   Molodenskij integration

Equation (5.11) can be envisaged as an integral equation for Δg; the solution of this equation can be written according to Molodenskij et al. [1962]:

$$\Delta g(\bar{\psi}, \bar{\alpha}) = -\gamma/R \, N(\bar{\psi}, \bar{\alpha}) +$$

$$+ \gamma/(4\pi R) \int_{\psi=0}^{\pi} \int_{\alpha=0}^{2\pi} M(\bar{\psi}, \psi, \bar{\alpha}, \alpha)(N(\psi, \alpha) - N(\bar{\psi}, \bar{\alpha})) \sin\psi \, d\psi \, d\alpha , \qquad (5.20)$$

where

$$M(\bar{\psi}, \psi, \bar{\alpha}, \alpha) = M(\beta) = -1/4 \, \csc^3(\beta/2) \qquad (5.21)$$

and β is the spherical distance between $(\bar{\psi}, \bar{\alpha})$ and (ψ, α). It is also possible to derive an expansion in Legendre polynomials for $M(\beta)$:

$$M(\beta) = \sum_{n=0}^{\infty} n(2n+1) \, P_n(\cos\beta) . \qquad (5.22)$$

This series is not converging in the ordinary sense and is not well suited for numerical computations. However, its use in the following derivations will prove useful because of the orthogonality properties of Legendre polynomials in $(0, \pi)$.

5.3.1.3 Derivation of the F-operator

Replacing the Δg values in the truncated Stokes integration (5.15) by Molodenskij's formula (5.20) we obtain

$$N^{\psi}{}_0 = 1/4\pi \int_{\psi=0}^{\pi} \int_{\alpha=0}^{2\pi} \bar{S}(\psi)[- N(\psi, \alpha) +$$

$$+ 1/4\pi \int_{\bar{\psi}=0}^{\pi} \int_{\bar{\alpha}=0}^{2\pi} M(\psi, \bar{\psi}, \alpha, \bar{\alpha})(N(\bar{\psi}, \bar{\alpha}) - N(\psi, \alpha)) \sin\bar{\psi} \, d\bar{\psi} \, d\bar{\alpha}] \sin\psi \, d\psi \, d\alpha. \qquad (5.23)$$

69

5. The Truncation.

Using elementary operations, replacing the integration kernels $\overline{S}$ and M by their series expansions (5.17) and (5.22), respectively, and utilizing the orthogonality of Legendre polynomials in $(0, \pi)$, we succeed in transforming eqn. (5.23) to

$$N^{\psi}{}_0 = 1/4\pi \int_{\psi=0}^{\pi} \int_{\alpha=0}^{2\pi} [\sum_{n=0}^{\infty} (2n+1)/2 \, (n-1)\alpha_n(\psi_o)P_n(\cos\psi)]N(\psi,\alpha)\sin\psi d\psi d\alpha . \qquad (5.24)$$

Replacing α_n by eqn (5.18) we get from eqn. (5.24)

$$N^{\psi}{}_0 = 1/4\pi \int_{\psi=0}^{\pi} \int_{\alpha=0}^{2\pi} \sum_{n=2}^{\infty} (2n+1)P_n(\cos\psi)N(\psi,\alpha) \sin\psi \, d\psi \, d\alpha \qquad (5.25)$$

$$-1/4\pi \int_{\psi=0}^{\pi} \int_{\alpha=0}^{2\pi} \sum_{n=0}^{\infty} (2n+1)/2 \, (n-1)Q_n(\psi_o)P_n(\cos\psi)N(\psi,\alpha) \sin\psi \, d\psi \, d\alpha .$$

Since the geoidal undulations do not contain any harmonics of degree 0 and 1, we can neglect the first two terms in the second summation. Furthermore, we note that the kernel

$$K_n = (2n + 1) \, P_n \qquad (5.26)$$

is a reproducing kernel for harmonics of degree n. Thus we write for eqn. (5.25)

$$N^{\psi}{}_0 = N - 1/2 \int_{\psi=0}^{\pi} R^{\psi}{}_0(\psi) \, 1/2\pi \int_{\alpha=0}^{2\pi} N(\psi, \alpha) \, d\alpha \sin\psi \, d\alpha \qquad (5.27)$$

with the new kernel function

$$R^{\psi}{}_0(\psi) = \sum_{n=2}^{\infty} (2n+1)/2 \, (n-1)Q_n(\psi_o)P_n(\cos\psi) . \qquad (5.28)$$

Equation (5.27) relates the "observed" geoid undulations N to the result of a truncated Stokes's integration with spherical cap ψ_o. It is the relation sought for according to eqn. (5.13).

5.3.2 Evaluation of the kernel function $R^{\psi}o(\psi)$

The series representation of the kernel function in Legendre's polynomials

$$R^{\psi}o(\psi) = \sum_{n=2}^{\infty} ((2n+1)(n-1))/2 \, Q_n(\psi_o) \, P_n(\cos\psi) \qquad (5.29)$$

is not suitable for a numerical evaluation for two reasons.

(1)   The series is not convergent in the ordinary sense, i.e.,

$$\lim_{n\to\infty} S_N \, , \quad S_N = \sum_{n=2}^{N} ((2n+1)(n-1))/2 \, Q_n P_n \qquad (5.30)$$

is not finite. To overcome this problem, special summation techniques have been used, called Cesàro summation [Lanzano, 1982, pp. 129 ff; Hobson, 1931, pp. 347 ff]. The Cesàro summation technique has been implemented on the SE/UNB HP 1000 minicomputer. Using double precision word length, numerical instabilities occurred for summations with upper limit N > 5000. Several tests indicated that for an accurate result the Cesàro summation had to be extended at least to N $\doteq$ 20 000. Thus the Cesàro summation technique had to be abandoned. However, this technique gave some insight into the shape of the kernel function and exhibited another disadvantage.

(2)   The kernel function $R^{\psi}o(\psi)$ behaves rather erratically when approaching the argument ψ_o (see Figure 5.2). From our numerical experience we cannot call this behaviour a singularity, as all numerical values remained finite. However, this strange behaviour is probably somehow connected to the non-convergence of the Legendre series (5.29).

5.3.2.1 · Regularization of the kernel function

For the evaluation of $R^{\psi}o(\psi)$ we write formally

$$R^{\psi}o(\psi) = R^{\psi}o(\psi) * f(\psi, \psi_o) * f^{-1}(\psi, \psi_o) \qquad (5.31)$$

with

71
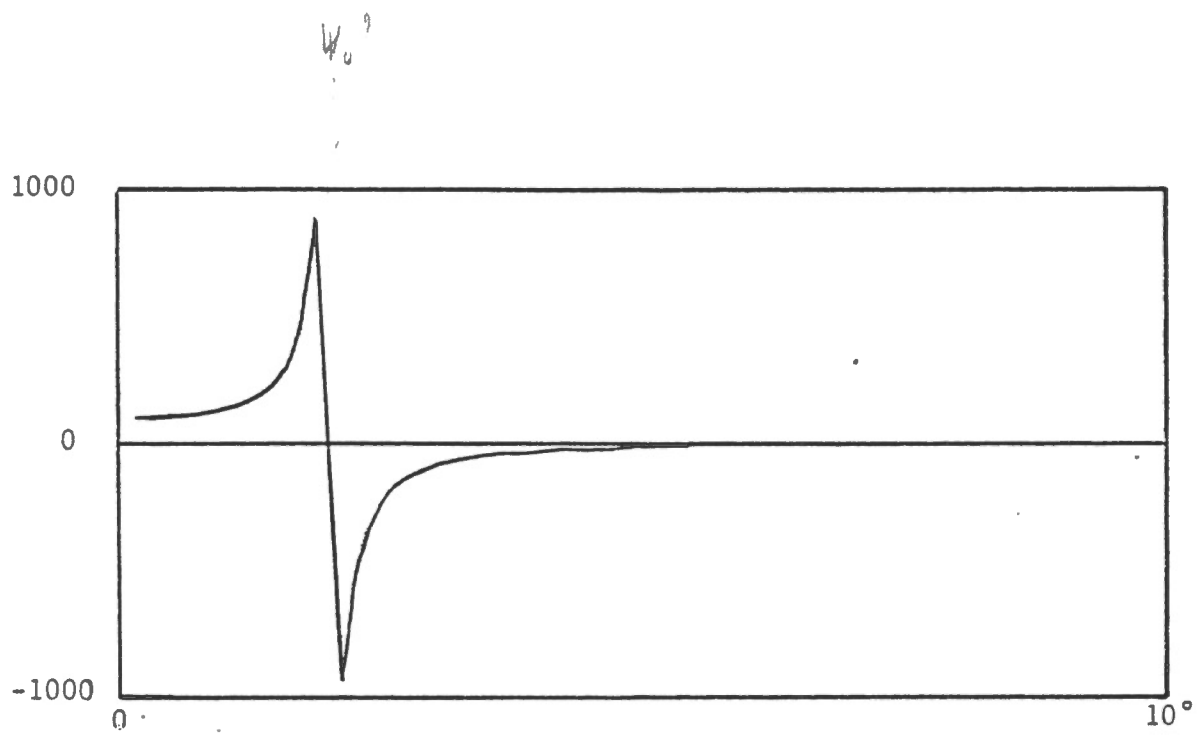
Figure 5.2: The kernel function $R^{\psi_0}(\psi)$

$$f(\psi, \psi_o) * f^{-1}(\psi, \psi_o) = 1. \tag{5.32}$$

We look for a function $f(\psi, \psi_o)$ of "simple" analytical structure that can be incorporated into the series expansion (5.29) for $R^{\psi}o(\psi)$ according to

$$R^{\psi}o(\psi)*f(\psi,\psi_o) = T^{\psi}o(\psi) = \sum_{n=0}^{\infty} (2n+1)/2\ t_n(\psi_o)\ P_n(\cos\psi) \tag{5.33}$$

and leads to convergence in the ordinary sense for the expansion (5.33).

As we do not know in advance which function $f(\psi, \psi_o)$ to choose, this is a trial and error process. Here we do not describe all the errors but proceed directly to the successful trials.

With

$$f(\psi, \psi_o) = \cos\psi - \cos\psi_o\ , \tag{5.34}$$

we obtain from eqns. (5.33) and (5.31)

$$T^{\psi}o(\psi) = \sum_{n=2}^{\infty} (2n+1)/2\ (n-1)Q_n(\psi_o)P_n(\cos\psi)(\cos\psi - \cos\psi_o)$$

$$= \sum_{n=2}^{\infty} (2n+1)/2\ (n-1)\{Q_n(\psi_o)P_n(\cos\psi)\cos\psi - \cos\psi_o Q_n(\psi_o)P_n(\cos\psi)\}. \tag{5.35}$$

We evaluate the first term in the brackets using (cf. Ben-Menahem and Singh [1981])

$$X\cdot P_l(x) = 1/(2l+1)((l+1)P_{l+1} + l\ P_{l-1}) \tag{5.36}$$

and obtain, after reorganizing, the summation

$$T^{\psi}o(\psi) = Q_2(\psi_o)P_1(\cos\psi)$$

$$+ \sum_{n=2}^{\infty} (2n+1)/2\ \{(n(n-2))/(2n+1)\ Q_{n-1}(\psi_o) + (n(n+1))/(2n+1)\ Q_{n+1}(\psi_o)$$

$$- (n-1)\cos\psi_o\ Q_n(\psi_o)\}\ P_n(\cos\psi) \tag{5.37}$$

73

$$T^{\Psi}0(\psi) = Q_2(\psi_0)P_1(\cos\psi) + \sum_{n=2}^{\infty} (2n+1)/2 \, t_n(\psi_0) \, P_n(\cos\psi) \,. \tag{5.38}$$

Comparing the coefficients of this expansion (see Figure 5.3a) with those of the original series (5.29) (see Figure 5.3b), we observe that the coefficients of $T^{\Psi}0(\psi)$ series decrease rapidly with increasing degree of expansion, whereas in the series expansion for $R^{\Psi}0(\psi)$ the coefficients increase with n. This rapid decrease of the coefficients of the $T^{\Psi}0(\psi)$ series expansion results in ordinary convergence for this series. A plot of the "regularized" kernel $T^{\Psi}0(\psi)$ is shown in Figure 5.4. For this plot, the series (5.37) was truncated at the degree N = 300; all numerical values obtained at this truncation level have converged to better than $1*10^{-5}$.

5.3.2.2 Test of the kernel $R^{\Psi}0(\psi)$

Assume the geoidal undulations N are given in an expansion of surface spherical harmonics

$$N(\psi, \alpha) = \sum_{n=2}^{\infty} N_n(\psi, \alpha) \,, \tag{5.39}$$

where n is the degree of the spherical harmonics. The gravity anomalies of the same gravity field are given by

$$\Delta g(\psi, \alpha) = \gamma/R \sum_{n=2}^{\infty} (n-1) \, N_n(\psi, \alpha) \,. \tag{5.40}$$

The Stokes integration applied to eqn. (5.40) within a spherical cap of radius ψ_0 can be written

$$N^{\Psi}0 = R/(4\pi\gamma) \int_{\psi=0}^{\psi_0} S(\psi) \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha) \, d\alpha \, \sin\psi \, d\psi \,, \tag{5.41}$$
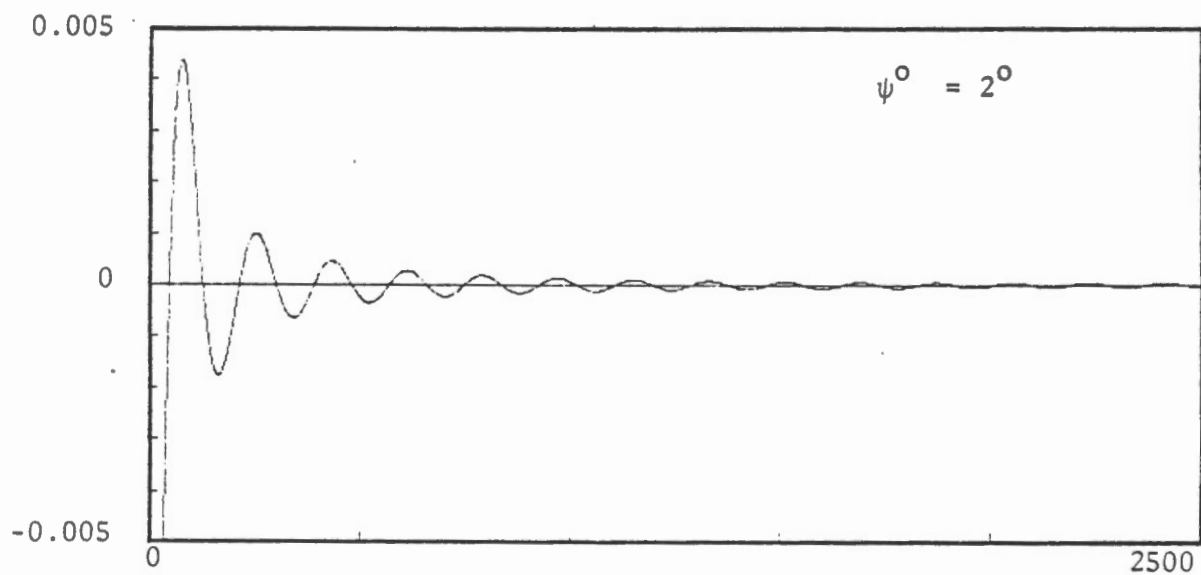
Figure 5.3a: Expansion coefficients for $T^{\psi_o}_o(\psi)$
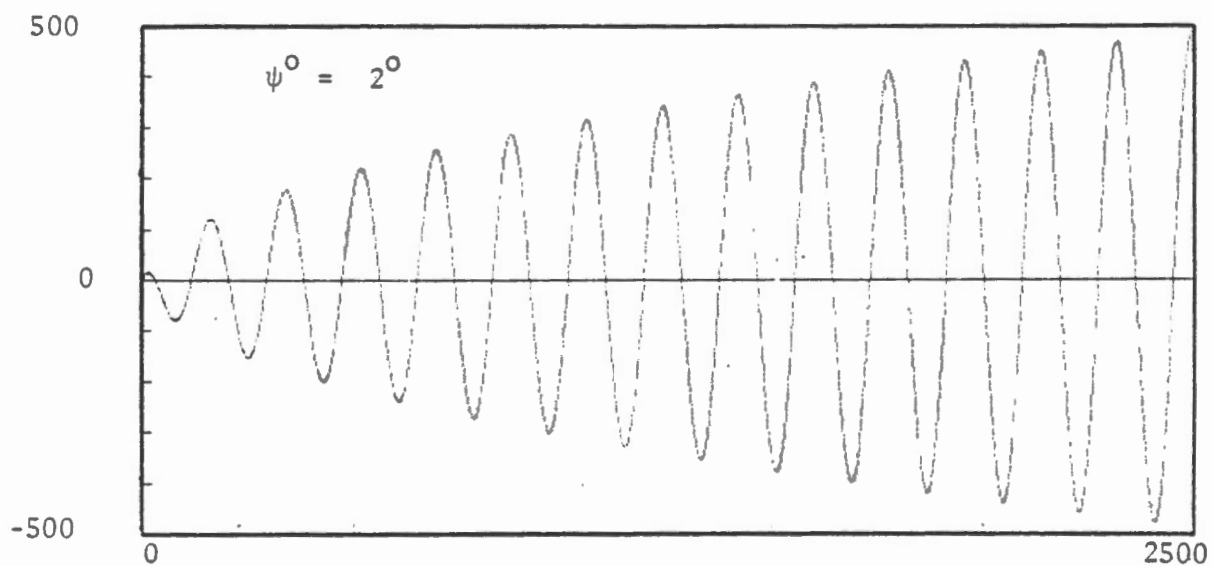
Figure 5.3b: Expansion coefficients for $R^{\psi_o}_o(\psi)$
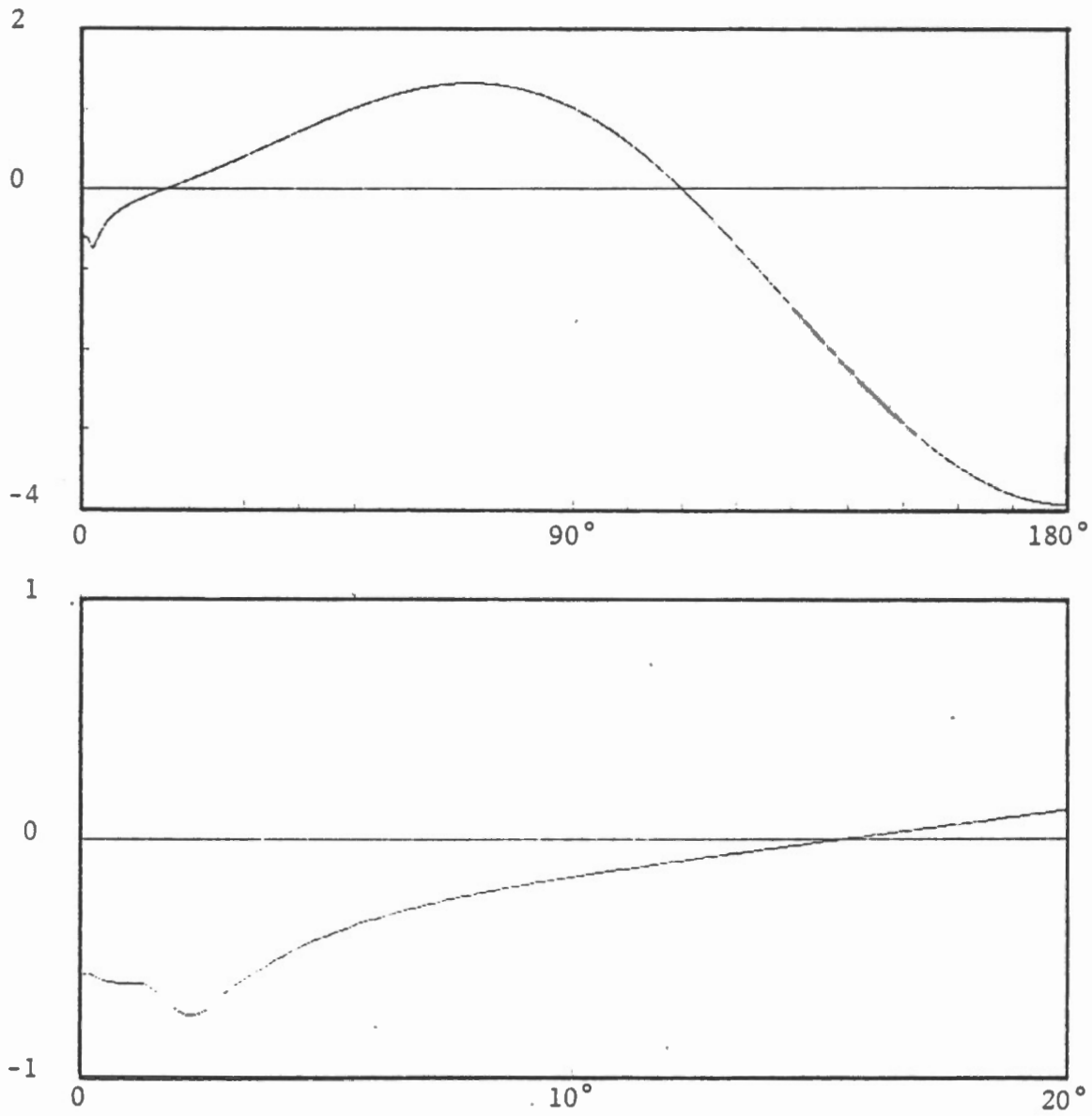
Figure 5.4:   Regularized kernel $T^{\psi_o}(\psi)$ $(\psi_o = 2^o)$

where we have chosen the spherical coordinate system (ψ, α) to be centred at the point of evaluation. Equation (5.41) can be evaluated using an expansion of the Stokes function $S(\psi)$ in Legendre's polynomials (e.g., Vaníček and Krakiwsky [1982])

$$S(\psi) = \sum_{l=2}^{\infty} (2l+1)/(l-1)\ P_l(\cos\psi) \tag{5.42}$$

and an expansion of the surface harmonics $N_n(\psi, \alpha)$ in associated Legendre's functions $P_l{}^m(\cos\psi)$ and trigonometric functions according to

$$N_l(\psi, \alpha) = \sum_{m=0}^{n} P_l{}^m(\cos\psi)(C_l{}^m \cos m\alpha + S_l{}^m \sin m\alpha)\ . \tag{5.43}$$

Inserting eqn. (5.43) in eqn (5.41) and performing the integration with respect to α, the orthogonality of the trigonometric functions in $(0, 2\pi)$ gives

$$1/2\pi \int_{\alpha=0}^{2\pi} \Delta g(\psi, \alpha)\ d\alpha \doteq \gamma/R \sum_{n=2}^{\infty} (n-1)\ C_n{}^o \cdot P_n{}^o(\cos\psi)$$

$$= \gamma/R \sum_{n=2}^{\infty} (n-1)\ C_n\ P_n(\cos\psi). \tag{5.44}$$

Equations (5.42) and (5.44) in eqn. (5.41) gives

$$N^{\psi_0} = 1/2 \int_{\psi=0}^{\psi_0} \sum_{l=2}^{\infty} (2l+1)/(l-1)\ P_l(\cos\psi) \sum_{n=2}^{\infty} (n-1)\ C_n P_n(\cos\psi)\ \sin\psi\ d\psi\ . \tag{5.45}$$

If the integration is extended over the interval $(0, \pi)$, the orthogonality relation of Legendre's polynomials

$$\int_{\psi=0}^{\pi} P_l(\cos\psi)\ P_n(\cos\psi)\ \sin\psi\ d\psi = 2/(2n+1)\ \delta_{nl} \tag{5.46}$$

77

gives

$$N^\pi = N = \sum_{n=2}^{\infty} C_n . \tag{5.47}$$

For $\psi_0 < \pi$, we exchange

$$\int_{\psi=0}^{\psi_0} S(\psi) \, f(\psi) \, \sin\psi \, d\psi = \int_{\psi=0}^{\pi} \overline{S}(\psi) \, f(\psi) \, \sin\psi \, d\psi \tag{5.48}$$

with

$$\overline{S}(\psi) = \begin{cases} S(\psi) & , \quad \psi \leq \psi_0 \\ 0 & , \quad \psi > \psi_0 . \end{cases} \tag{5.49}$$

$\overline{S}$ is evaluated in Legendre's polynomials according to eqn. (5.17)

$$\overline{S}(\psi) = \sum_{n=0}^{\infty} (2n+1)/2 \; \alpha_n(\psi_0) \, P_n(\cos\psi) \tag{5.50}$$

with

$$\alpha_0, \alpha_1 = - Q_0(\psi_0), - Q_1(\psi_0)$$

$$\tag{5.51}$$

$$\alpha_n = 2/(n-1) - Q_n \, , \, n \geq 2 .$$

Using now eqns. (5.48), (5.50) and (5.51) in eqn. (5.45), we obtain (again exploiting the orthogonality relations)

$$N^{\psi_0} = \sum_{n=2}^{\infty} (1 - (n-1)/2 \; Q_n(\psi_0)) \, C_n . \tag{5.52}$$

For $\psi_0 = \pi$, all truncation coefficients $Q_n(\pi)$ vanish and eqn. (5.52) reduces to eqn. (5.47). For $\psi_0 = 0$, the truncation coefficients are

$$Q_n(o) = 2/(n-1) \tag{5.53}$$

and, therefore,

$N^0 = 0$.

Once truncation coefficients $Q_n(\psi_o)$ have been computed, eqn. (5.52) provides an excellent method to check the results of numerical integration procedures for specific gravity field constituents C_n of degree n.

Figure 5.5 shows reference values computed from eqn. (5.52) for n=2,3,...,200, using only one coefficient C_n different from zero. This coefficient of degree n was given the value 1, i.e., the integration in $(0, \pi)$ according to eqn. (5.47) gives the integral value $N^\pi = 1$. Figure 5.5 should be interpreted in the following sense: For any gravity field constituent of degree n, which leads to a geoidal undulation of 1 m at the point of evaluation, the truncated Stokes's integration gives the value N^{ψ_o}. In the example shown in Figure 5.5, the truncated Stokes's integration estimates a geoidal undualtion of 1.4 m for a gravity field constituent of degree n = 50.

As the integration with the newly derived kernel $R^{\psi_o}(\psi)$ applied to the geoidal undulation computes the result of the truncated Stokes's integration, the $N^{\psi_o}(n)$ values of Figure 5.5 can be used as a reference.

Using the geoidal undulations (5.39), the surface spherical harmonics (5.43), the kernel evaluation (5.33), (5.31), we can write for the integration formulae derived in section 5.3.1:

$$N^{\psi_o} = N - 1/2 \int_{\psi=0}^{\pi} T^{\psi_o}(\psi) * [\cos\psi - \cos\psi_o]^{-1} *$$

$$* 1/2\pi \int_{\alpha=0}^{2\pi} \sum_{l=2}^{\infty} N_l(\psi, \alpha) \sin\psi \, d\alpha \, d\psi . \tag{5.54}$$

The integration with respect to α gives (cf. eqn. (5.44))

$$1/2\pi \int_{\alpha=0}^{2\pi} \sum_{l=2}^{\infty} N_l(\psi, \alpha) \, d\alpha = \sum_{n=2}^{\infty} C_n P_n(\cos\psi) \tag{5.55}$$
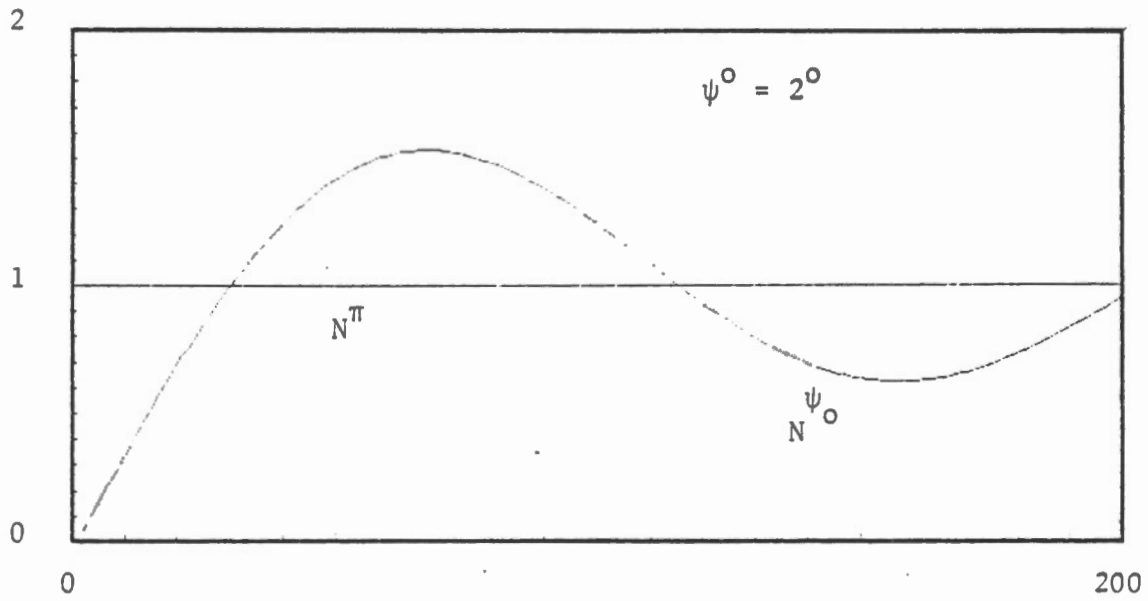
leading to

Figure 5.5:   Truncated Stokes integration results for different
gravity field constituents.

$$N^{\psi}{}_0 = N - 1/2 \int\limits_{\psi=0}^{\pi} T^{\psi}{}_0(\psi)[\cos\psi - \cos\psi_o]^{-1} \sum_{n=2}^{\infty} C_n P_n(\cos\psi) \sin\psi \, d\psi \, . \qquad (5.56)$$

Equation (5.56) has been discretized for the numerical integration

$$\overline{N}^{\psi}{}_0 = N - 1/2 \sum_{i=1}^{I} \overline{T}^{\psi}{}_0(\psi_i)[\cos\psi_i - \cos\psi_o]^{-1} \sum_{n=2}^{\infty} C_n P_n(\cos\psi_i) \sin\psi_i \, \Delta\psi_I \qquad (5.57)$$

with ·

$$\Delta\psi_I = 2\pi/I$$

$$\psi_i = (i - 1/2) \, \Delta\psi_I$$

and I is a suitably chosen number of intervals. $\overline{T}^{\psi}{}_0(\psi)$ is an approximation of $T^{\psi}{}_0(\psi)$ using the first 300 term in the series expansion (5.38). As previously described for the Stokes integration, the numerical integration (5.57) was again done for single gravity field constituents of degrees 2 through 200. The result of these integrations are undistinguishable from those shown in Figure 5.5 in the scale of that figure. Therefore, we plot in Figures 5.6 and 5.7 the absolute and relative deviations of the numerical integration results from the true values given in eqn. (5.52). These results are based on an interval width of $\Delta\psi_I = 0.1*2\pi/180°$.

5.3.2.3   Accuracy considerations

The error of the test integration results as shown in Figures 5.6 and 5.7 can be due to two reasons:
- insufficient accuracy when approximating the integral numerically;
- insufficient accuracy when evaluating the kernel function according to eqn. (5.37).

