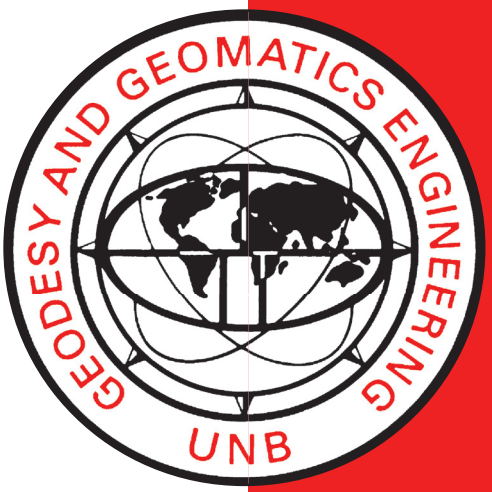
COMPUTATION OF A FILE OF GEOIDAL HEIGHTS USING MOLODENSKIJ'S TRUNCATION METHOD

P. VANICEK
C. ZHANG
P. ONG

September 1990



TECHNICAL REPORT
NO. 147

PREFACE

In order to make our extensive series of technical reports more readily available, we have scanned the old master copies and produced electronic versions in Portable Document Format. The quality of the images varies depending on the quality of the originals. The images have not been converted to searchable text.

COMPUTATION OF A FILE OF GEOIDAL HEIGHTS USING MOLODENSKIJ'S TRUNCATION METHOD

Petr Vaníček
Changyou Zhang
Peng Ong

Department of Surveying Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

September 1990
Latest Reprinting June 1993

PREFACE

This technical report is a reproduction of a final contract report submitted to the Geodetic Survey Division of Energy, Mines and Resources Canada, July 1990.

As with any copyrighted material, permission to reprint or quote extensively from this report must be received from the authors. The citation to this work should appear as follows:

Vaníček, P., C. Zhang, and P. Ong (1990). "Computation of a file of geoidal heights using Molodenskij's truncation method." Department of Surveying Engineering Technical Report No. 147, University of New Brunswick, Fredericton, New Brunswick, Canada, 106 pp.

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

In March 1986, we produced a (gravimetric) geoid for Canada under a DSS contract with the Geodetic Survey Division. The technique selected for the computation was the generalized Stokes approach treating satellite-derived potential coefficients as defining a reference field of higher order. The generalized Stokes integration kernel was then modified following the technique suggested by Molodenskij to minimize the radius of the truncated integration cap and thus minimize the computer time computation. This geoid was later improved and published and distributed under the name of the "UNB Dec. '86" geoid.

The present contract called for a recomputation of this geoid using the most up-to-date gravity anomaly values available in Canada. These values came in three guises: point values, 5' x 5' mean values, and 1° x 1° mean values. While the first set of values came directly from the federal gravity data bank, the second and third sets were prepared for this project especially by the Geodetic Survey Division under the supervision of Dr. André Mainville, the Scientific Authority for this contract.

The new geoid has been computed on a 10' x 10' grid covering the whole of Canada up to latitude 72°N. It uses the more recent GEM-T1 potential coefficients up to degree and order 20 replacing the older GEM-9 coefficients. The differences between the new and the UNB Dec. '86 solutions are at most of the order of 1 m and are predominantly of a longer wavelength character. Significant local differences, however, also can be seen; these are due to the gravity data improvement. The accuracy of the new solution is improved, compared to the old solution, mainly because of the better accuracy of the GEM-T1 coefficients.

2. INTRODUCTION

This report describes the work done under a sole source DSS Contract No. 23244-8-4564/01-SS let to the University of New Brunswick (UNB) on 9 March 1989. The first author of this report was the Principal Investigator, the second author was a graduate student at the UNB Department of Surveying Engineering until the end of August 1989, and the third author is still a graduate student at the same Department. All of the computations were done first by Changyou Zhang and, starting 1 September 1989, by Peng Ong under the guidance and supervision of the Principal Investigator.

The requirement of the contract was "to recompute a file of geoidal heights using the Molodenskij truncation method described in a previous contract report, *The Canadian Geoid*, 27 February 1986 by Vaníček et al." The geoidal heights had to be recomputed on a regular geographical grid, at a spacing of 10 arcmin covering Canada for a minimum area extending from latitude 42°N to 72°N and longitude 218°E to 318°E — a matrix of 181 by 601 values.

The geoid was to be computed with a new set of point free-air gravity anomaly values updated to July 1988 and two new sets of mean free-air gravity anomaly values on geographical 'rectangular' grids of 5 arcmin and 1 arcdegree. The Geodetic Survey Division was to provide the three sets in a computer-readable form. The reference spheroid was to be defined through the potential coefficients of the NASA Goddard Space Flight Center model GEM-T1.

The goals of the Contract have been achieved after some delays caused by problems in data acquisition and change in personnel. The computations were carried out by means of the original GIN program suite [Vaníček et al., 1987] modified to accept the new data sets. A new technique based on two-dimensional Fourier transform [Colombo, 1981; Wang and Rapp, 1989] was adopted for the computation of the topographic correction to geoidal heights. Also, a new approach was used for reducing the gravity anomalies to the reference spheroid. This approach is based on the generation of the GEM-T1 reduction values (using the Tscherning et al. [1983]

program POT) on a sparse regular grid and two-dimensional quadratic interpolation within the grid. It results in a considerable computer time saving.

The file of geoidal heights contains the required values of:

(a) total geoidal height referred to GRS 80;

(b) associated standard deviation;

(c) innermost-zone integration contribution;

(d) inner-zone integration contribution;

(e) outer-zone integration contribution;

(f) tropographic effect;

(g) indirect effect.

The technical details are described in the following seven chapters.

3. DATA SETS USED

3.1 Point Free-Air Gravity Anomalies

Two new files of point free-air gravity anomalies containing values available in July 1988 were prepared by the Geophysics Division of the Geological Survey of Canada and transmitted to us by the Contract Scientific Authority. The first file contains 266 065 values on land, the second 322 809 values at sea. These two files were merged creating one homogeneous file (of 588 874 records).

In our 1986 work, we used the file A.M1212072.EXMAIN.SORT for point gravity anomaly. Following the same practice this time and conforming to the input format of the GIN program, we created 20 sequential files. The program SORT.EXMAIN has been written to serve this purpose, and the following 20 sequential files were created:

1.	A.M1212072.REDUCED.TOPO.PGA1	Latitude:	40°	-	50°10'
		Longitude:	218°	-	238° 10'
2.	A.M1212072.REDUCED.TOPO.PGA2	Latitude:	40°	-	50°10'
		Longitude:	238°	-	258° 10'
3.	A.M1212072.REDUCED.TOPO.PGA3	Latitude:	40°	-	50°10'
		Longitude:	258°	-	278° 10'
4.	A.M1212072.REDUCED.TOPO.PGA4	Latitude:	40°	-	50°10'
		Longitude:	278°	-	298° 10'
5.	A.M1212072.REDUCED.TOPO.PGA5	Latitude:	40°	-	50°10'
		Longitude:	298°	-	320° 10'
6.	A.M1212072.REDUCED.TOPO.PGA6	Latitude:	50°	-	60°10'
		Longitude:	218°	-	238° 10'
7.	A.M1212072.REDUCED.TOPO.PGA7	Latitude:	50°	-	60°10'
		Longitude:	238°	-	258° 10'
8.	A.M1212072.REDUCED.TOPO.PGA8	Latitude:	50°	-	60°10'
		Longitude:	258°	-	278° 10'
9.	A.M1212072.REDUCED.TOPO.PGA9	Latitude:	50°	-	60°10'
		Longitude:	278°	-	298° 10'

10.	A.M1212072.REDUCED.TOPO.PGA10	Latitude:	50°	-	60°10'
		Longitude:	298°	-	320° 10'
11.	A.M1212072.REDUCED.TOPO.PGA11	Latitude:	60°	-	70°10'
		Longitude:	218°	-	238° 10'
12.	A.M1212072.REDUCED.TOPO.PGA12	Latitude:	60°	-	70°10'
		Longitude:	238°	-	258° 10'
13.	A.M1212072.REDUCED.TOPO.PGA13	Latitude:	60°	-	70°10'
		Longitude:	258°	-	278° 10'
14.	A.M1212072.REDUCED.TOPO.PGA14	Latitude:	60°	-	70°10'
		Longitude:	278°	-	298° 10'
15.	A.M1212072.REDUCED.TOPO.PGA15	Latitude:	60°	-	70°10'
		Longitude:	298°	-	320° 10'
16.	A.M1212072.REDUCED.TOPO.PGA16	Latitude:	70°	-	80°10'
		Longitude:	218°	-	238° 10'
17.	A.M1212072.REDUCED.TOPO.PGA17	Latitude:	70°	-	80°10'
		Longitude:	238°	-	258° 10'
18.	A.M1212072.REDUCED.TOPO.PGA18	Latitude:	70°	-	80°10'
		Longitude:	258°	-	278° 10'
19.	A.M1212072.REDUCED.TOPO.PGA19	Latitude:	70°	-	80°10'
		Longitude:	278°	-	298° 10'
20.	A.M1212072.REDUCED.TOPO.PGA20	Latitude:	70°	-	80°10'
		Longitude:	298°	-	320° 10'

The point gravity anomalies were all corrected by us for the atmospheric attraction effect [Vaníček et al., 1987]. They were also transformed to refer to the GRS 80, instead of the GRS 67 in which they were given, by means of the standard reduction

$$\Delta g^{1980} = \Delta g^{1967} - (0.8316 + 0.0782 \sin^2\phi - 0.0007 \sin^4\phi) \text{ mGal.}$$

Finally, the anomalies were reduced to the reference spheroid defined by the first 20, 20 GEM-T1 coefficients using the procedure described in Chapter 5.

3.2 5' x 5' Mean Free-Air Gravity Anomalies

This mean anomalies file was specifically created for this work by the Geodetic Survey Division under the supervision of the Contract Scientific Authority (Mainville and Véronneau

[1989], Appendix A), and delivered to us on magnetic tape. The structure, format, and technique used for creating the mean anomalies are described in Appendix B. These mean anomalies were already corrected for the atmospheric attraction effect and referred to GRS 80.

For all 5' x 5' mean gravity anomalies, one direct access file is created. This file is divided into 200, 8° by 8° blocks. Each block overlaps with the eastern and northern adjacent blocks by four degrees so that almost three times the necessary amount of data is stored to minimize the time needed for data manipulation. The following diagram shows the way it is done.

J \ I	1	2	I	7	8
1	40°-48° 214°-222°	44°-52° 214°-222°	* 214°-222°	64°-72° 214°-222°	68°-76° 214°-222°
2	40°-48° 218°-226°	44°-52° 218°-226°	* 218°-226°	64°-72° 218°-226°	68°-76° 218°-226°
3	40°-48° 222°-230°	44°-52° 222°-230°	* 222°-230°	64°-72° 222°-230°	68°-76° 222°-230°
J	40°-48° *	44°-52° *	* *	64°-72° *	68°-76° *
24	40°-48° 306°-314°	44°-52° 306°-314°	* 306°-314°	64°-72° 306°-314°	68°-76° 306°-314°
25	40°-48° 310°-318°	44°-52° 310°-318°	* 310°-318°	64°-72° 310°-318°	68°-76° 310°-318°

In each block, we store data in the order of increasing latitude and increasing longitude for points of equal latitude. Let M be the total record length in all the blocks preceding the I,J-block and let Min and Max denote the boundaries of this block in both latitude and longitude. The arrangement of records within the I,J-block is shown in the following diagram.

Latitude	Longitude: Min	Min + 4°	Min + 8° = Max
Min			
Min + 2.5'	M + 1		M + 2
Min + 7.5'	M + 3		M + 4
Min + 12.5'	M + 5		M + 6
:	:		
Min + 7°57.5' Max	M + 191		M + 192

Here, each record contains 48 anomalies. Since there are 192 records in the block, the total number of anomalies in the block is 9216.

The parameters I, J, M are computed from the following formulae:

$$I = \text{INT}\left(\frac{PHI - 44}{4} + 1.5\right),$$

$$J = \text{INT}\left(\frac{DLAM - 218}{4} + 1.5\right),$$

and

$$M = \frac{9216 * (I - 1) * 25 + 9216 * (J - 1)}{48}.$$

The block boundaries are evaluated as:

Latitude: Min = 40 + (I - 1) * 4

Max = Min + 8

Longitude: Min = 214 + (I - 1) * 4

Max = Min + 8.

As in 1986, empty cells in the 5' x 5' file were filled with values taken from the 1° x 1° file where these values were available. Where neither 5' x 5' nor 1° x 1° values existed, the geoid was not computed resulting in holes in the geoid map — see section 7.2.

3.3 1° x 1° Mean Free-Air Gravity Anomalies

This mean anomalies file was again specifically created for this work by the Geodetic Survey Division under the direction of the Contract Scientific Authority. Its description is contained in Appendix B.

For the outer zone Stokes's integration in our technique, 1° x 1° anomalies up to the spherical distance of 6° from the computation point are needed. The new file does not contain values in some peripheral areas in the North Atlantic needed for the proper integration. Thus the geoid in this region does not stretch all the way to the originally envisaged boundary. Empty 1° x 1° cells are treated as having a mean gravity anomaly equal to 0 mGal and a standard deviation of 50 mGal, the same way they were treated in 1986.

3.4 GEM-T1 Potential Coefficients

This set of potential coefficients up to degree and order 36 was produced by the NASA Goddard Space Flight Center research team from the analysis of orbits of 21 satellites [Marsh et al., 1988]. It was supposed to be the most accurate set of coefficients determined from pure satellite data at the beginning of 1989, and the decision to use this set in our work was based on this understanding. The accuracy of the GEM-T1 coefficients is estimated to be almost twice as high as that of the GEM-9 coefficients used in our 1986 work [Marsh et al., 1988].

There is one perceived problem with the GEM-T1 coefficients, however, that we had to deal with: the solution relies rather heavily on the use of Kaula's rule of thumb. The use of this rule allowed the GSFC team to estimate the coefficients all the way to (36, 36) while the physics of the orbital analysis suggests that perhaps (20, 20) should be the upper limit for the degree and order of a pure satellite-determined field. Indeed, the noise level of coefficients above this limit shows a rather steep increase for degrees and orders above 20.

This reason, together with the practical reason of keeping the degree and order of the reference spheroid the same as in our 1986 computation, led us to the choice of truncating the (36, 36)

GEM-T1 field to (20, 20). The truncation is permissible because of the global orthogonality of spherical harmonics.

The values of the GEM-T1 potential coefficients were obtained from the Department of Surveying Engineering computer library. They were made to refer to GRS 80 using the approach described by Vaníček et al. [1987].

The standard deviation of the GEM-T1 reference spheroid — one value for the whole of Canada — was obtained in the same way as the GEM-9 value was obtained in 1986 [Vaníček et al., 1988]. Denoting by σ_2^n the "degree error variances" of the GEM-T1 coefficients, obtained from March et al. [1988], we get the standard deviation σ_{N20} as:

$$\sigma_{N20} = R \sqrt{\sum_{n=2}^{20} (2n + 1)\, \sigma_n^2} \ ,$$

where R is the mean earth radius. Its value is

σ_{N20} = 85 cm.

This approach can be refined in the future by evaluating a position dependent standard deviation from a more general formula (cf. Vaníček et al. [1987, eqn. (2.32)]).

4. THE MODIFIED GIN PROGRAM SUITE

The original GIN (**Geoid IN**tegration) program suite was written in 1985 for the computation of the Canadian geoid undertaken under a DSS contract as described in Vaníček et al. [1987]. It was subsequently improved in 1986 to produce the "UNB Dec. '86" geoid used now extensively in Canada. The approach used in the GIN is that of 'generalized Stokes integration' for higher-order reference spheroid using Molodenskij's truncation method for modification of Stokes's integration kernel. The theory is described in detail by Vaníček et al. [1987] and Sjöberg and Vaníček [1990]; see also Appendix D.

By selecting the degree and order of the reference spheroid (defined by the first (20, 20) GEM-T1 coefficients) to be the same as those used in 1986 (defined by the GEM-9 coefficients) we have eliminated the necessity to recompute the modified Stokes integration kernel. Consequently, the integration routines of the GIN program did not require any changes. The only change required came from the difference in the structure of the old and new 5' x 5' mean anomaly files. While the old file used 5' x 5' grid for latitudes between 35°N and 50°N; 5' x 10' for latitudes between 52°N and 70°N; and 5' x 15' for latitudes between 70°N and 75°N; the new file is homogeneous, 5' x 5', throughout.

The only other modification to the GIN suite was the elimination of the routine for topographic correction to gravity anomalies. Using the two-dimensional Fourier transform approach — see Chapter 6 — topographic correction to geoidal heights are now computed directly, by-passing the necessity to correct gravity anomalies first, and resulting in computer time saving.

It should be reiterated here that the GIN program suite integrates over point gravity anomalies in the 'innermost zone' of 10' x 10', over 5' x 5' mean anomalies in the 'inner zone' of 2° x 2°, and over 1° x 1° mean anomalies in the 'outer zone' which extends to a spherical distance of 6° from the computation point. When there are not enough point values in the innermost zone, 5' x 5' mean anomalies are used instead. When there are 5' x 5' mean anomalies missing for the inner

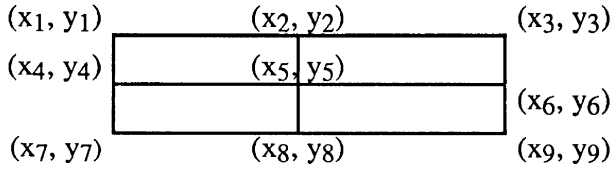
zone integration, they are replaced by 1° x 1° mean values. When even 1° x 1° mean values are missing, the integration is aborted.

We also note that the GIN suite produces an estimate of the standard deviation for the Stokes contribution (the result of generalized Stokes's integration). The mathematical derivation can be found in Vaníček et al. [1987]. This standard deviation is added quadratically to the standard deviation of the GEM-T1 contribution discussed in section 3.4 to give the standard deviation of the final geoidal height. This standard deviation is a part of the output file. The program suite is listed in external Appendix E.

5. REDUCTION OF GRAVITY ANOMALIES TO REFERENCE SPHEROID

Our technique for geoidal height computation requires that all the gravity anomalies be first reduced to the reference spheroid [Vaníček et al., 1987]. This is achieved by subtracting from the given anomaly values the 'reference field' values generated from the first (20, 20) GEM-T1 coefficients.

The generation of the reference field values for the approximately 600 000 irregularly spaced point anomalies is a time-consuming task, because each value is computed from a (truncated) series of spherical harmonics. To reduce the computational effort needed for the reduction of point gravity anomalies, we decided to use the following interpolation technique. The reference field values are first computed on the 5' x 5' grid (same as the one used for mean anomalies) using the POT program [Tscherning et al., 1983], and these values are used to correct the 5' x 5' mean anomalies. Then a quadratic surface is fitted to 9 adjacent grid values as shown in the diagram.



where x, y is the local Cartesian system defined as

$$x = R(\varphi - \varphi_5)\ ,\ y = R(\lambda - \lambda_5)\cos\phi_5\ .$$

This surface can be described as follows:

$$a_1 x^2 y^2 + a_2 x^2 y + a_3 x y^2 + a_4 x^2 + a_5 y^2 + a_6 x y + a_7 x + a_8 y + a_9 = \Delta g$$

where $a_1, a_2, \ldots, a_9$ are unknown parameters.