5.3.2.3.1 Accuracy of the numerical integration

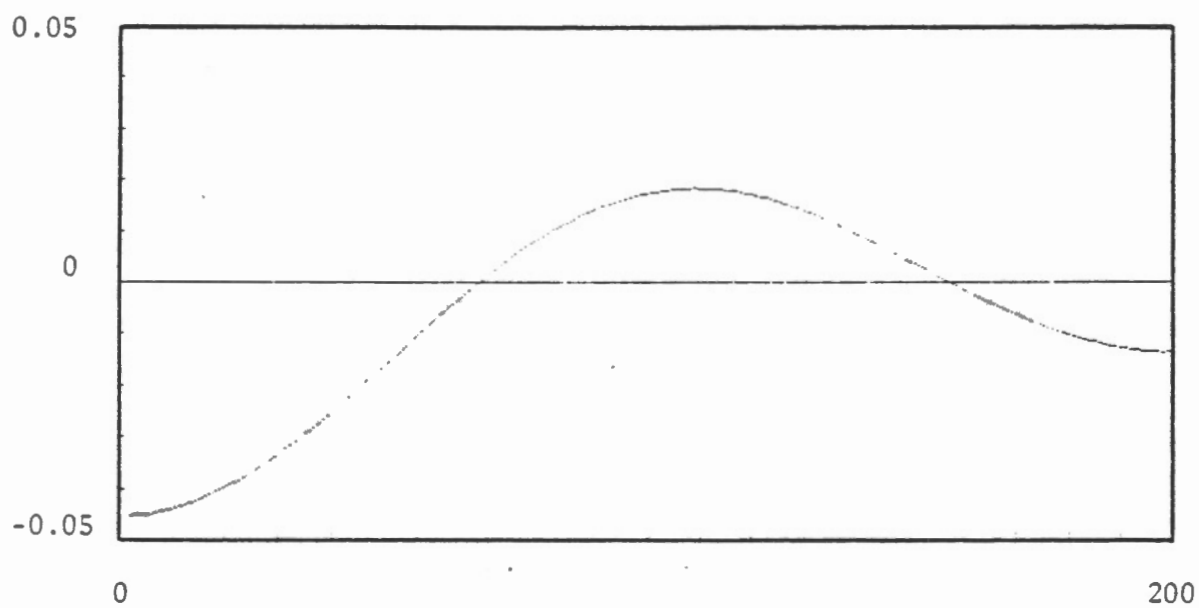The numerical integration procedure used here is a simple quadrature method according to

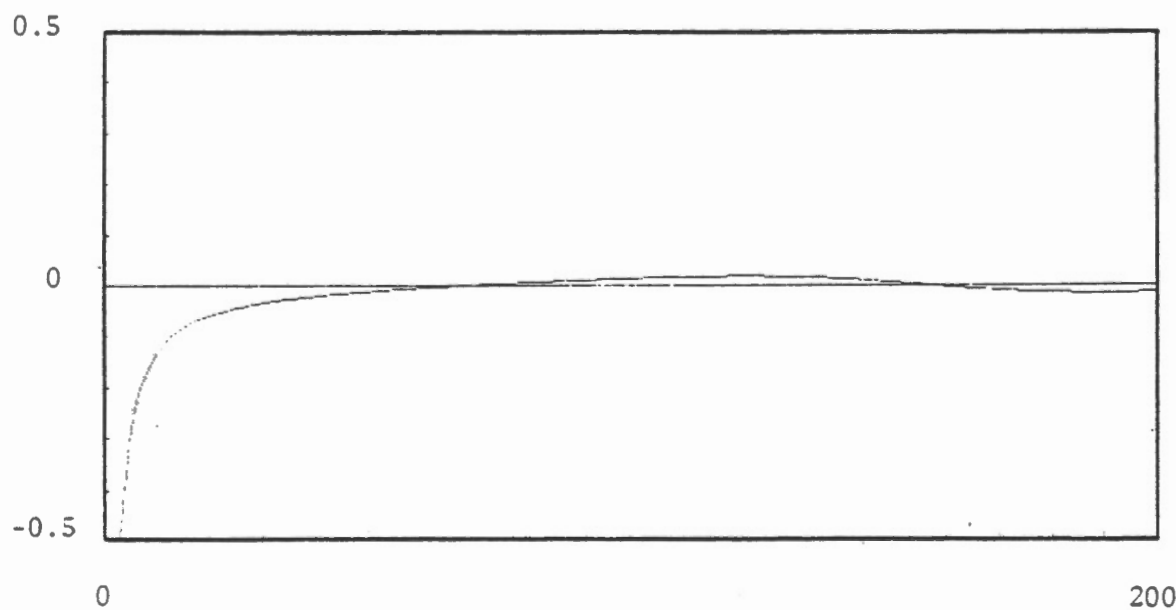Figure 5.6: Absolute error of numerical integration for $\Delta\psi = 0.1°$

Figure 5.7: Relative error of numerical integration for $\Delta\psi = 0.1°$

Figure 5.8. The integral value is approximated by a finite sum of rectangles; for well-behaved functions, the error of this procedure depends mainly on the interval width $\Delta\psi$. Figure 5.9 shows the results for integrations with $\Delta\psi = 0.2°$, $\Delta\psi = 0.1°$, and $\Delta\psi = 0.5°$. Obviously, there is no appreciable difference in the error curves of these integrations.

The function to be integrated here, however, is not as well behaved as shown in Figure 5.8. Recalling eqn. (5.56) and the shape of the regularized kernel $T^\psi o(\psi)$, the function to be integrated here looks as shown in Figure 5.10.

The symmetrical singularity dominating for ψ close to ψ_0 is introduced by the $(\cos\psi - \cos\psi_0)^{-1}$ term. Any non-symmetric arrangement of the integration intervals $\Delta\psi$ with respect to the symmetry line of the singularity will cause large errors.

Figure 5.11 shows errors of the integration result if we shift the subdivision of the total integration interval by small amounts. The nodal points of the error curve remain unchanged by this procedure; thus we can artificially remove the error for the "no shift" case by introducing small shifts.

Figure 5.12 shows results for deliberate shifting; for a shift of $0.0025°$, the integration error becomes negligibly small. Although this procedure leads to error-free integration results, it should not be used for two reasons: Firstly, there is no known mathematics or physics behind this procedure; and, secondly, the "correct" shift would have to be determined for every ψ_0 on an experimental basis. We conclude this subsection with the result that the errors in the integral values, as shown in Figures 5.6 and 5.7, are not due to the numerical evaluation of the integrals. However, methods have been found to remove these errors by using some unfounded tricks in the numerical integration procedure.

5.3.2.3.2 Accuracy of the kernel function

All computations described so far were based on an evaluation of the regularized kernel $T^\psi o(\psi)$ in terms of Legendre's polynomial up to degree 300. We evaluate now the Legendre series (5.38) up to degrees N=200, N=500, N=800 and compare the integration errors with those for N=300 used so far. The result presented in Figure 5.13 clearly indicates that there is an error in the
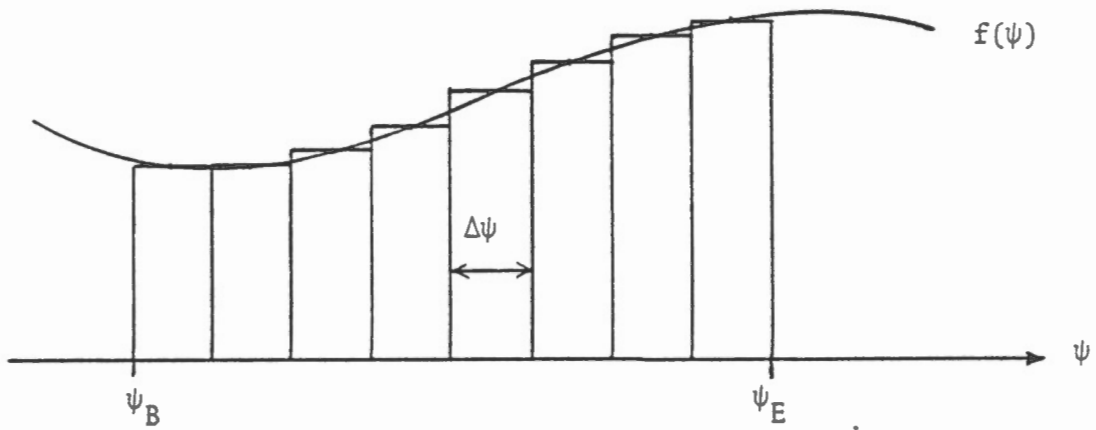
Figure 5.8: Numerical integration procedure (schematic)

OVERLAY OF ERRORS FOR NUMERICAL INTEGRATION
WITH INCREMENTS OF .05, .10 AND .20 DEG.
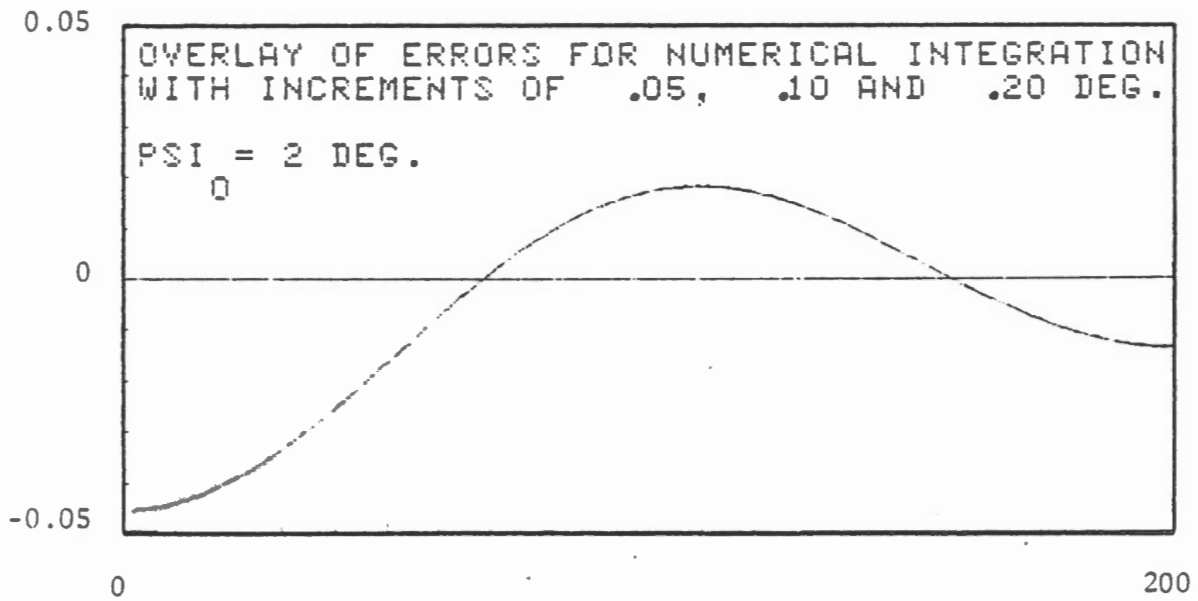
PSI = 2 DEG.
   0

Figure 5.9: Error curves for different $\Delta\psi$
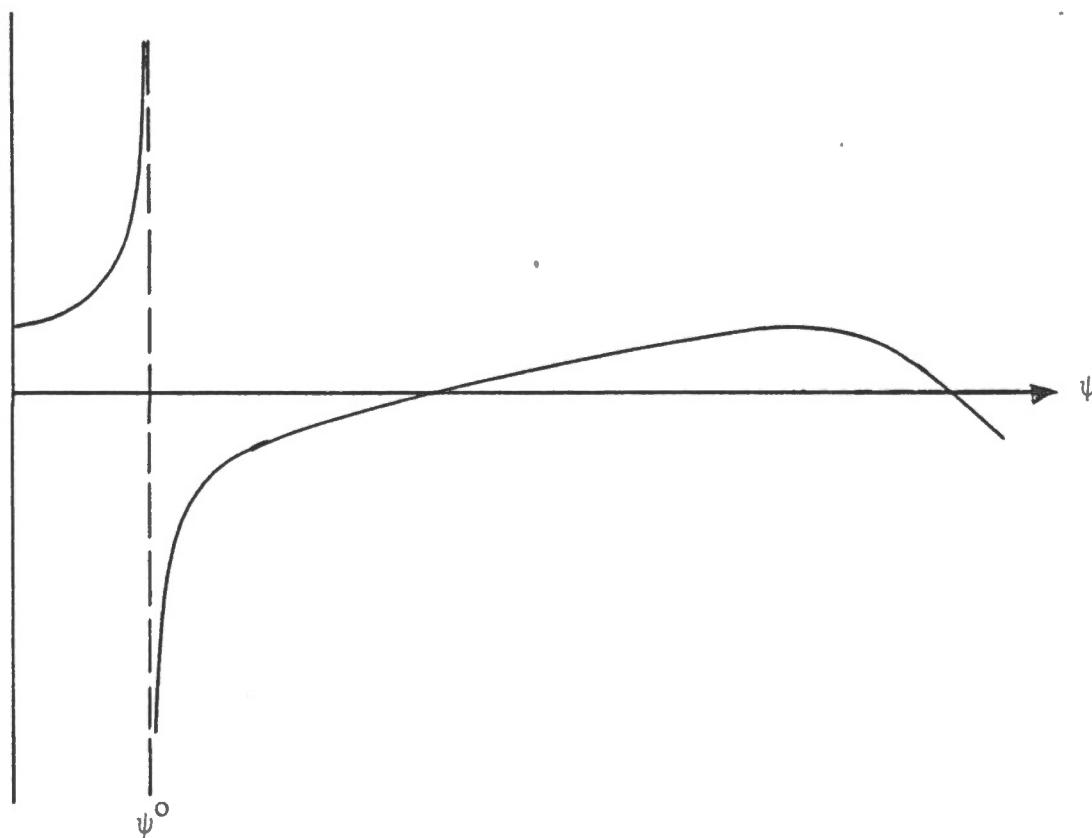
Figure 5.10: Shape of function to be integrated (schematic)
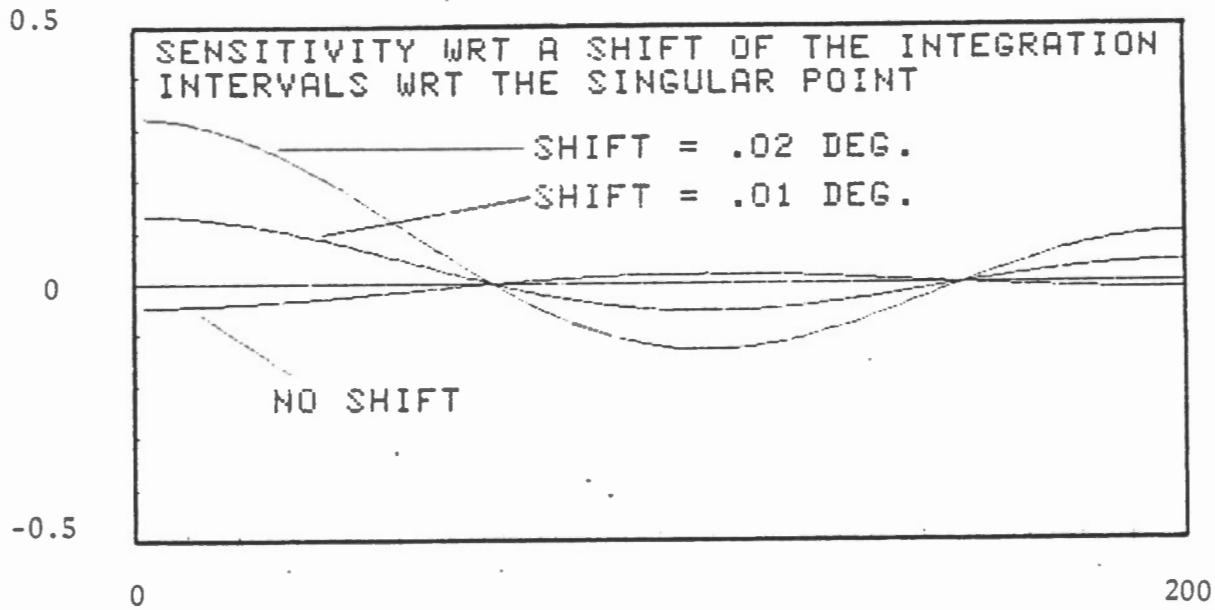
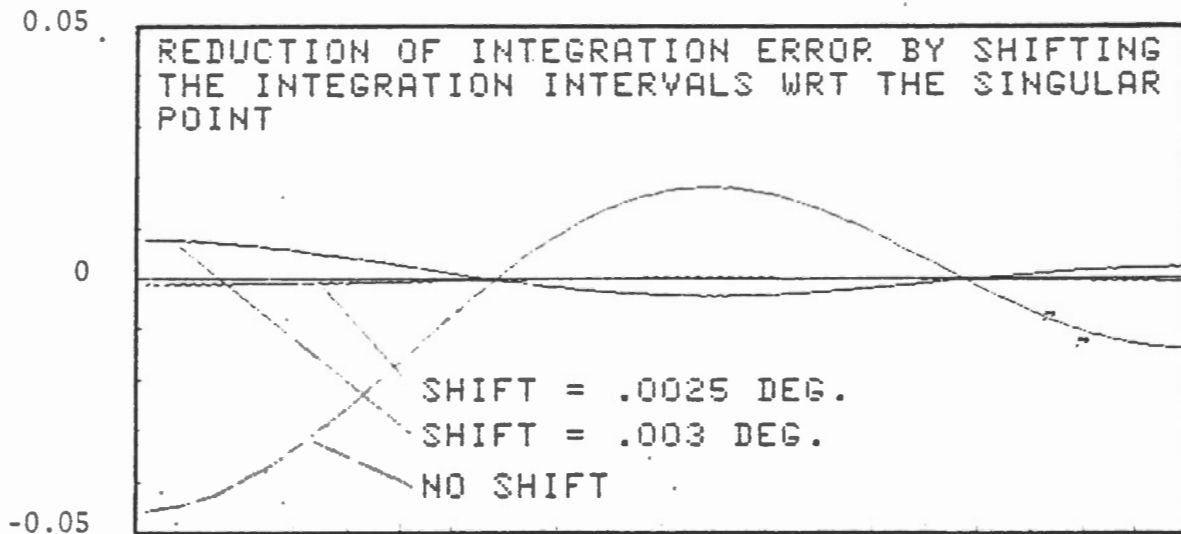Figure 5.11: Error curves for shifted integration intervals

Figure 5.12: Reduction of integration error by deliberate shifting of the integration intervals

statement in section 5.3.2.1 of this report. We stated in that section that all values of the kernel expansion truncated at degree 300 had converged to 10^{-5}. It was found upon closer inspection that the convergence is not that good for $\psi < 5°$. However, from Figure 5.13, we see that for any gravity field constituent of degree $n > 20$, the recovery of the truncated Stokes's integration by means of the new kernel is better than 5%, if we evaluate the Legendre series (5.38) up to degree 500. According to Kaula's rule of thumb [Kaula, 1966], the coefficients C_n of the terrestrial gravity field behave roughly as

$$C_n \sim \pm 10^{-5}/n^2 . \tag{5.58}$$

That relates to the amplitudes T_n of the corresponding potential field constituents according to

$$T_n \sim \pm GM/R \cdot 10^{-5}/n^2 . \tag{5.59}$$

Potential field constituents relate to geoidal undulations according to

$$N_n = 1/\gamma \, T_n \tag{5.60}$$

and with

$$\gamma = GM/R^2 \tag{5.61}$$

we obtain, from eqns, (5.59) and (5.60)

$$N_n \sim \pm 64 \; m \cdot n^{-2}$$

$$N_n \sim < \pm 0.2 \; m \; for \; n > 20 \tag{5.62}$$

indicating that a 5% integration accuracy is good for ± 1 cm in recovered geoidal height constituent of degree $n > 20$.

Figures 5.11 through 5.13 show two nodal points for the error curves at $n \sim 68$ and $n \sim 148$. These nodal points indicate that the error in the recovery of a gravity field constituent of degree 68 or 148 is zero, independent of the quality of our computation (Figure 5.13). The underlying reason is rather simple: Spherical harmonics of degree 68 and 148 have nodal lines at a spherical distance of 2°, see Figure 5.14. This leads to extremal contributions for the truncated Stokes's integration (cf. Figure 5.5), but the contribution to the integral value (5.57) is zero at the singular point $\psi = \psi_0$. This indicates that the integration error discussed in this section results mainly from inaccuracies in the kernel function evaluation for $\psi = \psi_0$.
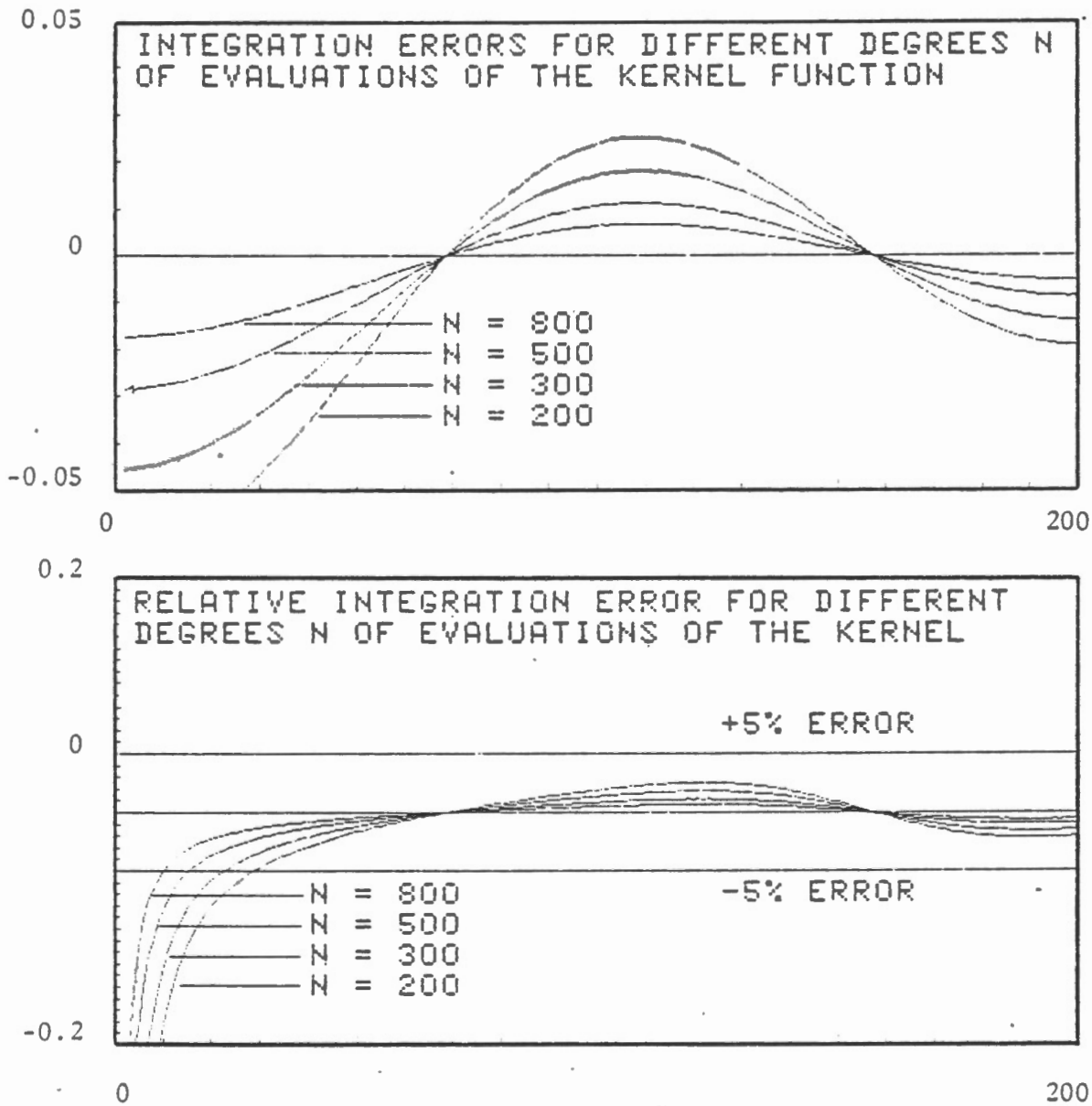
Figure 5.13: Integration errors for different degrees of kernel
function evaluation.

LEGENDRE POLYNOMIALS AT PSI=2 DEG
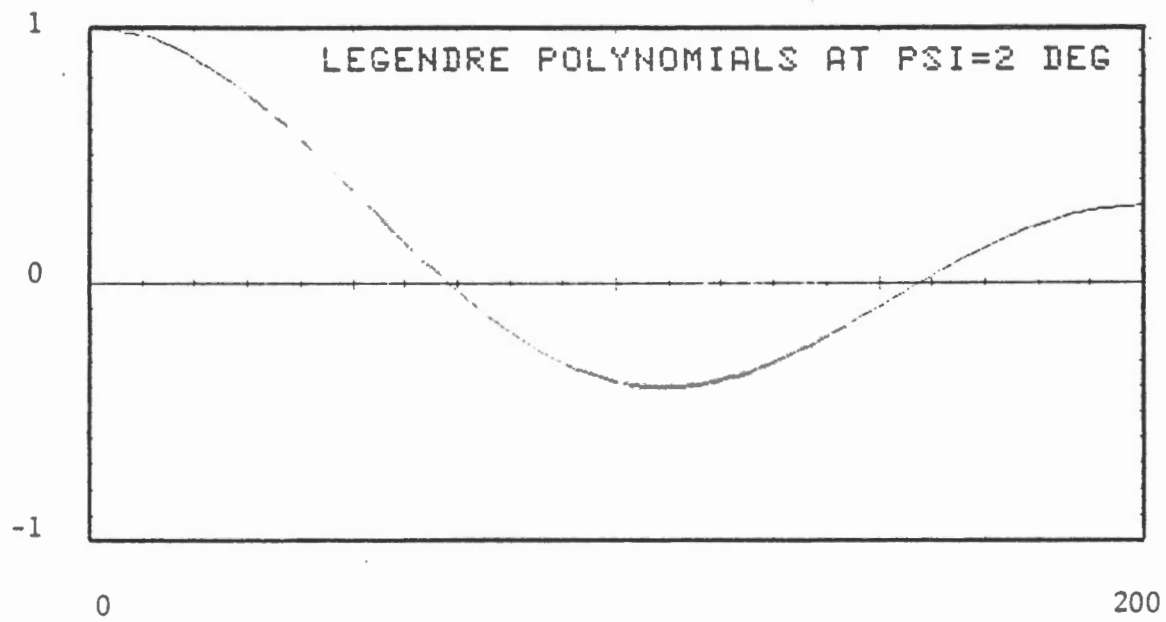
Figure 5.14:   Legendre polynomials for $\psi = 2^{\circ}$

5.3.2.4 Improvement and stabilization of the numerical integration procedure

Summarizing the preceding sections, we can state:

(a)   the integrant in eqn. (5.56) has a simple singularity for $\psi = \psi_o$;

(b)   the integral value is finite;

(c)   any geoid constituent of degree n, whose value is zero for $\psi = \psi_o$, does not produce any

integration error, independent of the slope of this geoid constituent at $\psi = \psi_o$ (Figures 5.13 and 5.14).

(a) and (b) indicate that we have treated the integration rather clumsily up to now, and (c) gives us some hint on how to treat it better.

Figures 5.13 and 5.14 show that no integration error occurs, if all gravity field constituents in the integration (5.56) are zero for $\psi = \psi_o$. (Of course, the actual gravity field coefficients C_n are not zero!)

We recall the original form of the integration kernel, eqn. (5.29)

$$R^{\Psi_o}(\psi) = \sum_{n=2}^{\infty} ((2n+1)(n-1))/2 \; Q_n(\psi_o) \, P_n(\cos\psi) . \tag{5.63}$$

The orthogonality of Legendre's polynomials in $(0, \pi)$ gives

$$\int_{\psi=0}^{\pi} R^{\Psi_o}(\psi) \sin\psi \; d\psi = 0 . \tag{5.64}$$

Thus we can write

$$\overline{N}(\psi_o) \int_{\psi=0}^{\pi} R^{\Psi_o}(\psi) \sin\psi \; d\psi = 1/2\pi \int_{\alpha=0}^{2\pi} N(\psi_o, \alpha) \, d\alpha \int_{\psi=0}^{\pi} R^{\Psi_o}(\psi) \sin\psi \; d\psi$$

$$= \sum_{n=2}^{\infty} C_n P_n(\cos \psi_o) \int_{\psi=0}^{\pi} R^{\Psi_o}(\psi) \sin\psi \; d\psi = 0. \tag{5.65}$$

Since the mean geoid undulation $\overline{N}(\psi_0)$ for ψ_0 is independent of ψ, we can write

$$1/2 \int_{\psi=0}^{\pi} R^{\psi}{}_0(\psi) \sum_{n=2}^{\infty} C_n P_n(\cos\psi_0) \sin\psi \, d\psi = 0 , \qquad (5.66)$$

and replacing $R^{\psi}{}_0(\psi)$ by

$$R^{\psi}{}_0(\psi) = T^{\psi}{}_0(\psi)(\cos\psi - \cos\psi_0)^{-1}$$

we obtain

$$1/2 \int_{\psi=0}^{\pi} T^{\psi}{}_0(\psi)(\cos\psi - \cos\psi_0)^{-1} \sum_{n=2}^{\infty} C_n P_n(\cos\psi_0) \sin\psi \, d\psi = 0 . \qquad (5.67)$$

Adding eqn. (5.67) to eqn. (5.56), we get

$$N^{\psi}{}_0 = N - 1/2 \int_{\psi=0}^{\pi} T^{\psi}{}_0(\psi) \, (\overline{N}(\psi) - \overline{N}(\psi_0))/(\cos\psi - \cos\psi_0) \sin\psi \, d\psi \qquad (5.68)$$

$$\overline{N}(\psi) = \sum_{n=2}^{\infty} C_n P_n(\cos\psi) = 1/2\pi \int_{\alpha=0}^{2\pi} N(\psi, \alpha) \, d\alpha . \qquad (5.69)$$

It remains to be shown that the integrant in eqn. (5.68) now remains finite for $\psi \to \psi_0$. Since the

kernel $T^{\psi}{}_0(\psi)$ is well behaved, this reduces to an examination of the function

$$F(\psi, \psi_0) = (\overline{N}(\psi) - \overline{N}(\psi_0)) / (\cos\psi - \cos\psi_0) . \qquad (5.70)$$

Using eqn. (5.69), we rewrite eqn. (5.70)

$$F(\psi,\psi_0) = \sum_{n=2}^{\infty} C_n F_n(\psi,\psi_0) = \sum_{n=2}^{\infty} C_n(P_n(\cos\psi)-P_n(\cos\psi_0))/(\cos\psi-\cos\psi_0) . \qquad (5.71)$$

91

With

$$\cos\psi = X \ ,$$

we write

$$F_n(X, X_0) = (P_n(X) - P_n(X_0))/(X - X_0) \ , \tag{5.72}$$

and a Taylor expansion of $P_n(X)$ at $X = X_0$ gives

$$F_n(X, X_0) = \sum_{i=1}^{\infty} 1/i! \ (\partial^i P_n(X))/\partial X^i \ \bigg|_{X=X_0} (X-X_0)^{i-1} \ . \tag{5.73}$$

Since all derivatives of Legendre's polynomials are finite, we have shown with eqn. (5.70) through eqn. (5.73) that the integrant in eqn. (5.68) remains finite for $\psi \rightarrow \psi_0$.

The measures described in this section improve the error characteristics and the stability of the numerical integration dramatically.

Figures 5.15 and 5.16 show the error of the numerical integration using the stabilized integrator (5.65) and a kernel $T^{\psi}o(\psi)$ developed up to degree $N = 800$. Thus the error plans should be compared to the "best curves" of Figure 5.13.

For low degree gravity field constituents ($n < 20$), the improvement is better than two orders of magnitude; for higher degree constituents, the improvement is still better than one order of magnitude.

The reason for the zig-zag behaviour of the error curve is unknown; as the errors are below any reasonable threshold, it is unnecessary to investigate this further.

The rather large integration errors shown in Figure 5.13 are completely due to the fact that we computed explicitly the singular integrant and eliminated this singularity by numerical integration, i.e., summation of the (almost) singular integrant. Any error in the numerical evaluation of the kernel was in this way very strongly amplified and resulted in large integration errors.

The same reason is underlying the error curves shown in Figure 5.17. Any small error introduced by a slightly asymmetric evaluation of the kernel function with respect to the singular point results in large integration errors. As the stabilized integrator has no singular integrant, a shift of the integration intervals with respect to ψ_0 does not result in integration errors different from those shown in Figure 5.15.
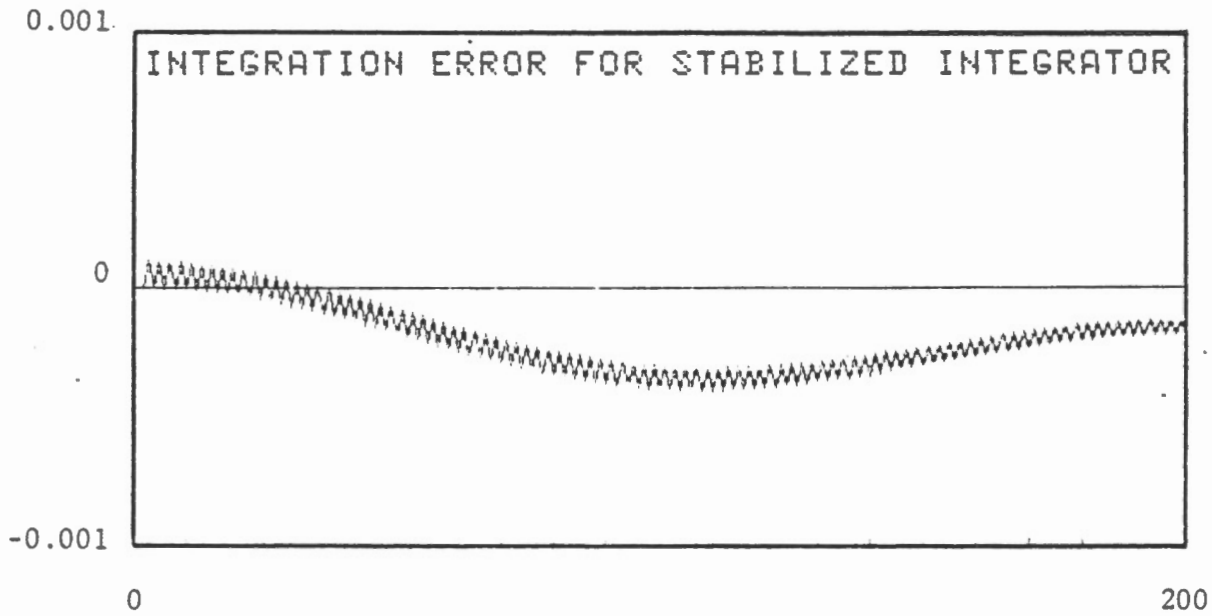
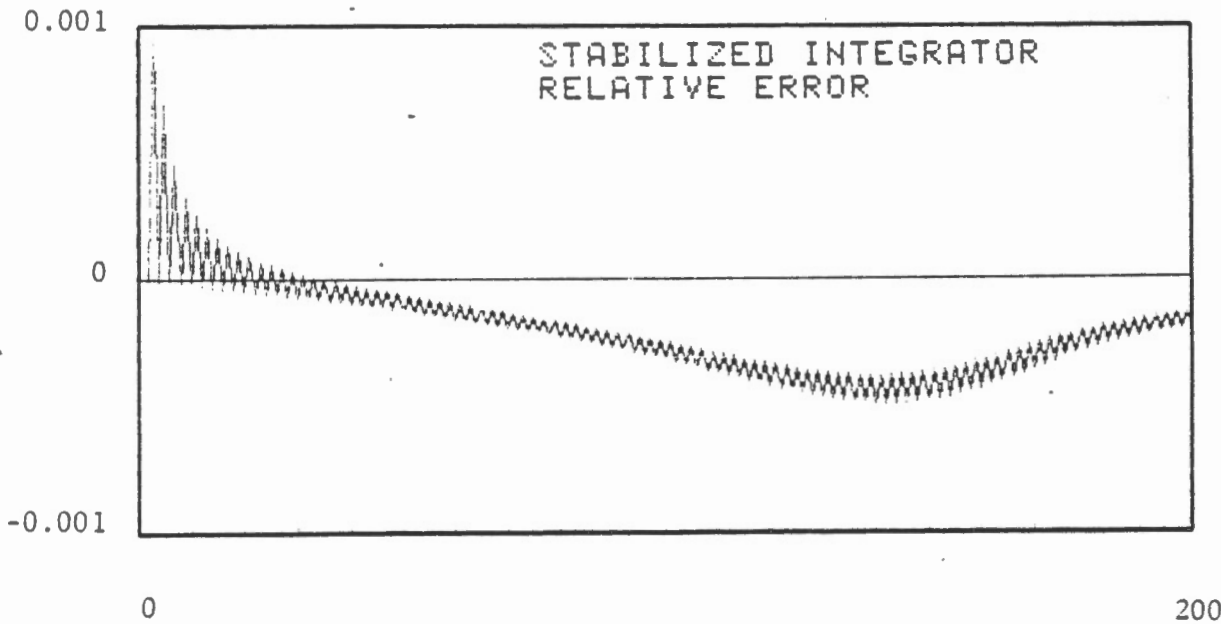Figure 5.15: Absolute integration error for stabilized integrator

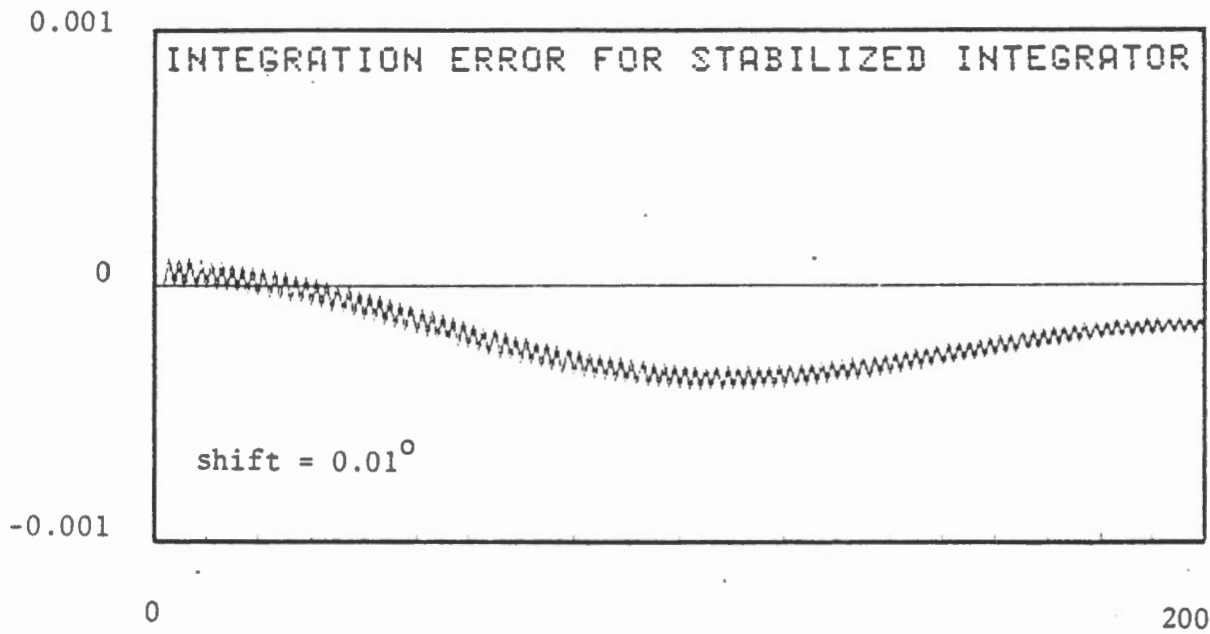Figure 5.16: Relative integration error for stabilized integrator

Figure 5.17: Absolute integration errors for stabilized integrator and shifted integration intervals.

5.3.3 Computational problems

Altimetry data is not available all over the globe, and to limit the computational burden we want to limit the integration domain to a spherical cap of some radius $\phi_0 (\phi_0 > \psi_0)$:

$$N^{\psi_0,\phi_0} = N - 1/2 \int\limits_{\psi=0}^{\phi_0} T^{\psi_0}(\psi) \, (\overline{N}(\psi) - \overline{N}(\psi_0))/(\cos\psi - \cos\psi_0) \sin\psi \, d\psi . \qquad (5.74)$$

For gravity field constituents up to degree n = 200, and several radii $\phi_0 = 10°, 20°, 30°, 60°, 90°,$ 120°, eqn. (5.74) has been evaluated. Results are plotted in Figure 5.18. It can be seen that results obtained for $\phi_0 = 10°, 20°, 30°$, are totally unacceptable. Results for $\phi_0 = 60°, 90°, 120°$, are closer to the true values ($\phi_0 = 180°$), but their errors are still too large. We see two nodal points for the error curves, i.e., for two gravity field constituents, we get the correct result independent of the integration boundary ϕ_0. The two nodal points appear at the same place as in Figure 5.13. We also observe in Figure 5.18, that all integration errors change their sign between $\phi = 60°$ and $\phi_0 = 90°$. Because of the structure of the error curves (nodal points), we can expect to have an errorless integration for some $60° < \phi_0 < 90°$. Do we have anything special between 60° and 90°? By expecting the regularized kernel function $T^{\psi_0}(\psi)$ (Figure 5.4), we find a maximum for $\psi \sim 71.5°$. Trying this value for the limit of the integration domain, ϕ_0, we obtain the results plotted in Figure 5.19. This figure is an overlay of the exact values ($\phi_0 = 180°$) and those obtained for $\phi_0 = 71.5°$. Only for very low degree gravity field constituents we see a difference between the two curves. All high-degree constituents are almost completely recovered.

The results shown in Figure 5.19 have been double checked by integrating eqn. (5.74) in 71.5° $< \psi \le 180°$. Results are shown in Figure 5.20 and prove that there is only a small contribution for low-degree gravity field constituents.

From Figures 5.19 and 5.20, we learn that integration up to the first positive maximum of the
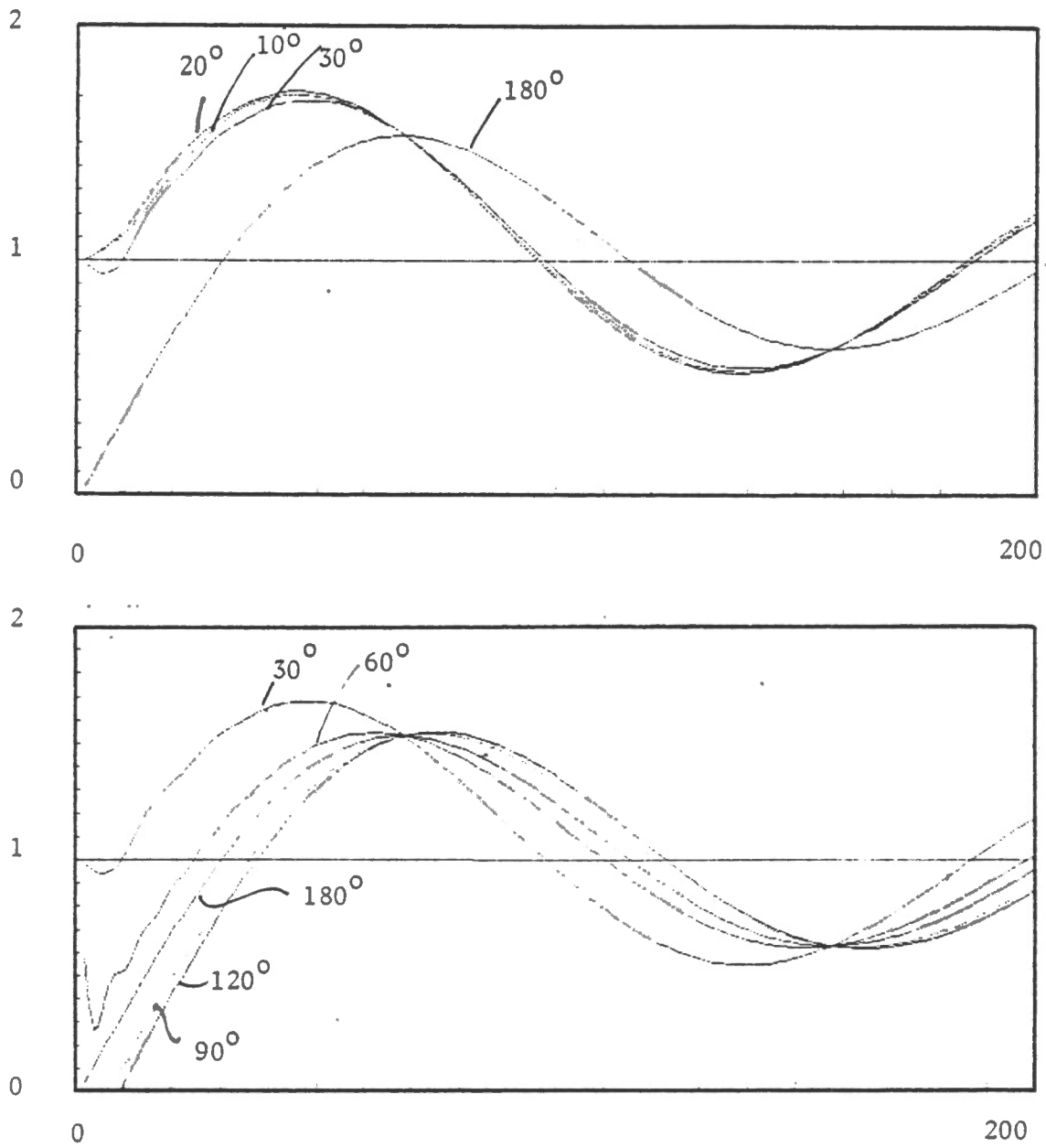
Figure 5.18:   Integration results for different truncation radii
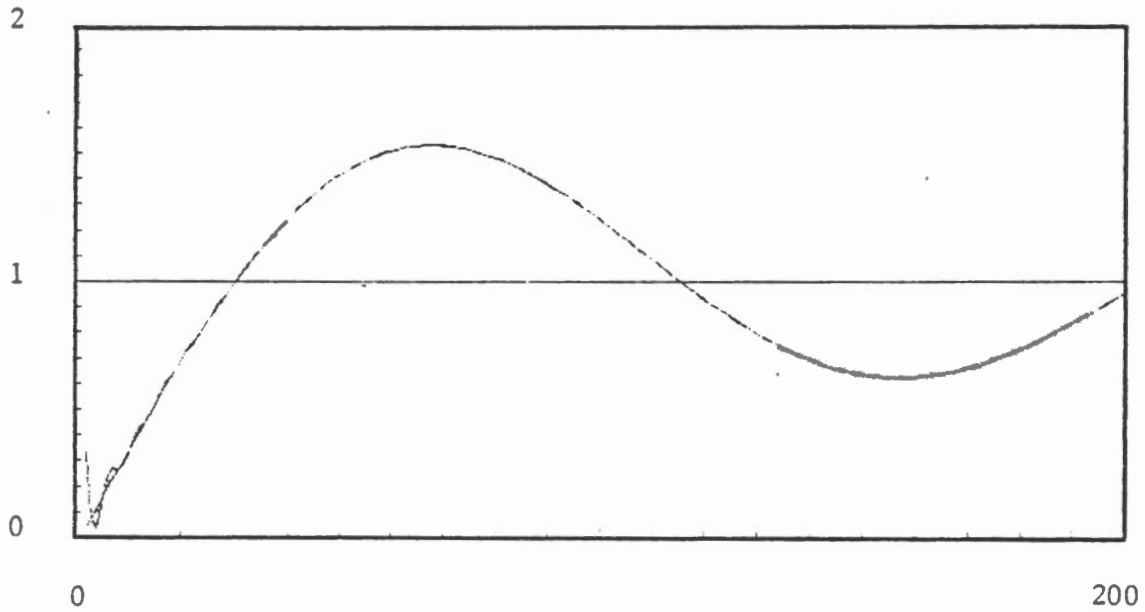
Figure 5.19: Integration results for truncated integration
$(\phi_o = 71.5^o)$ and non truncated integration.
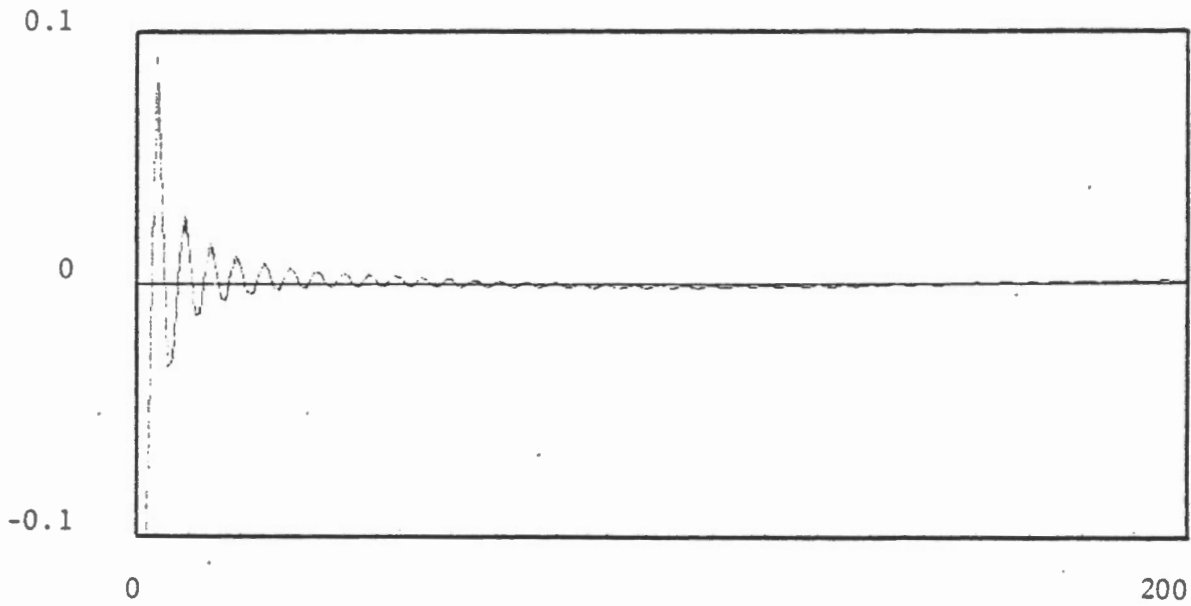
Figure 5.20: Integration results for the domain $71.5^o \leq \psi \leq 180^o$

kernel function $T^{\psi}o(\psi)$ gives reasonable results for high-degree gravity field constituents. The function $T^{\psi}o(\psi)$ is obviously dominated by low-degree constituents (cf. Figure 5.1). Since we cannot expect to obtain a reasonable result for low-degree gravity field constituents from any local integration, we may think of removing the low-degree harmonics from the kernel function. Thus we obtain from eqn. (5.27)

$$N^{l,\psi}{}_0 = N_l - 1/2 \int\limits_{\psi=0}^{\pi} R_l^{\psi}o(\psi) \; \overline{N}(\psi) \sin\psi \, d\psi \qquad (5.75)$$

with

$$R_l^{\psi}o(\psi) = \sum_{n=l+1}^{\infty} (2n+1)/2 \, (n-1) \, Q_n(\psi_o) \, P_n(\cos\psi) .$$

Regularizing this 'spheroidal' kernel $R_l^{\psi}o(\psi)$ according to the derivations of section 5.3.2.1, we obtain

$$R_l^{\psi}o(\psi) = 1/(\cos\psi - \cos\psi_o) \, T_l^{\psi}o(\psi) , \qquad (5.76)$$

with

$$T_l^{\psi}o(\psi) = \sum_{n=l+1}^{\infty} (2n+1)/2 \, t_n(\psi_o) \, P_n(\cos\psi) +$$

$$+ (l+1)/2 \, \{l \cdot Q_{l+1}(\psi_o) \, P_l(\cos\psi) - (l-1) \, Q_l(\psi_o) \, P_{l+1}(\cos\psi)\} \qquad (5.77)$$

and $t_n(\psi_o)$ according to eqns (5.37) and (5.38). This kernel $T_l^{\psi}o(\psi)$ has been computed for $l=20$, and its shape is shown in Figure 5.21. The promising results obtained when integrating up to the first maximum of the kernel $T^{\psi}o(\psi)$ ($\phi_0 \sim 71.5°$) led to an integration with $T_{20}^{\psi}o(\psi)$ up to $\phi_0=8.3°$, the first positive maximum of $T_{20}^{\psi}o(\psi)$.

The results of this integration are shown in Figure 5.22. For degrees $l > 40$, the gravity field constituents are recovered remarkably accurately by using the truncated integration up to $\phi_0 = 8.3°$.
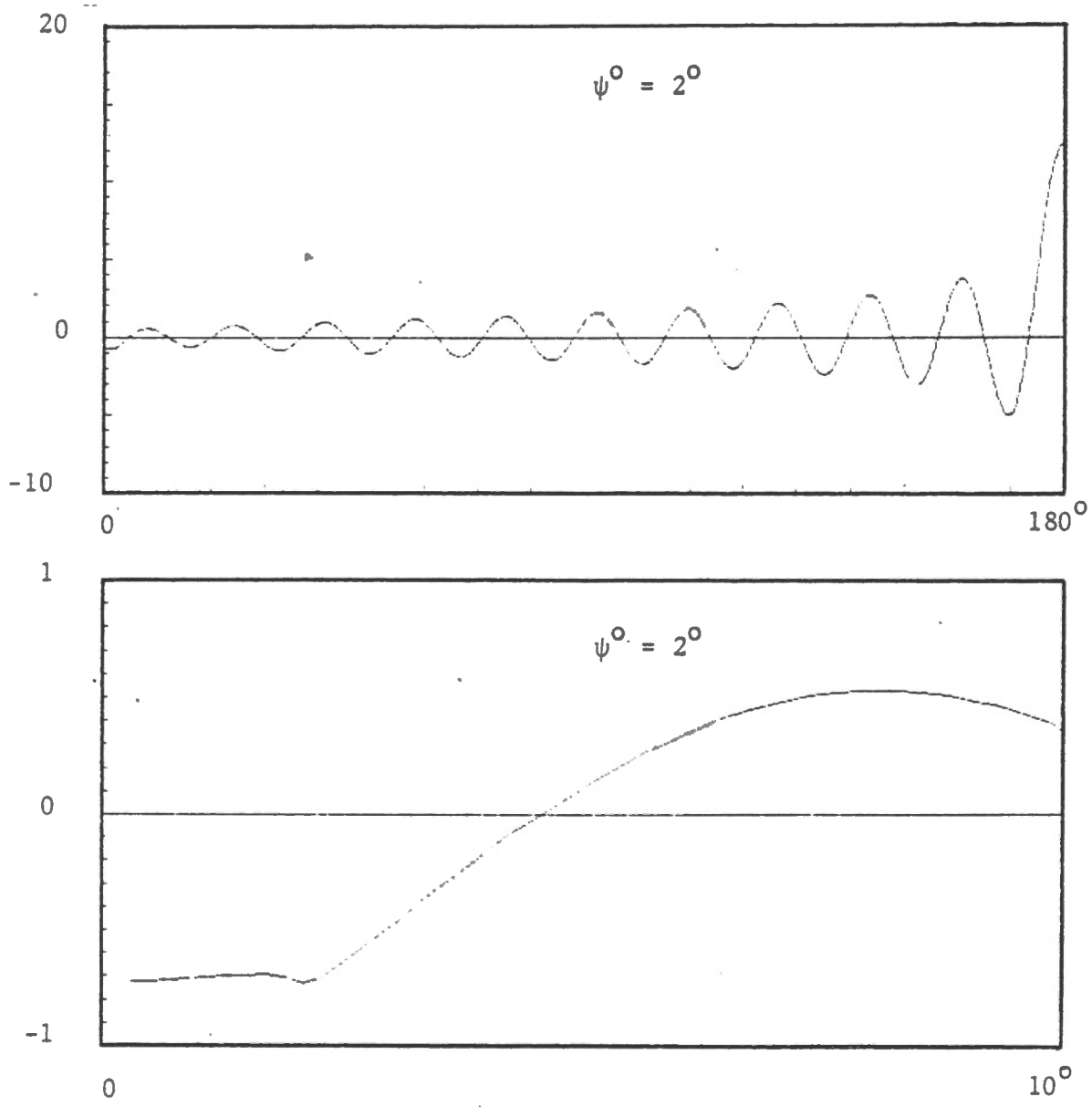
Figure 5.21: Shape of $T_{20}^{\psi_0}(\psi)$
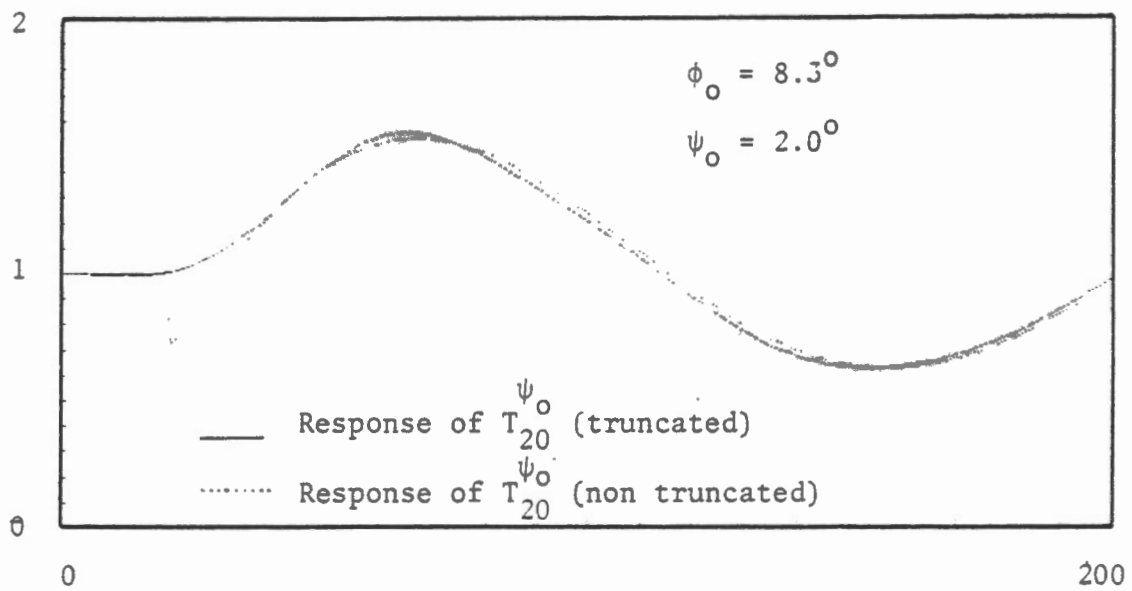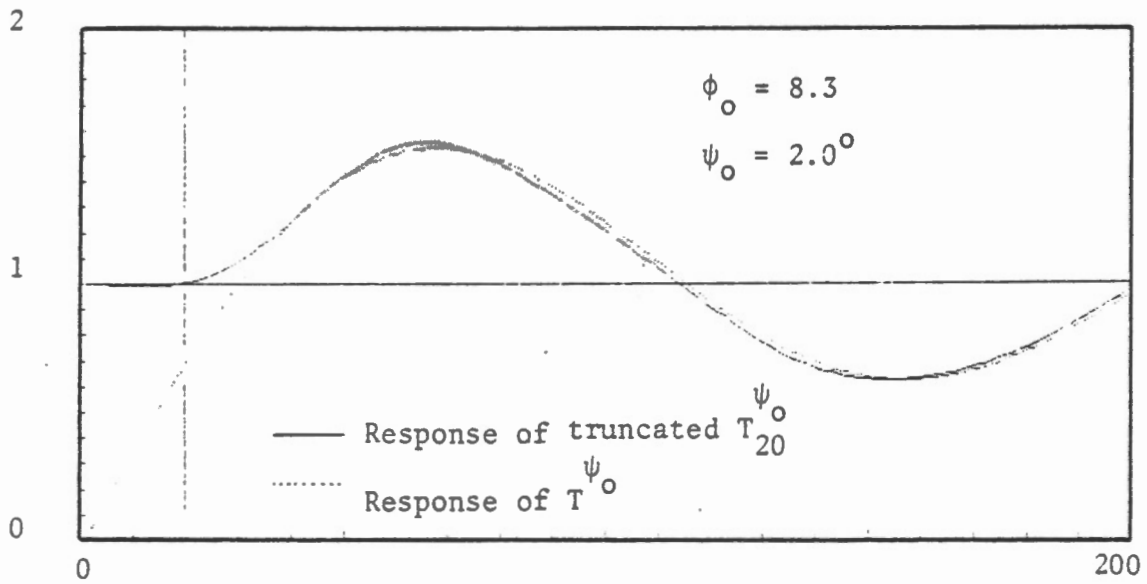
Figure 5.22:   Response of truncated and non truncated $T_{20}^{\psi_o} (\psi)$

The differences between the exact values and the truncated integration results may be due to slight errors in the computation of ϕ_0. We conclude:

> Integration with the 'spheroidal' kernel $T_{20}{}^{\Psi}o$ up to its first maximum ($\phi_0 = 8.3°$) gives acceptable results for gravity field constituents of degree $l > 40$. For $T_{10}{}^{\Psi}o$,
>
> the constituents of degree $l > 20$ are recovered when integrating up to $\phi_0 = 13.5°$ (cf. Figure 5.23).

If we look at the plot of the 'spheroidal' kernel $T_{20}{}^{\Psi}o(\psi)$ in Figure 5.21 we see a behaviour incompatible with Figure 5.3a, which shows the coefficients of the series expansion for $T^{\Psi}o$. A closer inspection of eqn. (5.77) shows that the kernel $T_l{}^{\Psi}o(\psi)$ is dominated by the term

$$(l+1)/2 \; \{l \; Q_{l+n}(\psi_0) \; P_l(\cos\psi) - (l-1) \; Q_l(\psi_0) \; P_{ln}(\cos\psi)\} . \tag{5.78}$$

This fact becomes obvious if we plot the first part of eqn. (5.77) only

$$\overline{T}_{20}{}^{\Psi}o(\psi) = \sum_{n=l+1}^{\infty} (2n+1)/2 \; t_n(\psi_0) \; P_n(\cos\psi) . \tag{5.79}$$

Figure 5.24 shows the rapid diminishing of this part of the kernel $T_{20}{}^{\Psi}o(\psi)$. However, since eqn. (5.78) is part of the kernel, we cannot simply neglect it but must look for ways to minimize it.

Since

$$P_l(-1) - P_{l+1}(-1) = \pm 2 , \tag{5.80}$$

the effect of eqn. (5.23) at $\psi = \pi$ can be minimized by

$$l \; Q_{l+1}(\psi_0) + (l-1) \; Q_l(\psi_0) = 0 . \tag{5.81}$$

Equation (5.81) determines a radius ψ_0 of truncation for the Stokes integration that minimizes the influence of the term (5.78) on the distant zones of the kernel $T_l{}^{\Psi}o(\psi)$.

For $l=20$, we obtain from eqn. (5.81)

$$\psi_0 = 2.936° . \tag{5.82}$$
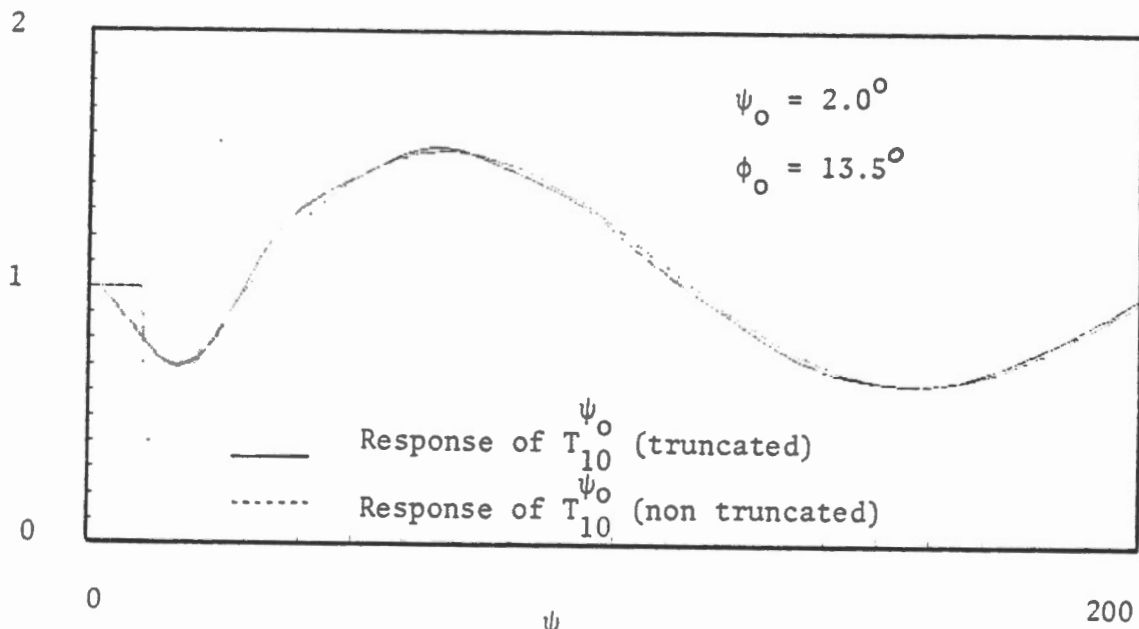
Figure 5.23: Response of $T_{10}^{\psi_o}$

$\psi_o = 2.0^o$

$\phi_o = 13.5^o$

Response of $T_{10}^{\psi_o}$ (truncated)

Response of $T_{10}^{\psi_o}$ (non truncated)
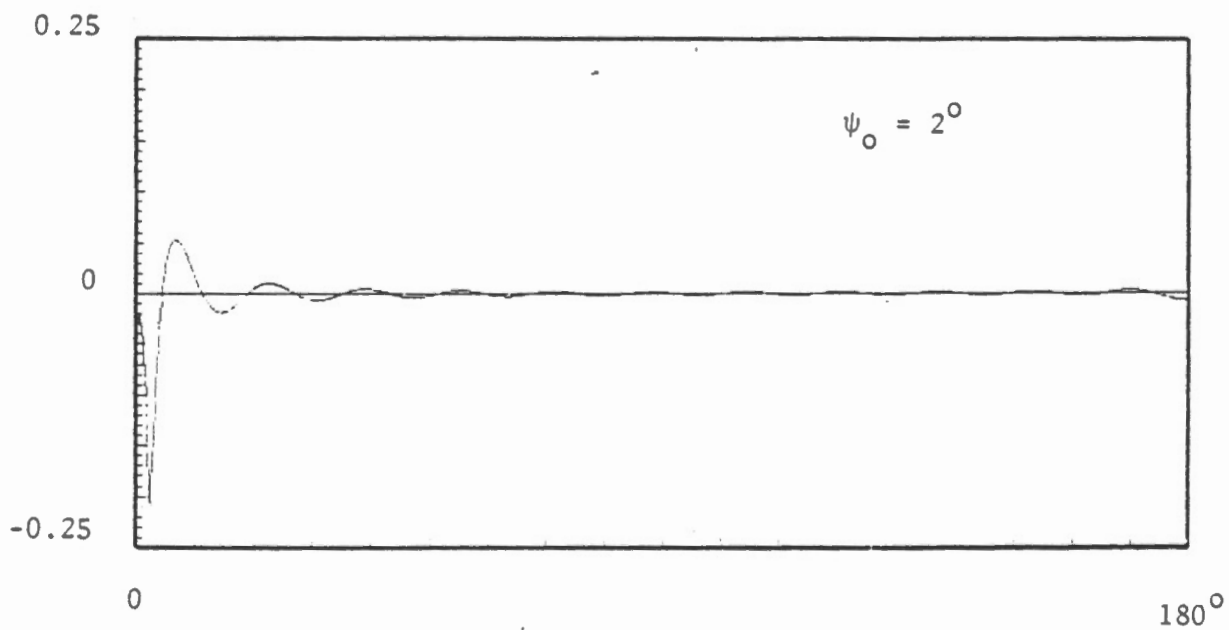
Figure 5.24: Plot of $\overline{T}_{20}^{\psi_o}(\psi)$

$\psi_o = 2^o$

Using this ψ_0, we obtain a $T_{20}\Psi o(\psi)$ shown in Figure 5.25.