We thus obtain nine equations as follows:

$$a_1 x_1^2 y_1^2 + a_2 x_1^2 y_1 + a_3 x_1 y_1^2 + a_4 x_1^2 + a_5 y_1^2 + a_6 x_1 y_1 + a_7 x_1 + a_8 y_1 + a_9 = \Delta g_1 = l_1$$

$$a_1 x_2^2 y_2^2 + a_2 x_2^2 y_2 + a_3 x_2 y_2^2 + a_4 x_2^2 + a_5 y_2^2 + a_6 x_2 y_2 + a_7 x_2 + a_8 y_2 + a_9 = \Delta g_2 = l_2$$

:

$$a_1 x_9^2 y_9^2 + a_2 x_9^2 y_9 + a_3 x_9 y_9^2 + a_4 x_9^2 + a_5 y_9^2 + a_6 x_9 y_9 + a_7 x_9 + a_8 y_9 + a_9 = \Delta g_9 = l_9 .$$

Defining $\underline{A}$, $\underline{L}$, $\underline{X}$ as follows:

$$\underline{A} = \begin{vmatrix} x_1^2y_1^2 & x_1^2y_1 & x_1y_1^2 & x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 & 1 \\ x_2^2y_2^2 & x_2^2y_2 & x_2y_2^2 & x_2^2 & y_2^2 & x_2y_2 & x_2 & y_2 & 1 \\ x_3^2y_3^2 & x_3^2y_3 & x_3y_3^2 & x_3^2 & y_3^2 & x_3y_3 & x_3 & y_3 & 1 \\ x_4^2y_4^2 & x_4^2y_4 & x_4y_4^2 & x_4^2 & y_4^2 & x_4y_4 & x_4 & y_4 & 1 \\ x_5^2y_5^2 & x_5^2y_5 & x_5y_5^2 & x_5^2 & y_5^2 & x_5y_5 & x_5 & y_5 & 1 \\ x_6^2y_6^2 & x_6^2y_6 & x_6y_6^2 & x_6^2 & y_6^2 & x_6y_6 & x_6 & y_6 & 1 \\ x_7^2y_7^2 & x_7^2y_7 & x_7y_7^2 & x_7^2 & y_7^2 & x_7y_7 & x_7 & y_7 & 1 \\ x_8^2y_8^2 & x_8^2y_8 & x_8y_8^2 & x_8^2 & y_8^2 & x_8y_8 & x_8 & y_8 & 1 \\ x_9^2y_9^2 & x_9^2y_9 & x_9y_9^2 & x_9^2 & y_9^2 & x_9y_9 & x_9 & y_9 & 1 \end{vmatrix}$$

$$\underline{L}^T = (l_1 \;\; l_2 \;\; l_3 \;\; l_4 \;\; l_5 \;\; l_6 \;\; l_7 \;\; l_8 \;\; l_9),$$

$$\underline{X}^T = (a_1 \;\; a_2 \;\; a_3 \;\; a_4 \;\; a_5 \;\; a_6 \;\; a_7 \;\; a_8 \;\; a_9),$$

we get

$$\underline{A}\underline{X} = \underline{L}.$$

The solution is obtained as

$$\underline{X} = \underline{A}^{-1}\,\underline{L}.$$

We note that the $\underline{A}^{-1}$ matrix remains the same for all cell foursomes and can thus be precomputed. Only the L-vector changes from location to location. The interpolation is thus very fast. We also note that the GEM-T1 (20,20) field is sufficiently smooth so that the second-degree surface used here approximates it to better than 1 μGal in any 10' x 10' area. We have confirmed this through numerical tests.

The 1° x 1° mean gravity values have been reduced directly, using values generated by the POT program.

The program for the interpolation is listed in external Appendix F.

6. THE TOPOGRAPHIC CORRECTION

In our 1986 computation, we used our own technique [Vaníček et al., 1987; Vaníček and Kleusberg, 1987] to compute the topographic correction to geoidal heights. This was done by correcting the gravity anomalies first and then integrating these corrected anomalies. This roundabout way can be short-circuited by using the two-dimensional Fourier transform as described by Colombo [1981].

Wang and Rapp [1989] derived the formula for the topographic correction to geoidal height δN_T as

$$\delta N_T = \frac{1}{2\pi\gamma} F^{-1}(F(\delta g_T)\, F(d^{-1}))\, \Delta x \Delta y,$$

where

$$\delta g_T = \frac{1}{2} G\sigma[F^{-1}(F(d_0^{-3})\, F(h^2)) - h_p^2\, F^{-1}(F(d_0^{-3})\, F(1))]\, \Delta x \Delta y\ .$$

Here, γ stands for average value of gravity, σ for average lithospheric density, G for the gravitational constant, h for heights supplied on a grid with steps equal to Δx and Δy, h_p is the height of the computing point, F and F^{-1} denote the discrete Fourier transform and its inverse, and d_0 is defined as

$$d_0 \begin{cases} \to \infty & \text{for } d = 0 \\ = d & \text{for } d \neq 0 \end{cases},$$

where d is the distance of the grid point from the computation point. The derivation of the above equations is given in Wang and Rapp [1989] which we reproduce in Appendix C for convenience. The program for the topographic correction is listed in external Appendix G.

We have acquired Wang's program to replace the original 'topographic correction' routine in the GIN suite and used it to generate the geoidal height corrections directly for the whole of Canada. To compute this correction, mean topographic heights on a 5' x 5' grid have been used.

These topographic data have been supplied to us on magnetic tape by the Contract Scientific Authority. We note that for grid points at sea, the topographic height h equals to zero.

It was during the period of this contract that Wang and Rapp pointed out the difference that exists between our approach to topographic correction and the original Helmert approach, restated later by Heiskanen and Moritz [1967]. Wang and Rapp claim that our approach is inappropriate — see Appendix C. Up until now, we have not been able to convince ourselves that they are right. Our investigations are continuing. Should Wang and Rapp be right, it will be a simple matter to correct the herewith presented geoid at a later date. It should be noted that this correction is really important only in the mountains.

7. THE RESULTS

7.1 The Geoidal Height File

We have produced the file of geoidal heights on a geographical grid of 10' x 10' covering the area 42° to 72°N by 218° to 318° E. These are referred to the GRS 80 reference ellipsoid [Moritz, 1980] as were the geoidal heights computed in 1986. In addition to these geoidal heights, we have estimated the standard deviations of these, using the capability of the GIN program. The pertinent formulae are given in Vaníček et al. [1987]. The new standard deviations are significantly smaller mainly because of the better accuracy of the GEM-T1 coefficients. This improvement has little effect on the accuracy of geoidal height differences which is more critical for positioning applications. Any improvement in the accuracy of the higher-frequency components is due to the more accurate and denser gravity anomaly values now available.

To facilitate the comparison with the "UNB Dec. '86" solution, we have also produced the individual contributions from the innermost, inner, and outer zone integrations. The sum of these three values, plus the topographic correction and the correction for the indirect effect (which are also filed for all the grid points), yields the total Stokes contribution, or the geoidal height above the reference spheroid. This high frequency part of the total geoidal signal can be used separately for studying the density distribution within the upper strata of the earth [Christou et al., 1989].

7.2 The Plots

The solution was sought in rectangles 10° x 20°, for the reasons of computation management. The following plots reflect this strategy, while the numerical file described above does not. The file was created by merging all the rectangles.

For final display, the whole numerical file should be plotted in the appropriate cartographic projection to allow an overlay with a topographical or other map, following a procedure similar to the one so successfully used for the UNB Dec. '86 geoid.

7.3 Testing

The results have been tested visually against the UNB Dec. '86 results. The differences seem to conform in broad features with the shape shown in Figure 16. This figure displays the expected difference for the UNB Dec. '86 geoid when the GEM-9 reference spheroid is replaced by the GEM-T1 reference spheroid. There are, however, sizeable differences of many metres between the two solutions. The areas of large differences coincide with the areas of large differences between the old and new gravity data sets shown in Figure 5 of Appendix A.

More thorough testing, including the plots of differences between this and other solutions, and comparisons with GPS/levelling results, is needed. These tests, however, were beyond the scope of this contract.

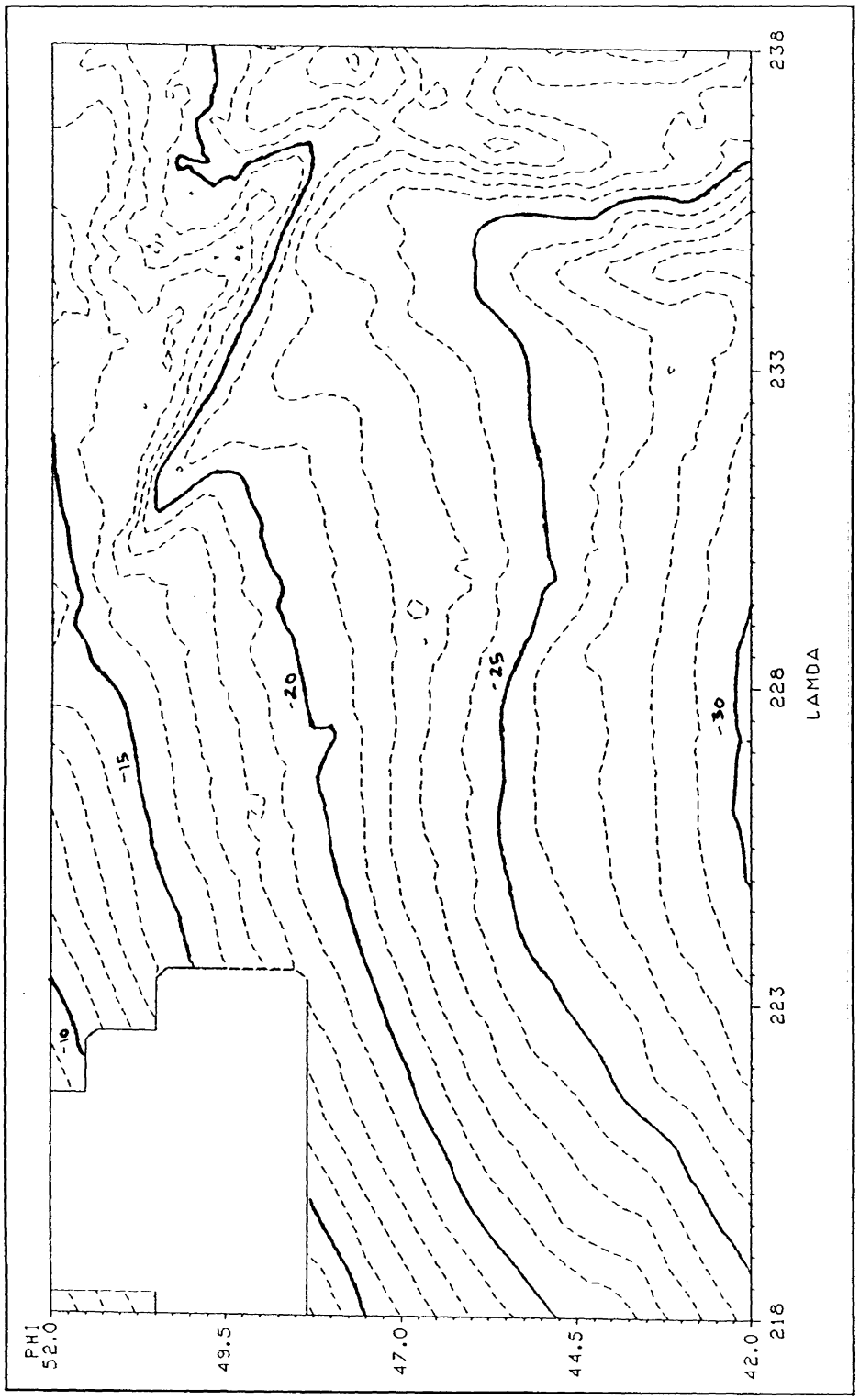


Figure 1
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 1
Region: Lat. = 42°, 52°; Long. = 218°, 238°
Range: - 30.0 m to -4.0 m; Contour interval: 1.0 m

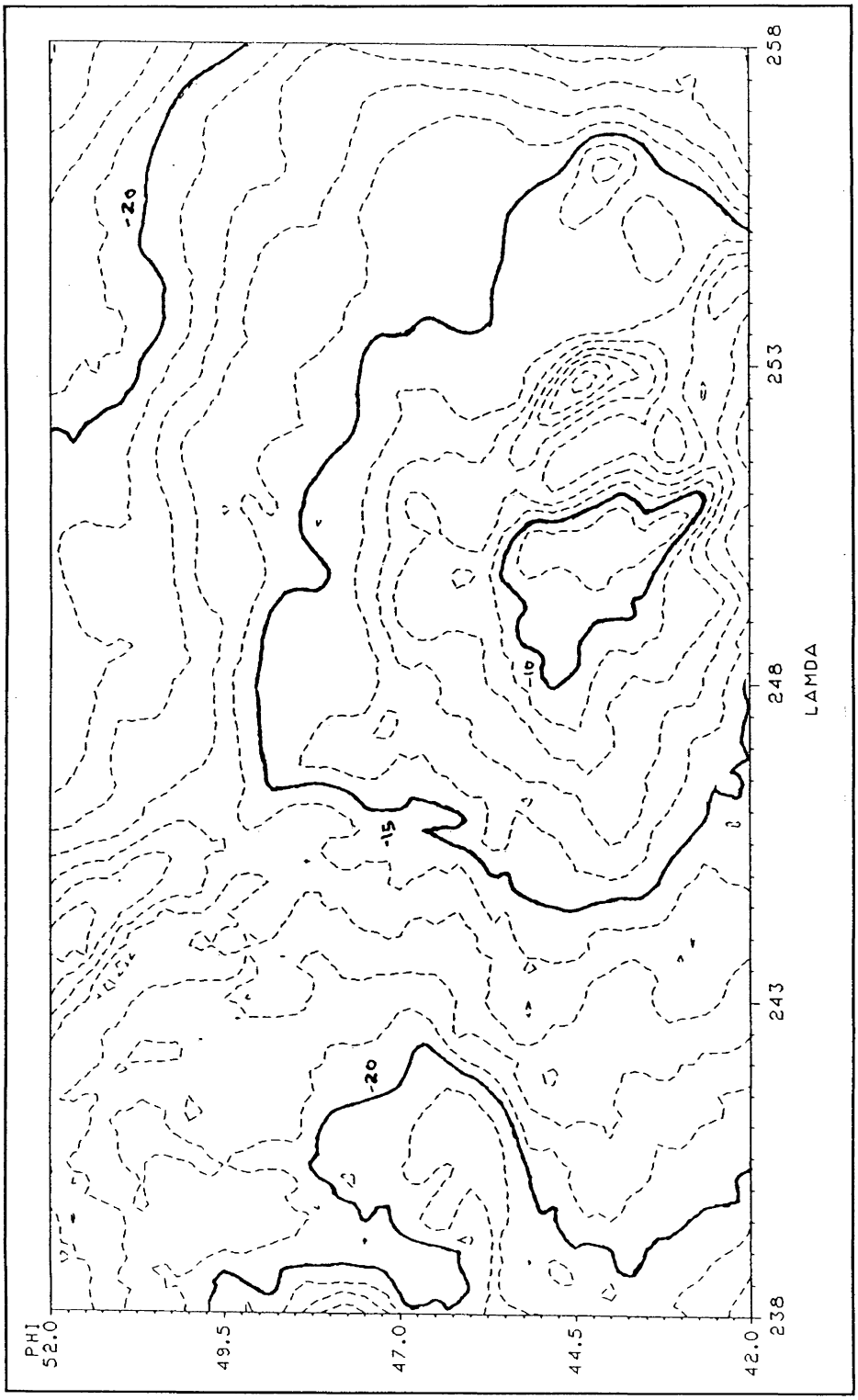


Figure 2
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 2
Region: Lat. = 42°, 52°; Long. = 238°, 258°
Range: - 24.0 m to -9.0 m; Contour interval: 1.0 m

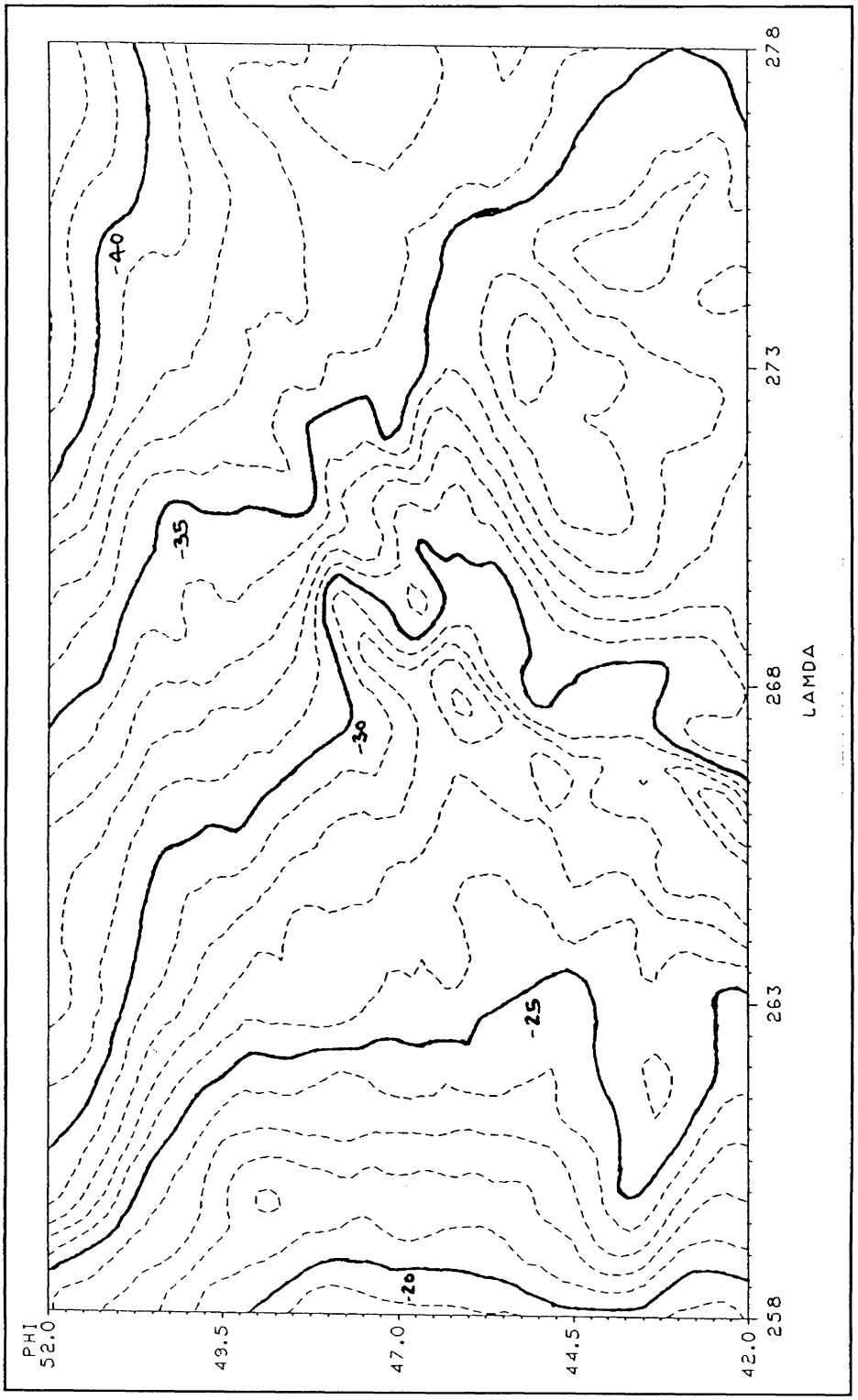


Figure 3
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 3
Region: Lat. = 42°, 52°; Long. = 258°, 278°
Range: - 42.0 m to -19.0 m; Contour interval: 1.0 m