Comparing Figure 5.25 with Figure 5.21, we see

- the dominating term eqn. (5.78) has been dramatically reduced;

- the shape and size of the kernel for small ψ is not very much different; the first maximum has

  been shifted from $\phi_0 = 8.3°$ to $\phi_0 = 10.2°$.

Integrating with this kernel, once in the interval (o, π) and once up to $\phi_0 = 10.2°$, we obtain the results shown in Figure 5.26. Comparing Figures 5.26 and 5.22, we see that (besides a change of the general shape of the response function) the main result of minimizing the term (5.78) is a removal of the discontinuity of the response function in the frequency domain! This is an interesting fact, but has no significance for our derivations, since it is related to the response function of the non-truncated integration.

The comparison of Figures 5.26 and 5.22 exhibits another feature: the interval of poor approximation $(20 < l < 40)$ in Figure 5.22 has been reduced to $(20 < l < 30)$ in Figure 5.26. This tendency suggests that we may get even better approximations with a truncated integration, if the kernel is computed for a larger Stokes's truncation radius ψ_0.

The previously described tests were again performed for $\psi_0=3.5°$. As expected, the term (5.78) again dominates the plot (cf. Figure 5.27), and the first positive maximum is shifted to about $12.2°$. As can be seen from Figure 5.28, now the errors resulting from truncation of the first maximum of the kernel function are almost invisible at the figure's scale. These errors are plotted in Figure 5.29. These error curves are based on an integration with the 'spheroidal' kernel $T_{20}\Psi o(\psi)$ up to the exact maximum of the kernel function $\phi_0 = 12.2107°...$. It is obvious that the errors of all recovered gravity field constituents above degree 20 are less than 2%.

5.3.4 Selection of kernel functions for numerical applications

For use on real altimetry data in selected test areas, the following problems have been considered. For the purpose of checking the marine gravity data through Stokes's integration and comparison with the corresponding quantity obtained from integration over altimetry data, an
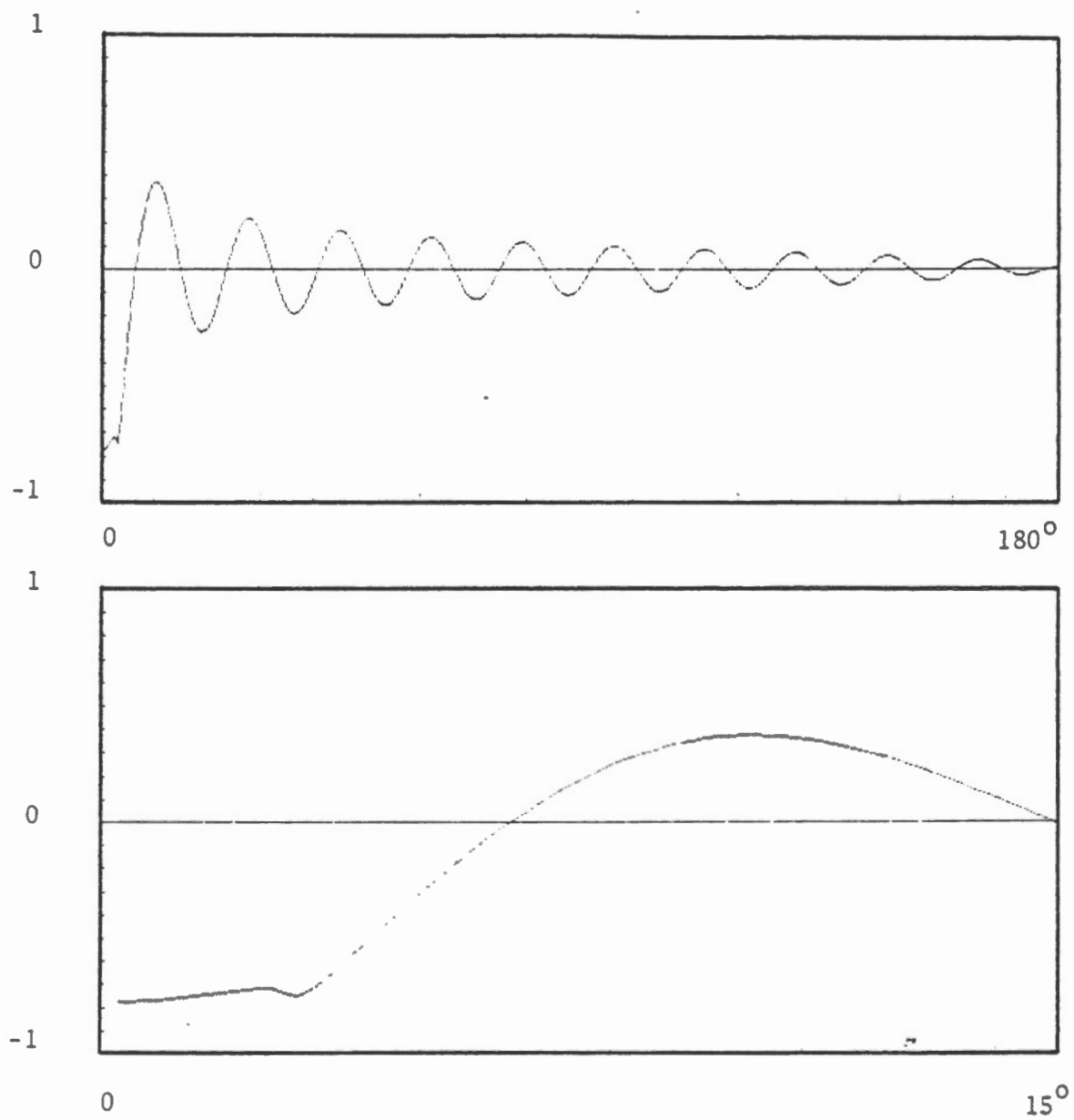
Figure 5.25:   Shape of $T_{20}^{\psi_o}(\psi)$ for $\psi_o = 2.936^o$
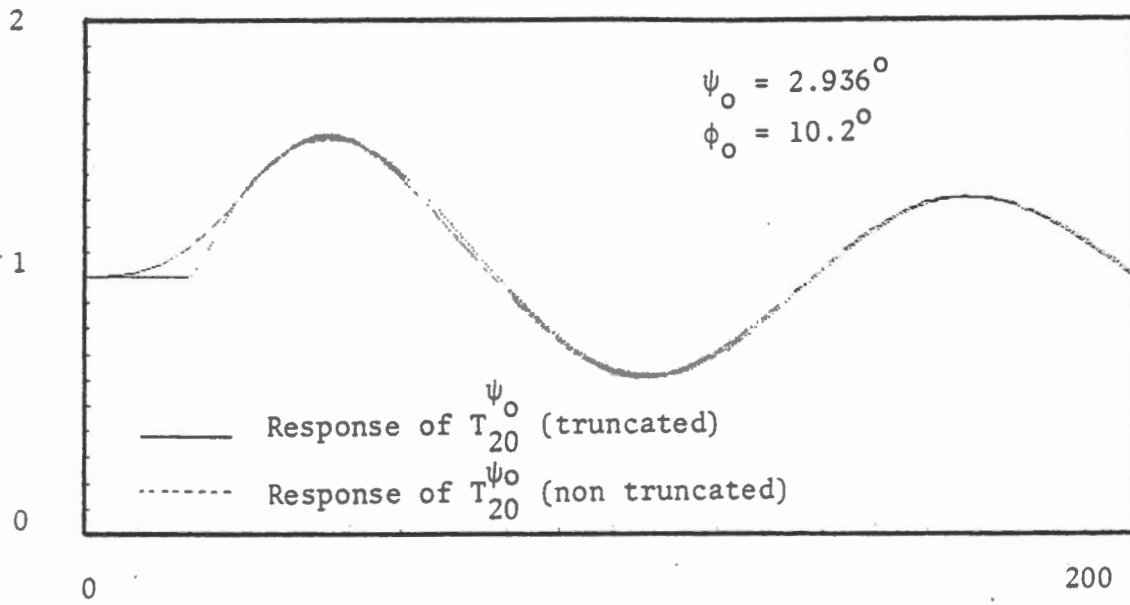
Figure 5.26: Response of $T_{20}^{\psi_o}$ (ψ) for $\psi_o = 2.936^o$
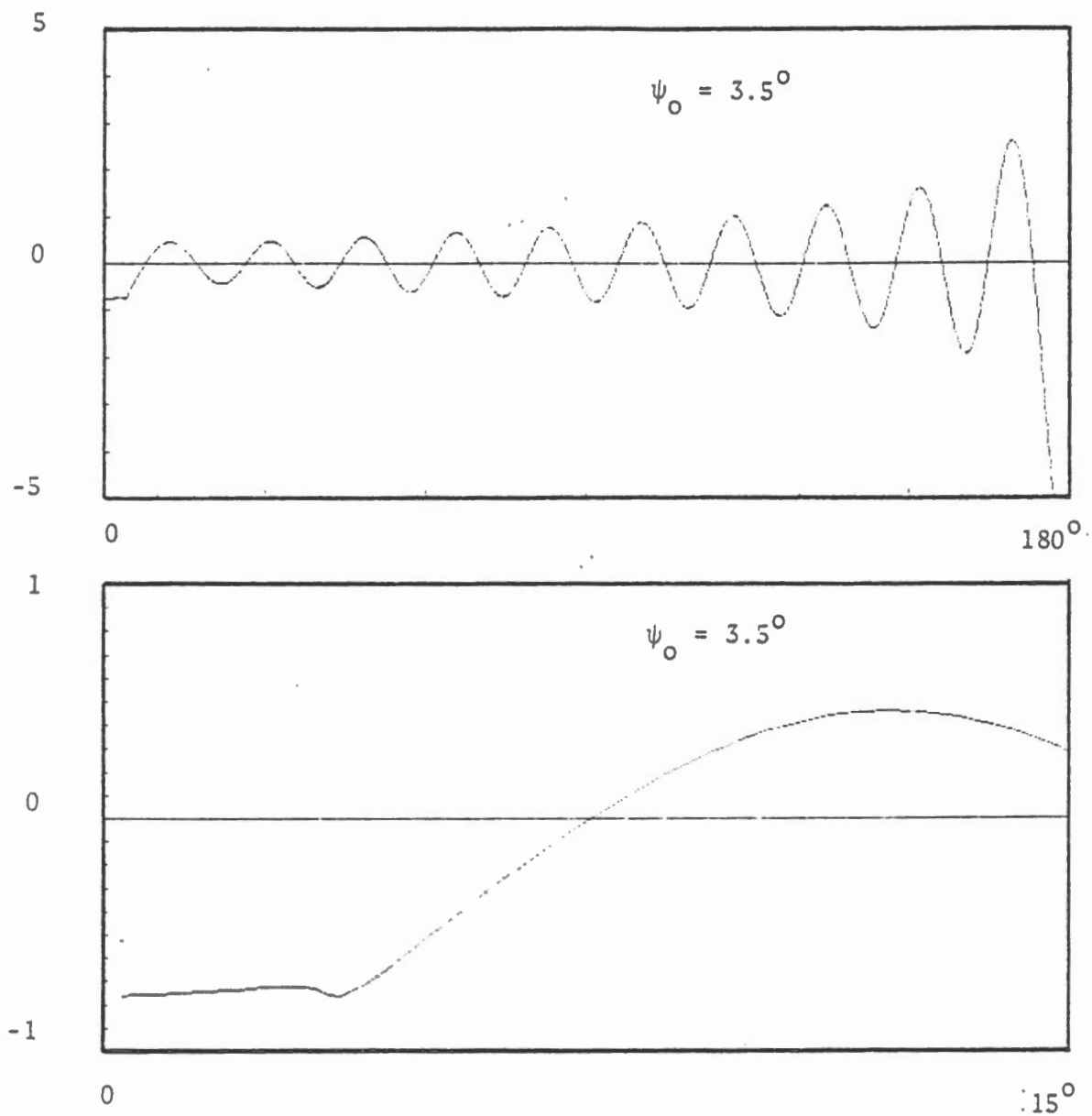
Figure 5.27: Shape of $T_{20}^{\psi_o}(\psi)$ for $\psi_o = 3.5^o$
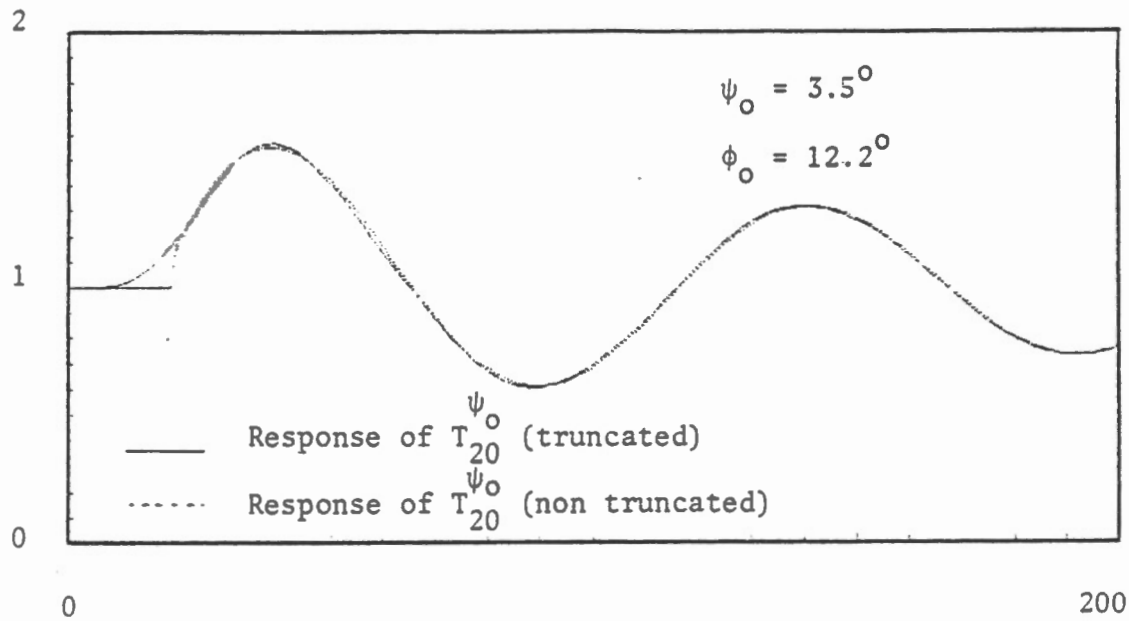
Figure 5.28: Response of truncated $T_{20}^{\psi_o}(\psi)$

Response of $T_{20}^{\psi_o}$ (truncated)

Response of $T_{20}^{\psi_o}$ (non truncated)

$\psi_o = 3.5^o$

$\phi_o = 12.2^o$
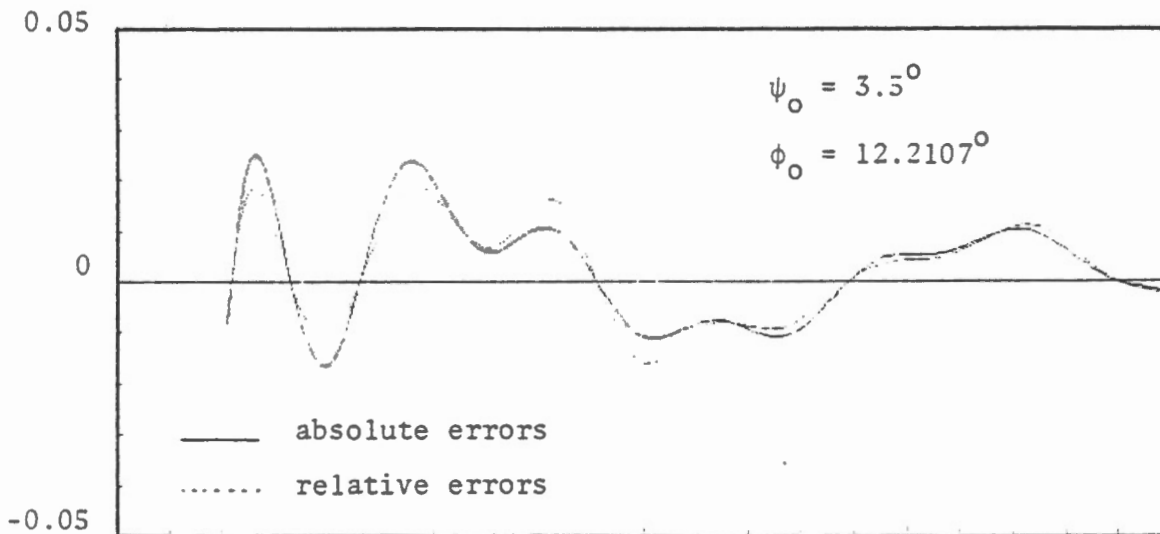
Figure 5.29: Errors for truncated integration with $T_{20}^{\psi_o}(\psi)$

absolute errors

relative errors

$\psi_o = 3.5^o$

$\phi_o = 12.2107^o$

integration radius as small as possible would be desirable. This would enable us to identify certain very small spatial areas of possible disagreement and thus indentify faulty gravity or altimetry measurements. On the other hand, the altimetric surface used in the integration is locally rather smooth, partly due to the gridding algorithm described in section 3.2.5. That means that for very small integration radii, the altimetric surface within the integration domain will be mainly flat without any remaining structure. This fact leads to a lower limit for the choice of an integration radius. It can be shown that the integration on the right-hand side of eqn. (5.74) yields zero, if the geoid N (represented by the altimetric surface) is a plane surface, i.e., no information can be extracted from this type of surface.

For the purpose of this contract, we decided to work with two kernel functions $T_{90}^{1°}$ and $T_{90}^{0.5°}$ and to discuss the integration results in comparison with spheroidal Stokes's integration using the Stokes kernel within the truncation radii of 1° and 0.5°, respectively. The corresponding truncation radii for the T kernels are 3.640 35° and 2.984 97°, respectively.

5.4 Numerical Integration Procedure for the Truncation Kernel

The altimetry derived partial geoidal height δN_A is given by (see, for instance, eqn. (5.74))

$$N^{\psi_0, \phi_0} = N - 1/2 \int_{\psi=0}^{\phi_0} T^{\psi_0}(\psi)(\bar{N}(\psi) - \bar{N}(\psi_0))(\cos\psi - \cos\psi_0)^{-1} \sin\psi \, d\psi , \qquad (5.83)$$

where $T^{\psi_0}(\psi)$ is the truncation kernel, ψ is the geocentric angle between the point of interest and the 'dummy' point, ψ_0 is the Stokes radius of integration, and $\bar{N}(\psi)$, $\bar{N}(\psi_0)$ are the mean geoidal heights (reduced for long wavelength geoidal undulations) of the 'dummy' points which are $\psi \pm d\psi$ and $\psi_0 \pm d\psi$ respectively, away from the point of interest.

For the numerical integration of eqn. (5.83), the area around the point of interest up to a maximum radius ϕ_0 is divided into rings of thickness $\Delta\psi$, as is shown in Figure 5.30. The rings
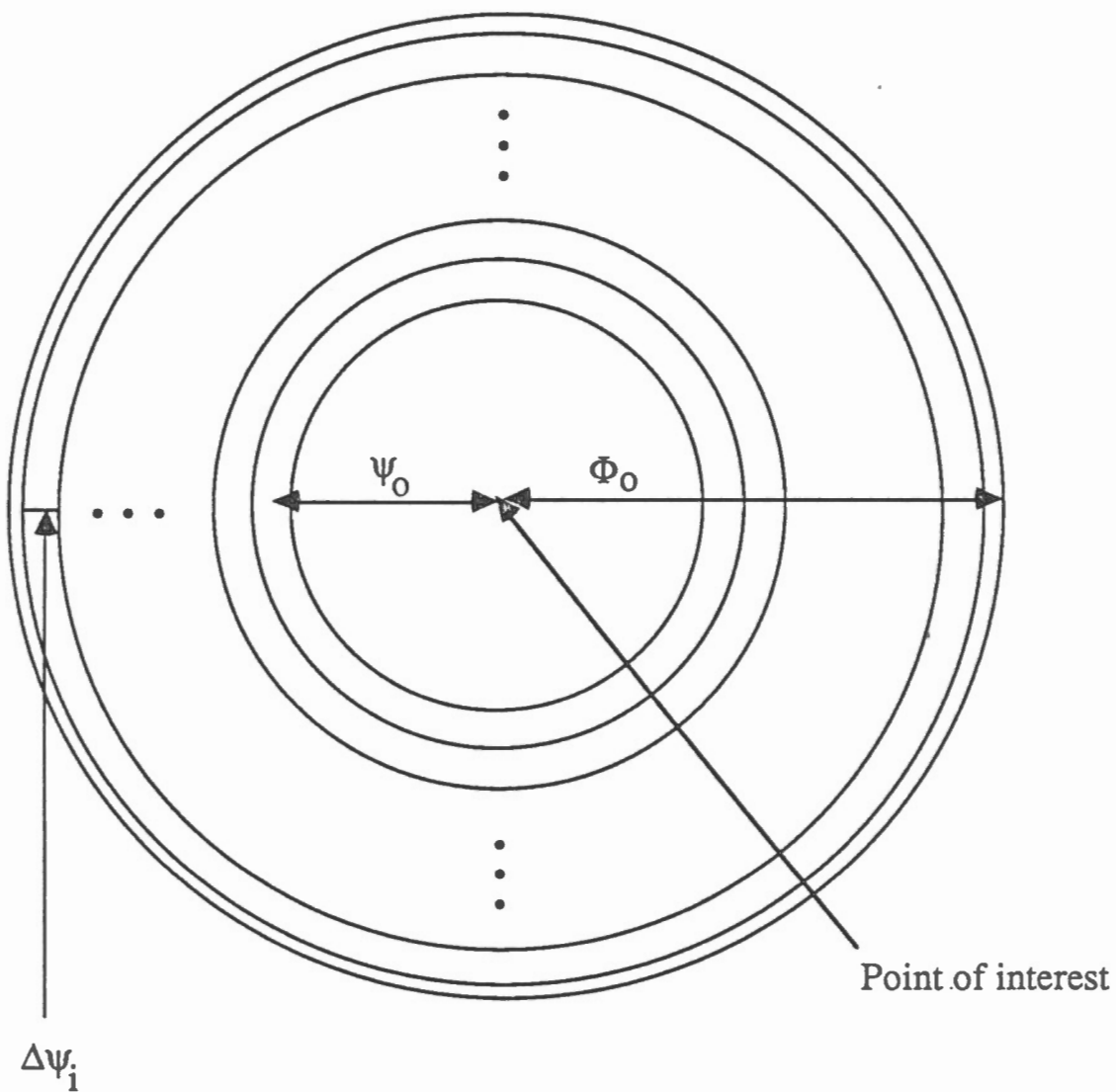
Figure 5.30: Sketch of the integration procedure for the truncation kernel

The figure contains the following labels: ψ_O, Φ_O, $\Delta\psi_i$, and "Point of interest".

are arranged in such a way that one of them will have a mean radius of ψ_0. Consequently, the mean radius of the circular area around the point of interest, as well as the thickness of the outer ring, will be less, in general, than $\Delta\psi$. Consequently, eqn. (5.83) can be written as:

$$N^{\psi_0,\phi_0} = N - 1/2 \sum_{i=1}^{NRINGS} T_i(\bar{N}(\psi_i) - \bar{N}(\psi_0))(\cos\psi_i - \cos\psi_0)^{-1} \sin\psi_i \Delta\psi_i , \qquad (5.84)$$

where NRINGS is the number of rings around the point of interest, and i is the ring index.

Since ψ_0, ϕ_0, ψ_i, and $\Delta\psi_i$ have predetermined constant values, and by setting

$$R_i = 1/2 \, T_i(\cos\psi_i - \cos\psi_0)^{-1} \sin\psi_i \Delta\psi_i , \qquad (5.85)$$

eqn. (5.84) then becomes

$$N^{\psi_0,\phi_0} = N - \sum_{i=1}^{NRINGS} R_i\{ \bar{N}(\psi_i) - \bar{N}(\psi_0)\} . \qquad (5.86)$$

For computational speed, the different values of R_i are stored and $\bar{N}(\psi_0)$ is evaluated initially for a specified point of interest and kept constant throughout the integration. $\bar{N}(\psi_i)$ is evaluated for each ring for a specified point of interest. Equation (5.86) indicates that the numerical integration is equivalent to a linear filter in the altimetric geoid domain with R_i being the transfer function.

5.5 Determination of Mean Gravity Anomalies from Altimetry

According to the derivations in sections 5.1 and 5.3, the result of the integration over the geoidal heights (altimetry) is equal to the corresponding truncated Stokes's integration over gravity data (cf. eqn. (5.1))

$$\delta N^{l,\psi_0,\phi_0} = R/(4\pi\gamma) \int_{\psi=0}^{\psi_0} S(\psi) \int_{\alpha=0}^{2\pi} \delta\Delta g^l(\psi, \alpha) \sin\psi \, d\psi \, d\alpha . \qquad (5.87)$$

110

The mean gravity anomaly, $\delta \overline{\Delta g}^l$, in the integration area, $\psi \leq \psi_0$ (not to be confused with the mean gravity along a constant radius, as used in eqn. (5.2)), is by definition constant within $\psi \leq \psi_0$. This means that eqn. (5.87) can be inverted to yield mean gravity anomalies within $\psi \leq \psi_0$ according to

$$\delta \overline{\Delta g}^l = \delta N^{l,\psi_0,\phi_0} \cdot 4\pi\gamma / \{R \int_{\psi=0}^{\psi_0} S(\psi) \int_{\alpha=0}^{2\pi} \sin\psi \, d\psi \, d\alpha\}$$

$$= \delta N^{l,\psi_0,\phi_0} \cdot 4\pi\gamma / \{R \int_{\psi=0}^{\psi_0} F(\psi) \, d\psi\} = C^{\psi_0} \cdot \delta N^{l,\psi_0,\phi_0} . \qquad (5.88)$$

Obviously, the conversion factor is a constant depending only on the truncation radius ψ_0:

$$C^{\psi_0} = 4\pi\gamma \{R \int_{\psi=0}^{\psi_0} F(\psi) \, d\psi\}^{-1} . \qquad (5.89)$$

5.6 Test of the Integration Programs

To verify the theoretical derivations of the last sections and to assess the accuracy of our approach, a geoidal surface and the corresponding gravity anomalies were computed, using only the spherical harmonic coefficients of degree 91 of the Wenzel [1985] expansion. Both the geoidal height and the gravity anomalies were evaluated on a 10' grid covering an area of 10° by 5°.

The result of a truncated Stokes's integration over the gravity anomalies is shown in Figure 5.31. The corresponding truncation integration over the geoidal heights gave results shown in Figure 5.32. At the scale of these figures, the main difference between the results seems to be that the Stokes integration gives a slightly smoother surface. Plotting the difference between the two integration results in Figure 5.33 reveals a small systematic error proportional to the integration result. This error amounts to about 5% and is attributed to the combined effects of kernel approximation and data discretization.
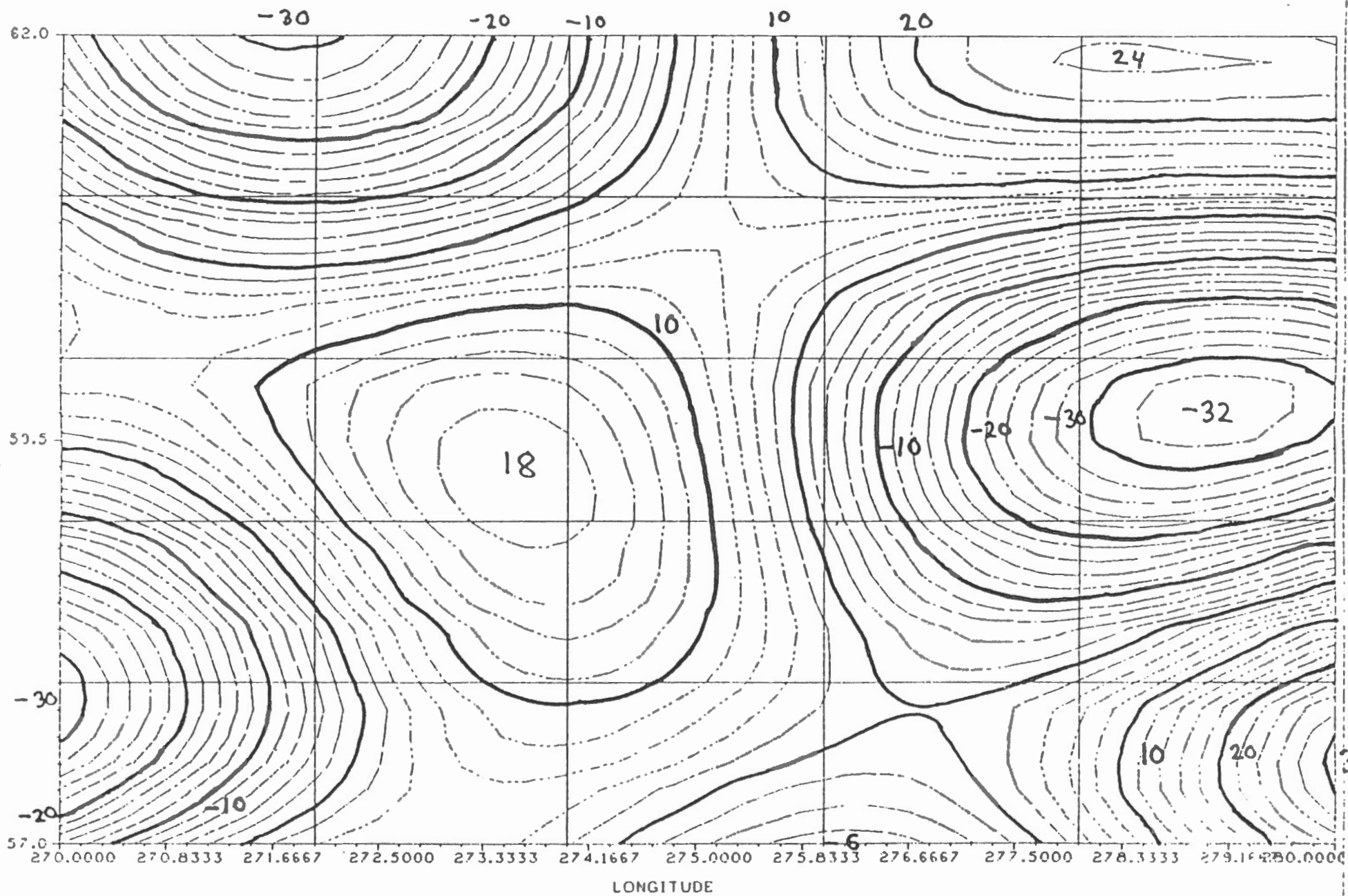
Fig. 5.31: Truncated Stokes's integration over gravity anomalies of degree 91 only.
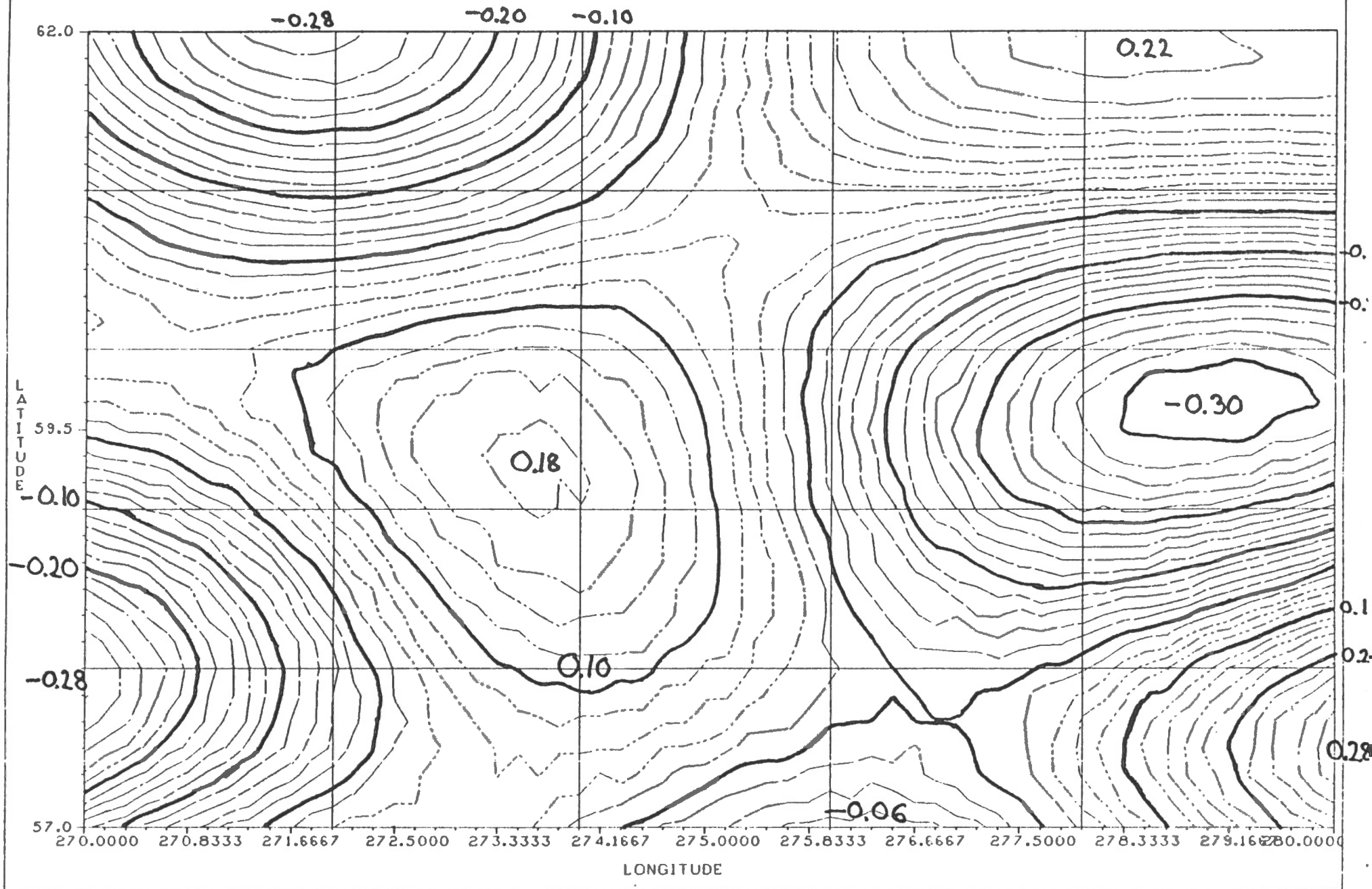
Fig. 5.32: Truncation integration over geoidal heights of degree 91 only.
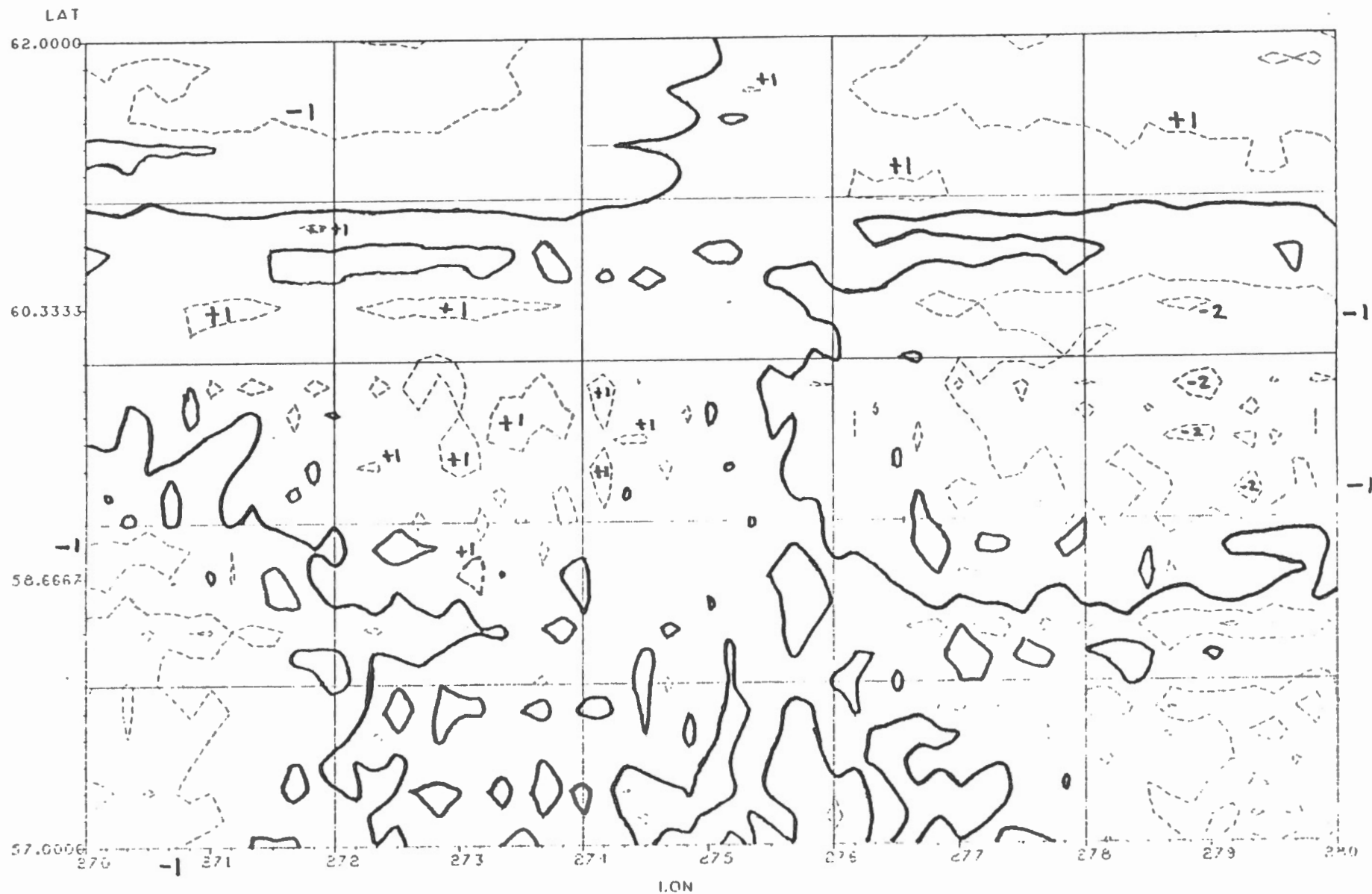
Fig. 5.33: Differences between Stokes's integration and truncation integration for a gravity field of degree 91 only.