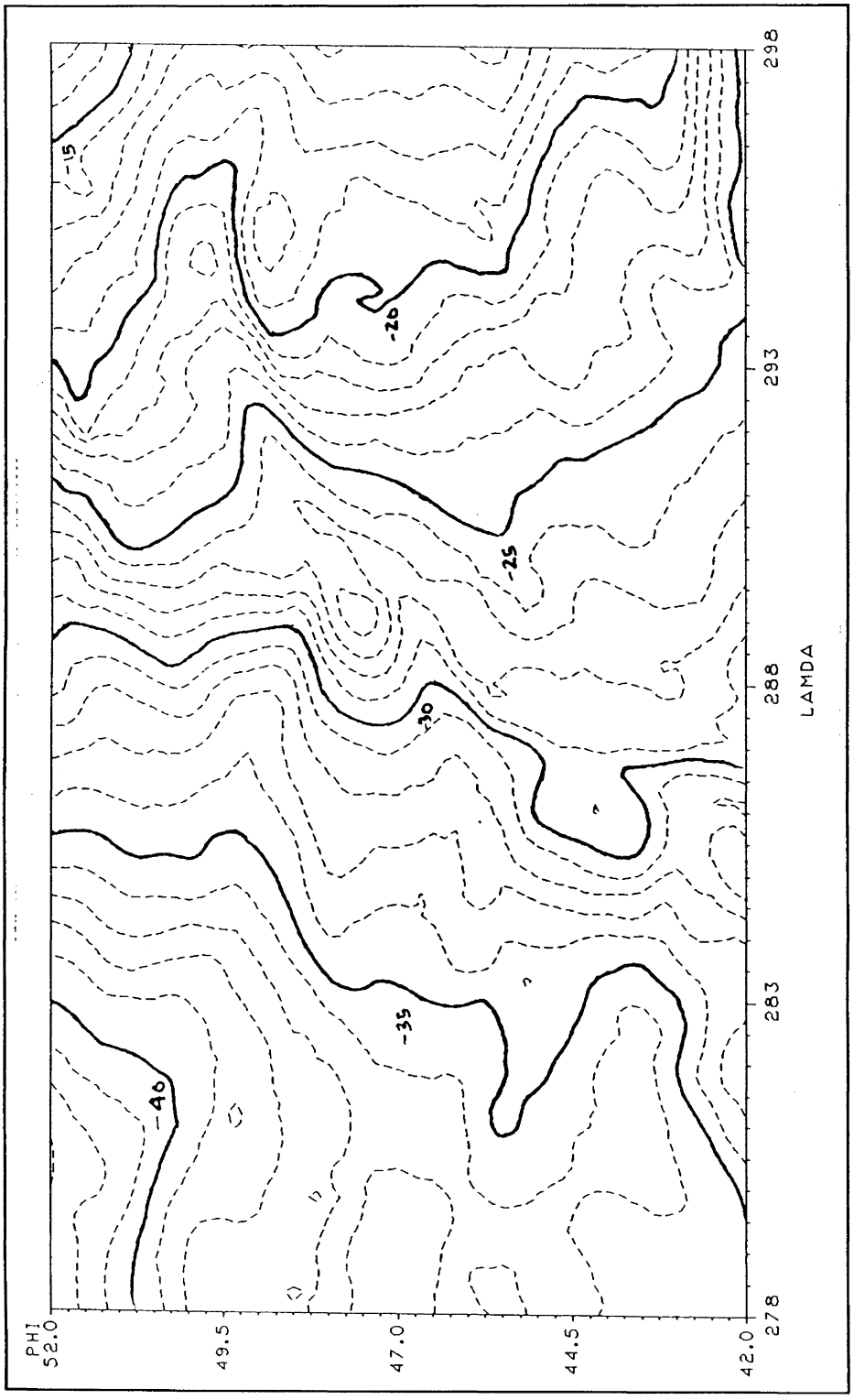


Figure 4
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 4
Region: Lat. = 42°, 52°; Long. = 278°, 298°
Range: - 43.0 m to -13.0 m; Contour interval: 1.0 m

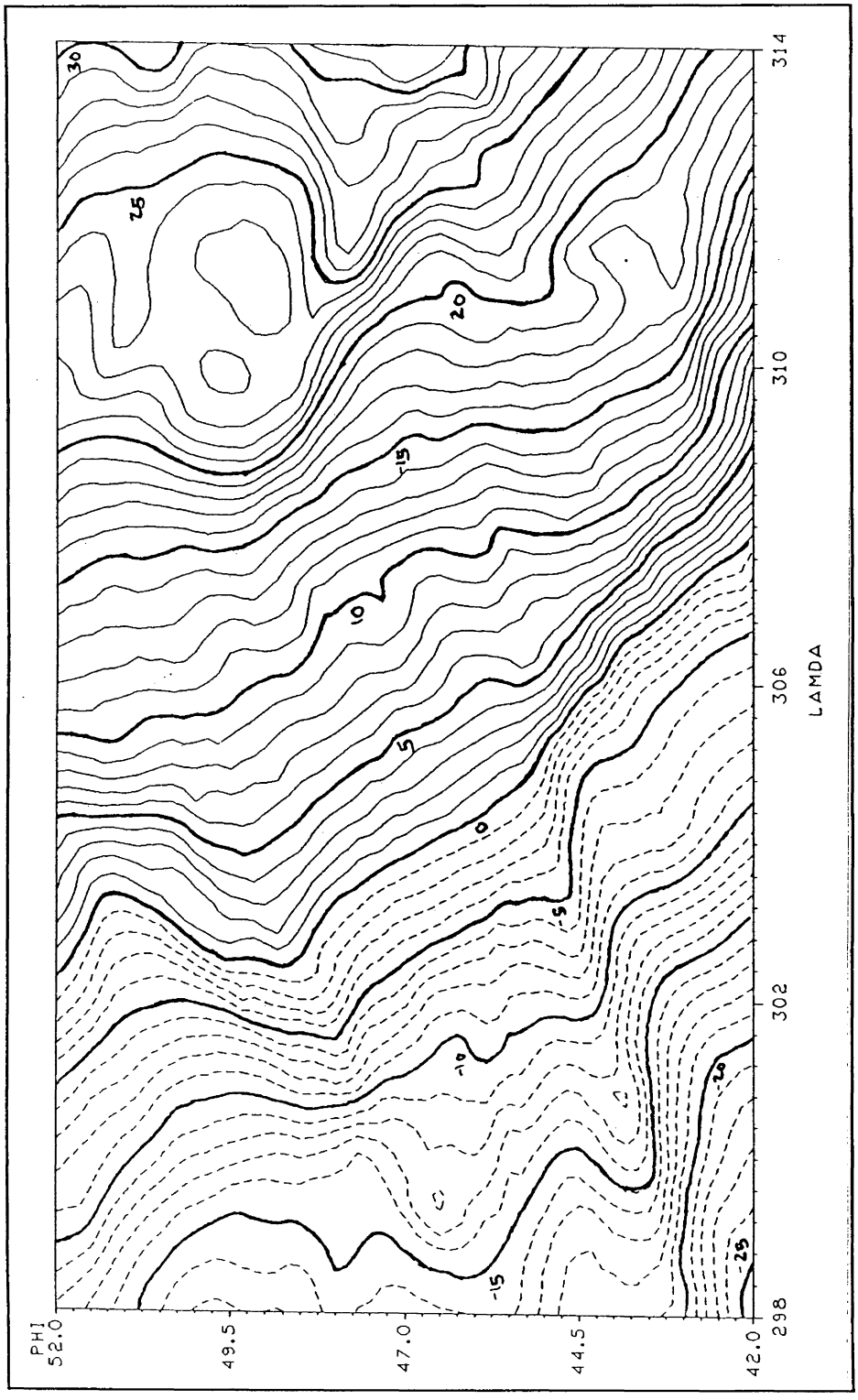


Figure 5
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 5
Region: Lat. = 42°, 52°; Long. = 298°, 314°
Range: - 26.0 m to 32.0 m; Contour interval: 1.0 m

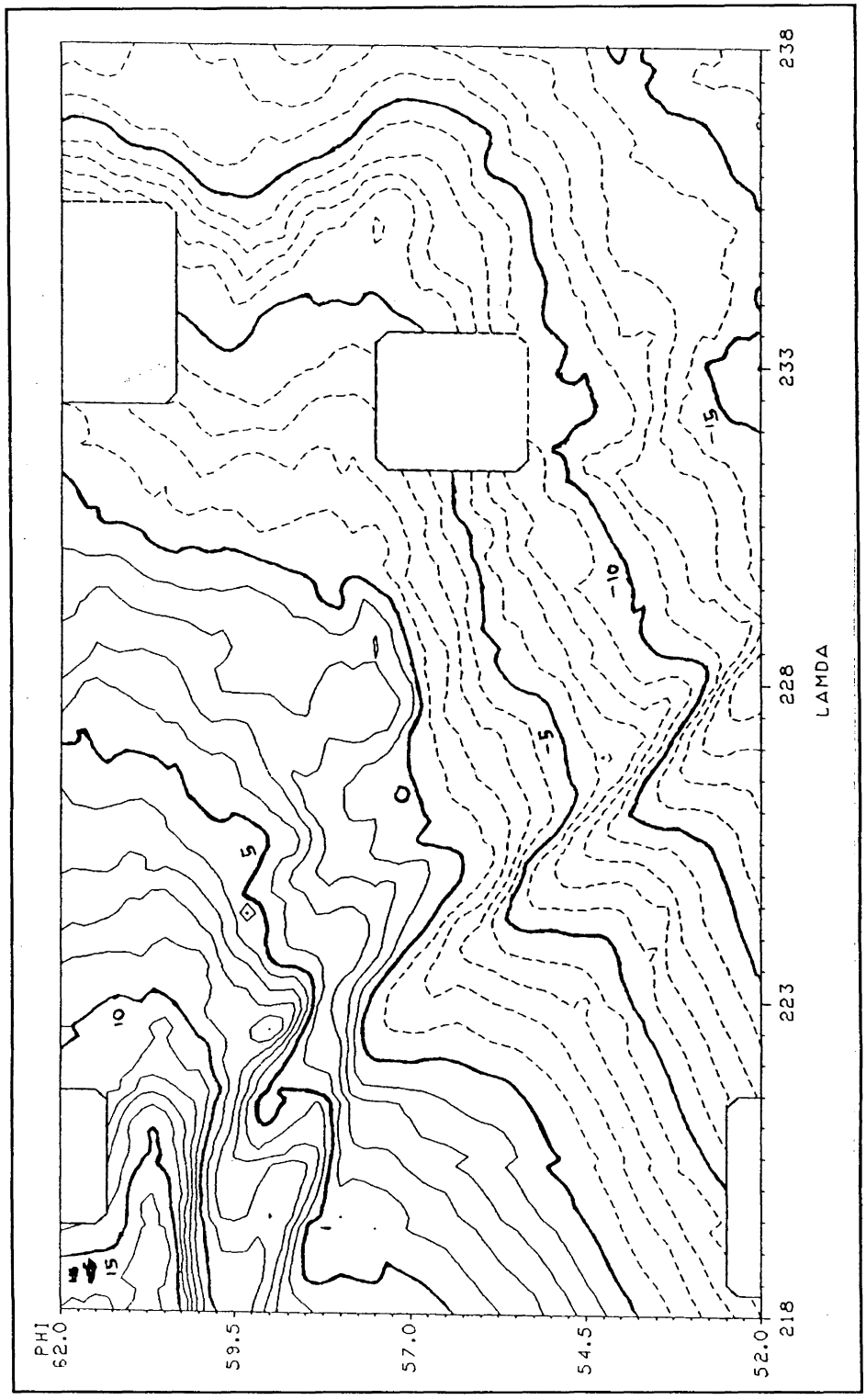


Figure 6
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 6
Region: Lat. = 52°, 62°; Long. = 218°, 238°
Range: - 16.0 m to 16.0 m; Contour interval: 1.0 m

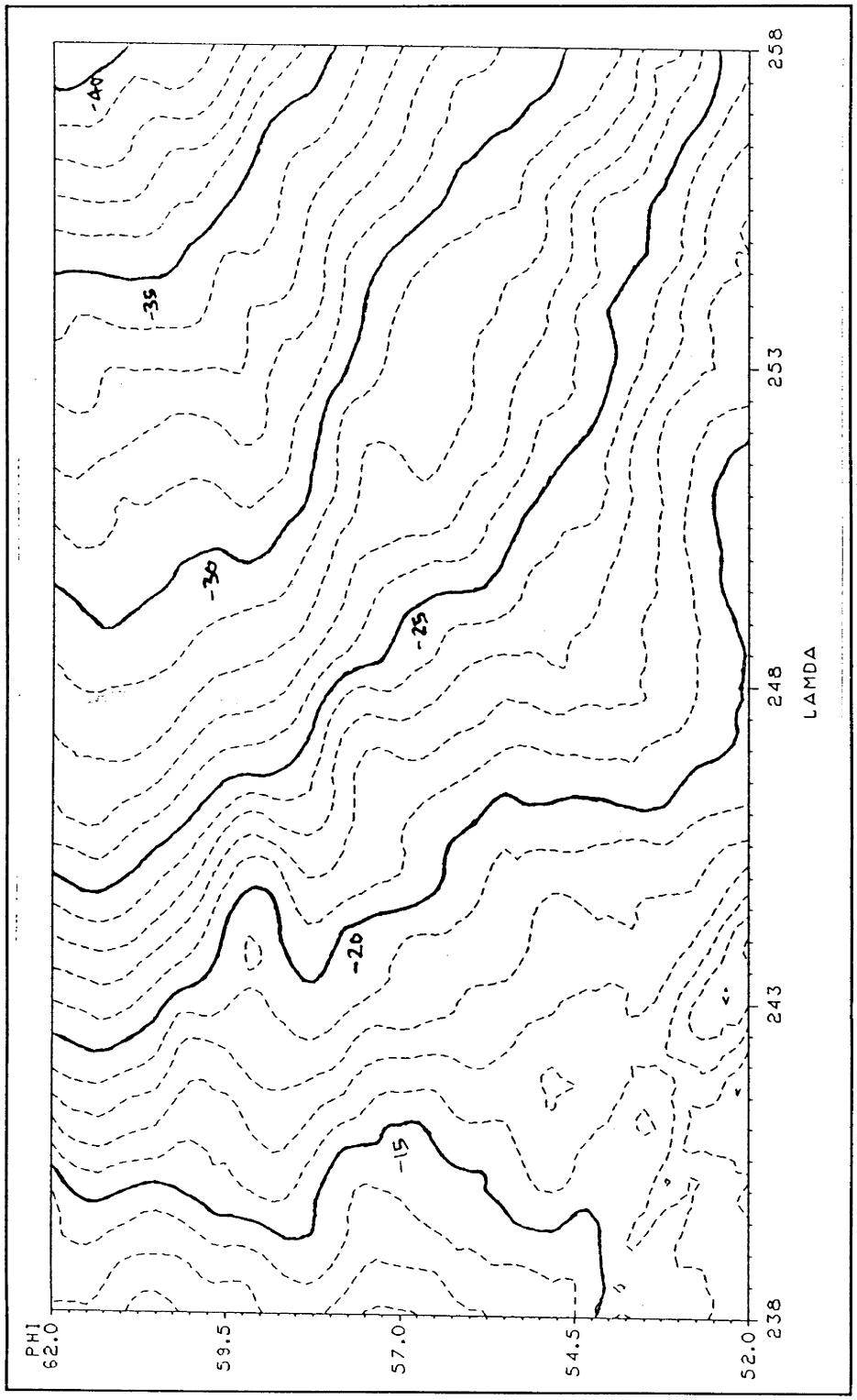


Figure 7
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 7
Region: Lat. = 52°, 62°; Long. = 238°, 258°
Range: - 40.0 m to -12.0 m; Contour interval: 1.0 m

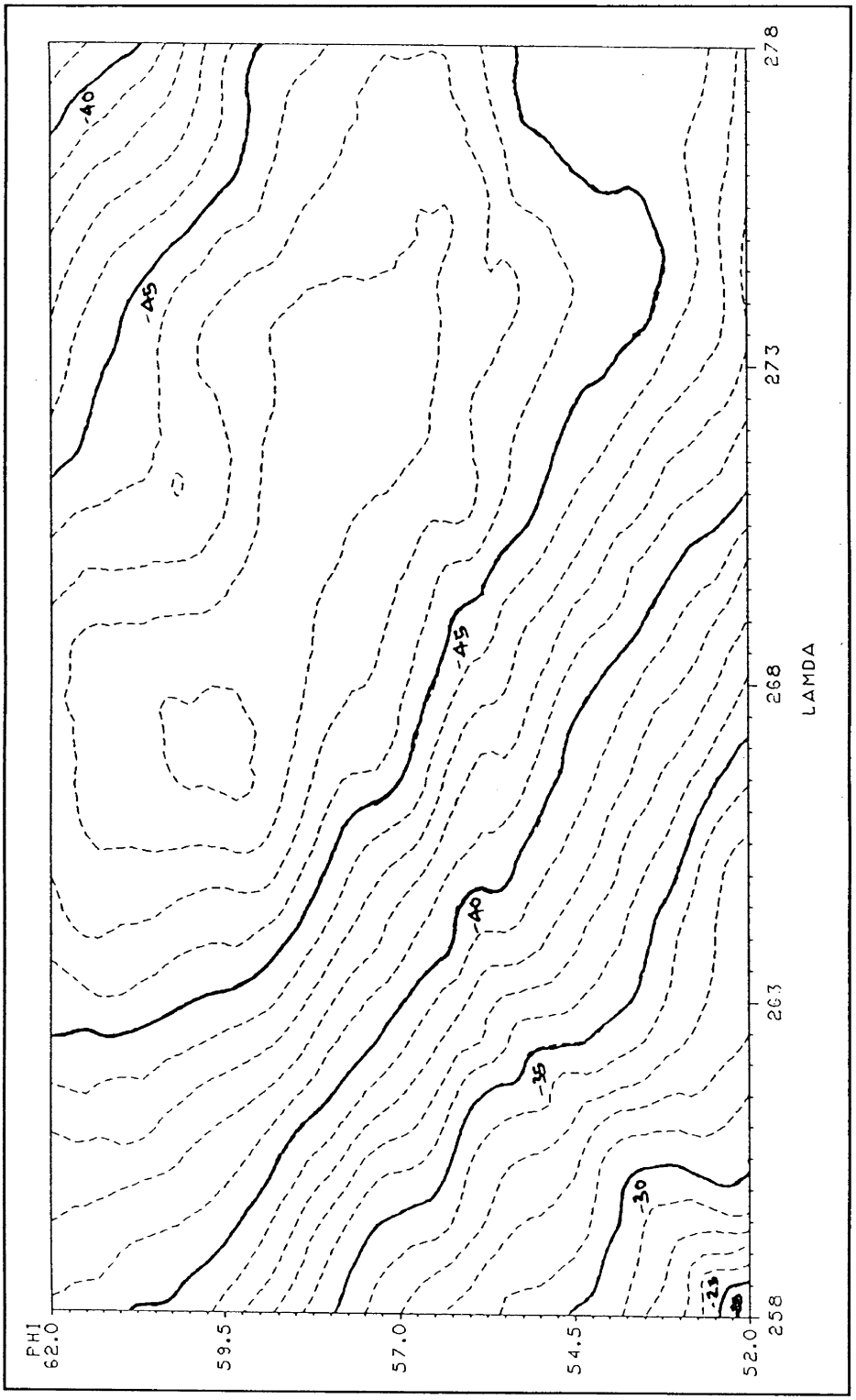


Figure 8
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 8
Region: Lat. = 52°, 62°; Long. = 258°, 278°
Range: - 49.0 m to -25.0 m; Contour interval: 1.0 m

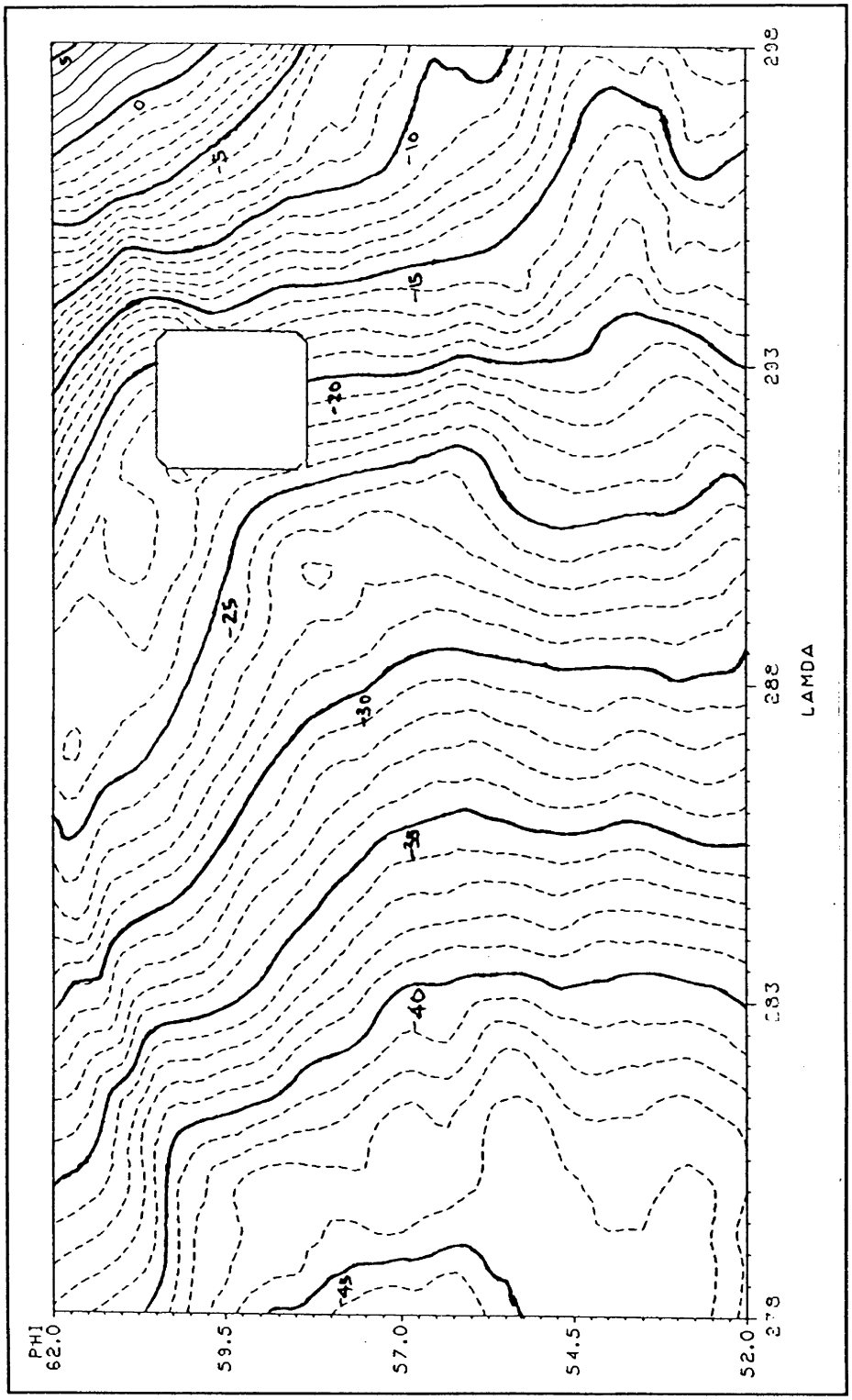


Figure 9
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 9
Region: Lat. = 52°, 62°; Long. = 278°, 298°
Range: - 46.0 m to 6.0 m; Contour interval: 1.0 m

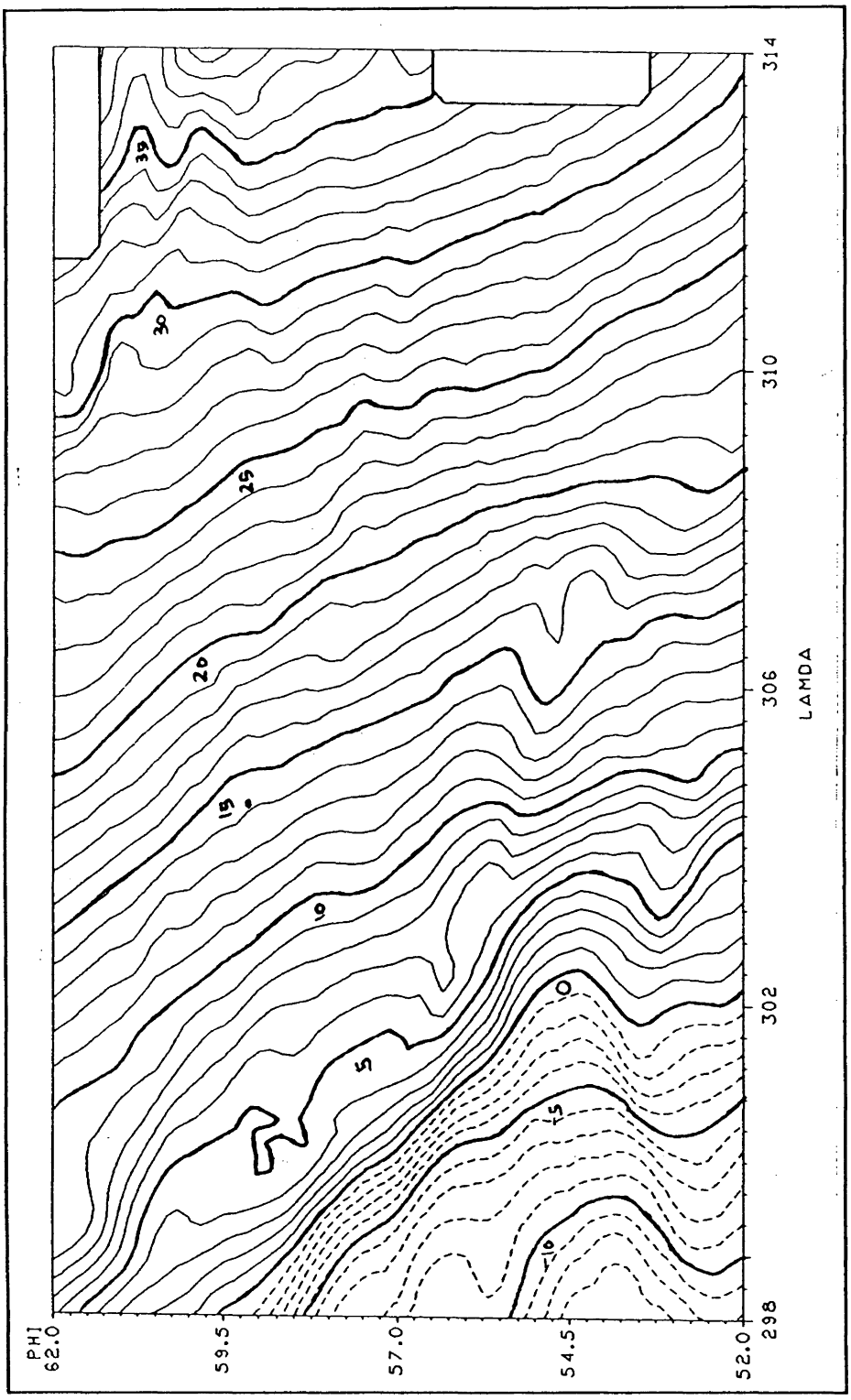


Figure 10
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 10
Region: Lat. = 52°, 62°; Long. = 298°, 314°
Range: - 13.0 m to 39.0 m; Contour interval: 1.0 m

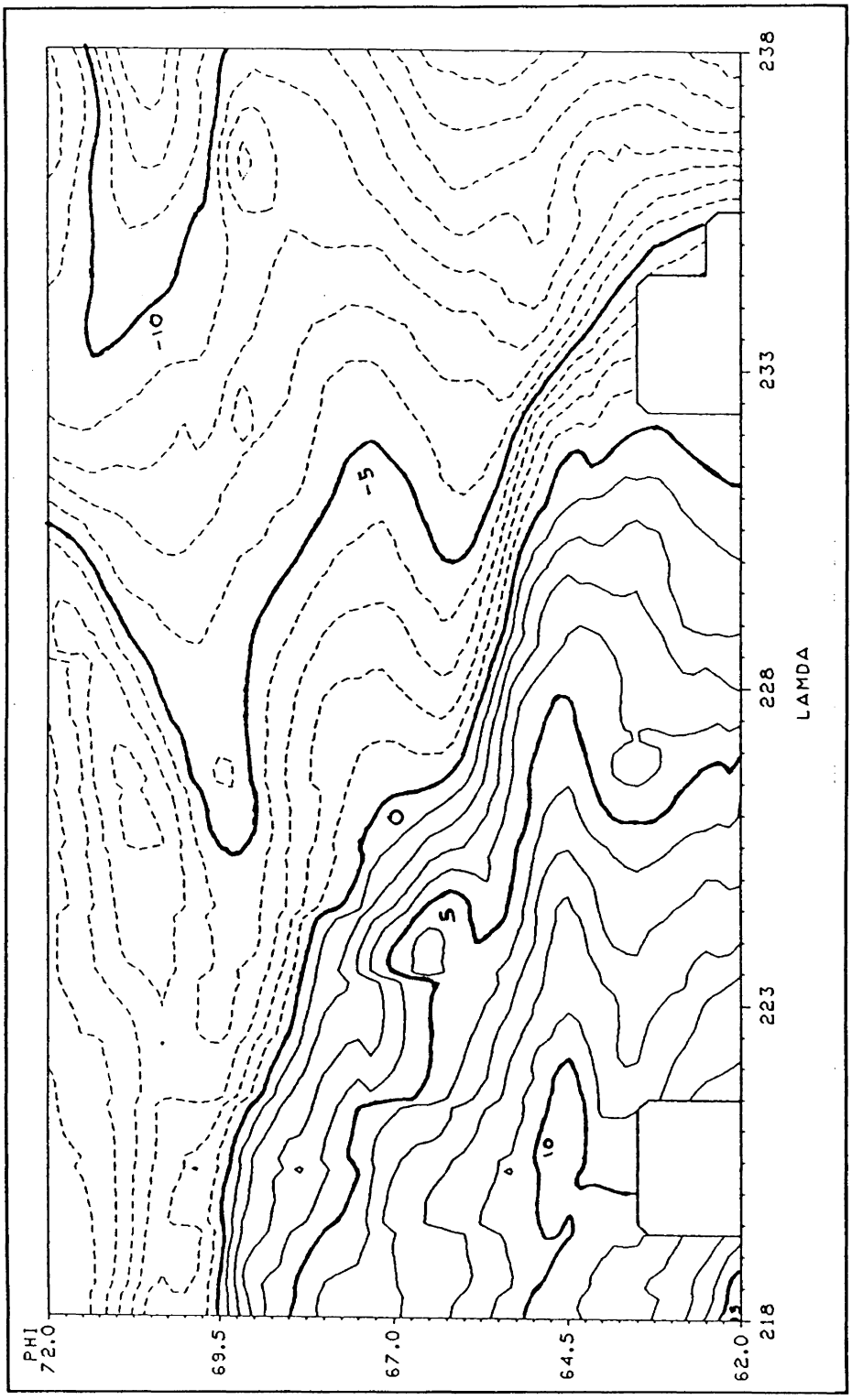


Figure 11
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 11
Region: Lat. = 62°, 72°; Long. = 218°, 238°
Range: - 13.0 m to 15.0 m; Contour interval: 1.0 m

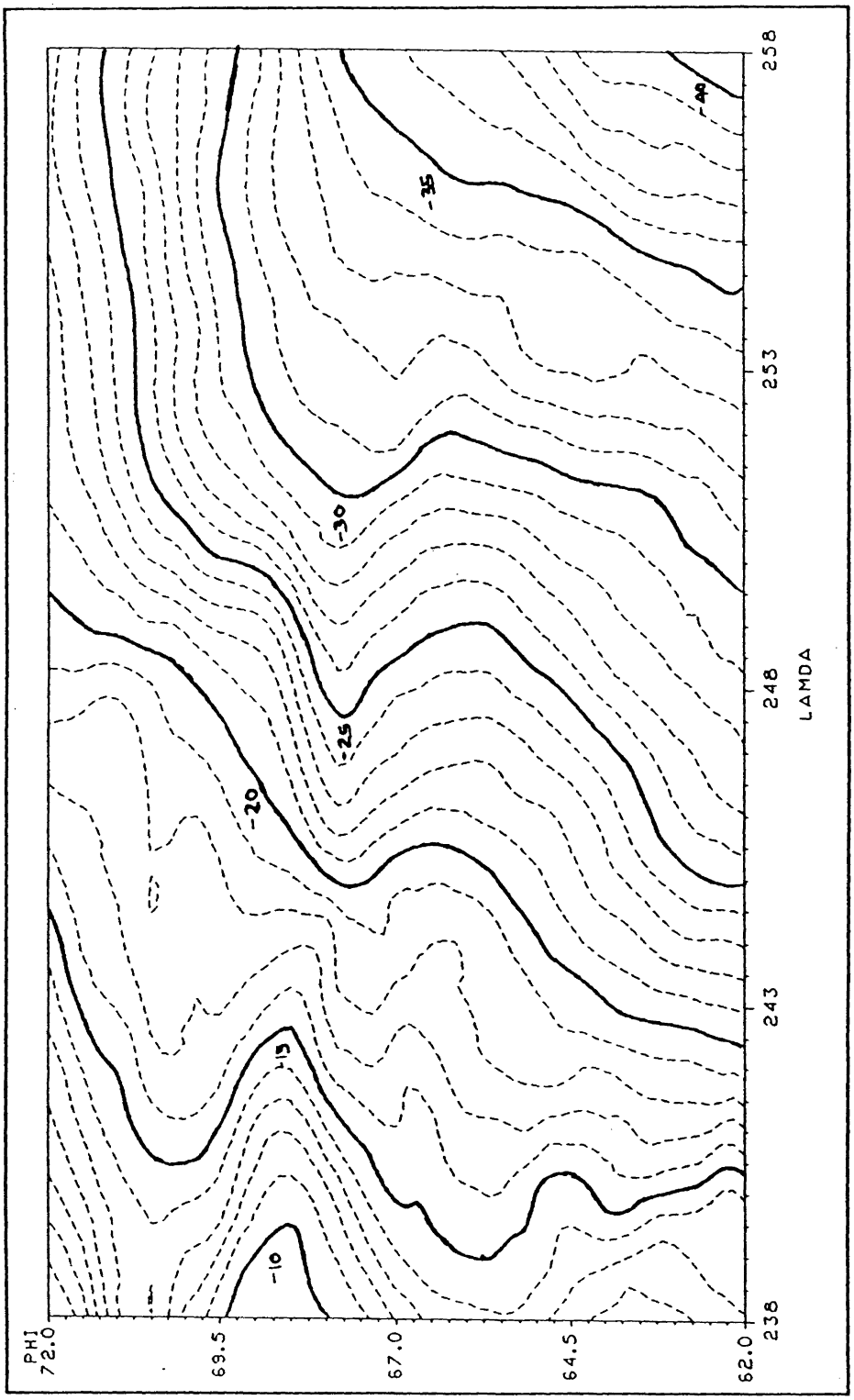


Figure 12
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 12
Region: Lat. = 62°, 72°; Long. = 238°, 258°
Range: - 9.0 m to 40.0 m; Contour interval: 1.0 m

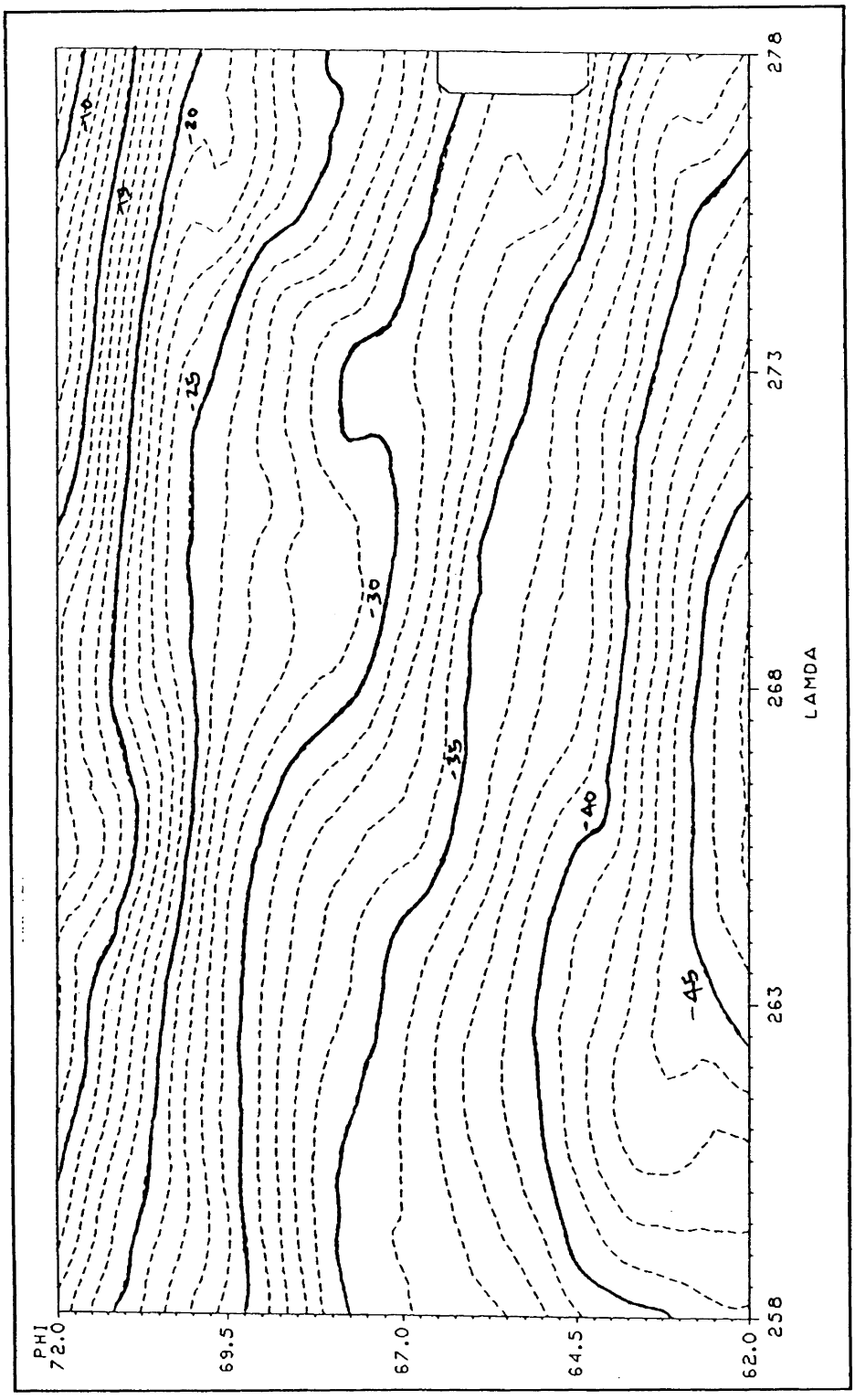


Figure 13
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 13
Region: Lat. = 62°, 72°; Long. = 258°, 278°
Range: - 47.0 m to -9.0 m; Contour interval: 1.0 m