6. APPLICATIONS

For application of the derived mathematical relations between geoidal surface and gravity anomalies to real data, two test areas were selected according to the research contract. Both test areas have an extension of 5° and 10° in latitude and longitude, respectively. The location of the test areas is shown in Figure 6.1. The Labrador Sea area was chosen to cover some of the local variations of sea surface topography, as discussed in section 3.

6.1 Input Data

Marine gravity data for Hudson Bay and the Labrador Sea (see section 3.1) was reduced for a spherical harmonic expansion of the gravity field up to degree 90 using Wenzel's [1985] set of coefficients. The Hudson Bay test area is more or less completely covered by gravity measurements. The lower left part of the Labrador test area is on land. Obviously, no marine gravity data is available in this part.

The input for integration with the 'truncation' kernel is primarily the altimetric surface interpolated on a 10' x 10' grid, according to section 3.2.5. The integrator requires data within a certain radius ϕ_0 around the point of interest. To facilitate the integration in coastal regions, the altimetric surface had to be augmented by the UNB gravimetric geoid, also available on the 10' x 10' grid. The merging of the two data sets required a decision on where to use which set. The first obvious choice would be to merge the two data files along the coastline. Unfortunately, the altimetry is rather unreliable in the coastal areas, as explained in section 3.4.2. Thus we decided to draw the merging line along latitude/longitude lines, somewhat off the coast as shown in Figure 6.2. The exact location of the merging line was chosen to coincide as well as possible with the lines of zero difference between the gravimetric geoidal heights and the altimetric sea surface heights (see section 3.4.2). Thus large discontinuities in the merged surface were avoided. Figure 6.3 shows a contour plot of the combined geoid-altimetry surface. This surface was subsequently reduced for long wavelength components using the Wenzel [1985] spherical harmonic expansion up to degree 90. A colour plot of the reduced geoidal surface is attached as an external appendix to this report.

FIG. 6.1: Location of test areas in Hudson Bay and the Labrador Sea.

FIG 6.2: Merging of altimetric sea surface and gravimetric geoid.

Merged Grav. Geoid and SEASAT Altimetry

(10 x 10 minute grid)
Reference Ellipsoid GRS60
Contour Interval 1 m

FIG. 6.3: Merged gravimetric geoid and SEASAT altimetry.

6.2 Hudson Bay Results

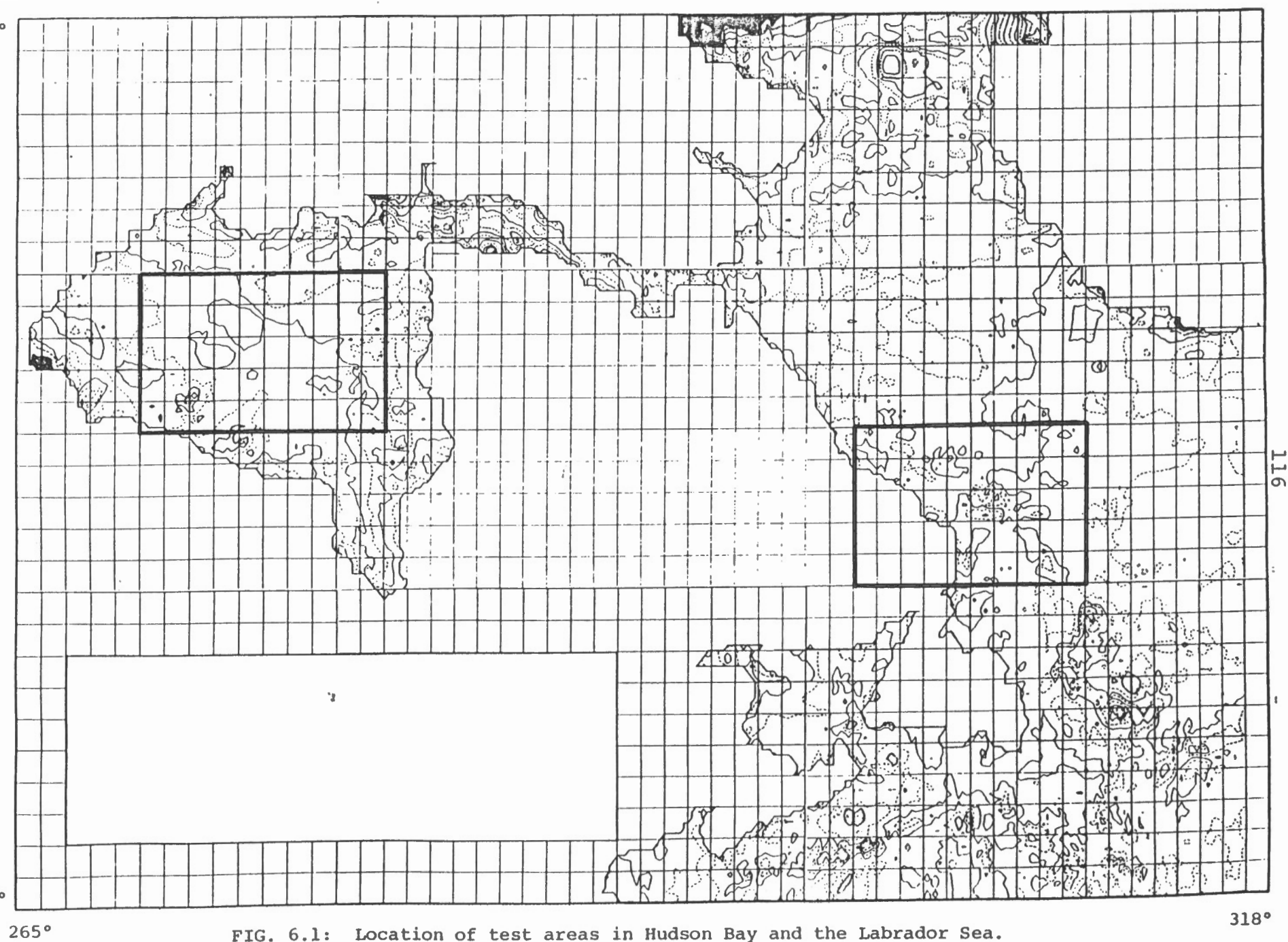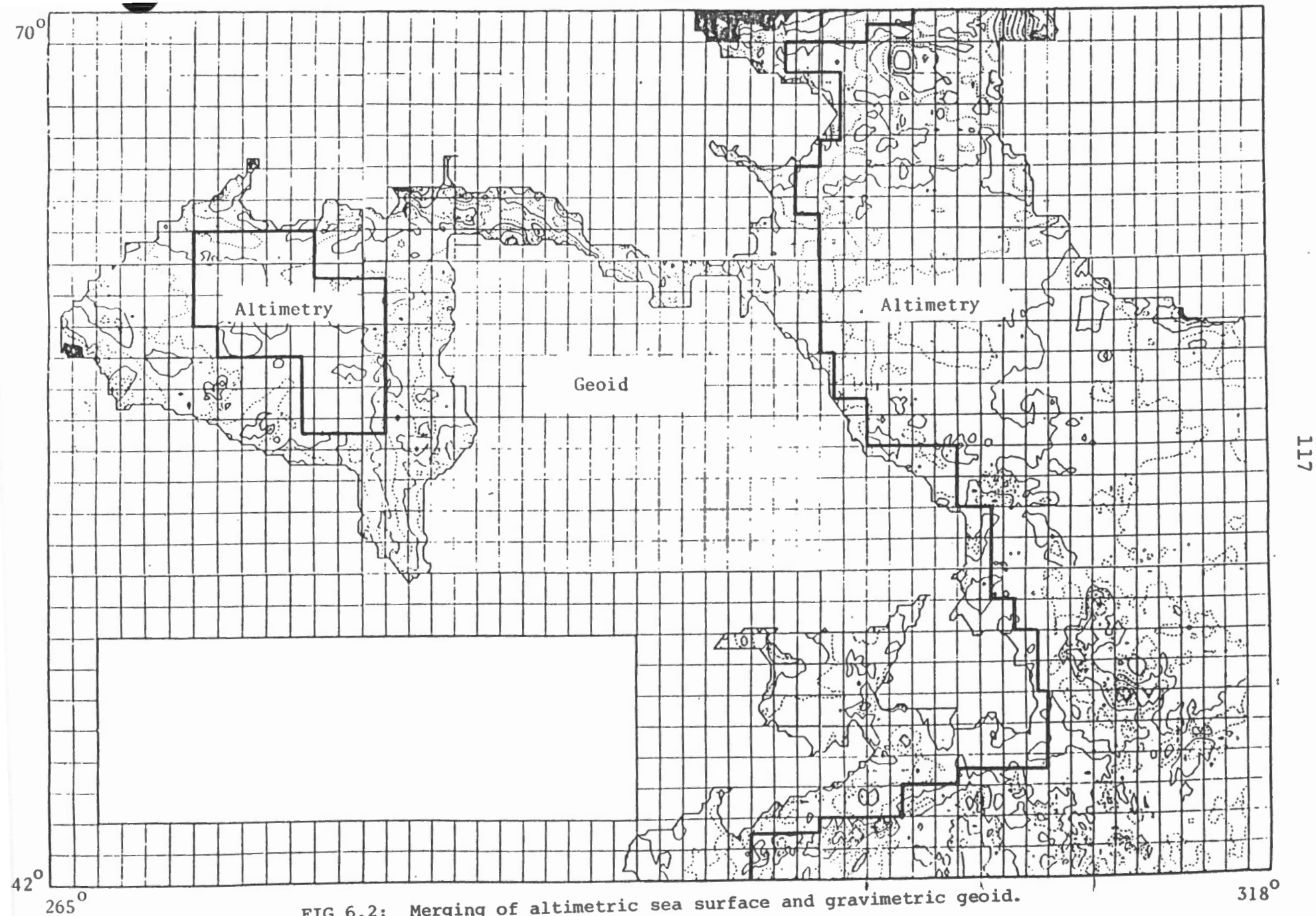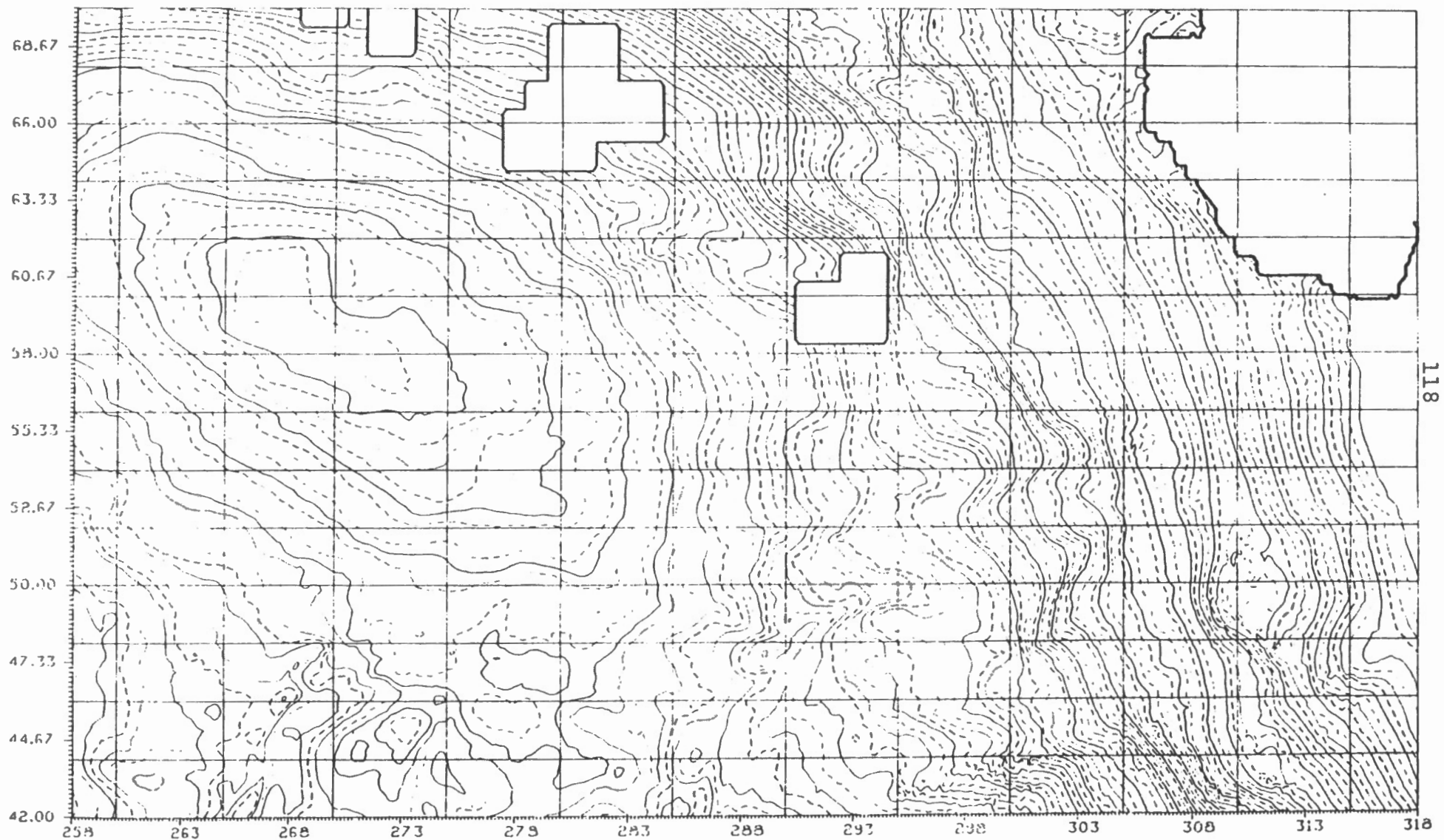Figures 6.4 and 6.5 show results pertaining to integrations with Stokes's truncation radius $\psi_0 = 1°$ in units of centimetres. The Stokes integration in Figure 6.4 shows very detailed structures in the eastern part reflecting mainly the local variations in gravity. Comparing this figure with EPB [1980], we see a mimicing of the relative gravity variations in Hudson Bay. This is obviously a consequence of the very localized integration. Both plots show flat local highs and lows in the western and middle part of the test area and rather large anomalies along the eastern boundary.

The integration with the truncation kernel for the same area is shown in Figure 6.5. The blank in the top part is due to empty areas in the input data file (see Figure 5.3). Clearly, the match of the two surfaces is quite good, not only as far as the basic shape is concerned but also as far as the amount of detail portrayed. It is encouraging to see that, apparently, the altimetric surface used for the 'truncation' contains enough high frequency information (wavelengths shorter than 100 km) as to be comparable with that contained in the collected marine gravity. There seems to be, however, an overall positive bias in the Stokes integration compared to the 'truncation' integration.

To compare the two results better, we have plotted the difference of the two surfaces (Figure 6.6). Now the overall positive bias in the Stokes integration shows quite clearly. Realizing that a constant error of 1 mGal in the gravity data produces a change of 12 cm in the Stokes integration with radius of 1°, this positive bias would indicate an average bias in the gravity data of the order of about 1 to 5 mGal (in the absence of any bias in the altimetry) in the western part of this area of Hudson Bay. We note that differences of several decimetres may be caused by random (uncorrelated) errors in the marine gravity. For the typical case of an average density of Δg values of 5' and their standard deviation of 5 mGal, the standard deviation of the Stokes integration would be 20 cm.

There is a sizeable localized negative difference of 1.2 m in the eastern part of Hudson Bay. It appears that it might be caused again by a local (within the 1° integration cap) bias of about 10 mGal in the marine gravity, although this hypothesis is difficult to sustain when we realize that the anomalous difference is outside the area of altimetry.

Just how much of this difference could be reflecting the presence of the SST in the altimetry is difficult to evaluate. The effect of merging the geoid with the altimetry, however, can be seen along the seams (shown in Figure 6.6 by − − − lines—cf. Figure 6.2) where rather abrupt steps of up to several decimetres can be discerned. The effect of random (uncorrelated) errors in the (gridded) altimetry is less than 7 cm in the Hudson Bay test area. These errors do not affect the
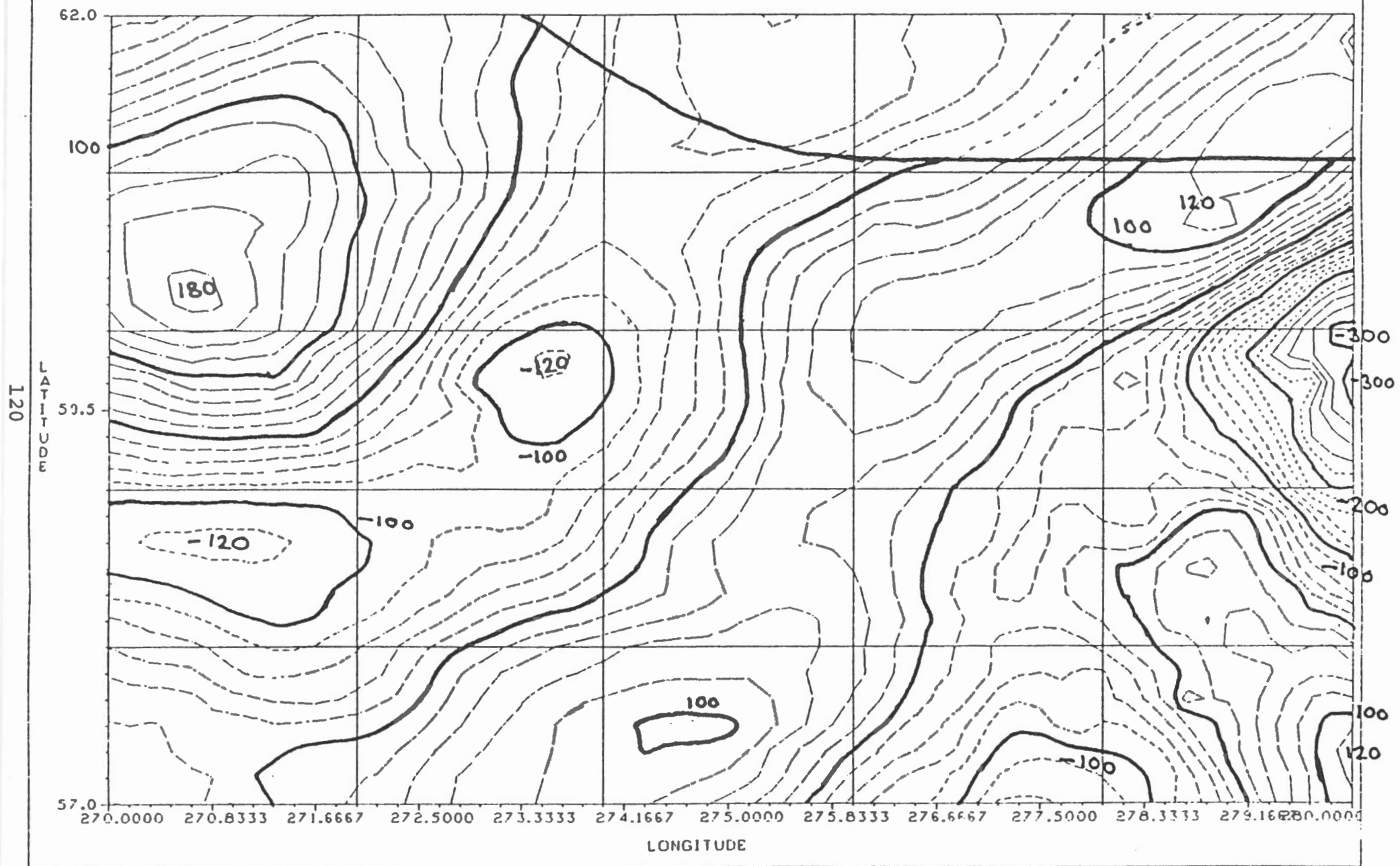
Fig. 6.4:   Stokes's integration result ($\psi_0 = 1°$) for Hudson Bay.

Fig. 6.5: Truncation integration result ($\psi_0 = 1°$) for Hudson Bay.

Fig. 6.6: Difference between Stokes's and truncation integration ($\psi_0 = 1°$) for Hudson Bay.

comparison in a significant way.

We should note that any longer wavelength differences between the two solutions (wave number 91, 92, ...) may be coming from the inaccuracy of the truncation kernel used in the actual convolution which manifests itself particularly in the low frequency domain (cf. section 5). These differences, however, should be of the order of one decimetre or less.

For the integration radius of 0.5°, the situation is very similar. The Stokes integration (Figure 6.7) and the 'truncation' integration (Figure 6.8) show the same features as well as the same amount of detail. Obviously, the altimetric surface possesses enough high frequency information to match the marine gravity information content.

The difference surface (gravity-altimetry) displayed on Figure 6.9 agrees with the 1.0° integration radius difference surface in broad features but differs in detail as one might expect. If these differences were caused by biases in marine gravity (here 1 mGal bias causes a 6 cm effect in the 'truncation' integration) then the change from one difference surface to the other would reflect the change in these biases from 1° cap averaging to 0.5° cap averaging.

Here, again, the role of SST is difficult to assess. But once more, the seams between altimetry and the gravimetric geoid are visible, causing steps of several decimetres.

To see if anything can be learned from the difference between 1° radius truncation and 0.5° radius truncation, the difference surface was generated (Figure 6.10). It is interesting to note that this surface is relatively smooth (wavelength of 100 km and longer) and it mimics, with a negative sign, the shape of the higher order geoid constituents (cf. the external appendix: geoid-(Wenzel (90, 90) field) two-colour plot)). It probably represents the contribution of deeper seated density anomalies so that the smaller radius truncated geoid (Figure 6.8) reflects the more localized (more focused) contribution of density anomalies located closer to the surface of the earth. This hypothesis, however, has to be tested in an area where the subsurface density distribution is relatively well known.

6.3 Labrador Sea Results

Figures 6.11 and 6.12 show results pertaining to the integration with Stokes's truncation radius 1°. The Stokes integration over gravity anomalies in Figure 6.11 shows again detailed structures of local gravity variations and compares quite well with the gravity maps [EPB, 1980]. Especially pronounced are the large positive anomalies along the ridge of the continental shelf and local lows at about (54.5°, 307.5°) and (55.5°, 304.5°). The corresponding truncation integration

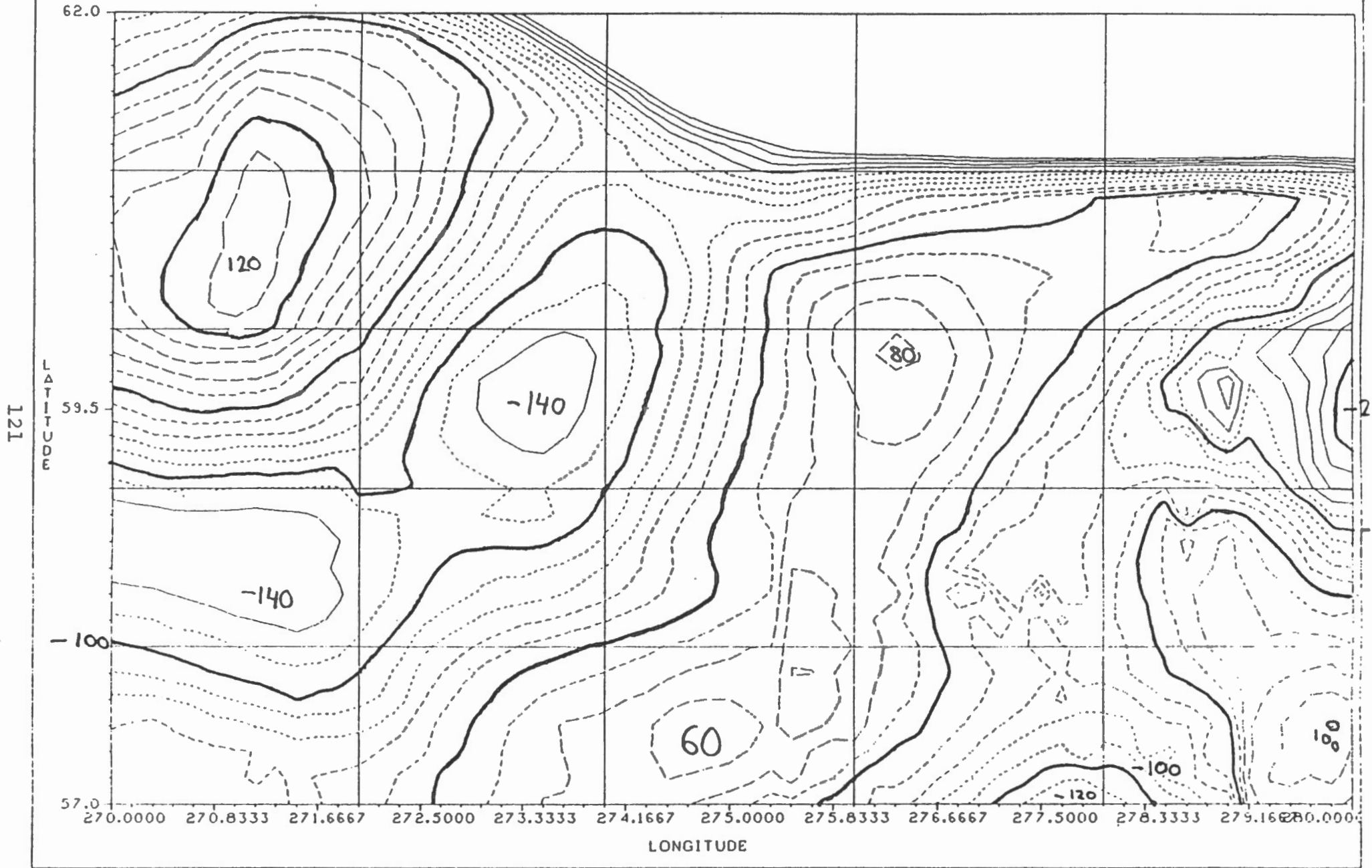Fig. 6.7: Stokes's integration result ($\psi_0 = 0.5°$) for Hudson Bay.

124

Fig. 6.8: Truncation integration result ($\psi_0 = 0.5°$) for Hudson Bay.