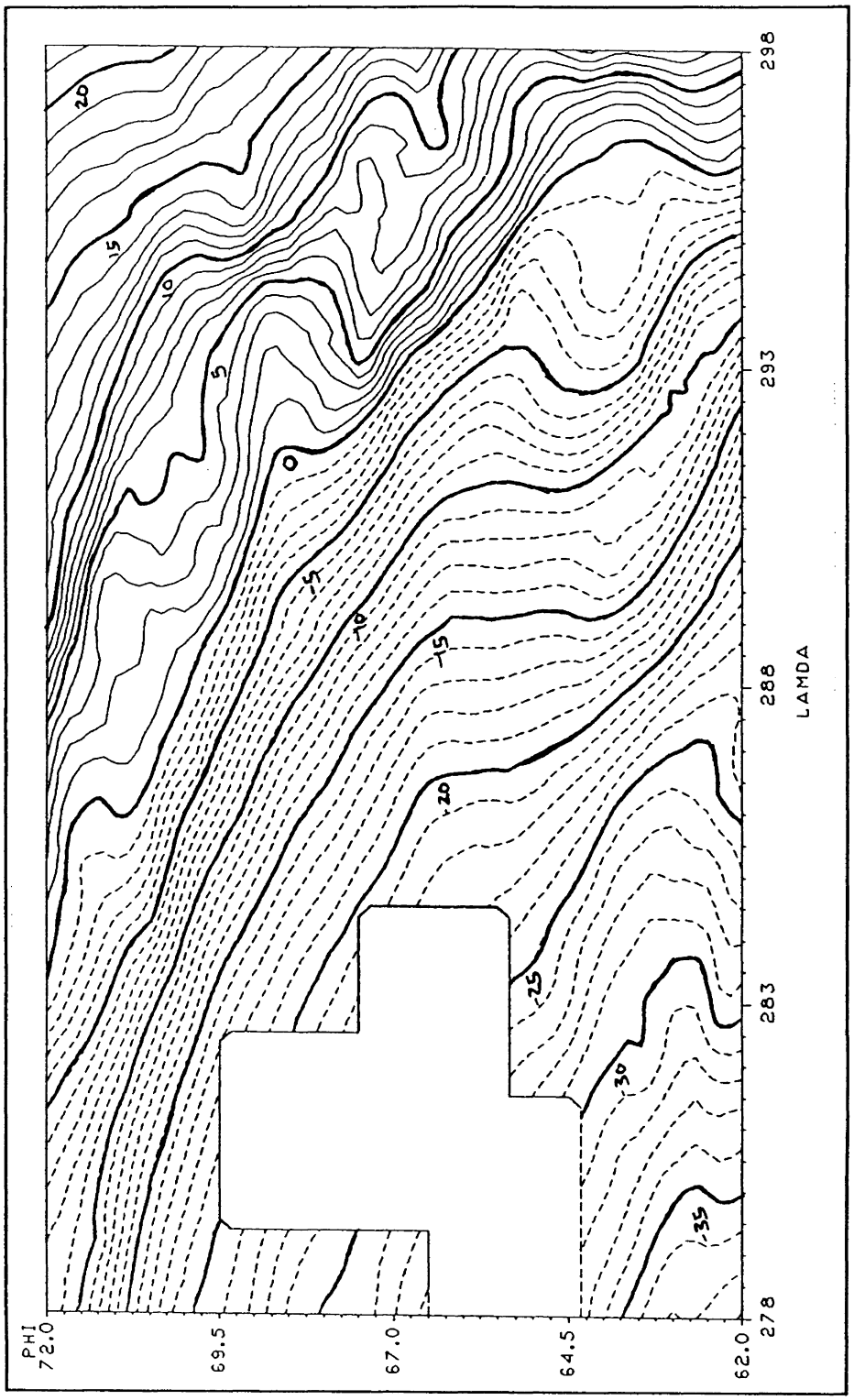


Figure 14
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 14
Region: Lat. = 62°, 72°; Long. = 278°, 298°
Range: - 38.0 m to 21.0 m; Contour interval: 1.0 m

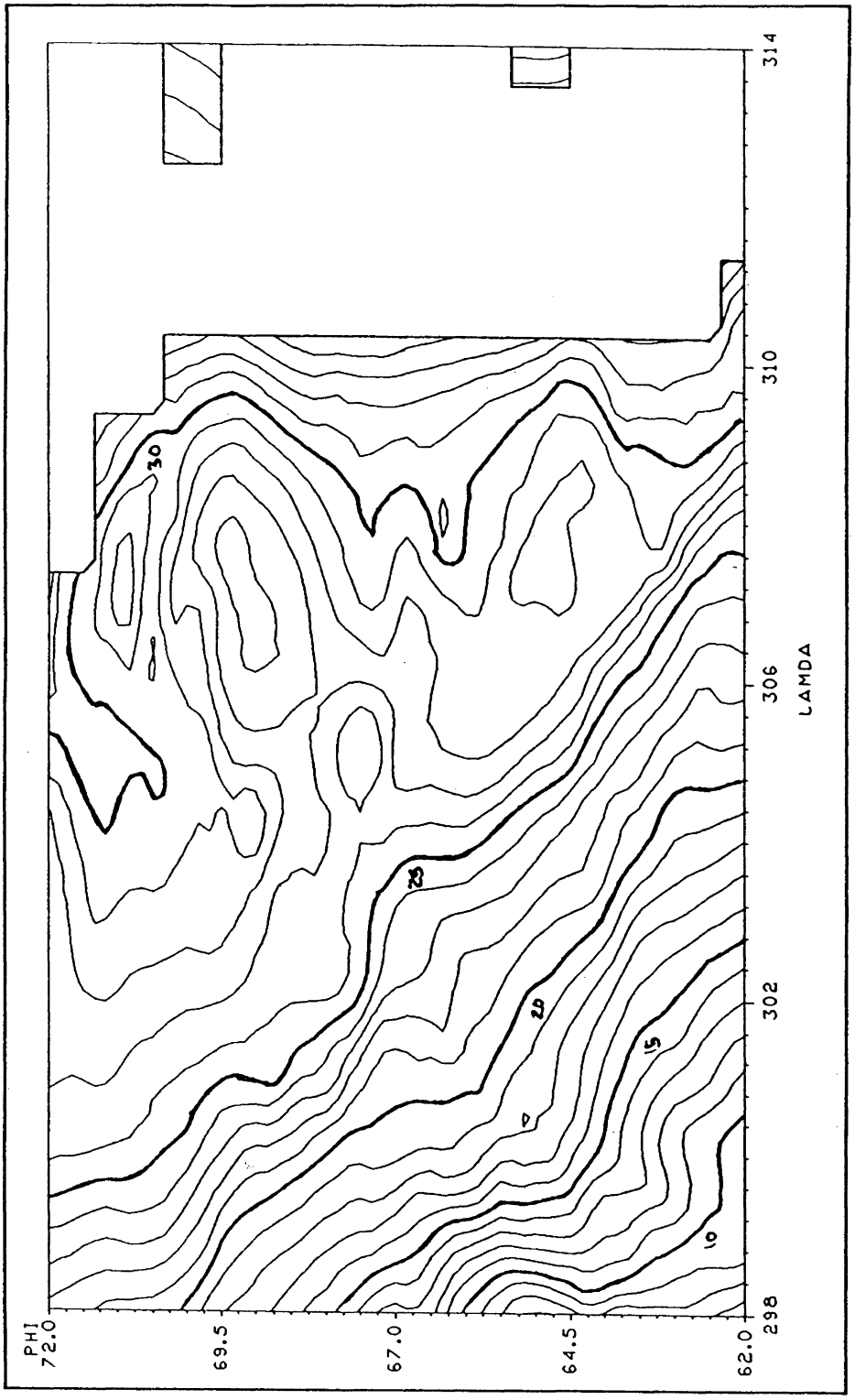


Figure 15
Geoid height (UNB90) with respect to GRS80 reference ellipsoid.
Block 15
Region: Lat. = 62°, 72°; Long. = 298°, 314°
Range: 7.0 m to 34.0 m; Contour interval: 1.0 m

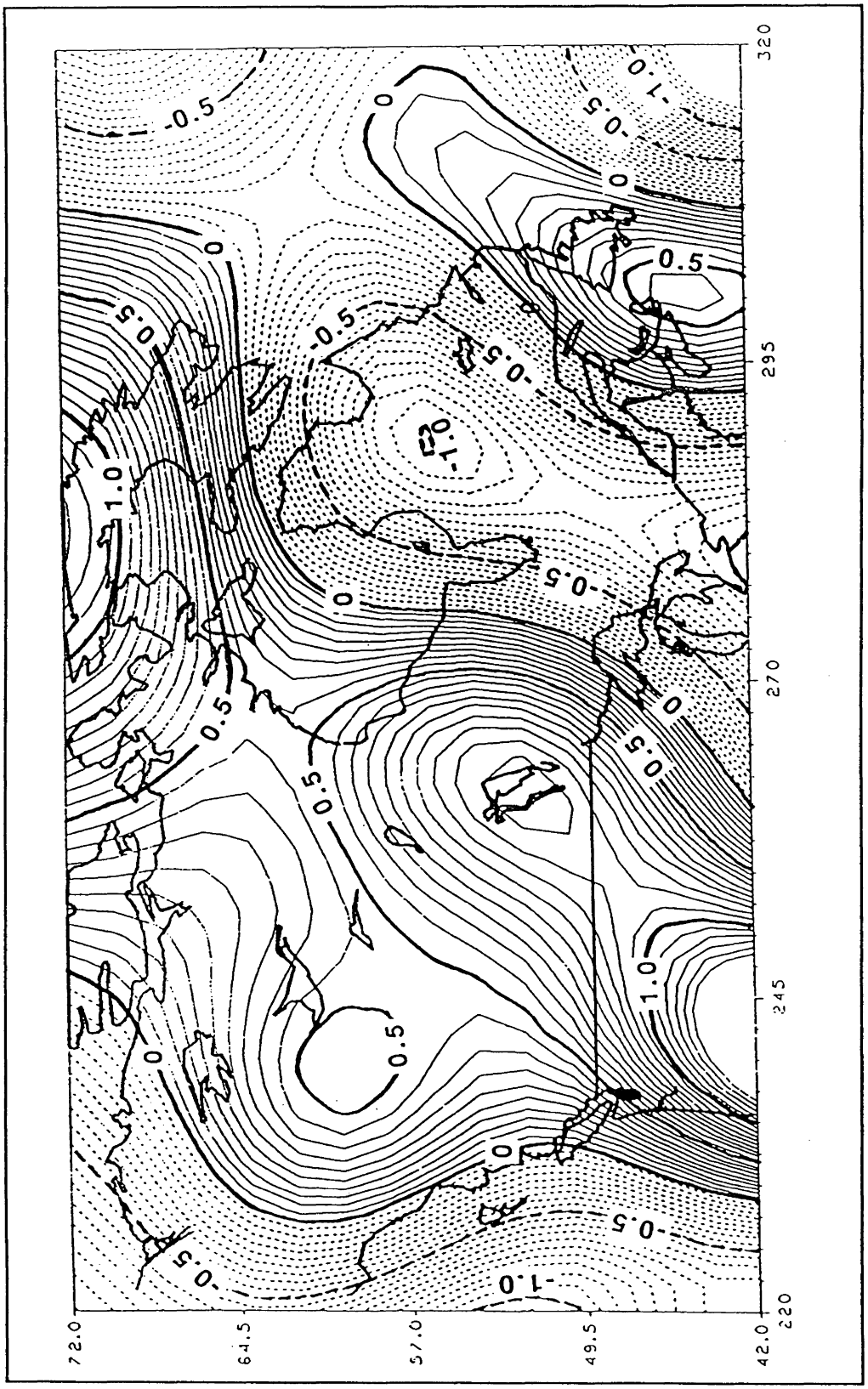


Figure 16
Correction to "UNB Dec. '86" geoid arising from the replacement of GEM-9 by GEM-T1 reference field. Contour labels in metres.

8. CONCLUSIONS

We have fulfilled the contract and produced a geoid which should be a marked improvement over the 1986 one. Even though it does not depart from the old one very significantly, every indication is that the new geoid is better because of the improvement in the satellite reference spheroid (GEM-T1 compared to GEM-9) and the improvement in terrestrial gravity coverage and quality. It will now require many tests against independent data, such as the combination of GPS and levelling and satellite altimetry, to establish its objective accuracy. These tests have been considered outside the scope of this contract.

The question of appropriate terrain correction to be applied should be resolved. Should theoretical resolution fail, it may be possible to resolve the question experimentally. Given the present accuracy of the geoid, GPS, and levelling, the experimental verdict seems to be now reachable.

We feel that the new geoid will be good enough, certainly in the flat and gently rolling areas of Canada, to be used with GPS for lower-order levelling. We expect the standard deviation of geoidal height differences computed from the grid values through some kind of interpolation, to be shown to be close to $10^{-6}.S$ (where S is the distance between the end points), i.e., 10 cm per 100 km, over most of Canada.

9. ACKNOWLEDGEMENTS

We wish to take this opportunity to thank the Scientific Authority, Dr. André Mainville, for his exemplary cooperation and repeated help. We are grateful to Mr. Yan Ming Wang, of The Ohio State University, for sharing with us his program for topographic correction computation. Dr. A. Kleusberg and Mr. N. Christou, of UNB, have contributed to the success of this undertaking by

offering advice which we wish to acknowledge gratefully. Last, but not least, we thank Ms. Wendy Wells for her editing and wordprocessing of this report.

Bibliography

Christou, N., P. Vaníček, and C. Ware (1989). "Geoid and density anomalies." *EOS, Transactions of the American Geophysical Union*, 30 May, pp. 625. 631.

Colombo, O. (1981). "Numerical methods for harmonic analysis on the sphere." Department of Geodetic Science Technical Report No. 310, The Ohio State University, Columbus, Ohio, 140 pp.

Heiskanen, W.A., and H. Moritz (1967). *Physical Geodesy*. Freeman and Co., San Francisco.

Mainville, A. and M. Véronneau (1989). "Creating a gravity grid over Canada using Bouguer anomalies and a digital elevation model." Presented at the Canadian Geophysical Union Annual Meeting, Montreal, Québec, May, 22 pp.

Marsh, J.G., F.J. Lerch, B.H. Putney, D.C. Christodoulidis, D.E. Smith, T.L. Felsentreger, B.V. Sanchez, S.M. Klosko, E.C. Pavlis, T.V. Martin, J.W. Robbins, R.G. Williamson, O.L. Colombo, D.D. Rowlands, W.F. Eddy, N.L. Chandler, K.E. Rachlin, G.B. Patel, S. Bhati, and D.S. Chinn (1988). "A new gravitational model for the earth from satellite tracking data: GEM-T1." *Journal of Geophysical Research*, Vol. 93, No. B6, pp. 6169-6215.

Moritz, H. (1980). "Geodetic Reference System 1980." *Bulletin Géodésique*, 54, pp. 395-405.

Sjöberg, L.E., and P. Vaníček (1990). "Reformulation of Stokes's theory for higher than second-degree reference field and modification of integration kernels." Department of Geodesy, The Royal Institute of Technology, Stockholm, Sweden. Submitted to *Journal of Geophysical Research*, January.

Tscherning, C.C., R.H. Rapp, and C. Goad (1983). "A comparison of methods for computing gravimetric quantities from high degree spherical harmonic expansions." *Manuscripta Geodaetica*, 8, pp. 249-272.

Vaníček, P. and A. Kleusberg (1987). "The Canadian geoid — Stokesian approach." *Manuscripta Geodaetica*, No. 12, pp. 86-98.

Vaníček, P., A. Kleusberg, R.G. Chang, H. Fashir, N. Christou, M. Hofman, T. Kling, and T. Arsenault (1987). "The Canadian geoid." Department of Surveying Engineering Technical Report No. 129, University of New Brunswick, Fredericton, N.B., December.

Wang, Y.M. and R.H. Rapp (1989). "Terrain effects on geoid undulation computations." Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio, May, accepted for publication in *Manuscripta Geodaetica*.

APPENDIX A

"Creating a gravity grid over Canada using Bouguer anomalies and a digital elevation model."

André Mainville
Marc Véronneau

Creating a Gravity Grid Over Canada Using Bouguer Anomalies and a Digital Elevation Model. *

By André Mainville and Marc Véronneau

Geodetic Survey Division, Canada Centre for Surveying, Surveys, Mapping and Remote Sensing Sector, Department of Energy, Mines and Resources, 615 Booth St., Ottawa, K1A 0E9.

Abstract

The gravity data contained in the National Gravity Data Base of the Geophysics Division of the Geological Survey of Canada are being used to produce two new grids of gravity covering Canada. The main application of these grids in geodesy is the computation of precise geoidal heights. The geoidal height (N) is essential in the computation of the common orthometric height (H) when using satellite positioning techniques, because it is the ellipsoidal height (h) that is obtained from satellite observations (h=H+N). Because some computational techniques use a cartesian grid while others use a geographical grid two new grids, with 5 km and 5 arcmin spacing respectively, are required. The spacing of gravity observations in Canada is about 10 km. In order to produce a more representative Digital Gravity Model (DGM) than the existing ones, Bouguer anomalies over land and free-air anomalies over water were used in the interpolation. The technique of least-squares collocation was utilized to interpolate. A description of the interpolation procedure is given. When Bouguer anomalies are used, the mean height of the cell is obtained from a Digital Elevation Model (DEM). These grids contain more information than the previous ones and will permit computation of a more precise geoid. Standard errors and other statistics are produced to examine the quality of the DEM and of the DGM in various regions of Canada.