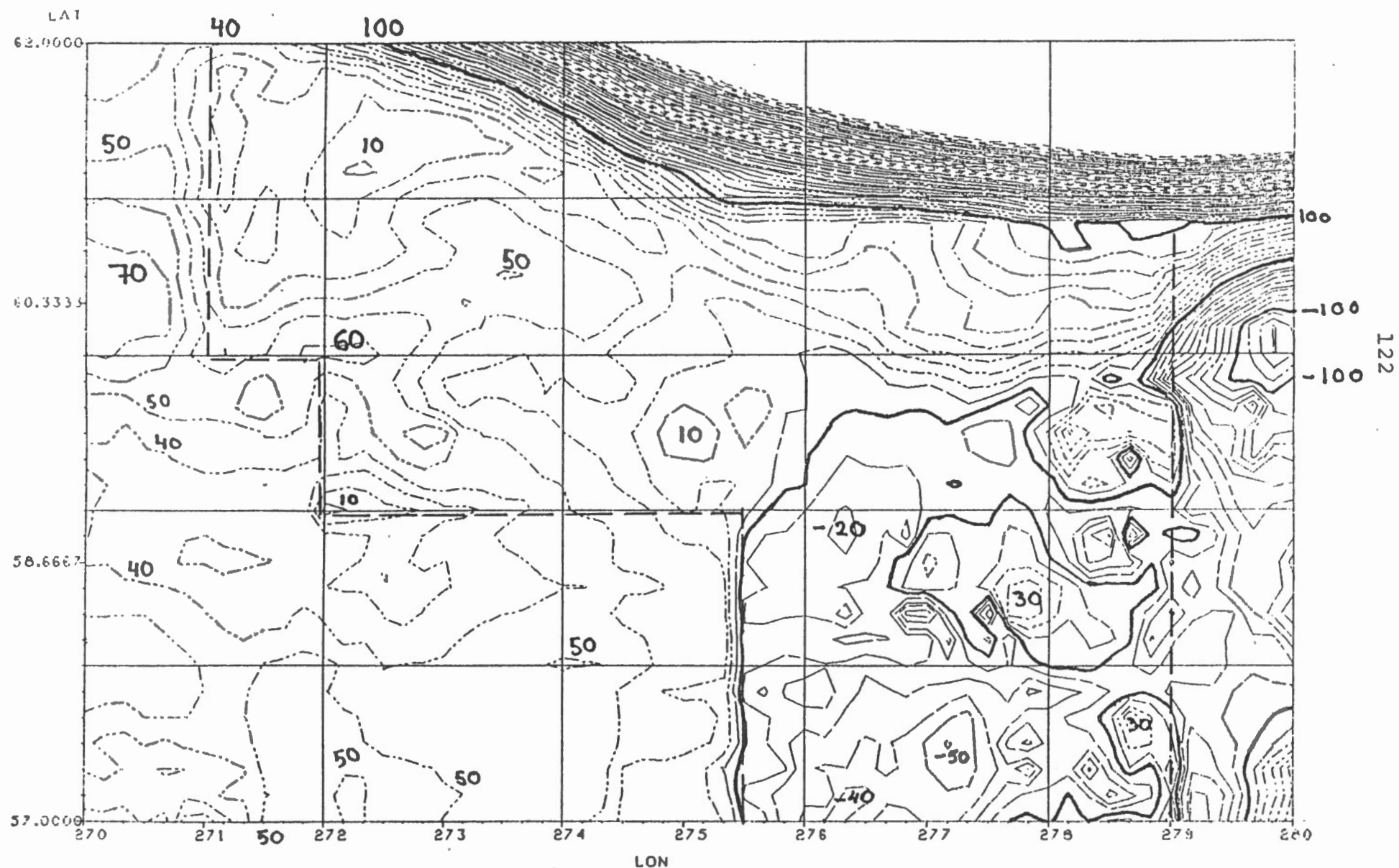Fig. 6.9: Difference between the Stokes and truncation integration ($\psi_0 = 0.5°$) for Hudson Bay.

Fig. 6.10: Difference between the truncation integration results for $\psi_0 = 1°$ and $\psi_0 = 0.5°$ in Hudson Bay.
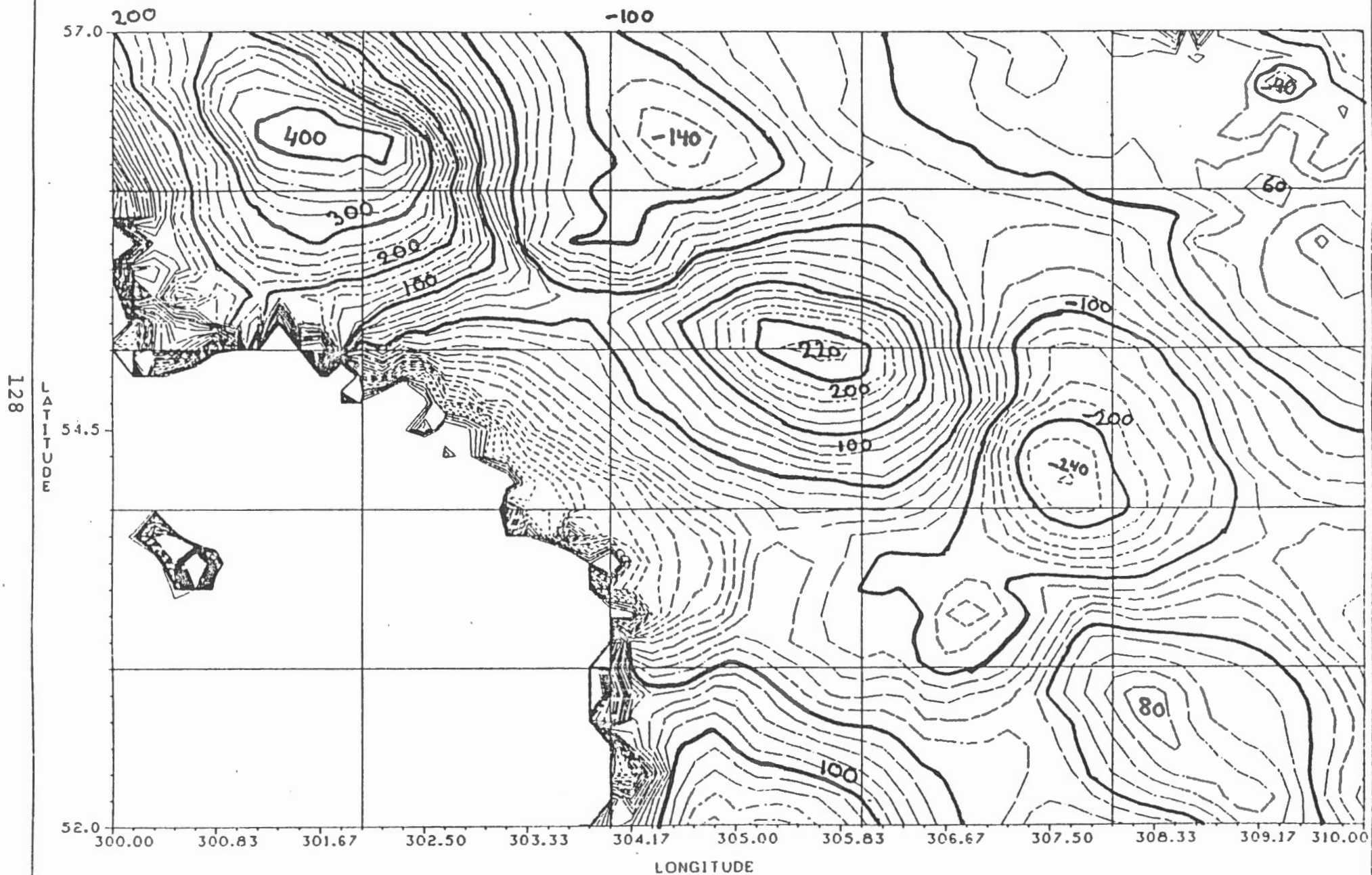
Fig. 6.11: Stokes's integration result ($\psi_0 = 1°$) for Labrador Sea.
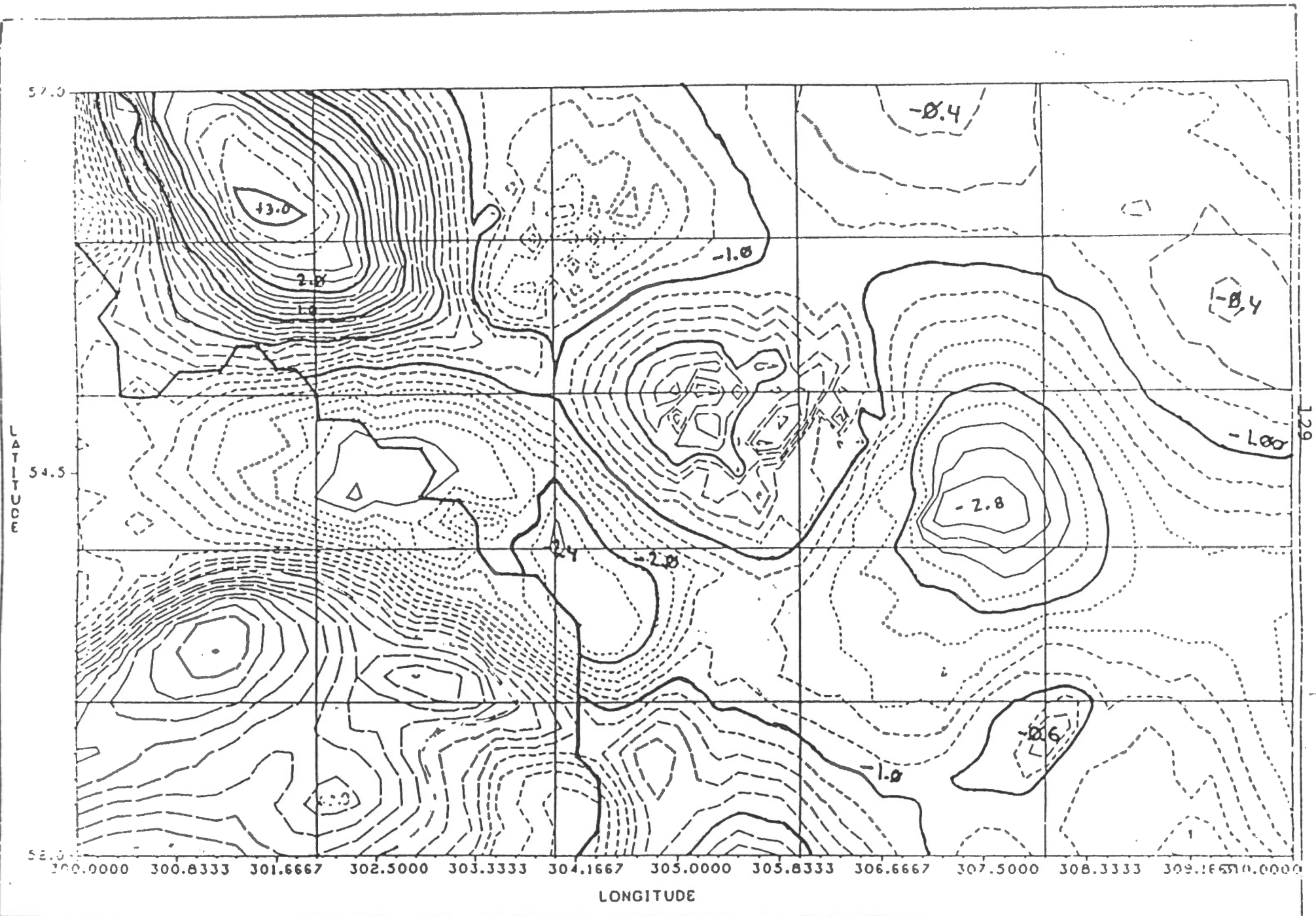
Fig. 6.12: Truncation integration result ($\psi_0 = 1°$) for Labrador Sea.

result shown in Figure 6.12 remodels all the main features but shows also differences in the areas around (55°, 306°) and (53°, 308°). The structure of the discrepancies becomes more obvious in the plot of the differences between Stokes's integration and the truncation integration in Figure 6.13. We see a general positive bias of 74 cm and extremal values of 2 m. The large differences are approximately colocated with the maxima of the Stokes integration shown in Figure 6.11, but they are not proportional to these. In contrast to the Hudson Bay results, here only small distortions seem to be caused by the merging of the gravimetric geoid and altimetric surface. The merging line is indicated by the broken line in Figure 6.13.

Results for the truncation radius $\psi_0 = 0.5°$ are shown in Figures 6.14 and 6.15. Again the general agreement is good, but the rather large differences at several locations already seen in Figure 6.13 persist. The mean difference is +85 cm and extremal values are about +210 cm.

Comparing the plot of the differences for $\psi_0 = 0.5°$ (Figure 6.16) with those for $\psi_0 = 1.0°$ (Figure 6.13) we see that the size of the features shown shrinks when reducing the integration radius. This is most obvious in the upper left corner of the plots and at the maximum differences at about (55°, 306°). In the upper left corner, only a rather small area remains above the 1.0 m level after reducing the truncation radius. The second feature mentioned above dissolves into several separate features. This is seen as a consequence of the focusing effect, i.e., reducing the comparison to less input data points. The plot of the differences between truncation integrations with $\psi_0 = 1.0°$ and $\psi_0 = 0.5°$ is shown in Figure 6.17. This figure depicts a more detailed structure than the corresponding plot for Hudson Bay although the range of the differences is more or less the same. The more localized structure may be due to the rougher input data along the ridge of the continental shelf (cf. Figure 6.3).
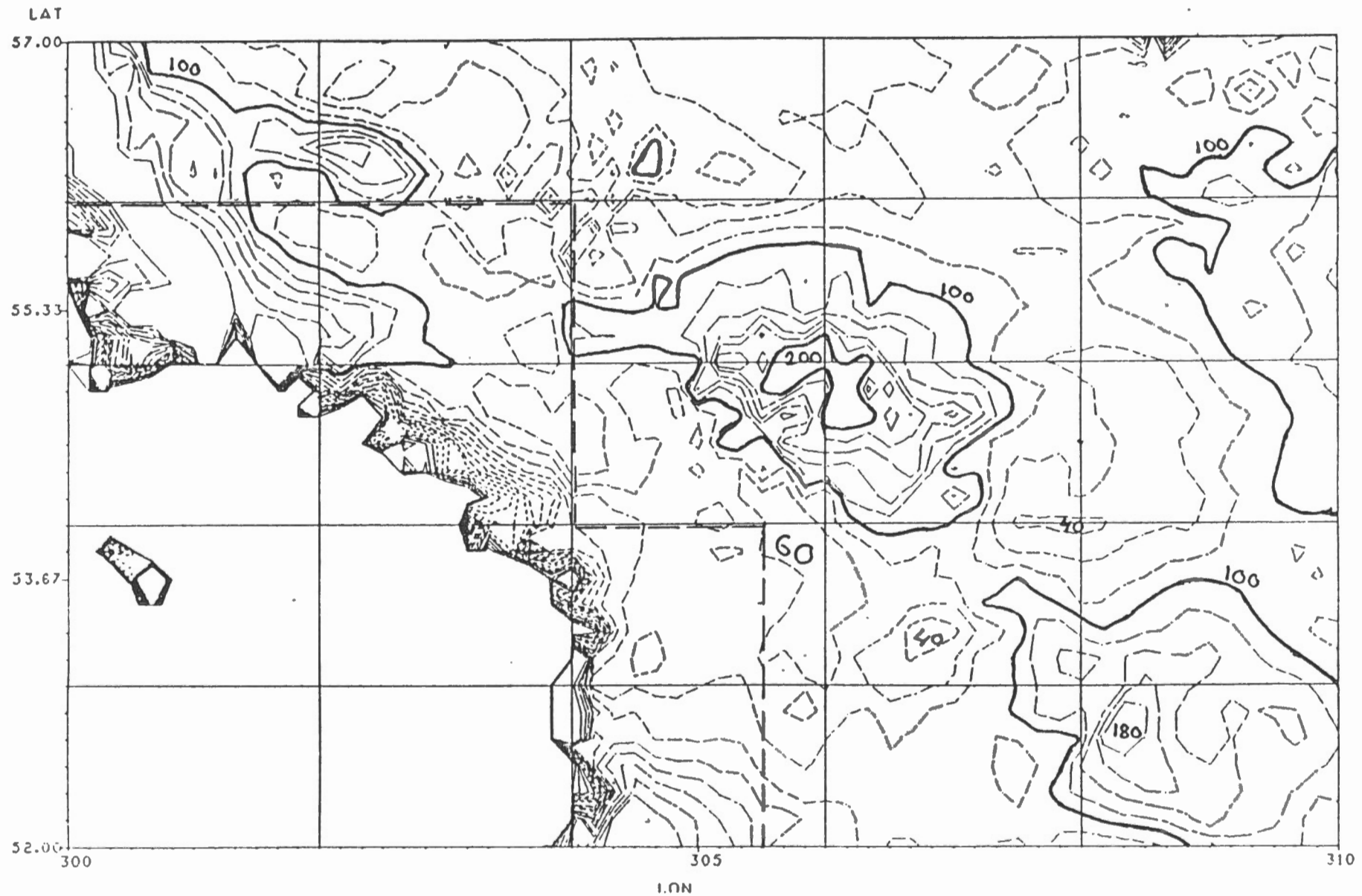
Fig. 6.13: Difference between the Stokes and truncation integration ($\psi_0 = 1°$) for Labrador Sea.
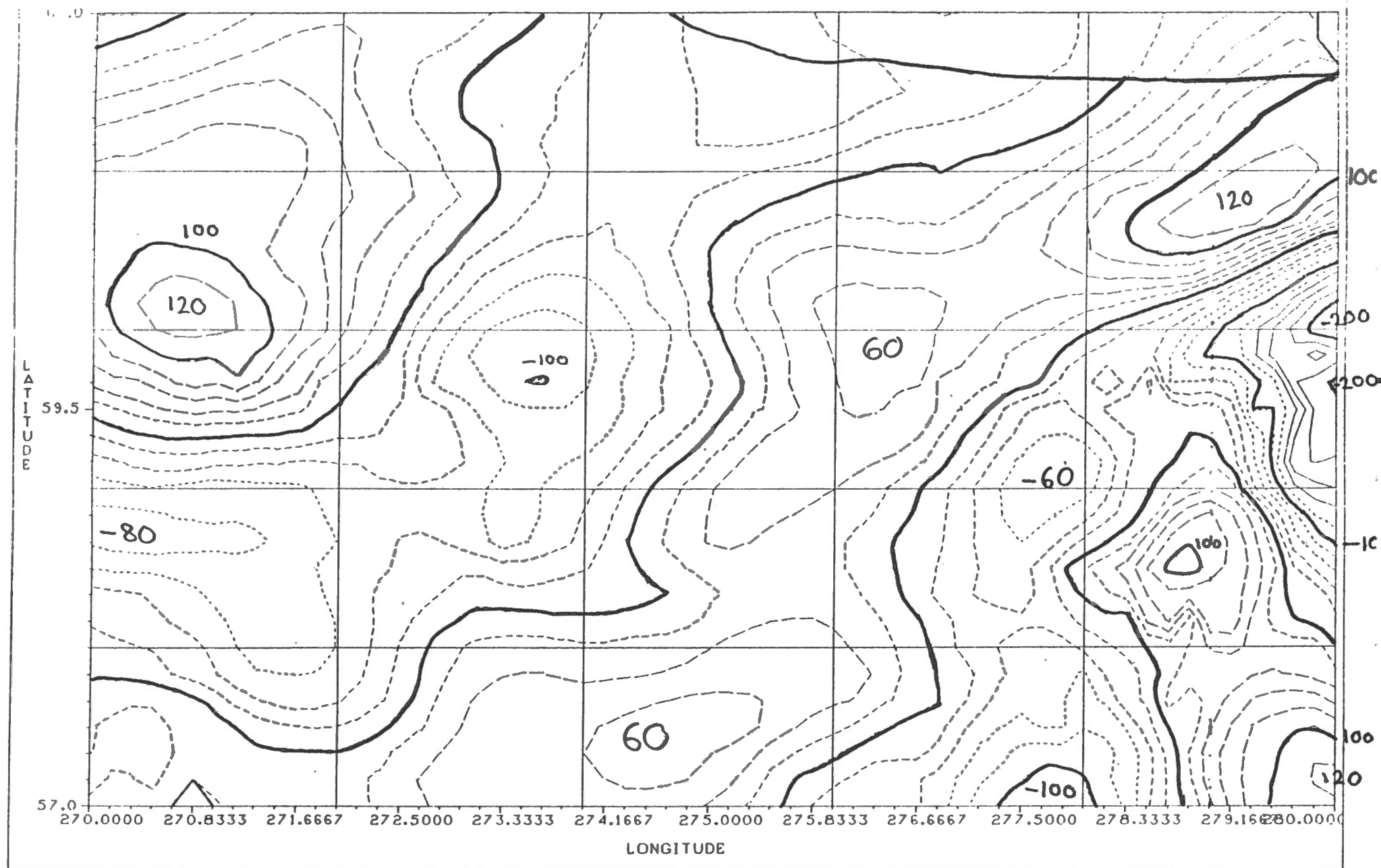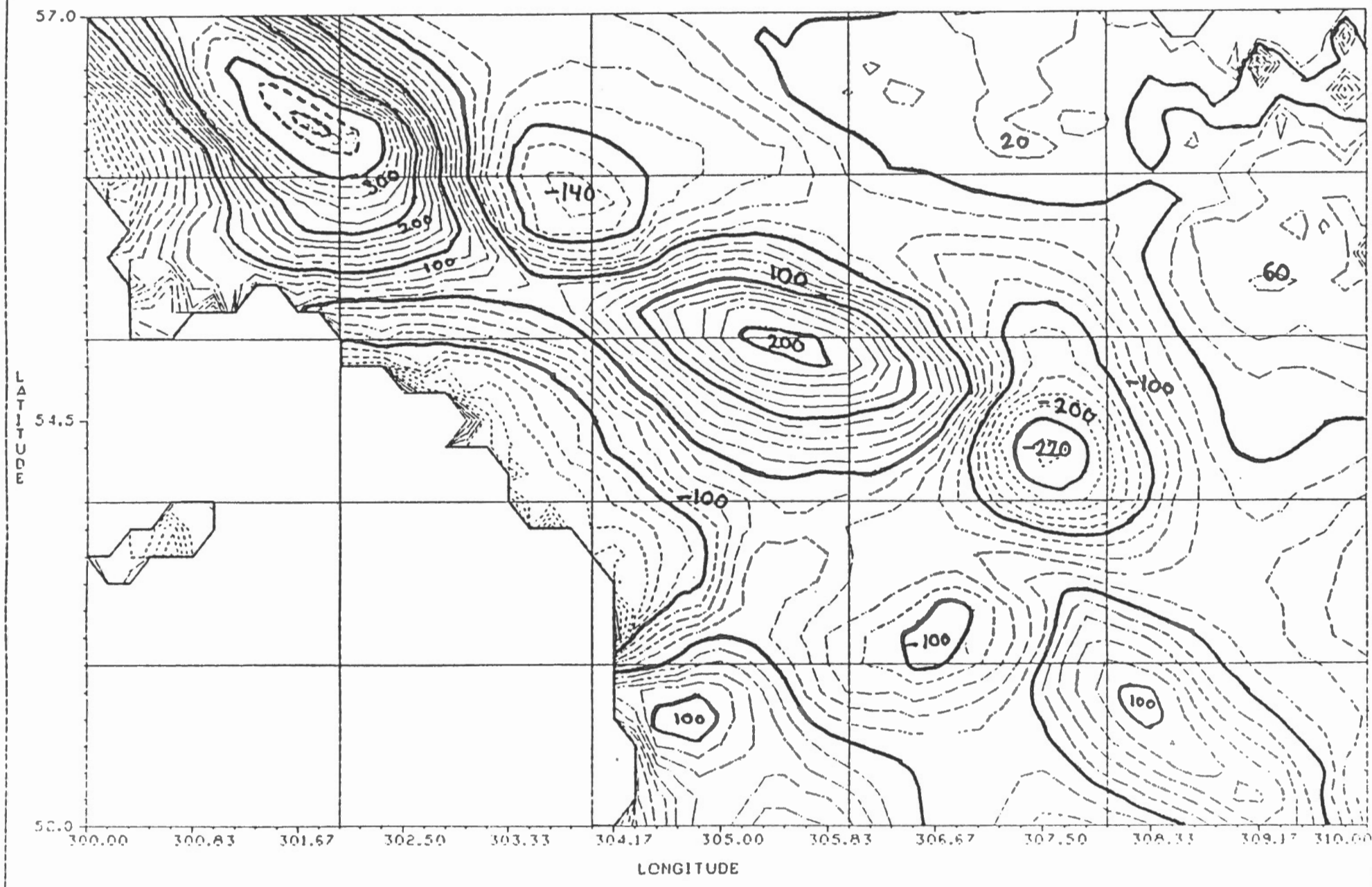
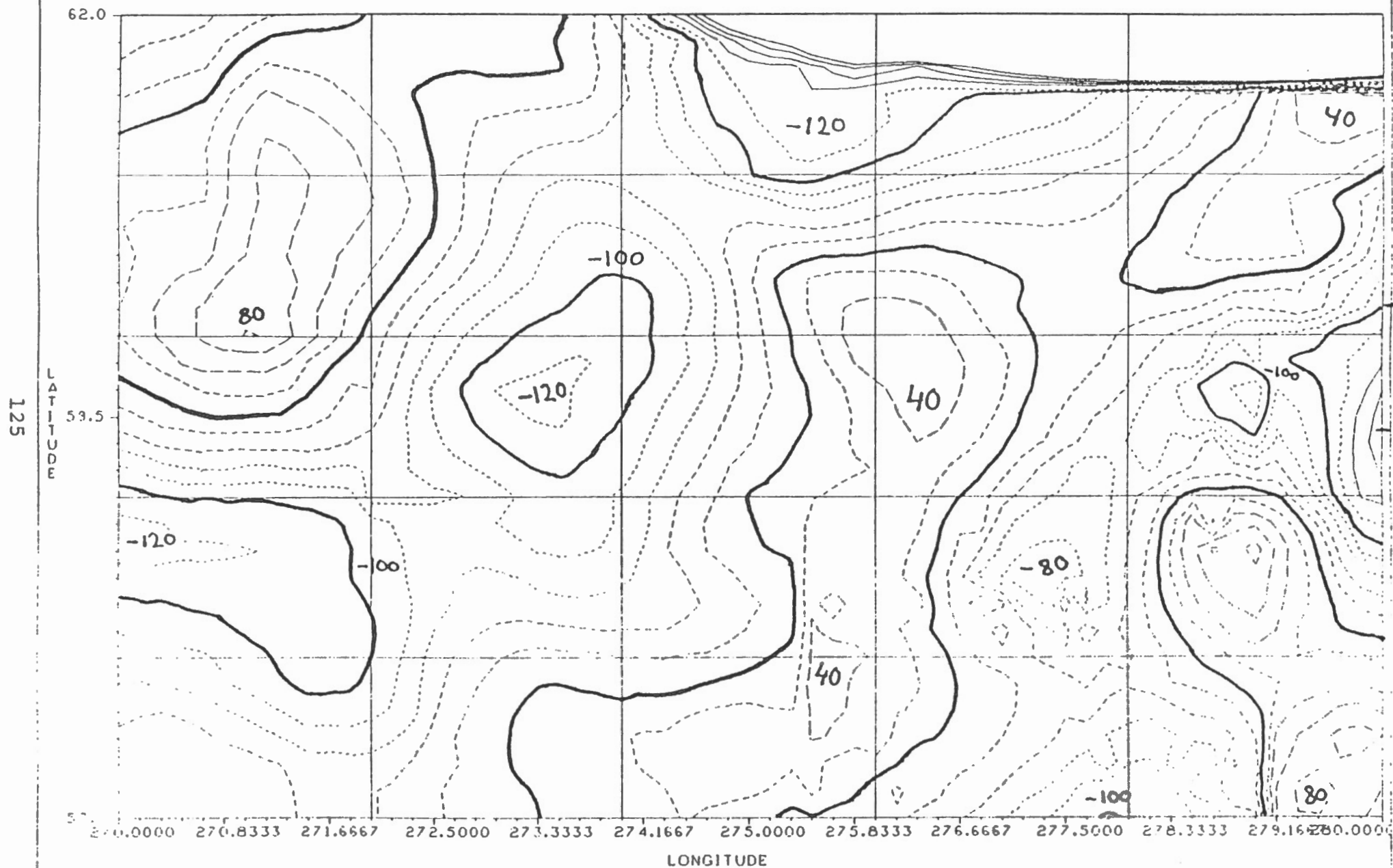Fig. 6.14: Stokes's integration result ($\psi_0 = 0.5°$) for Labrador Sea.
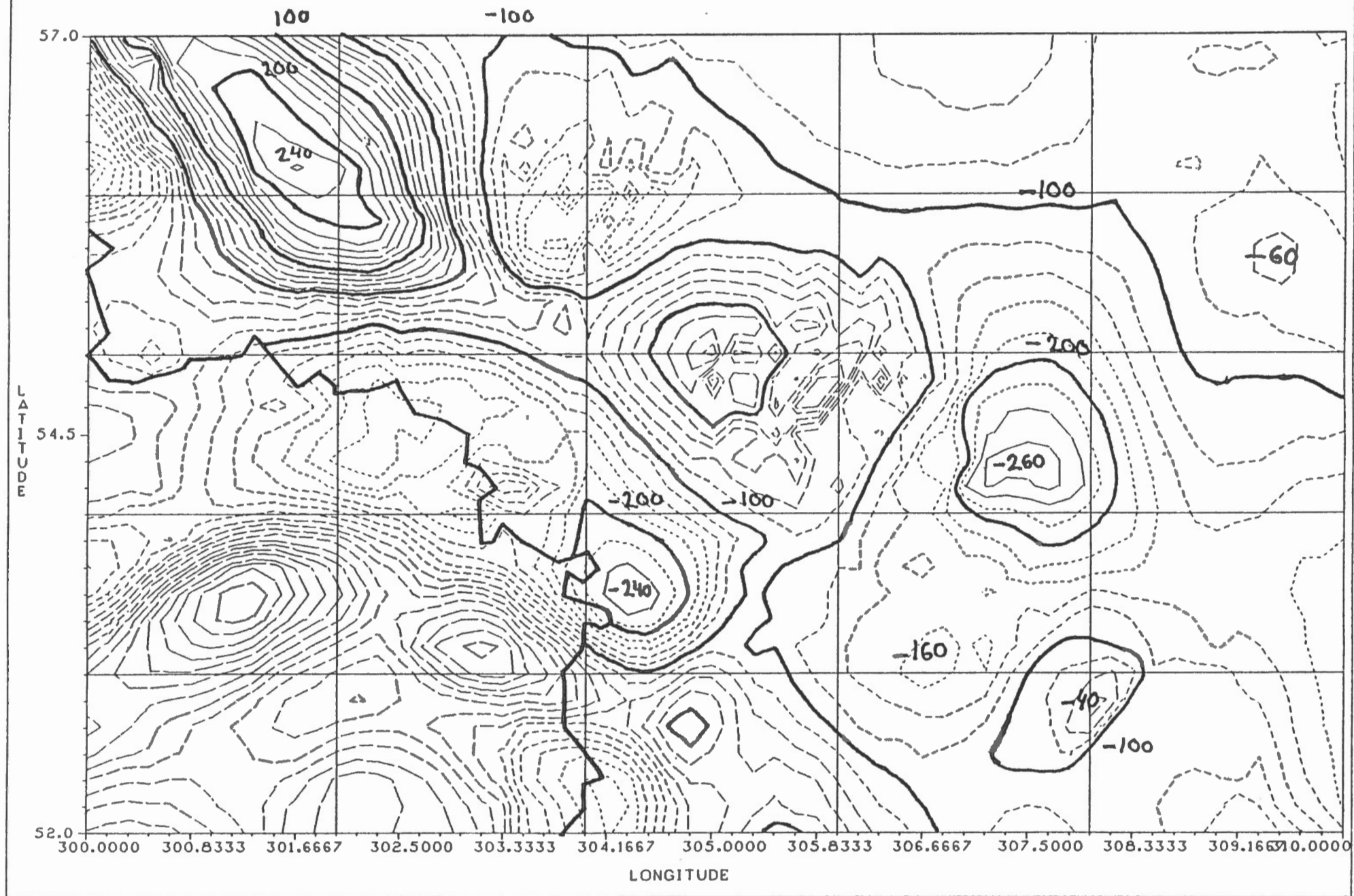
Fig. 6.15: Truncation integration result ($\psi_0 = 0.5°$) for Labrador Sea.
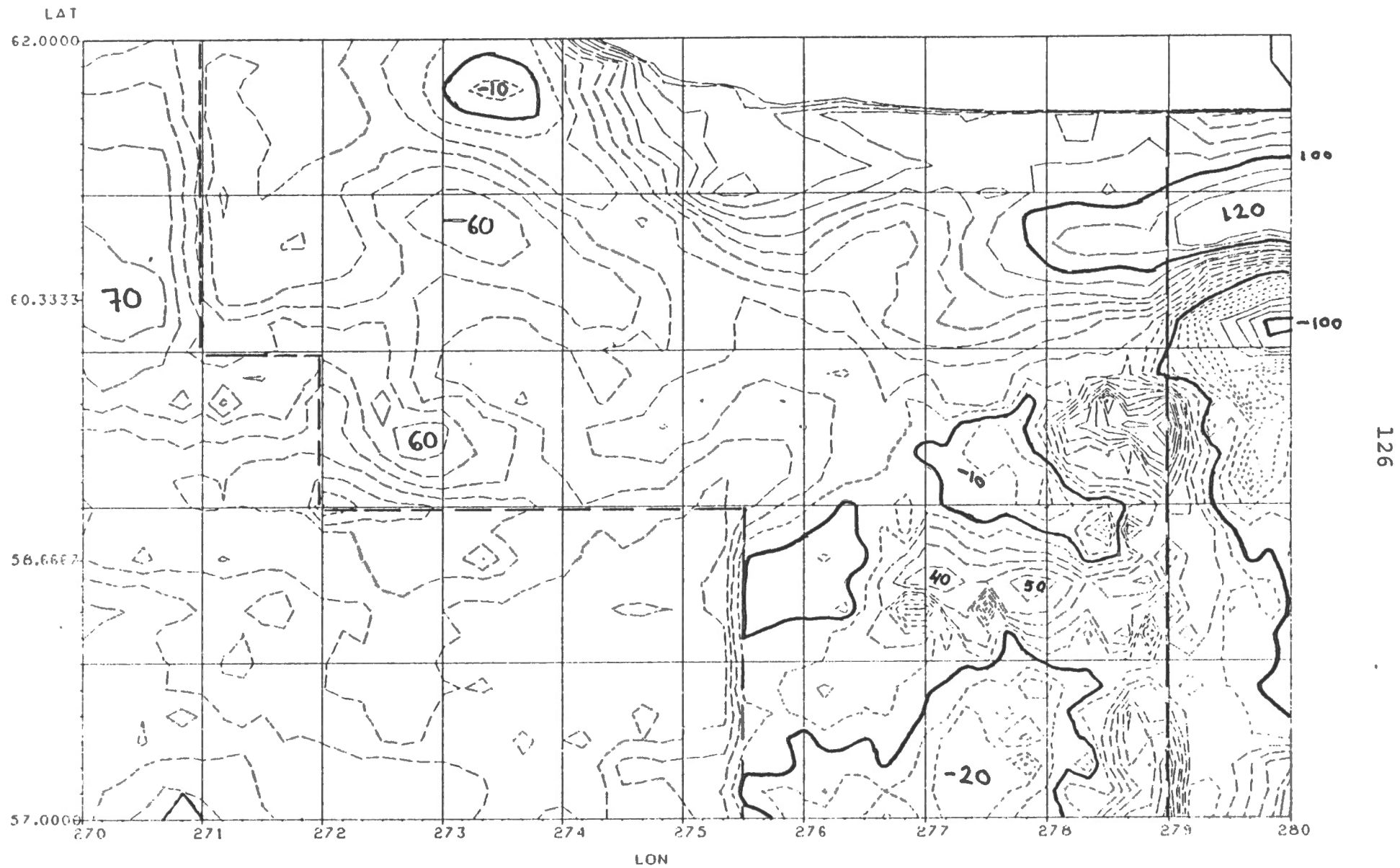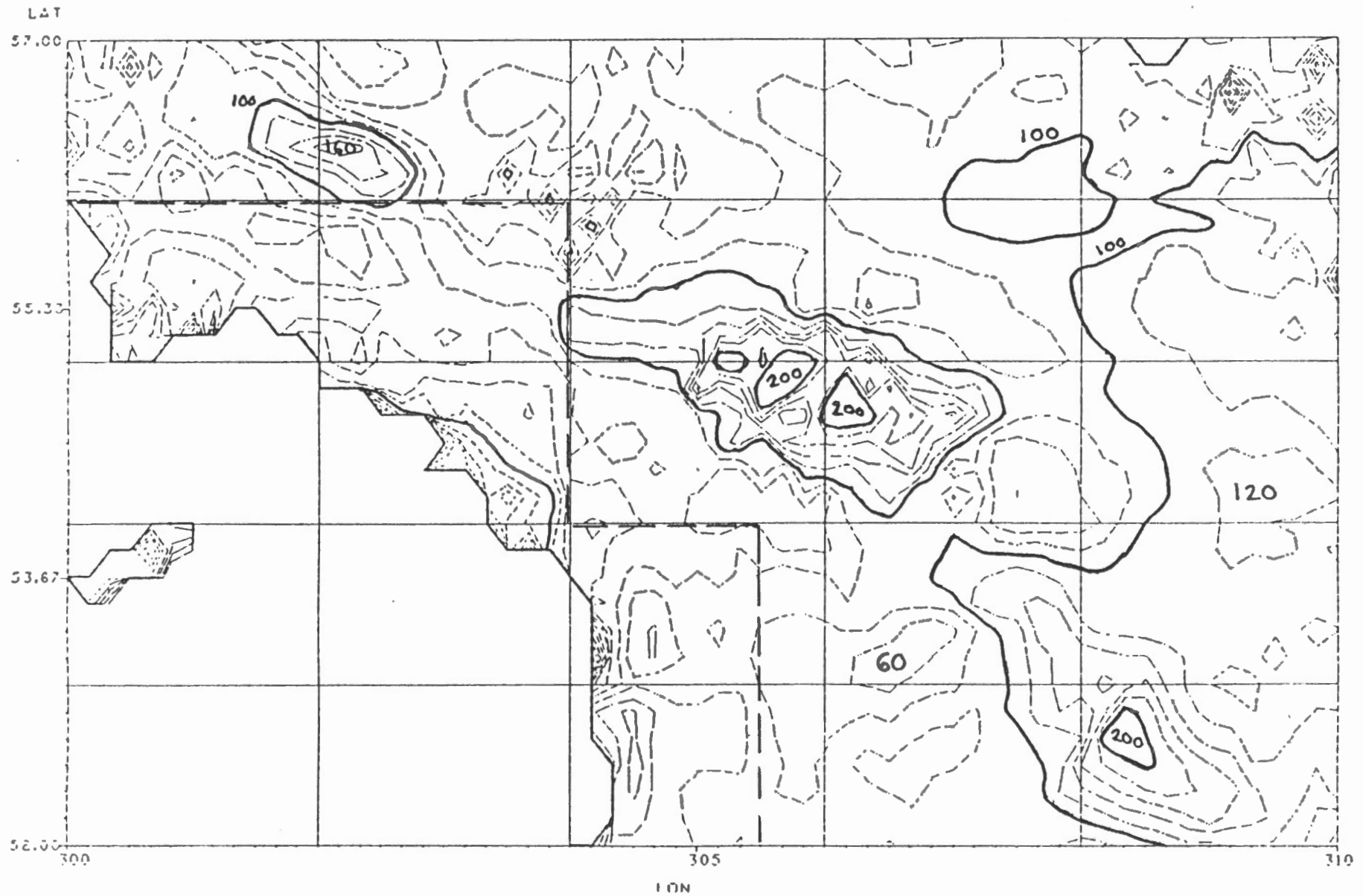
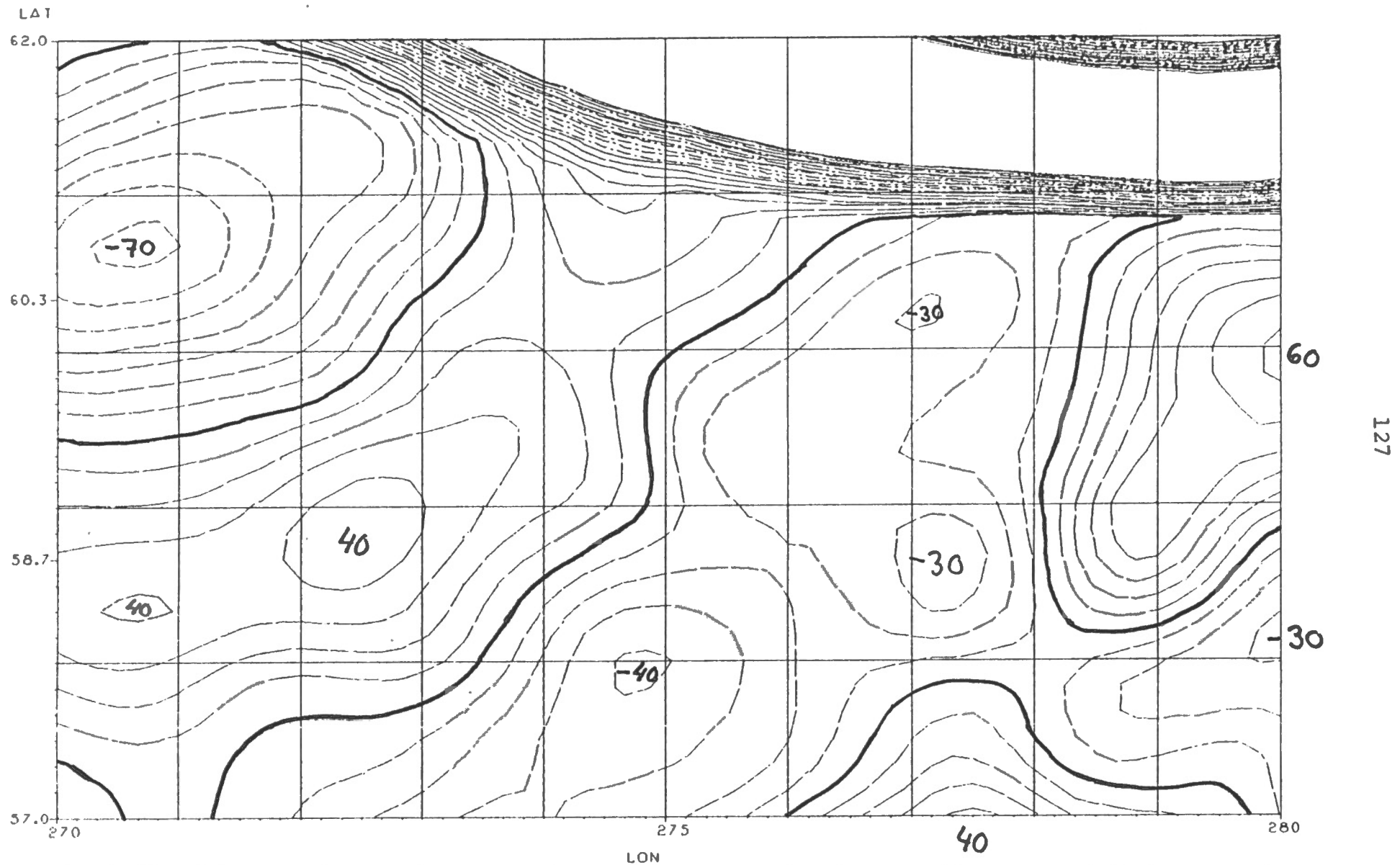Fig. 6.16: Difference between the Stokes and truncation integration ($\psi_0 = 0.5°$) for Labrador Sea.
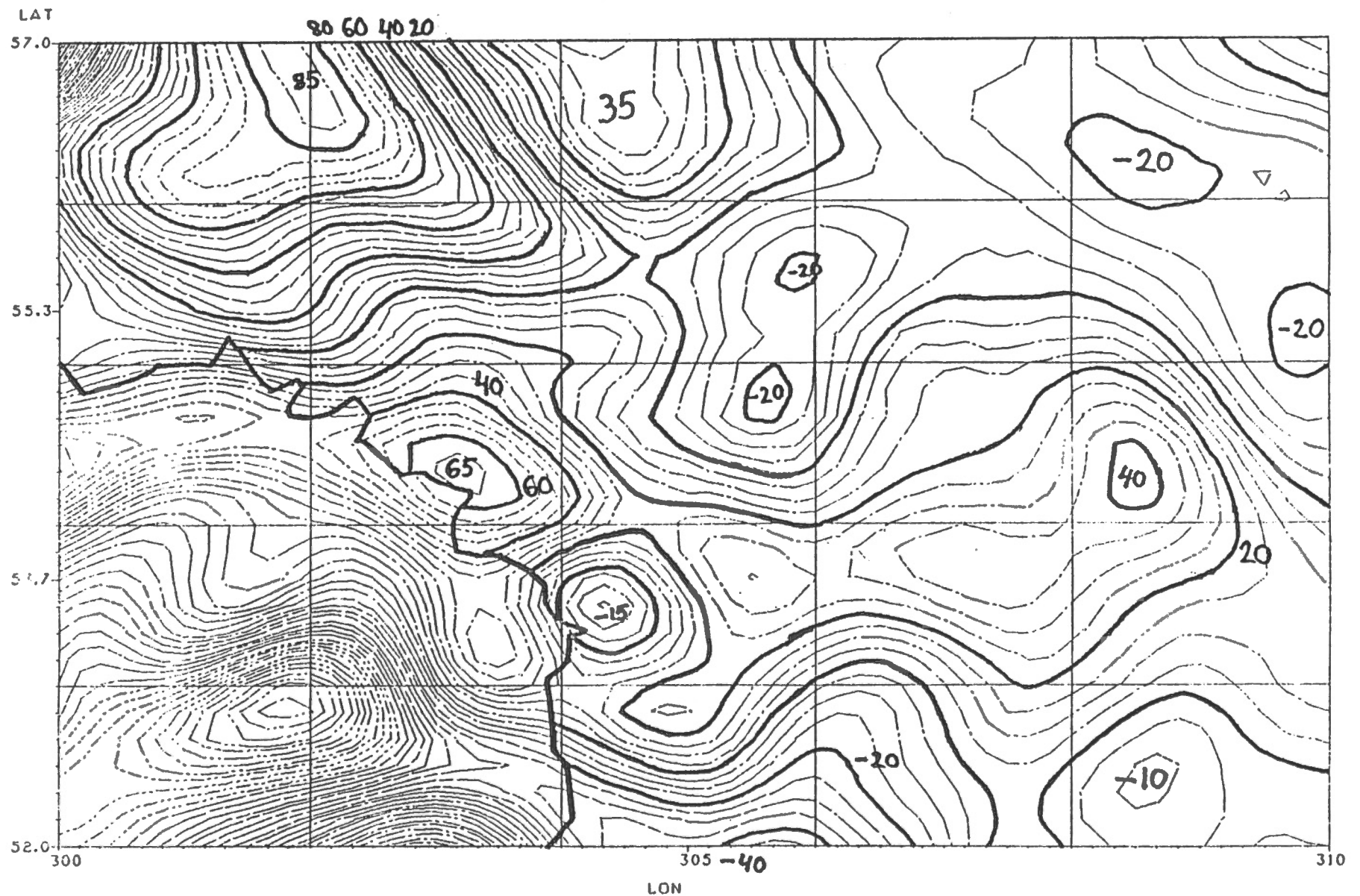
Fig. 6.17: Difference between truncation integration results for $\psi_o = 1°$ and $\psi_o = 0.5°$ in Labrador Sea.

7. UTILIZATION OF DIGITAL IMAGE ANALYSIS AND DISPLAY TECHNIQUES

7.1 Objectives

The overall goal of the work, as stated in Chapter 7 of the contract, was the

> Provision of expertise and software for digital image analysis and for display techniques using colour graphics, as part of Items 1 to 6 where applicable, and as an extension of them.

Specific operations proposed were:

1) representation of contours and raw or gridded data by colour coding;
2) display of data set overlays in colour composite and using colour mixing;
3) superimposition of processing flags on data sets;
4) mutual registration of data sets from different sources;
5) resampling raster data to change the pixel size for display purposes;
6) relief shading of data with directional illumination;
7) gridding data using interpolation which exploits data trends (principal component transformation);
8) manipulation of large images as collections of smaller images;
9) compression and decompression of image data (for storage, input, output);
10) Fourier filtering of image data (e.g., for cross-correlation, for cross-spectral techniques to analyse the relationship of bathymetry/elevation to gravity);
11) derivation of horizontal and vertical gradients, second derivatives (e.g., Laplacian);
12) correlation of anomalies using associated geophysical data like magnetic and gravity anomalies (e.g., searching for regions of preselected amplitude ranges for several parameters, using eigen-pictures in which pixels are presented by vectors associated with amplitudes of parameters defining the axes, comparing pseudo-gravity anomalies derived from magnetic anomalies using Poisson's relation);
13) multiple image operations (e.g., addition, subtraction, ratios).

7.2 Method of Approach

BIO and the contractor own a different type of data processing facility which does not permit a direct transfer of operational procedures. This impediment was recognized in the contract by the clause:

> Differences between the contractor's and the scientific authority's image analysis and graphics systems (and continuing development of this capability at AGC) make direct transfer of software problematic. Specification of the means of transfer must be decided within the spirit of the contract through mutual agreement between the contractor and the scientific authority.

According to the above guidelines, work on the contract proceeded in the following steps:

1) A comparative study of the two data processing systems was undertaken by means of hands-on experience and a review of the operator's manual, to identify operational and structural differences.
2) Problems and special needs associated with the processing of the data in question were identified and appropriate action taken.
3) General procedures were compiled for the implementation of the operations detailed in section 7.1.
4) Computer programs were developed for operations not available on the BIO facility.
5) The data sets available for or generated within this contract were processed in the contractor's facility, according to the operations proposed by the Scientific Authority and detailed in section 7.1. This processing was done subject to the limitations in variety, type, and size of the data sets. The results of this process were transferred to and stored in the BIO facility to serve as examples of the desired operations.

7.3 Comparison of the Two Image Processing Facilities

The data processing facility installed at BIO is the EASI/PACE (Expert Analysis System Interface/Pattern Analysis and Classification Expert) software that operates on a DEC VAX/VMS 750 hardware. It was developed by Perception Computing Inc. of Downsview, Ontario.

The contractor's facility is a Dipix Systems Ltd. ARIES II (Applied Resource Image Exploitation System) software that operates on a DEC LSI 11/23 hardware.

EASI/PACE is in many ways similar to ARIES II in its operating procedure and structure. A principal difference between the two is in the hierarchical nature of the EASI/PACE operating environment. EASI links the operator to PACE image analysis tasks via menu-prompts and a Parameter File, much like OIS in ARIES, to alleviate programming requirements. However, no EASI/PACE Parameter File equivalent exists on ARIES.

EASI/PACE parameters are global, unlike ARIES parameters, thus easing the tedious job of specifying parameters when repeating a task.

EASI/PACE does not have batch tasks like ARIES, with associated batch and queue files, to execute several processing tasks simultaneously or during off-hours.

EASI/PACE has a Procedure Level which acts like a fourth-level language, allowing the user to set up a sequence of pre-defined tasks that can be invoked as if it were a single-function task. This capability is missing in ARIES II, which is menu driven and presents a variety of options in each pass through a task.

No distinction exists between image and non-image files in EASI/PACE. It has, however, a more convenient graphics file handling approach by regarding both themes and graphics as bitmaps. An EASI/PACE inconvenience is the way images are loaded onto disk and into video display. The operator must assign space on disk and in video, whereas ARIES handles much of this requirement. EASI/PACE displays are limited to the screen size. In ARIES the size of the video memory is the limiting factor, and the opertor can roam in the image even in zoom one.

The software packages of both facilities serve essentially the same purpose and can produce similar results. There is, however, a difference in the variety of processing operations available and in the procedures that must be followed.

7.4 Preprocessing Operations

Two basic requirements imposed by both processing facilities are that:
- the data must be in a gridded or rasterized format, and
- it must be stored as a positive eight bit integer per grid cell, commonly referred to as a pixel.

Both the GSFC1 and the OSU sea surface heights are stored in profiles. Therefore, an interpolation procedure was developed and coded to rasterize these data sets (see section 3.1.4). The grid cell size chosen was 10 minutes of arc in geographic coordinates. Therefore, a shift of +100 m was applied to convert the whole range into positive values in units of metres, then a

multiplication factor of 100 was employed to convert all positive values from metres to centimetres.

Geoidal and adjusted sea surface heights ranged from -60 metres to +70 metres in the study area and were stored in one centimetre increments. Therefore, a shift of +100 m and a scale factor of 1/66 was applied, then rounded to the nearest integer to obtain all positive values and to fit the data within the eight bit (0 to 255) operating range. Each increment now represented a 66 cm change in height, which resulted in a substantial generalization. To partially alleviate this problem, new software was developed for operations which yield results in a range considerably narrower than that of the original data. These routines can handle values of any magnitude. The results are then reformatted into the eight bit pixel range for display purposes. (See section 7.6 for further particulars.)

Finally, the record length and the block size of each data set was changed to a dimension compatible with the input requirements of image processing systems.

7.5 Operational Procedures

Both the EASI/PACE and the ARIES II softwares were designed primarily for the processing and analysis of digital image data. However, both are capable of handling any data subject to the requirements and limitations stated in section 7.4. The method of approach is to select the particular processing task(s) from the software package which can perform a desired operation with acceptable results. Therefore, in answering the questions raised in the contract, the image processing activities which best fit a desired operation are described in general terms. Based on this explanation, a person familiar with the EASI/PACE system can identify the proper PACE task. In many instances, the required PACE tasks are listed as a guide. The contractor's three visits to BIO were insufficient to become intimately familiar with PACE. In fact, the analyst in charge was herself unsure of several aspects of the operation.

1. Representation of contours by colour coding. This can be achieved, in an elementary way, by the pseudo-colour display option DCP>PO. The 0 to 255 data range will be displayed in 36 system assigned colours. The colours approximately represent a seven unit contour interval.

Contouring with user assigned colours and interval can be achieved with density slicing. ARIES has a black and white and a colour density slicing option. PACE tasks THR, data base

image to bitmap THResholding, used in conjunction with LUT, image enhancement via Look-Up Table, perform the same operations.

2. Display of data sets in colour composites and colour mixing. This is provided by the multi-channel route of image display: DCP>ID 1 2 3 which superimposes image planes 1, 2,3 in the order of red, green, and blue.

A user defined pseudo-colour display of a single channel data can be created by defining a set of three colour-transformation functions for the three primary colours. The single channel data must then be mapped through these functions into three channels and displayed by DCP>ID 1 2 3. STR, data base image contrast STRetching, LUT, data base Look-Up Table transformation ,are the most likely tasks to use.

3. Superimposition of processing flags. This requires that the flags be stored in a bitmap, which can be created by THR, data base image to bitmap THResholding. In ARIES this can also be achieved by the image classification and annotation tasks. Tasks SCC, Single-Class Classification of image data, MCC, Multi-Class Classification of image data, followed by MAP, encode data base bitmaps in image channel (theme MAP creation) are the likely operations to use. It was not possible to obtain sufficient information on these tasks during the visits to BIO.

4. Mutual registration of data sets. This is a geometric registration problem which can be solved with task GCA, Ground Control point Acquisition (image to image). It must be followed by a resampling of the data. Probability task IRG, data base Image Recreation serves this purpose, but this could not be confirmed.

5. Resampling raster data to change the pixel size. This is an integral part of geometric registration and is a fundamental operation in any digital image analysis system. PACE must also have this capability and probability task IRG performs this function.

6. Relief shading with directional illumination. This task is not available in either image

processing facility and a program to perform this operation was developed by the contractor (see section 7.6).

7. Gridding data using interpolation. Neither system has the necessary software included to perform this operation. A program, using bilinear surface polynomial, was developed by the contractor (see section 3.1.4). The principal component transformation referred to under item 7 in section 7.1 can be performed by task HOT, data base image transforms: HOTelling, tasseled cap, etc., but this operation is not associated with gridding.

8. Manipulation of large images as collection of smaller images. This is probably possible via the task MER, MERge data base vector segments.

9. Compression and decompression of image data. This is not normally done in image processing system since most operations require direct access files.

The data should be stored in direct access files, rather than in sequential files, to perform image processing and display operations in an efficient manner. Therefore, it is not feasible to use compression and decompression techniques.

10. Fourier filtering. This is performed with task FFT, Fast Fourier Transform.

11. Derivation of horizontal and vertical gradients. This can be realized with spatial filtering. Task FLT, data base image spatial FiLTering has a number of preset low-pass, high-pass, and directional filters which can be convolved with the data base. Provision also exists for constructing user specified filter kernels.

12. Correlation of anomalies using associated geophysical data, etc. This can be accomplished with a variety of tasks depending on the nature of the correlation. MCC, the Multi-Class Classsification and ARI, two-channel data base image ARIthmetic tasks appear the

most useful. In ARIES, correlation matrices and area statistics can also be generated. The contractor was unable to ascertain whether this is also possible in PACE.

13. Multiple image operations. These are performed through task ARI, referred to above.

7.6 New Computer Programs

The following computer programs were written as part of the contract.

1. AS: Record Assembler. The record length of the input data to image processing systems must be equal to the number of pixels along one line. In other words, each line in the data set is a separate record. In AS, the output from the interpolation (gridding) program is re-assembled to satisfy this requirements. An optional scaling operation is also included which degrades data to the eight bits per pixel.

2. MT: Tape Format. This program writes the data to tape in a format which is compatible to MIB, the tape read and file creation task used at BIO.

3. FA: Differencing. In this program the arithmetic difference between two data sets is computed, pixel by pixel. It runs without the restriction of the eight bits per pixel range. Input to FA should be in a format generated by AS, without the scaling option. FA includes a scaling to eight bit pixels.

4. IL: Directional Illumination. The function of this program is to enhance the data by analytical relief shading operations. Relief shading is the creation of a three-dimensional impression by employing variations in light and dark. It is based on the principle that the lighting of a three-dimensional form will result in varying amounts of illumination on the varying slopes.

The method described by Yoeli [1965] and Woodham [1980] was used, whereby the illuminance of a surface area element is taken proportional to the cosine of the angle between the light ray and the normal to the area element. In other words, the illuminance varies with the direction of the illumination and the orientation of the surface. The direction of the illumination in turn is a function of the elevation angle of the light source and its azimuth. Both variables are controlled interactively during the execution of the program.

The input should be in a format generated by AS, without scaling. The output from IL is scaled to eight bit pixel.

5. IV: Inverting. Image processing operations assume that the origin of a data set is at the top left pixel. IV inverts a data set about the x-axis, whereby line n becomes line 1, line n-1 becomes line 2, and so on.

7.7 Examples

Data sets GSFC-1, OSU, and GEM-9 were processed in the contractor's Dipix ARIES II facility to demonstrate the utilization of digital image analysis and display techniques and to provide examples for the evaluation of the utility of processing techniques.

The following is a listing of the name and description of the files created with a short explanation of the processing techniques employed. Reference is made to the operations listed in section 7.1.

1. GCAFFGEM9. Black and white display of GEM-9 geoid.
2. GCAFFOHIO. Black and white dissplay of OSU sea surface heights.
3. GCAFFNAS1. Black and white display of GSFC-1 sea surface heights.
4. GCAFFSTND. Black and white display of the standard deviations of the OSU sea surface heights.

Files 1 to 3 were produced by applying a constant shift of +100 m and a multiplication factor of 100 to convert the gridded data to positive integers in units of centimetres. The range of values in the above files is from 40 to 235. Each increment corresponds to a change of 0.666 m in height. The standard deviations are stored in 5 cm increments. These files can be displayed as a black and

white image on the video monitor of any image analysis sytem, using readily available image display tasks.

5. GCAFFCOR1, GCAFFCOG1 and GCAFFCOB1. Colour density sliced OSU sea surface heights.

6. GCAFFCNR1, GCAFFCNG1 and GCAFFCNB1. Colour density sliced GSFC-1 sea surface heights.

In each case the total range of the data was sliced into 32 levels of equal spacing and then a different colour was assigned to each slice. A slice represents approximately 4.0 m change in height. For display purposes, a separate file is required for each primary colour. These examples demonstrate operation #1.

7. GCAFFPOR2, GCAFFPOG2 and GCAFFPOB2. Pseudo-colour display of OSU sea surface heights.

Transformation functions were applied to create an image file for each of the primary colours. A pseudo-colour image appears when the three files are displayed as a colour composite. This is an example of operation #2.

8. GCAFFDON. Difference of OSU and GSFC-1 sea surface heights (OSU-GSFC).

9. GCAFFDO12. Difference of OSU sea surface heights and GEM-9 geoidal heights.

10. GCAFFDN12. Difference of GSFC-1 sea surface heights and GEM-9 geoidal heights.

These files were created with the differencing task FA (see section 7.6) and serve as an example for operation #13.

11. GCAFFHOTL, GCAFFHOBR, GCAFFHOTR and GCAFFHOBL. Analytical relief shading enhancements of the OSU sea surface heights. Assumed direction of illumination is from the north-west, south-east, north-east, and south-west, respectively, at an elevation angle of 30°.

12. GCAFFDITL, GCAFFDITR, GCAFFDIBR and GCAFFDIBL. Analytical relief shading enhancements of OSU and GEM-9 difference (file GCAFFDO12). Same illumination parameters as for item 11.

13. GCAFFHGTR. Analytical relief shading enhancement of GEM-9 geoidal heights. Assumed direction of illumination is from the north-east at 30° elevation angle.

The above files were generated by task IL and demonstrate operation #6.

14. GCAFFPDOR, GCAFFPDOG and GCAFFPDOB. Pseudo-colour enhancement of differences between OSU and GEM-9 data (file GCAFFDO12).

Same explanation as for item 7.

15. GCAFFSTD2. Classification and colour coding of the standard deviations of the OSU sea surface heights (file GCAFFSTND) into the following ranges:

blue: 0.8 cm - 1.6 cm

magenta: 1.6 cm - 2.5 cm

cyan: 2.5 cm - 4.5 cm

green: 4.5 cm - 12.5 cm

red: larger than 12.5 cm

This is a theme file (bitmap) produced by the parallelepiped classification task which can be displayed by itself or superimposed on the GCAFFOHIO image. The 'searching for regions of preselected amplitude range' which forms part of operation #13 is illustrated.

16. GCAFFGEOL. Zero height line for GEM-9 data.

17. GCAFFOHOL. Zero height line for OSU data.

18. GCAFFNAOL. Zero height line for GSFC-1 data.

19. GCAFFDOGO. Lines of zero difference between OSU and GEM-9 data.

20. GCAFFDONO. Lines of zero difference between OSU and GSFC-1 data.

21. GCAFFDNGO. Lines of zero difference between GSFC-1 and GEM-9 data.

22. GCAFGRZD. A geographic coordinate grid at 2.5° spacing.

The above items are theme files (bitmaps) produced by classification which can be superimposed in colour coding on appropriate data sets. They serve as examples for operation #3.

23. GCAFFCOS2. Coastline.

This is a generalized version of the North American coastline, resampled to fit the pixel size of the other data sets, colour coded, and converted into a theme file (bitmap). It demonstrates operations #4 and #3.

All files listed in this section have been installed in the image processing system at BIO.

7.8 Evaluation of the Examples

In this section, the visual appearance and potential information content of data displays, produced by various techniques, are compared. No attempt is made to interpret the displays from a geoscience stand point.

A gray scale display of unenhanced data sets is not very useful since the general trend of the geoidal and sea surface heights is barely visible. It shows a gradual change from large negative values in the Hudson Bay area to large positive values in mid-Atlantic Ocean. Local undulations of a few metres are not discernible with the human eye at the 66 cm gray scale increments. This is

especially so towards the lower and upper ends of the gray scale, due to a non-linear response of the vidicon tube of the display monitor.

Density slicing provides an accentuated view of the changes in surface heights, especially when colour coding is introduced. Unfortunately, the contour-like slices create the impression of step-like changes. Pseudo-colour composites prove more useful. They preserve the impression of the global change in the surface heights and, at the same time, also depict major local undulations.

The relief shading enhancement creates the most dramatic effect. Since the slope map is computed from the original data set at one centimetre increments, local anomalies become clearly visible. Furthermore, by changing the direction of the illumination, undulations with different orientation could be viewed.

The difference images are also generated from the original data sets whereby small variations between the two surfaces are clearly visible.

8. CONCLUSIONS

The primary goal of this contract report was the investigation of possibilities to extract gravity field information from satellite altimetry to check and/or replace marine gravity observations. Since altimetry refers to the sea surface and not to an equipotential surface of the gravity field (geoid), we have first summarized in section 2 the **definitions** of all quantities involved in the relation between altimetry and gravity.

Several data sets were acquired to facilitate the investigations. **Marine gravity** was made available by Energy, Mines and Resources Canada in the form of single point data for Hudson Bay and the North Atlantic, mainly on the Canadian continental shelf. Three different SEASAT **altimetry** data sets were acquired from the Ohio State University, the Goddard Space Flight Center, and the U.S. National Geodetic Survey. The OSU data was found to be best suited for our research since, for the areas of interest, it was corrected for tides and orbital errors in a consistent way. The GSFC altimetry was not adjusted in a consistent way, and the NGS altimetry consisted of more or less raw data. Gridding algorithms have been developed to interpolate sea surface heights on a regular geographic grid from measurements along satellite tracks. A **gravimetric geoid** was computed based on satellite derived potential coefficients, point gravity anomalies, and mean gravity anomalies. For the North Atlantic regions, 5' x 5' mean gravity anomalies had to be computed from the point gravity data.

Sea surface topography has been computed in three different ways. The Levitus dynamic topography was found to represent mainly global features off the continental shelf. The SST as determined as the difference between the altimetry sea surface and the gravimetric geoid was shown to be most accurate for features of local and regional extensions in the case of the continental shelf. Long wavelength features are distorted by errors in both the altimetry and the gravimetric geoid. Several very pronounced anomalies in this SST were detected along the ridge of the continental shelf. These anomalies certainly deserve more research, especially in determining their causes. The SST obtained from the zero frequency response technique was restricted to points along the coastline. The altimetric/gravimetric SST is unreliable in these areas due to spurious reflections from the land. The main conclusion regarding the SST is that the three methods discussed represent different features in different areas. They should be seen as complementing each other.

Section 4 discussed the **transformation between different functionals of the gravity field** of the earth, mainly the equipotential surface (geoid) and the gravity anomaly. It was shown

that the direct transformation of geoidal heights into gravity anomalies as given by Molodenskij's integration is a numerically unstable problem, i.e., small changes in the geoidal heights may produce very large changes in the gravity anomaly result.