Introduction

In Canada, about 600,000 gravity observations are currently available from the Gravity Data Centre of the Geophysics Division (Geological Survey of Canada, Department of Energy, Mines and Resources). The gravity data as of July 1988 are here used to produce two new regular grids of gravity covering Canada. While such grids can be used to reproduce maps of gravity or for geophysical studies, our main intent is to compute precise geoidal heights (N) which when

* Presented at the Canadian Geophysical Union Annual Meeting, Montreal, Quebec, May, 1989.

combined with ellipsoidal heights (h) determined by satellite techniques provide us with orthometric heights (H=h-N).

Surface gravity anomalies (in a circular area of radius of up to 10°) are used to estimate according to various methods the contribution of the local gravity field to the geoidal height. Some of these methods are described in Lachapelle (1978), Schwarz and Sideris (1985), Vanicek et al. (1986), (1987) and Mainville (1987). The gravity grid mostly utilized in Canada for previous geoid computations was created in 1978 (Lachapelle, 1978) and last updated in August 1979 (Lachapelle and Mainville, 1980). The object of this paper is to recompute such grid and present the preliminary results of the computation for an improved free-air gravity anomaly grid to be used for geoid computation.

Producing a grid of gravity values is not absolutely necessary. Some numerical methods (Tscherning, 1985) (Landau et al., 1988) which use least-squares collocation are successfully applied to irregularly spaced data to obtain various geoid components. There are, however, advantages to produce gridded data. Numerical integration and the manipulation of data are often simplified with gridded data. Numerical problems encountered with irregularly spaced data are eliminated. Computations are often faster with gridded data. Computer storage is reduced since one does not have to save the coordinates of the values. This is especially important with large data sets.

For theoretical developments based on a sphere or an ellipsoid, the geographically gridded data is more practical. This was the case for Lachapelle (1978), Vanicek et al. (1986) and Mainville (1987). Other developments use a planar reference surface. This was the case for Schwarz and Sideris (1985) who used an integration technique based on the fast Fourier method which requires data on a planar cartesian grid. For mapping purpose, a cartesian grid is also required. Two grids, a geographical and a cartesian one, are being produced to satisfy both of the above applications. The cartesian grid will be computed after completion of the geographical grid. Here preliminary results of the later grid are presented.

The objective is to compute as small as practical a grid in view of recovering as many frequencies as possible from the gravity field. The average spacing of the surface gravity measurements on land in Canada is around 10 km (Figure 1). The spacing is smaller in populated areas and even smaller over oceans (Figure 1). It was decided to produce a cartesian grid with cells of 5 km and a geographical grid with cells of 5 arcmin. A 5 arcmin in latitude represents 9 km. A 5 arcmin in longitude represents 7 km at latitude 42°, 5 km at latitude 57° and 2 km at latitude 82°. Obviously

we are trying to compute a denser grid than the observed one. For this reason, our objective is to create cells with representative gravity values and not with mean values. It may happen that a mean value is the most representative value of a cell. This is the case when for example, there are 10 or more values in a small cell of 9 by 7 km. This was mostly the case over oceans. There, the most representative value was in effect obtained by computing the mean of the 10 or more free-air anomalies in a 5' cell. This, however, happened only for 3 % of the grid. For more than 50% of the grid there was no gravity observation in the cell. And, in another 25% of the cases, there was only 1 observation in the cell. In both cases, the mean value is obviously not the most representative value one can find. For these cases, which represent more than 75 % of the grid, an interpolation solution was found preferable. After various tests which are described below, an interpolation procedure using the 5 nearest observations within 30 km to the cell center was selected.

To find the most representative value of a cell a digital elevation model (DEM) was used in the following way. Simply put, if a cell contains one gravity observation but the mean elevation of the cell is higher than the elevation of that one observation (e.g. made in the valley), then one can assume with some confidence that the most representative gravity value for the cell is smaller than that one observed value. The DEM is also used to add more gravitational information into the gridded gravity set herein called digital gravity model (DGM). If a cell contains no gravity observation but its mean elevation is higher than the surrounding cells, then one can assume with some confidence that its gravity value is smaller than the surrounding cells. As one can already understand, the accuracy of the DEM is very important for a successful application of the above principles.

The Digital Elevation Model

The main characteristics of the DEM sought are the following ones. It must cover all of Canada, be as dense as possible and as accurate as possible. The densest DEM available and covering all of Canada has a 5 arcmin spacing. This drive our choice for a 5' spacing DGM. According to Schwarz and Sideris (1985), gravity observations at 3 km spacing is required in order to recover geoidal heights to an accuracy of 0.3 ppm. While gravity observation can certainely not be replaced by simple interpolation solutions even using a DEM like herein proposed, a 3 km DEM would have been more adequate than the 5 arcmin DEM available.

The DEM used is the worldwide gridded data set called ETOPO5 of the american National Geophysical Data Center. It was obtained through the canadian Geophysics Division. ETOPO5 contains sea-floor elevations compiled by the U.S. Naval Oceanographic Office and land elevations supplied by the american Defense Mapping Agency (DMA). The accuracy of the bathymetric data was of no concern since the depth values will not be required here. The portion of ETOPO5 over Canada contains mostly the same data that Lachapelle (1978) used in computing his DGM over Canada ten years ago. It was verified that the elevations in both files are exactly the same except for some small updated areas. It was discovered that the definition of the grid of ETOPO5 does not agree with the original DMA file used by Lachapelle (idem). After verification with the 1 km DEM (Geophysics Division, 1984) over the canadian Rockies, the ETOPO5 grid has wrong coordinates for the portion over land and had to be shifted by 2.5 min south and 7.5 min west to agree with the original DMA file.

An estimated accuracy is published with the DMA file. The 68,927 cells (18% of the DEM) with an elevation estimated less accurate than 9 metres are shown on Figure 2. The other 308,296 cells (82% of the DEM) have accuracy estimates better than 9 metres. There are also 346,817 cells (92% of the DEM) more accurate than 20 metres. Elevation errors of 9 and 20 metres correspond to Bouguer plate corrections (equations (3) and (4) below) of 1 and 2 mgal respectively. This introduces errors in the DGM of 1 to 2 mgal. These are errors smaller than the interpolation error and are of the order of the accuracy of most free-air gravity values since their elevations are often accurate to a few metres. Excluding the DEM over oceans and lakes which are not used here, and updating the elevations in B.C. and Yukon with the above mentioned 1 km DEM, then the principal areas of concern are the large inaccurate regions of Ontario and Quebec (Figure 2) and along the coasts. These areas and the overall quality of the DEM were verified against the mean height of the gravity observations. The map of the differences agrees fairly well with Figure 2. The largest differences between the mean height of the gravity observations and the DEM are mostly in rolling and mountainous areas and are mostly due to gravity being measured in valleys. A good quality DEM computed from the average of many point values within a cell should be more representative than the average computed from few observations. This forms the basis for using the DEM in the computation method described below. It is however acknowledged that a more accurate DEM is required. For example, some discrepancies are found along the coasts. There a carefull examination would be required. The concerns left are the incertitude of the accuracy of the DEM, especially in the areas shown on Figure 2, and the general concern that a densier DEM throughout Canada is required.

Computations

To form the geographical DGM the gravity was interpolated at the elevation given by the 5' DEM. The surface gravity observation (g) - not corrected for the atmospheric effect (Moritz, 1980a) - is first reduced by the theoretical normal gravity (γ) - here based on the GRS67 (IAG, 1967) - and the free-air correction (0.3086 H) to obtain a smaller quantity than the gravity value, the free-air anomaly (Δg_F) in mgal.

$$\Delta g_F = g - \gamma + 0.3086\ H \tag{1}$$

H is the elevation, in metres above mean sea level, of the gravity observation. It is obtained by barometric, inertial survey or spirit leveling methods during the gravity survey. The five Δg_F nearest to the cell center are then chosen for the interpolation. Their corresponding Bouguer anomaly (Δg_B) is calculated only if none of the five Δg_F were observed over fresh or salted water.

$$\Delta g_B = \Delta g_F - 0.1119\ H \tag{2}$$

Because of this reduction, the range of the five Δg_B is, most of the time, when on land, smaller then the range of the five Δg_F. The idea being that smaller the range, better the interpolation, the Bouguer anomaly is preferred in the interpolation over land. Most of the time, when over water, the range of the Δg_F is smaller then the range of the Δg_B. In this case the free-air anomaly is preferred in the interpolation. If at least one of the five Δg_F was measured over water it was verified that it is preferable to perform the interpolation with the Δg_F. Before interpolating with least-squares collocation the five Δg_B or Δg_F must be reduced by their mean (Moritz, 1980b). The mean value is then added back to the interpolated value. If a Bouguer anomaly was interpolated ($\overline{\Delta g}_B$), the Bouguer plate due to the DEM is added back.

$$\Delta g_F = \overline{\Delta g}_B + 0.1119\ H_{DEM} \tag{3}$$

If a free-air anomaly was interpolated ($\overline{\Delta g}_F$) because at least one observation was over water, the Bouguer plate between the mean height of the five Δg_F and the elevation at the point of interpolation given by the DEM is added back.

$$\Delta g_F = \overline{\Delta g}_F + 0.1119\ (H_{DEM} - \bar{H}) \tag{4}$$

where $\bar{H}$ is the mean height, in metres, of the five Δg_F and H_{DEM} is the elevation, in metres, of the DEM. Details of the derivation of equations (3) and (4) are given in the appendix. Even the interpolated free air anomaly is corrected for the Bouguer plate - equation (4) - since it is assumed that the height of the DEM is more representative then the mean height of the 5 Δg_F. Obviously that statement is very dependent on the accuracy of the DEM at various places. In any event, when H_{DEM} and $\bar{H}$ are both zero then - see equation (4) - no correction is applied. In all other cases a correction is applied.

Before gridding the whole 572,176 gravity observations using the above procedure, the mean, the weighted mean and the least-squares collocation solutions were tested. With these three solutions statistics were obtained showing the reason to adopt the collocation solution. The covariance function used here was the one derived by Schwarz and Lachapelle (1980). We have also tried to reduce the magnitude of the Δg_B (or Δg_F over water) with the gravity anomalies obtained from the geopotentiel model OSU86F (Rapp and Cruz, 1986). It did not reduced the magnitude of the anomalies and did not ameliorate the estimation. This shows that while OSU86F type models are useful to model geoidal heights, they are not yet accurate enough to model the high frequencies of the Earth gravity field needed for this application of gravity anomalies.

Tests

The different techniques of interpolation, the number of observations and the many ways to smooth the observations represent many parameters that one can fine tune in order to find out the most representative gravity value in a cell. The final DGM contains a mixture of these parameters which optimal values were selected by the following tests.

Over land, Bouguer anomalies are much smoother than free-air anomalies; for this reason it is preferable to interpolate gravity using Bouguer anomalies. This was tested at 7 typical regions across Canada (Figure 4) where the interpolation was performed at points of known gravity value. This test is summarized in Table 1. Each region covers a 1° by 1° cell with the coordinates of the south-west corner given in column 3 and the number of observations in the cell given in column 4. With as few as 5 observations the distance of the furthest observation from the interpolation point is 10 to 18 km as seen in column 5 of Table 1. It was verified that the interpolation deteriorates if observations were chosen too far from the interpolation point. A

distance of 10 to 18 km is already far enough to get a representative gravity value at a 5' (9 km) spacing. Finally it was decided to always use the 5 closest observations in the interpolation procedure but to a maximum radius of 30 km. The root mean square (RMS) difference when interpolating with free-air or Bouguer anomalies at known points are given in the last 2 columns of Table 1. Over Mount Columbia and in the Okanagan Valley, two mountainous regions of Brithish Columbia, where the topography varies from 354 metre to 3372 metres, gravity values are recovered with an RMS of 14.4 and 8.6 mgals when using Bouguer anomalies. RMS of 58.8 and 76.4 mgal are obtained when using free-air anomalies. It was then verified that it is always preferable over the land area to use Bouguer instead of free-air anomaly. It was tested in typical regions, in Edmonton, a rolling region at an altitude of 1000 m, in northern Quebec, a rolling region at an altitude of 600 m and in northern Ontario, a flat region at an average altitude of 200 m. The results of the last two regions of Table 1 demonstrate that free-air anomalies should be used over the oceans. One of the region over the ocean was 5 times denser in observations then the other one. This explains its 1 mgal RMS.

The results in Table 1 are those obtained using a simple arithmetic mean. As seen from Table 2 the results are basically the same with three other techniques that were tested. The weighted mean based on the inverse of the distance, the weighted mean based on the distance and on the standard error of the observations, and the least-squares collocation technique, all gave, for all practical purpose, identical results. Kassim (1980, p.76) had concluded that the weighted mean gave better results than collocation. His weighted mean used Bouguer anomalies while his collocation solution used free-air anomalies the correlation with the elevation removed as suggested by Lachapelle and Schwarz (1980). If Kassim (idem, p.34) would have used the correlation coefficient (0.1119) of the Bouguer plate, the weighted mean and the collocation technique would have given identical results. This was later proven by Mainville (1982, p.128) where the best interpolation procedure testing various correlation coefficients always converged to the Bouguer coefficient. These results are summarized in Figure 4 and Table 3. The RMS difference is always smaller when using Bouguer anomalies instead of the anomalies reduced by the correlation coefficients b and h_o (Table 3).

In addition, it is here demonstrated that the use of the straight mean gives identical results to the weighted mean, and collocation (Table 2). One should not forget however that it is the mean of the Bouguer anomaly with the correction of equation (3) or the mean of the free-air with the correction of equation (4), and not of the free-air or of the gravity itself that must be used.

As demonstrated in Kassim (idem, p.76) the least-squares collocation is however the technique that performed best in estimating precision. Kassim (idem, p.27) also mentionned that the weighted mean has sometimes given unacceptable results when data coverage is sparse. For these two reasons the least-squares collocation is here preferred, first to get the estimated accuracy and then to get the optimum value when the observations are distant from the interpolation point.

Based on these tests, the procedure chosen to compute the 5' DGM is the following one. When there was 10 or more observations in a 5' cell the straight mean was computed from the Bouguer anomalies (case 1) obtained with equation (2), or from the free-air anomalies (case 2) if at least one point is over water. The RMS of the standard deviations was also computed. Equation (3) is used with case 1 to calculate the free-air value at the DEM elevation. Equation (4) is used with case 2. For the cells with 1 to 9 points the collocation technique was adopted to interpolate and compute the estimated accuracy using the 5 closest points within a 30 km radius. In case (1), Bouguer anomalies and equation (3) were used. In case (2), free-air anomalies and equation (4) were used.

Analysis

Figure 3 shows the standard errors of the DGM, estimated by collocation, larger than 3 mgal. That map also shows the extent of the preliminary DGM. The final DGM will cover the full map of Figure 3 i.e. from latitude N35° to N90° and longitude W150° to W40°. Table 4 summarizes the computations. All of the 572,476 surface gravity observations contained in the National Gravity Data Base were used in the computations. The full base map on Figure 3 contains 871,200 5' cells. The extent of the preliminary DGM embraces approximatively 456,000 5' cells. Approximatively 288,000 cells contain no observation; 84,390 cells contain 1 observation; 71,229 cells have 2 to 10 observations; 11,921 cells contain 10 or more observations; 416,538 cells were interpolated using the 5 nearest observations within 30 km; and, 22,552 cells were interpolated/extrapolated using 1, 2, 3 or 4 nearest observations within 30 km. Thus, approximatively 5,000 cells were left without free-air values. They are shown on Figure 3 as black regions in B.C., Yukon, Ungava Bay, Foxe Bassin, Ellesmere Islands and in the south-east corner. Over the oceans the grid could be filled by satellite radar altimeter derived data, and by values derived from topographical masses (Sideris et al, 1988) in B.C. and Yukon. These numbers show that 63% of the grid contains no gravity observation and 82% have none or only one observations. It demonstrates that a procedure to find a representative value rather than

to find an average value is here required. Taking the averages would smooth the computed gravity field. Finally the above numbers demonstrates that an interpolation technique (here the collocation technique described in the previous sections) was required to predict 91% of the grid.

The map of the estimated standard error (Figure 3) agrees well with the varying distribution (spacing) of the gravity observations (Figure 1). Compared to the results obtained in Table 1, the estimated standard errors might be a bit pessimistic, and over the mountains in south B.C., a bit optimistic. While the covariance function used here performed reasonably well and was more representative of the situation over Canada than the one used to produce the old gridded data set in 1978, a more representative covariance function could be found from the reduced anomalies used in our computations.

The new gridded free-air data set was compared to the old 1978 data set. Their differences, when larger than 10 mgals, are shown in Figure 5. The two data sets are significantly different. It is not surprising that there are large differences over the mountains of B.C., southern Alberta, Ellesmere Island and the Appalachians. It is due to the better use of the Bouguer anomaly and the interpolation technique used here. The main differences however occur over all three oceans and also the Great Lakes, Lake Winnipeg and most other lakes (see Figure 5). The reason being that a terrain correction was used in 1978 and that it was here shown that simple free-air anomalies should be used over oceans (see Table 1) and lakes. The Hudson Bay was computed in 1979, during the last update to the 1978 data set. A terrain correction was not applied at that time because it was already felt that it was wrong. Statistics of the differences between the two gridded data sets are given in Table 4. The RMS difference between the 2 data sets is 16.1 mgal for 172,062 cells; 31% of the cells have differences larger than |5| mgal; and, 17% have differences larger than |10| mgal. Histograms of the differences are also shown in Figure 6. The old data set consists of 5' by 5', 5' by 10', 5' by 15' and 5' by 20' cells. 69% of the old data set, mostly in central Canada (see Figure 5), agrees with the new data set. That explains the successes and usefulness that the old data set has brought to its many users. On the other hand it explains the misfortune to other projects, around the Great Lakes, the Winnipeg lake and many other regions shown on Figure 5.

The cells computed by collocation were at the same time computed using the simple arithmetic mean of the same 5 nearest observations. Figure 7 show the histograms of the differences. The grid was divided in 3 regions extending from latitude 35° to 55°, 55° to 75° and 75° to 90°. The northern region does not show any particular problem which could have been expected because of the convergence of the meridians creating a denser grid requiring more interpolation. It should

prove very useful in the future to have a denser grid in the north to help computing the geoid. Other statistics are given in Table 4. A RMS difference of 3.1 mgal for 451,011 cells was obtained between the collocation and the average technique. 93% of the cells agree within |5| mgal and 98.2% of the cells have differences smaller than |10| mgal. The largest differences are being investigated. In any event, it shows the relative agreement that exist between the different technique previously tested (see Table 2 and the Tests section).

Conclusion

In Canada, more than half a million gravity observations are currently available from the Gravity Data Centre of the Geophysics Division (Geological Survey of Canada, Department of Energy, Mines and Resources). The local distribution of gravity observations currently available is fairly irregular. In order to avoid numerical problems, save computer time, use fast techniques such as the fast Fourier technique it is important to use a gravity distribution that is fairly regular. It is most appropriate to use a regular grid, either a cartesian grid that follows a projection such as the Lambert Conformal Conical or a geographical grid that follows the geographic meridians and parrallels. The production of the gravity map in the Canadian Geophysical Atlas and the FFT technique require the cartesian grid. Other methods require the geographical grid. Both are required and both are being developed in this project. Here the results of the preliminary geographical grid are presented.