Section 5 contains the main mathematical contribution of this report. A kernel function, called the **truncation kernel,** is derived for the computation of partial spheroidal Stokes's contributions by integrating over geoidal heights in a local area. This technique can be used to compare integrals over gravity anomalies with integrals over geoidal heights in local areas. The integrals are numerically stable even for small integration areas. Some non-standard mean gravity anomalies can be derived from the integral over geoidal heights simply by multiplying with a conversion factor given in eqn. (5.89). It should be emphasized that this technique is completely new, and that it has been developed particularly for this research contract.

Section 6 contains **applications** of the 'truncation integration' in the two test areas: Hudson Bay, and the Labrador Sea. For these applications, the SEASAT altimetry was augmented by the UNB gravimetric geoid for coastal and land areas to facilitate truncation integrations all over the test areas. No SST was subtracted for the marine regions. For each area, integrations were performed with Stokes's truncation radii of 0.5° and 1°. The comparison of Stokes's and the truncation integrations show that the technique developed for this contract is feasible.

The differences between the two results are compatible with possible regional biases in marine gravity data. For both the 1° and 0.5° integration caps, the spectral content of both the truncated Stokes and truncated altimetric geoid appear to be about the same, showing—at least at this level of resolution—that the altimetry possesses enough high-frequency information for meaningful comparison with marine gravity.

This conclusion is based, so far, solely on visual inspection. A more formal answer to the question of spectral content would have to be obtained through a formal analysis of the two kinds of results. This would call for the development of a formalism (e.g., double Fourier trigonometric series, averaging of one-dimensional spectra along different directions) and a time investment incongruent with the present contract.

Generally, the results of 'truncating' the altimetric geoid are intuitively pleasing and should be considered good. The magnitude of the truncated geoid, even when it is referred to a (90, 90) field, is sizeable, and the sensitivity to regional biases in gravity is quite high. To further develop this tool, it would be advisable to derive truncation kernels corresponding to integration radii of less than 0.5°, i.e., $\psi_0 = 1/3°$, 0.25°, 0.1°, to establish to what extent the 'truncated' geoid can be 'focused' to very local features of the gravity field.

Discussing the spectral content, there is another aspect which should be clarified. This is the

relation between the 'truncated' geoid computed here and a 'spectral window' (combination of certain harmonics in the development of the gravity field into spherical harmonic functions) in higher frequencies. From a visual inspection, there appears to be some correspondence between these two ways of treating the geoid. Understanding this correpondence would go a long way in the understanding and physical interpretation of the 'truncated' geoid. This understanding, however, is beyond the scope of this project.

Section 7 deals with the **utilization of digital image analysis and display techniques.** This investigation has confirmed that digital image analysis systems are valuable tools for manipulating and displaying such data. The interpretability of displays created in this manner is superior to that of contour maps. Most of the data processing can be accomplished with the existing software packages of such systems. A skilled image analyst, in close collaboration with a geoscientist, can generate the desired results. Unfortunately, the constraint imposed by the eight bit pixel often requires that the data be degraded. This loss in spatial resolution can be offset, to a certain extent, by the introduction of colour and various other enhancements. Analytical relief shading is especially effective. It depicts localized, low-amplitude undulations of the surface. It is important that data processing operations which yield a range of values considerably narrower than that of the original data sets be performed on the undegraded data. Most image processing packages cannot handle pixel values larger than eight bits, therefore, customized software is needed for this purpose. Such a program was written for the differencing and analytical relief shading. The three data sets available for the experiments were insufficient to demonstrate all operations proposed under this contract. Nevertheless, sufficient evidence was provided to prove the utility of digital image analysis and display techniques to marine geodesy.

Turning now to the as yet unanswered specific questions of the contract, we can offer the following answers:

(i)   What is the best level of agreement between the observed marine gravity and altimetrically-derived gravity? The comparison is done in the 'geoid space' by comparing the 'truncated' altimetric geoid with the truncated Stokes integration over marine gravity. The differences are plotted in Figures 6.6 and 6.9 for the two areas of interest, and they seem to indicate that regional biases over 1° and 0.5° caps of up to 10 mGal may exist in the observed marine gravity. It is doubtful whether by shrinking the (integration) cap further better accuracy than some 5 mGal can ever be expected in the bias detection.

(ii)   What level of confidence can be put on altimetrically-derived geoid for filling in data gaps in marine gravity coverage, and over gaps of what size? The altimetrically-derived sea surface has a very good short wavelength accuracy (standard deviation of about 5 cm) and this accuracy is retained by the 'truncated' geoid. To what extent this surface is representative of the geoid depends

8. Conclusions.

on the SST which is poorly known. Assuming that the SST is predominantly of a long wavelength character, the truncated geoid referred to a (90, 90) reference field would be relatively insensitive to SST effects. On the other hand, the truncation kernel, as developed here, is inaccurate for wave numbers just above 90 and thus medium wavelength errors of the order of 10 cm may creep in. Further, the smaller the truncation radius, the less sensitive the truncated geoid is to regional gravity biases ($\psi_0 = 1°$: 1 mGal ... 12 cm, $\psi_0 = 0.5°$: 1 mGal ... 6 cm). The 'Stokesian average' Δg^* over a 1° cap can be obtained to an accuracy of some 2 mGal, over 1/2° cap to some 4 mGal, etc.

(iii)   How well can we bridge the coastal gap between marine and land surveys by interpolation using satellite-derived gravity? Unfortunately, not very well, since altimetry data are weak along coastlines as well due to contamination from land reflections. Even when an accurate gravimetric geoid is used, the coastal zone remains a problem. It seems to us that this problem could be more readily solved in the 'geoid space' than the 'gravity space'. In the 'geoid space', the problem reduces to the development of an algorithm for realistic merging of altimetric sea surface with gravimetrically derived geoid. Attempts along these lines are now being made at UNB under the terms of an NSERC operating grant.

(iv)   What are the spectral characteristics of different data sets? This question has been answered, to a large extent, earlier in this section. The tricky part of this question concerns the 'raw' altimetry data in the satellite track form. How does one define a spectral characteristic of two-dimensional anisotropic data? We feel that only the gridded data can be investigated spectrally, and the suggestions given above would then pertain.

(v)   How realistic is the predicted error of the altimetrically-derived gravity? This question has been basically answered under (ii). Generally, the larger the area over which the 'Stokensian average' Δg^* is computer the better the accuracy that can be realistically obtained.

REFERENCES

Abramowitz, M. and I.A. Stegun (1964). *Handbook of Mathematical Functions*. Dover reprint 1965.

Ben-Menaham, A. and S.J. Singh (1981). *Seismic Waves and Sources*. Springer, N.Y.

Chelton, D.B. and D.B. Enfield (1985). "Ocean signals in tide gauge records." Submitted to *Journal of Geophysical Research*.

Christou, N. and R. Yazdani (1986). "Processing of SEASAT altimetry data." Paper presented at the Annual Canadian Geophysical Union Meeting, Ottawa.

Coleman, R., and R.S. Mather (1976). "Computational procedures for the use of the inverse of Stokes's operator." University of New South Wales, Unisurv G24, Sydney, Australia, pp. 123-139.

Dehlinger, P. (1978). *Marine Gravity*. Elsevier Oceanography Series, 22, Elsevier Scientific Publishing Company, Amsterdam.

Earth Physics Branch (1980). "Gravity anomaly map of Canada." Earth Physics Branch of Energy, Mines and Resources Canada, Ottawa.

Earth Physics Branch (1985). "Anomaly data base and image output description." Divison of Gravity, Geothermics and Geodynamics, Earth Physics Branch of Energy, Mines and Resources Canada, Ottawa.

Earth Physics Branch (1986). "Integration of Atlantic Geoscience Centre marine gravity data into the National Gravity Data Base." Joint publication of Earth Physics Branch and Atlantic Geoscience Centre, EPB Open File 85-32, Ottawa.

Fomin, L.M. (1964). *The Dynamic Method in Oceanography*. Elsevier, Amsterdam.

Gorny, A. (1977). "World Data Bank II: General users guide." Vol. 1, Central Intelligence Agency, Office of Geographic and Cartographic Research, Washington, D.C., U.S.A.

Haxby, W.F., G.D. Karner, J.L. LaBrecque and J.K. Weissel (1983). "Digital images of combined oceanic and continental data sets and their use in tectonic studies." *EOS Transaction*, Vol. 64, No. 52, pp. 995-1004.

Hearty, B. (1986). Personal communication. Division of Gravity, Geothermics and Geodynamics, Earth Physics Branch of Energy, Mines and Resources Canada, Ottawa.

Hobson, E.W. (1931). *The Theory of Spherical and Ellipsoidal Harmonics*. Cambridge University Press.

Kaula, W. (1966). *Theory of Satellite Geodesy*. Blaisdell, Waltham, Ma.

Kassim, F.A. (1980). "An evaluation of three techniques for the prediction of gravity in Canada." Department of Surveying Engineering Technical Report No. 73, University of New Brunswick, Fredericton.

Lanzano, P. (1982). *Deformations of an Elastic Earth.* Academic Press.

Levitus, S. (1982). "Climatological atlas of the world oceans." NOAA Professional Paper 13, U.S. Department of Commerce, Washington, D.C., U.S.A.

Lorell, J., M.E. Parke and J.F. Scott (1980). "SEASAT Geophysical Data Record (GDR) users handbook (altimeter)." Document 622-97, Revision A, Jet Propulsion Lab, Pasadena, Ca, U.S.A.

Marsh, J.G. (1985). Personal communication. Goddard Space Flight Center, Greenbelt, Md, U.S.A.

Marsh, J.B. and T.V. Martin (1982). "The SEASAT altimeter mean sea surface model." *Journal of Geophysical Research,* Vol. 87, No. C5, pp. 3269-3280.

Marsh, J.G., A.C. Brenner, D.B. Beckley and T.V. Martin (1984). "Global mean sea surface based on the SEASAT altimeter data." Preprint.

Mather R.S., R. Coleman and O.L. Colombo (1976). "On the recovery of long wave features of the sea surface topography from satellite altimetry." University of New South Wales, Unisurv G24, Sydney, Australia, pp. 21-26.

Merry, C.L. and P. Vaníček (1982). "Investigation of local variations of sea-surface topography. *Marine Geodesy,* 7, pp. 101-126.

Molodenskij, M.S., V.F. Eremeev and M.I. Yurkina (1960). *Methods for Study of the External Gravitational Field and Figure of the Earth.* Translated from Russian by the Israel Program for Scientific Translations for the Office of Technical Services, U.S. Department of Commerce, Washington, D.C., U.S.A., 1962.

Montgomery, R.B. (1969). "Comments on oceanic levelling." *Deep Sea Research,* Supplement to Vol. 16, Pergamon Press, pp. 147-152.

Olbers, D.J. and J. Willebrand (1984). "The level of no motion in an ideal fluid." *Journal of Physical Oceanography,* 14, pp. 203-212.

Paul, M. (1973). "A method of evaluating the truncation error coefficients for geoidal heights." *Bulletin Géodésique,* pp. 413-425.

Pond, S. and G.L. Pickard (1983). *Introductory Dynamical Oceanography.* 2nd ed., Pergamon Press, Oxford, England.

Rapp, R.H. (1981). "The earth's gravity field to degree and order 180 using SEASAT altimeter data, terrestrial gravity data, and other data." The Ohio State University Report No. 322, Columbus, Oh, U.S.A.

Rapp, R.H. (1982a). "A global atlas of sea surface heights based on adjusted SEASAT altimeter

data." Department of Geodetic Science and Surveying Report No. 333, The Ohio State University, Columbus, Oh, U.S.A.

Rapp, R.H. (1982b). "A summary of the results from the OSU analysis of SEASAT altimeter data." Department of Geodetic Science and Surveying Report No. 335, The Ohio State University, Columbus, Oh, U.S.A.

Rapp, R.H. (1983). "The determination of geoid undulations and gravity anomalies from SEASAT altimeter data." *Journal of Geophysical Research*, Vol. 88, No. C3, pp. 1552-1562.

Rowlands, D. (1981). "The adjustment of SEASAT altimeter data on a global basis for geoid and sea surface height determinations." Department of Geodetic Science and Surveying Report No. 325, The Ohio State University, Columbus, Oh, U.S.A.

Schwiderski, E.W. (1980). "Global ocean tides, a detailed hydrodynamical interpolation model." *Marine Geodesy*, 3, pp. 161-217.

Stewart, R.H. (1985). *Methods of satellite oceanography*. University of California Press, Berkeley, Ca, U.S.A.

Talwani, M. (1970). "Gravity." In: *The Sea*, ed. A.E. Maxwell, Vol. 4, Part 1, Wiley-Interscience, N.Y., U.S.A., pp. 270-297.

Tapley, B.D., G.H. Born and M.E. Parke (1982). "The SEASAT altimeter data and its accuracy assessment." *Journal of Geophysical Research*, Vol. 87, No. C5, pp. 3179-3188.

U.S. National Research Council (1985). *Geodesy: A Look to the Future*. Committee on Geodesy, National Academy Press, Washington, D.C., U.S.A.

Vaníček, P. and E.J. Krakiwsky (1982). *Geodesy: The Concepts*. North-Holland, Amsterdam.

Vaníček, P., J.D. Boal and T.A. Porter (1972). "Proposals for a more modern system of heights for Canada." Surveys and Mapping Branch Technical Report No. 72-3, Department of Energy, Mines and Resources, Ottawa.

Vaníček, P., A. Kleusberg, R.G. Chang, H. Fashir, N. Christou, M. Hofman, T. Kling and T. Arsenault (1986). "The Canadian geoid." Final contract report to the Geodetic Survey of Canada of Energy, Mines and Resources Canada, Ottawa.

Wenzel (1985). Personal communication.

Woodham, R.I. (1980). "Using digital terrain data to model image formation in remote sensing." SPIE Vol. 238 - *Image Processing for Missile Guidance*, pp. 361-369.

Wunsch, C. (1978). "The North Atlantic circulation west of 50°W determined by inverse methods." *Rev. Geophys.* 16, pp. 583-620.

Yazdani, R., N. Christou and E. Derenyi (1986). "Processing of SEASAT altimetry data on a digital image analysis system." Presented at the Canadian Symposium on Remote Sensing, Edmonton, May.

154

Yoeli, P. (1965). "Analytical hill shading (A cartographic experiment)." Surveying and Mapping, Vol. XXV, No. 4, pp. 253-259.

Zlotnicki, V. (1984). "On the accuracy of gravimetric geoids and the recovery of oceanographic signals from altimetry." *Marine Geodesy,* 8, pp. 129-157.

APPENDIX A

OPERATOR INTERACTION SEQUENCE FOR RUNNING IMAGE PROCESSING PROGRAMS INSTALLED AT BIO BY CONTRACTOR

PROGRAM IL: DIRECTIONAL ILLUMINATION

Given a raster digital file and an assumed light source, the program creates a file of terrain reflectance values. The azimuth, the elevation angle of the light source, and the pixel size are inputs to the program. The prompts for running the program with sample answers are as follows:

1.	Enter name of input file:	(INPUT.DAT)
2.	Enter number of rows:	(211)
3.	Enter number of columns:	(541)
4.	Enter grid spacing:	(50)
5.	Enter azimuth of the light source in deg, min, sec:	(45,0,0)
6.	Enter elevation angle in deg,min,sec:	(20,0,0)
7.	Do you wish to use previously calculated reflectance values if slope is within 5% (Y/N)	(N)
8.	Enter name of output file:	(OUTPUT.DAT)

PROGRAM FA: DIFFERENCING

This program calculates the difference of two VMS files. The option in the program allows the creation of an image file for display with proper offset and scale values. The input to the program is two files, and the output is either an image file scaled, or a VMS file. The step by step procedure is as follows:

1.	Enter name of input file (F1):	(F1.DAT)
2.	Enter name of input file (F2):	(F2.DAT)
3.	Enter number of rows:	(211)
4.	Enter number of columns:	(541)
5.	Do you wish to create the pixel file (Y/N)?	
6.	Enter name of output file:	(OUT.DAT)
7.	Enter minimum value for the difference:	(-20)
8.	Enter maximum value for the difference:	(20)

PROGRAM AS: RECORD ASSEMBLER

This program assembles a number of records into one record. It is necessary for creating a master file in which each line is a separate record, as required by programs IL and FA. Step by step procedure is as follows:

1. Enter name of input file: (INPUT.DAT)
2. Enter number of records: (11605)
3. Enter number of columns: (541)
4. Do you wish to create the image file (Y/N)?
5. Enter name of output file   (OUTPUT.DAT)
6. Enter number of records per line for assembly: (55)
7. Please enter the scale factor: (0.01)

PROGRAM IV: INVERTING

This program inverts data about the x-axis. This means that the last record of the input data becomes the first record of the output data.

1. Enter name of input file: (INPUT.DAT)
2. Enter number of records: (11.605)
3. Enter number of columns: (541)
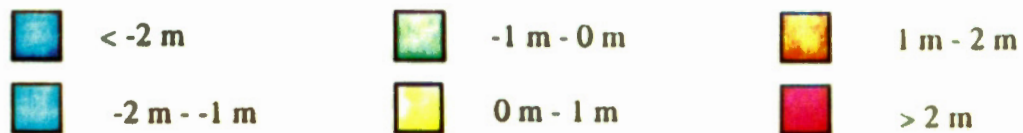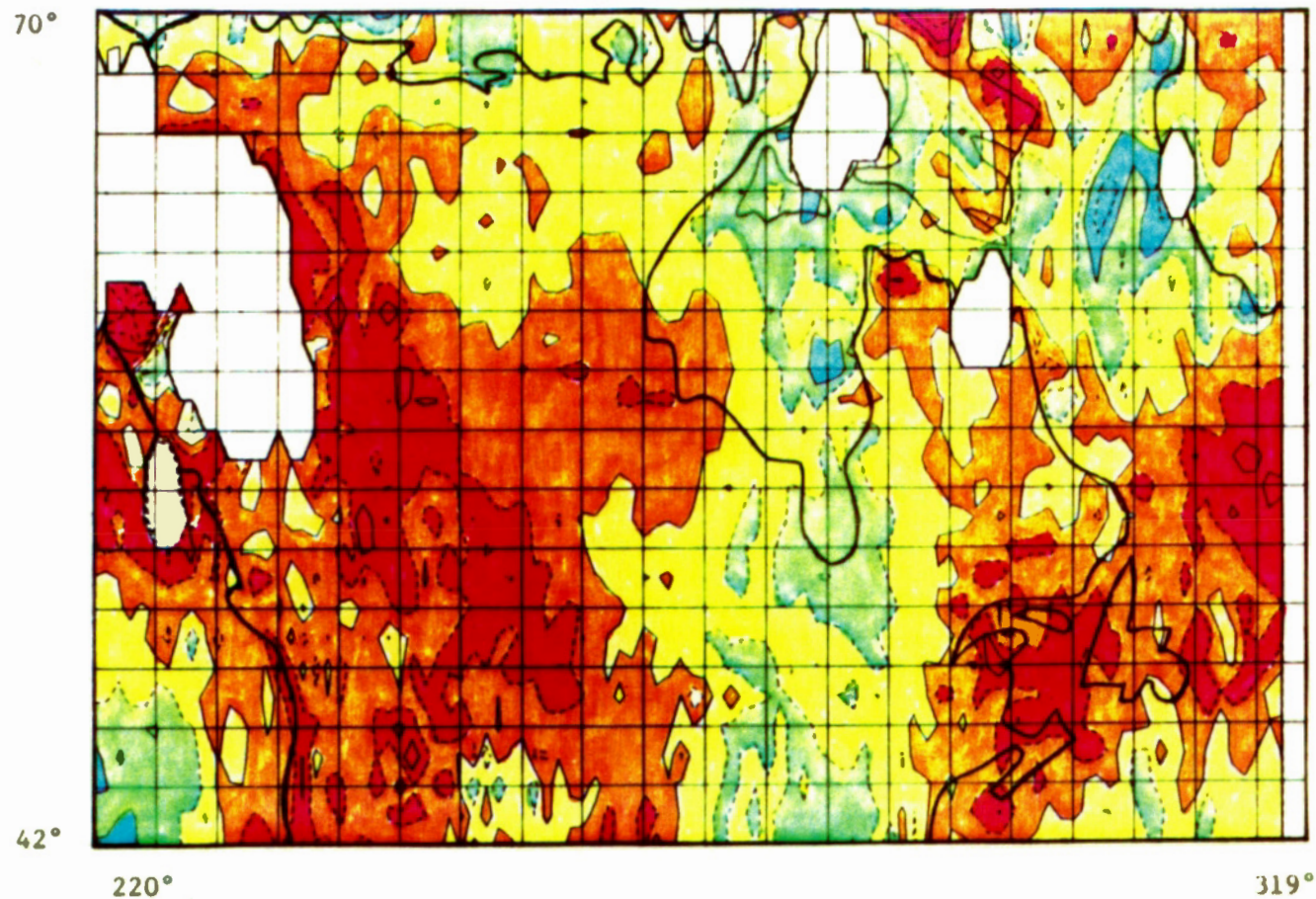4. Enter name of output file: (OUTPUT.DAT)

PROGRAM NT: TAPE FORMAT

This program writes the image file (8 bits) onto a tape which is compatible with the MLB format.

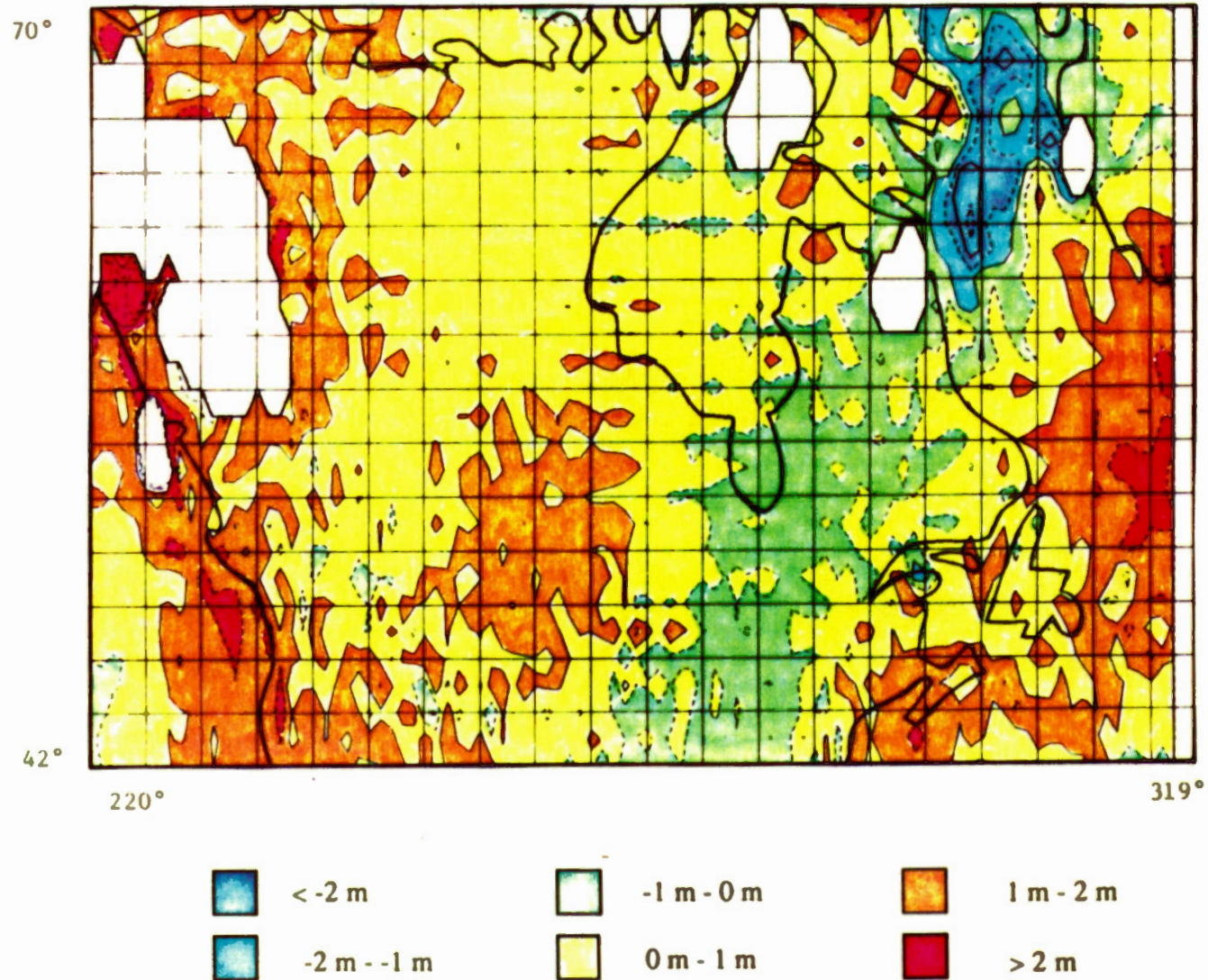1. Enter name of input file: (INPUT.DAT)
2. Enter number of rows: (211)
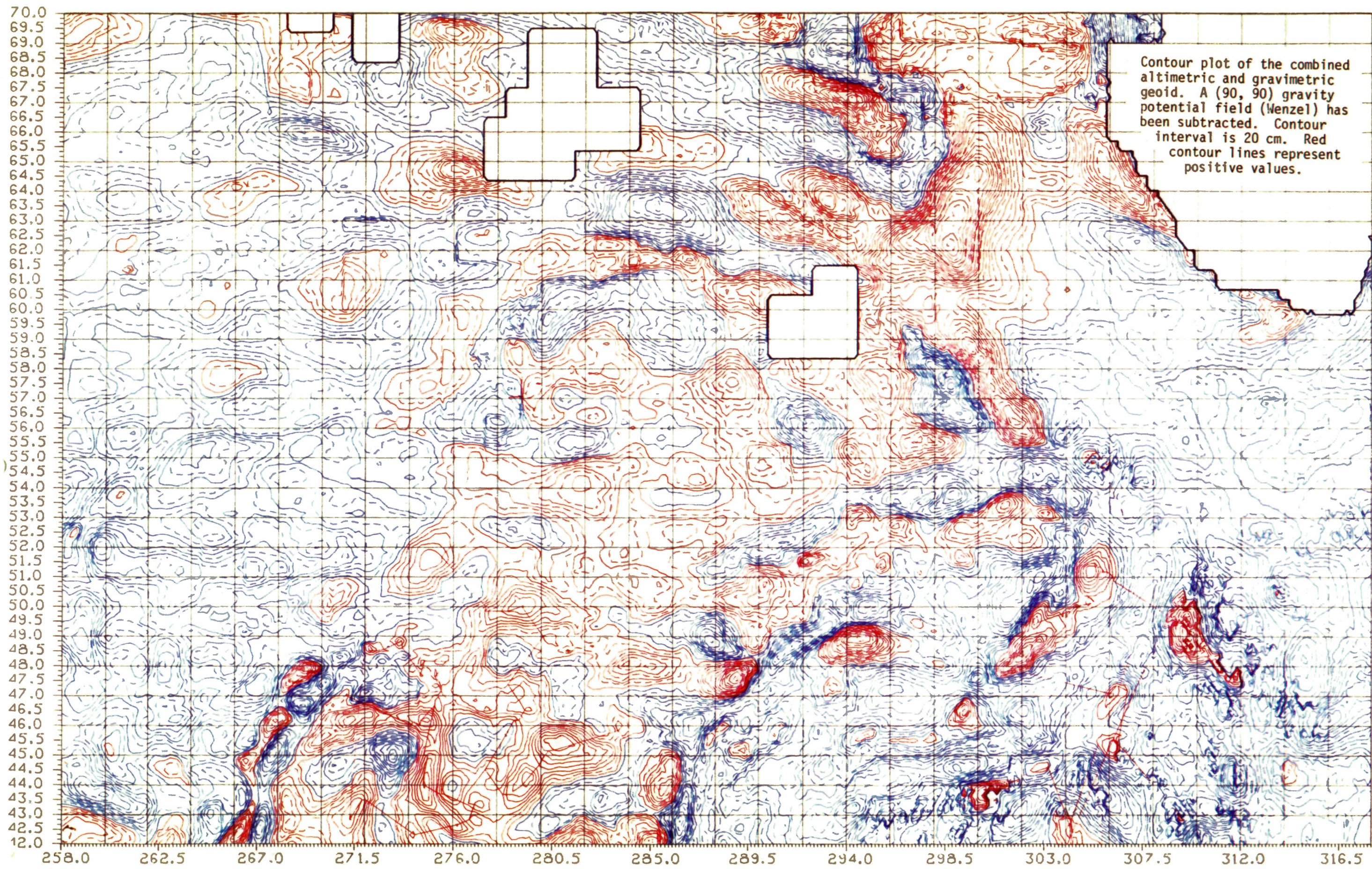3. Enter number of columns: (541)

4. Enter the number of channels: (1)
5. Enter a title for data: (data)
6. Enter the channel descriptor: (channel 1)

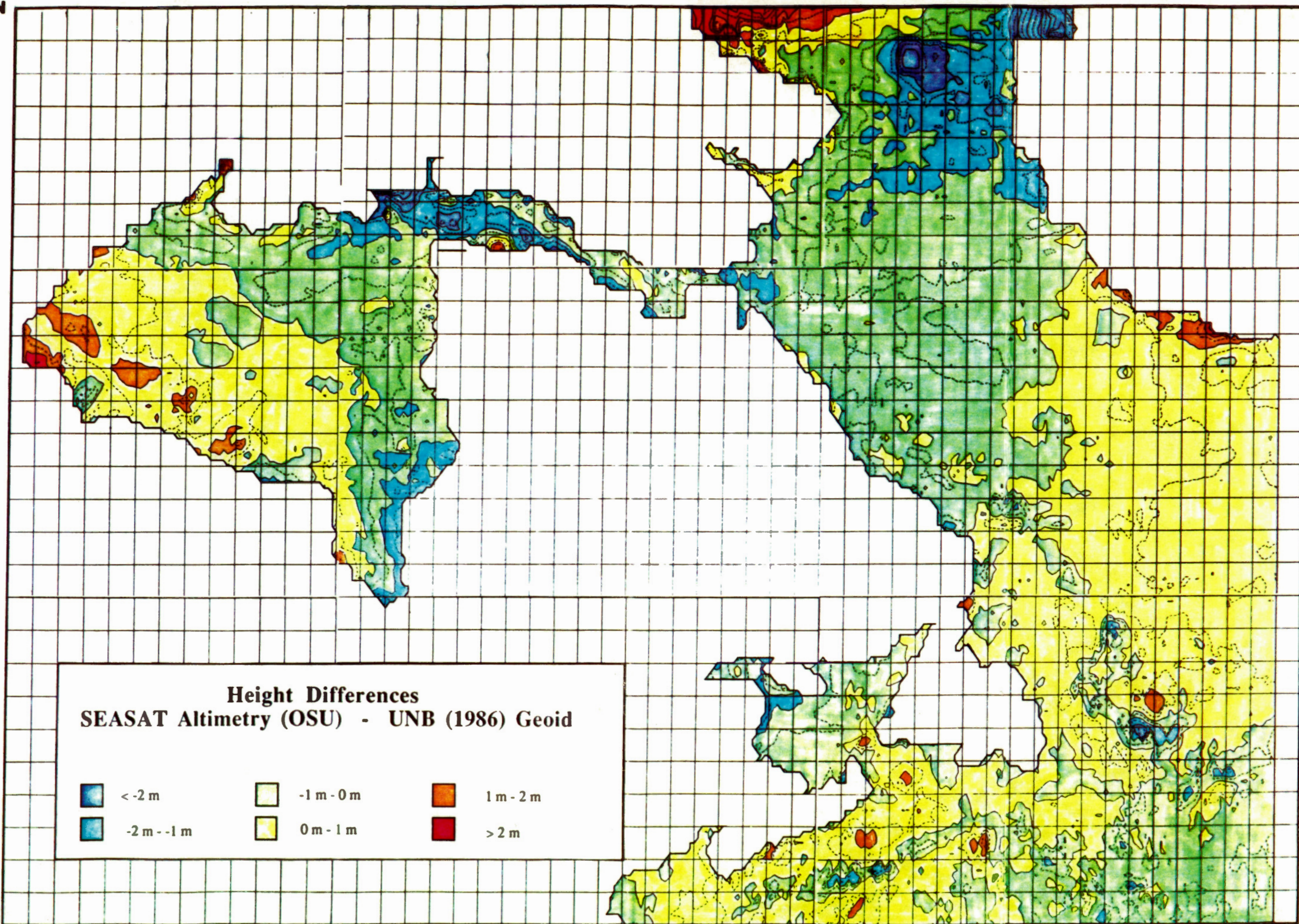**Geoid Height Differences**
RAPP (1983) – UNB (1986)

Legend:

- < -2 m
- -2 m – -1 m
- -1 m – 0 m
- 0 m – 1 m
- 1 m – 2 m
- > 2 m

Geoid Height Differences
WENZEL (1984) - UNB (1986)

Contour plot of the combined altimetric and gravimetric geoid. A (90, 90) gravity potential field (Wenzel) has been subtracted. Contour interval is 20 cm. Red contour lines represent positive values.

Height Differences
SEASAT Altimetry (OSU) — UNB (1986) Geoid

▣	< -2 m	▢ -1 m - 0 m	▢ 1 m - 2 m
▢	-2 m - -1 m	▢ 0 m - 1 m	▢ > 2 m

70° N

42° N

265° E

318° E