It is well known that better interpolation is achieved when the range of the magnitude of the anomalies used in the interpolation is small. This statement was verified by various tests described here which resulted in using Bouguer anomalies to interpolate between anomalies on land and using free-air anomalies to interpolate between anomalies when at least one anomaly was observed over the ocean or a lake. The new gridded gravity data file contains many ameliorations. (1) The cells are smaller, being 5' by 5' instead of the 10' by 10', 5' by 15' and 5' by 20' cells which were being used. This will allow to compute a more detailed geoid above latitude 50°. (2) The free-air anomaly computation over oceans and lakes was corrected and this will improve the geoid solutions over and around these areas shown on Figutre 5. (3) The use of the closest observations to the cell center and of the Bouguer anomaly has improved the griddec set especially over the mountains but also over all Canada. (4) Since a more representative covariance function than the one used in 1978 was used, the estimated standard deviation are more representative than before. The covariance function could however be recomputed using the residual anomalies used here. (5) A 5' by 5' digital elevation model (DEM) was used to

correct the interpolated gravity values to obtain more representative values. While the mean of many elevations within a 5' cell as obtained from a DEM is more representative than a mean of the elevation at a few gravity stations, there are some areas where the quality of the DEM used is uncertained (Figure 2). The DEM used needs additional verifications and ameliorations. A densier and more precise DEM would be welcomed.

References

Geophysics Division (1984) **Digital Terrain File, Open File 83-25**, Internal Report of the Gravity Data Centre, Geophysics Division, Geological Survey of Canada, 21 pp., January 1, 1984.

Goodacre, A.G., R.A.F. Grieve, and J.F.Halpenny (1987) **Gravity Maps of Canada**, Geological Survey of Canada, Canadian Geophysical Atlas, Maps 1 to 5, scale 1:10 000 000.

IAG (1967) **Geodetic Reference System 1967**, International Association of Geodesy, Bulletin Geodesique, Paris, 1967.

Kassim, F.A., (1980) **An Evaluation of Three Techniques for the Prediction of Gravity Anomalies in Canada**, Technical Report No.73, Department of Surveying Engineering, University of New Brunswick, 82 pp., September, 1980.

Kearsley, A.H.W., M.G. Sideris, Jan Krynski, René Forsberg and K.P. Schwarz (1985) **White_Sands Revisited - A Comparison of Techniques to Predict Deflections of the Vertical**, Report No. 30007 of the Division of Surveying Engineering, The University of Calgary, Calgary, 166 pp., Alberta.

Lachapelle, G. (1978) **Evaluation of 1° x 1° Mean Free Air Gravity Anomalies in North America**, Collected Papers, Geodetic Survey Division, Canada Centre for Surveying, Dept. of Energy, Mines and Resources, p. 183-213, Ottawa.

Lachapelle, G. and A. Mainville (1982) **Disturbing Potential Components Software at the Geodetic Survey of Canada**, Internal Report, Geodetic Survey Division, Canada Centre for Surveying, Dept. of Energy, Mines and Resources, Ottawa.

Landau, H., K. Hehl, B. Eissfeller, G.W. Hein and W.I. Reilly (1988) **Operational Geodesy Software Packages**, Institute of Astronomical and Physical Geodesy, University FAF Munich, Federal Republic of Germany.

Mainville, A. (1982) **Évaluation des déviations de la verticale et des ondulations du géoïde à l'aide de données hétérogènes dans les régions montagneuses**, M.Sc. Thesis presented at the Graduate School of "l'Université Laval", 184 p., Quebec, March 1983.

Mainville, A. (1987) **Intercomparison of Various Geoid Computational Methods at GPS Stations**, Presented at the XIX General Assembly of the International Union of Geodesy and Geophysics, 17 p., Vancouver, August 1987.

Moritz, H. (1980a) **Geodetic Reference System 1980**, The Geodesist Handbook 1980, Bulletin Geodesique, Vol.54, No.3, pp.395-405.

Moritz, H. (1980b) **Advanced Physical Geodesy**, Abacus Press, Tunbridge Wells Kent, 500pp., U.K.

Rapp, R.H., and J.Y. Cruz (1986) **Spherical Harmonic Expansion of the Earth's Gravitational Potential to Degree 360 Using 30' Mean Anomalies**, Report No.376 of the Dept. of Geodetic Science, The Ohio State University, Columbus, Ohio, 22 pp., December 1986.

Schwarz, K.P. and M.G. Sideris (1985) **Precise Geoid Heights and Their Use in GPS Interferometry**, Contract Report No. 85-004 of the Geodetic Survey Division, Canada Centre for Surveying, Dept. of Energy, Mines and Resources, Ottawa, 38 pp., September 1985.

Schwarz, K.P., and Lachapelle (1980) **Local Characteristics of the Gravity Anomaly Covariance Function**, Bulletin Géodésique, Vol.54, No.1, Paris.

Sideris, M., K.P. Schwarz and A.C. Rauhut (1988) **The Geoid in Northern British Columbia**, Contract report No. 88-004 of the Geodetic Survey Division, Canada Centre for Surveying, Dept. of Energy, Mines and Resources, Ottawa, 41 pp., June 1988.

Tscherning, C.C. (1985) **Geocol - a Fortran-program for Gravity Field Approximation by Collocation**, Technical Note, Geodaetisk Institut, 3.ed., March 25, 1985.

Vanicek, P., and A. Kleusberg (1987) **The Canadian Geoid - Stokesian Approach**, Manuscripta Geodaetica, Springer-Verlag, Vol.12, No.2, pp.86-98.

Vanicek, P., A. Kleusberg, R-G. Chang, H. Fashir, N. Christou, M. Hofman, T. Kling and T. Arsenault (1986) **The Canadian Geoid**, Contract Report No. 86-001 of the Geodetic Survey Division, Canada Centre for Surveying, Dept. of Energy, Mines and Resources, Ottawa, 117 pp., February 1986.

Appendix

The intention of this appendix is to show the source of equations (3) and (4). After an interpolation that used 5 Bouguer anomalies one obtain a Bouguer anomaly Δg_{1B} with a corresponding gravity value g_1, and an associated height H_1 (perhaps the mean height of the 5 Bouguer anomalies). The DEM gives the elevation H_2 for the point of interpolation, to which corresponds a gravity value g_2. Neglecting the detailed terrain correction, but taking into account the Bouguer plate and the free-air correction, the relation between the 2 gravity values is

$$g_2 = g_1 - 0.1119 (H_1 - H_2) + 0.3086 (H_1 - H_2) \quad . \tag{A}$$

Since the corresponding free-air anomalies are defined as

$$\Delta g_{2F} = g_2 - \gamma_2 + 0.3086 H_2 \tag{B}$$

and

$$\Delta g_{1F} = g_1 - \gamma_1 + 0.3086 H_1 \tag{C}$$

and since $\gamma_1 = \gamma_2$, then

$$\Delta g_{2F} = \Delta g_{1F} + 0.1119 (H_2 - H_1) \tag{D}$$

which is equation (4). Following equation (1), one has

$$\Delta g_{1B} = \Delta g_{1F} - 0.1119 H_1 \tag{E}$$

which inserted into equation (D) gives

$$\Delta g_{2F} = \Delta g_{1B} + 0.1119 H_2 \tag{F}$$

which is equation (3).

Table 1. Gravity interpolation in Canada using free-air and Bouguer anomaly.

Block number	Regions	S-W corner of 1° block		# of points estimated	Radius (km)	RMS difference (mgal) free-air	Bouguer
1	Columbia Mt, B.C.	50°	241°	116	10	58.9	14.4
2	Okanagan, B.C.	50°	241°	84	12	76.2	8.6
3	Edmonton, Alta.	52°	245°	119	10	3.7	2.6
4	North Ontario	53°	272°	41	18	5.2	5.3
5	North Quebec	53°	290°	45	16	3.6	2.5
6	Atlantic Ocean	42°	295°	548	4	2.2	14.4
7	Atlantic Ocean	48°	308°	549 out of 2245	2	1.1	4.0

Table 2. Gravity interpolation in Canada using the arithmetic mean, the weighted mean based on distance, the weighted mean based on distance and standard deviation, and the least-squares collocation solution.

Block number	Arithmetic mean	Weighted mean distance	dist. & std. dev.	Least-squares collocation
1	14.4	14.2	14.2	14.7
2	8.6	8.6	8.6	9.2
3	2.6	2.6	3.2	2.7
4	5.3	5.2	5.2	4.9
5	2.5	2.5	2.5	2.5
6	2.2	1.9	1.7	1.8
7	1.1	1.0	1.0	1.0

Table 3. Comparison of gravity interpolation in North American Rockies using the Bouguer coefficient (0.1119) and other elevation correlation coefficients (b and h_o).

Block number	S-W corner of block		# of points estimated	b	h_o	RMS difference (mgal) using $-b(h-h_o)$	using -0.1119h
1	30°	241°	489	.094	701	6.2	6.0
2	30°	246°	655	.044	1247	7.0	4.6
3	30°	251°	419	.069	1589	4.9	3.5
4	35°	241°	677	.071	1501	6.6	5.7
5	35°	246°	573	.084	1686	8.9	3.6
6	35°	251°	638	.062	2044	7.9	4.4
7	40°	236°	288	.062	908	8.9	7.6
8	40°	241°	225	.064	1389	7.9	5.9
9	40°	246°	891	.090	1763	5.0	5.6
10	40°	251°	271	.062	1507	4.9	4.7
11	45°	236°	330	.073	878	12.4	6.5
12	45°	241°	392	.066	1193	15.5	6.8
13	45°	246°	371	.078	1248	5.6	7.3
14	50°	231°	336	.062	877	24.4	12.7
15	50°	236°	420	.085	1326	10.0	6.4

Table 4. Statistics of the gridded gravity set.

Region :	Latitude Longitude	N35° to N90° W40° to W150°
Number of gravity observations :		572,476
Number of 5' by 5' cells :		≈456,000
# of cells with no observation :		≈288,000
# of cells with 1 obs.:		84,390
# of cells with 2-10 obs.:		71,229
# of cells with 10 and more obs.:		11,921
# of cells interpolated using 5 nearest obs. within 30 km :		416,538
# of cells interpolated using 1, 2, 3 or 4 nearest obs. within 30 km :		22,552
Comparison between the new and the old gridded gravity sets:		
# of 5', 10' 15' and 20' cells :		167,540
Maximum Difference (mgal):		174
Minimum Difference (mgal):		-387
Mean Difference (mgal):		0.8
RMS Difference (mgal):		16.1
% Difference smaller than | 5 | mgal :		69 %
% Difference smaller than | 10 | mgal :		93 %
Comparison between the collocation and mean solutions, both using a maximum of 5 nearest observations within 30 km :		
# of cells interpolated :		451,011
Maximum Difference (mgal):		94
Minimum Difference (mgal):		-36
Mean Difference (mgal):		-0.1
RMS Difference (mgal):		3.1
% Difference larger than | 5 | mgal :		7 %
% Difference larger than |10 | mgal :		1.8 %

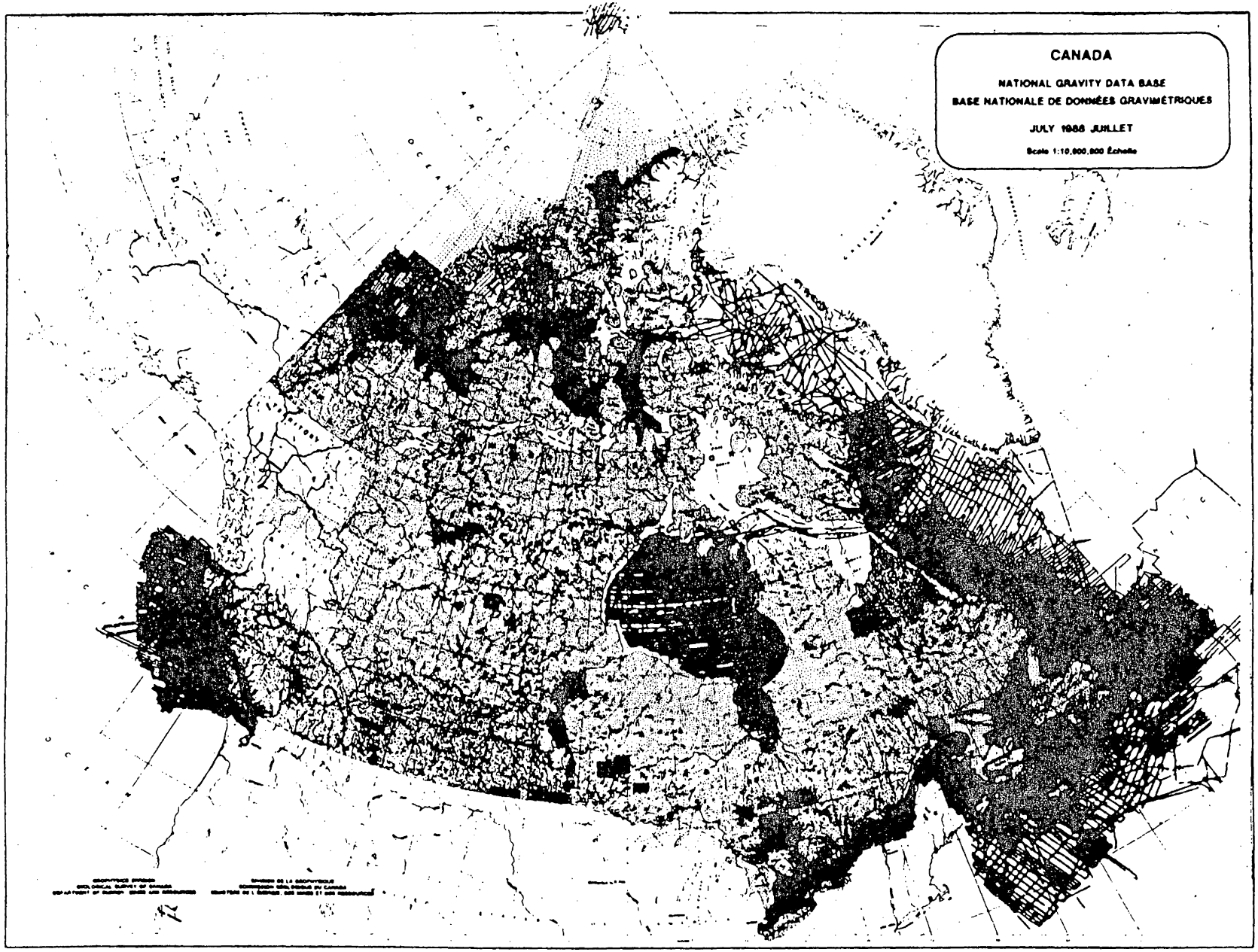


Figure 1. Surface gravity observations distribution in Canada used to create the gravity grid.

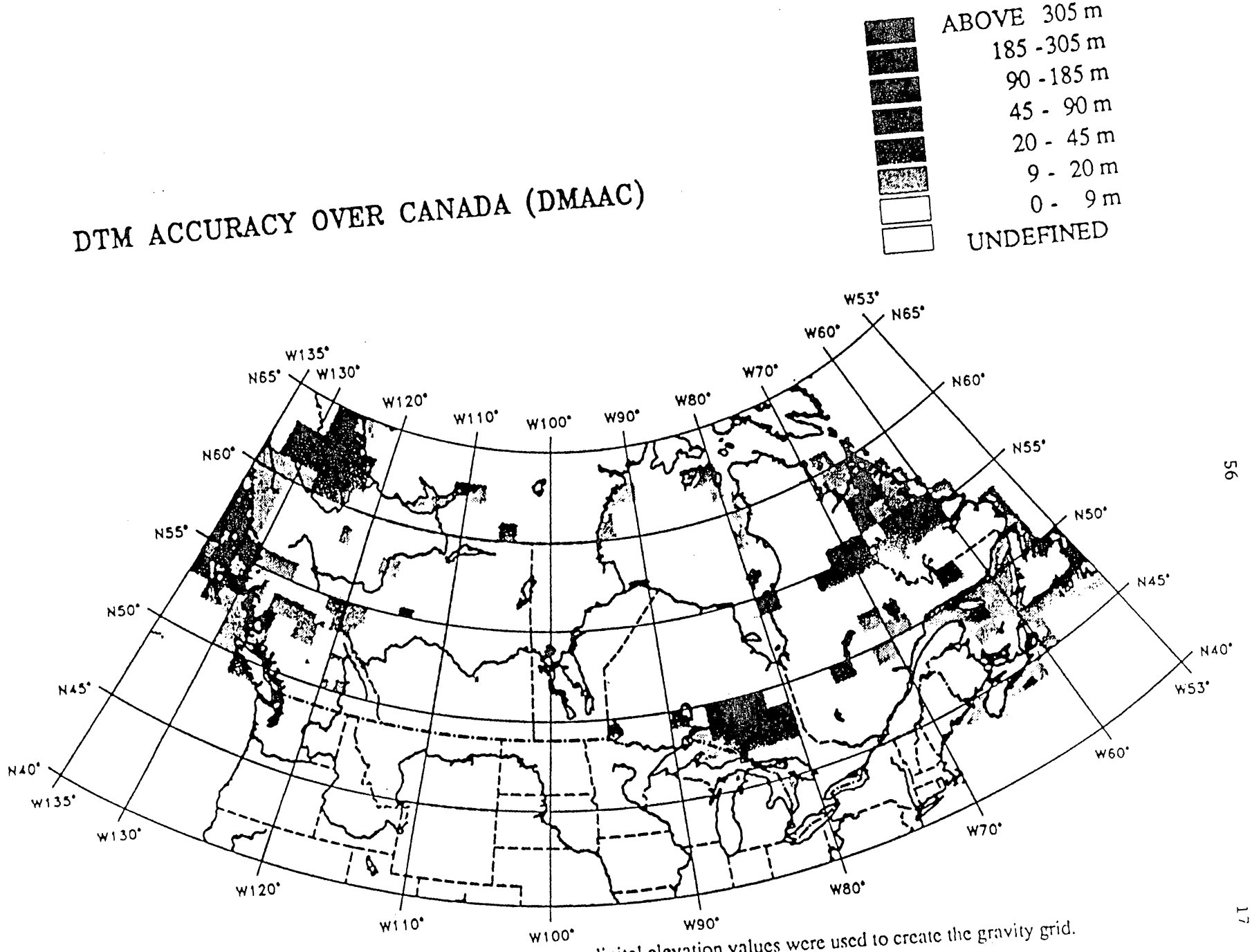


Figure 2. Regions in Canada where less accurate digital elevation values were used to create the gravity grid.

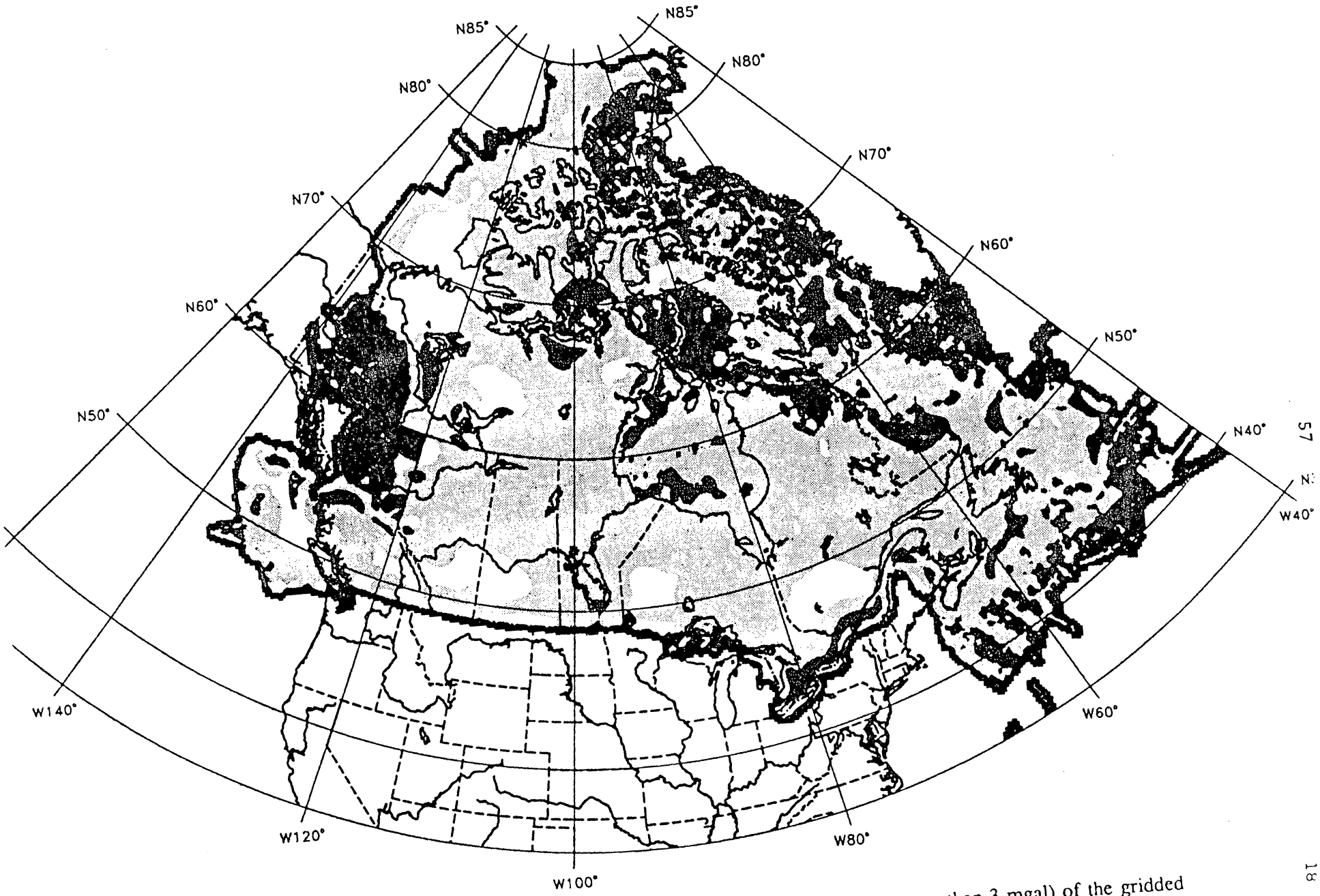


Figure 3. Preliminary gridded gravity set coverage, and standard errors (larger than 3 mgal) of the gridded gravity set estimated by least-squares collocation.

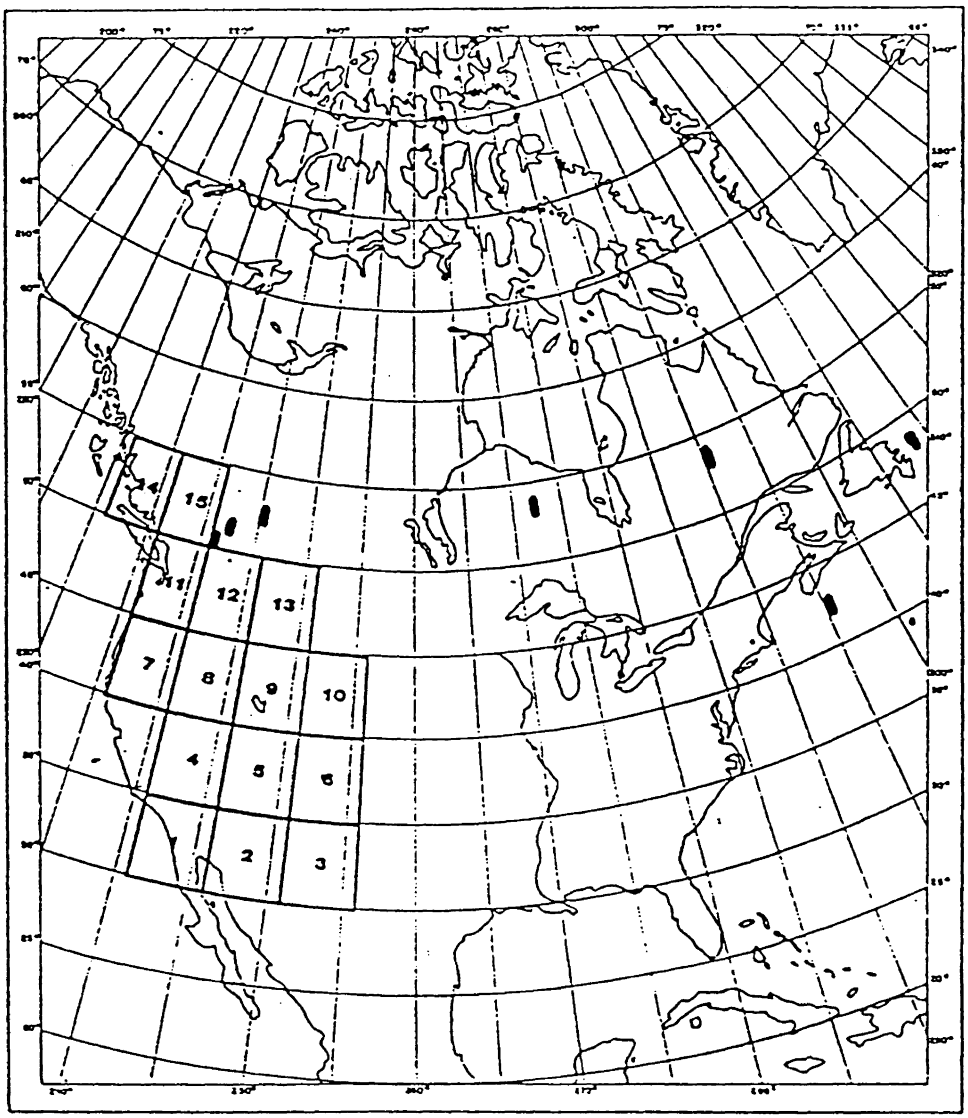


Figure 4. Fifteen 5 by 5 arcdegree regions in North American Rockies where interpolation techniques were tested with the Bouguer plate coefficient and other elevation correlation coefficients; and, seven 1 by 1 arcdegree regions in Canada where interpolation techniques were tested using Bouguer and free-air gravity anomalies.

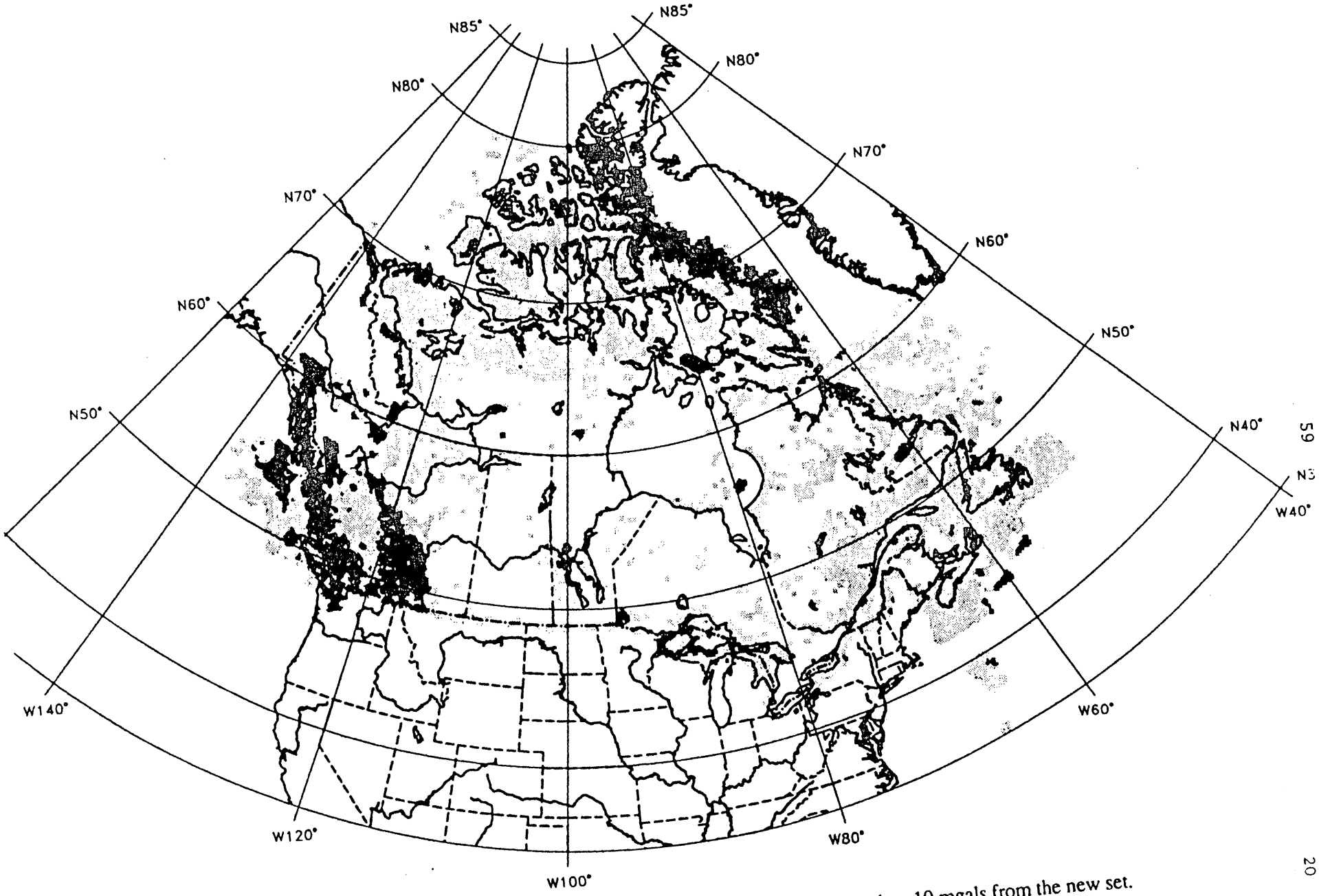


Figure 5. Areas where the old gridded free-air gravity set differs by more than 10 mgals from the new set.

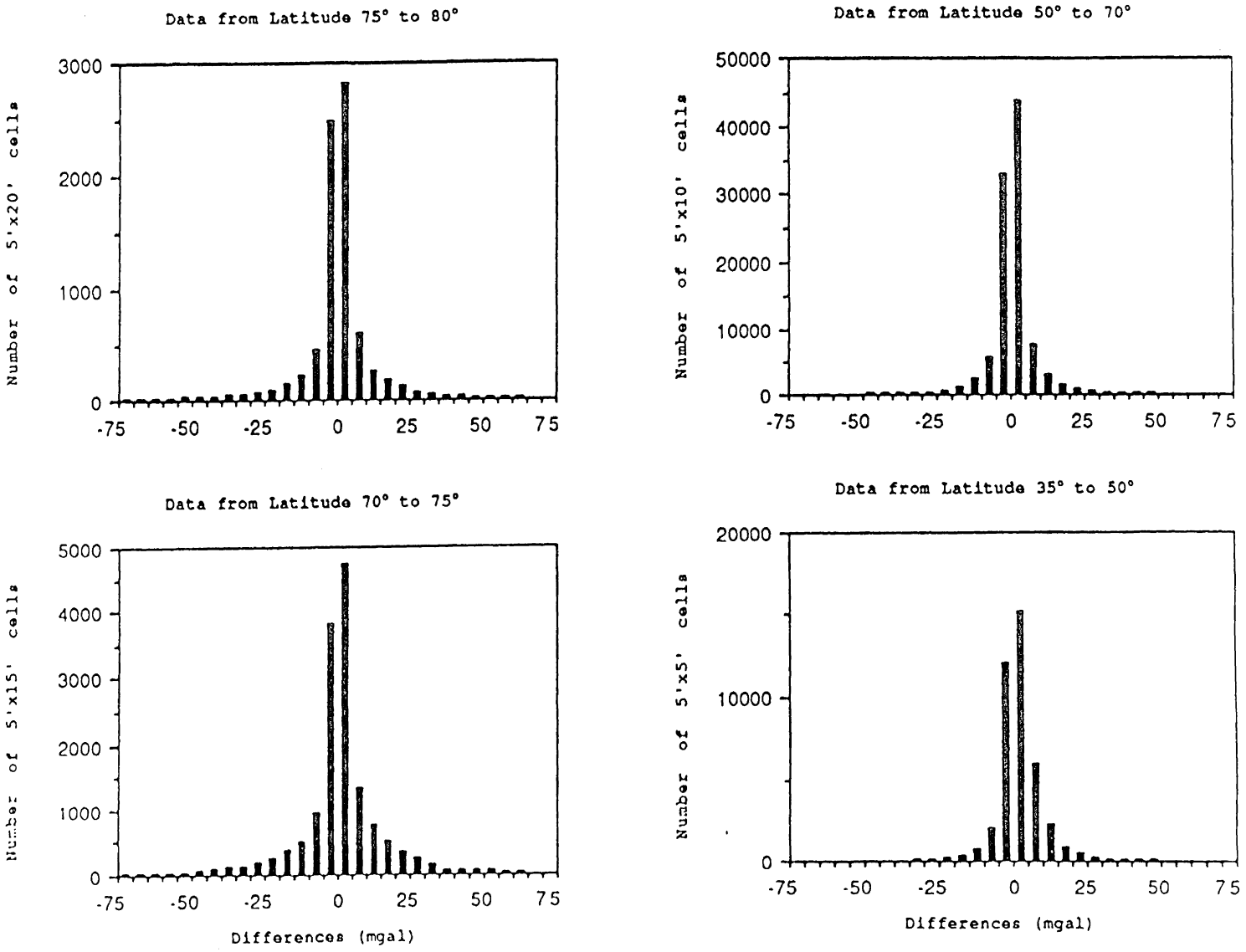


Figure 6. 5 mgal interval histograms of the new minus the old gridded gravity sets.

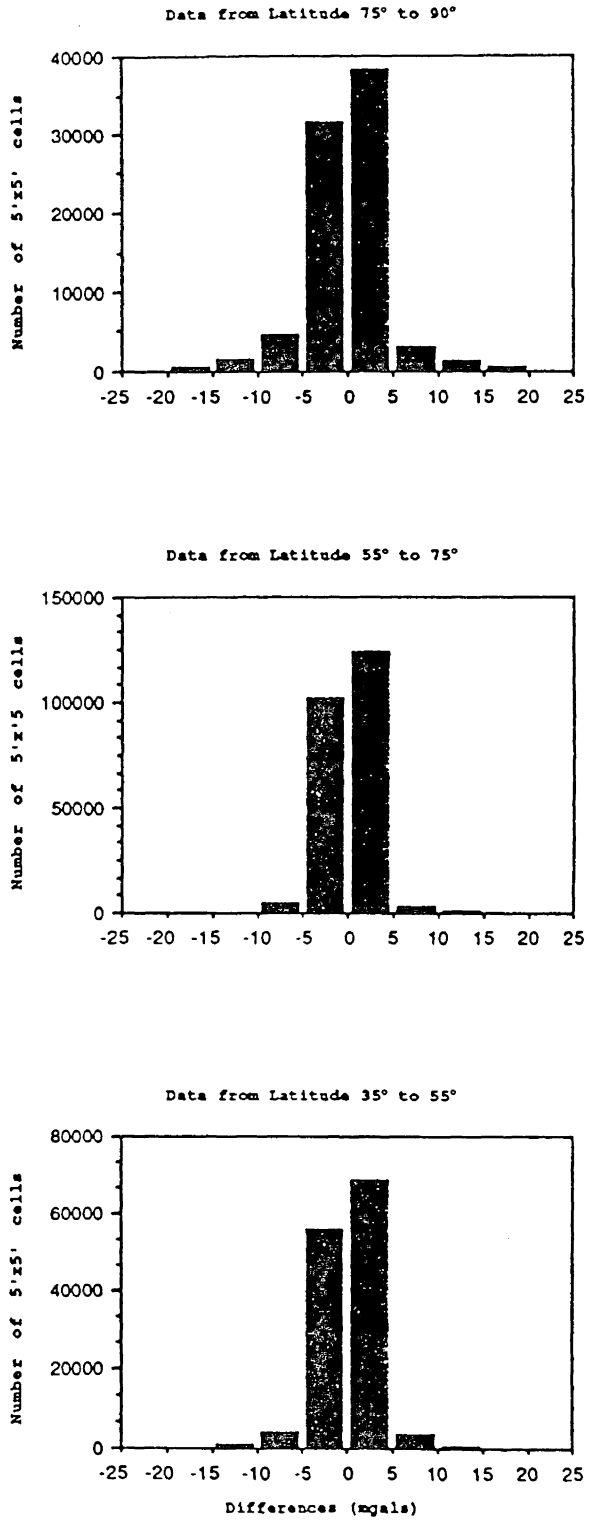


Figure 7. 5 mgal interval histograms of the least-squares collocation minus the arithmetic mean solution.

APPENDIX B

Excerpt from a letter SM2800-337703 of 25 October 1989

Dr. A. Mainville
Geodetic Survey Division

EXCERPT FROM A LETTER SM2800-337703 OF 25 OCTOBER 1989 BY
DR. A. MAINVILLE, GEODETIC SURVEY DIVISION, EMR, OTTAWA.

In support to the geoid contract, ~~please find enclosed~~ a magnetic tape containing the 5' X 5' and the 1° X 1° grids of gravity anomalies covering Canada. The contents and format are described below. There are 4 files stored as matrices. The edges of the 2 maps enclosed show the coverage of the grids. The coverage of the 2 grids is basically from latitudes N90° to N35° and longitudes E210° to E320°.

The gridded anomalies refer to GRS80, and the atmospheric gravity correction (.87 to .54 mgal) has already been applied.

The first grid is stored in 3 files, the second grid in one file. The first point of each file is the centre of the most north-west cell . You then read from west to east for each latitude down to the south limit of the region (e.g. the first point in file #1 is lat. 55° 2.5' and long. 210° 2.5').

FILE #1 : Spacing: 5' by 5', Latitude: 55° to 35°, Longitude: E210° to E320°
Size: 33,413 blocs or 17 mbytes (240 rows x 1320 columns)

FILE #2 : Spacing: 5' by 5', Latitude: 75° to 55°, Longitude: E210° to E320°
Size: 33,413 blocs or 17 mbytes (240 x 1320)

FILE #3 : Spacing: 5' by 5', Latitude: 90° to 75°, Longitude: E210° to E320°
Size: 25,060 blocs or 13 mbytes (180 x 1320)

FILE #4 : Spacing: 1° by 1°, Latitude: 90° to 35°, Longitude: E210° to E320°
Size: 733 blocs or 0.4 mbytes (55 x 110)

Canada

The 5' by 5' Free Air Anomaly Grid

FILES 1 to 3: RECORD = 53 / BLOCK = 5300 / DENSITY = 6250, ASCII.

Column # 1 : Free air gravity anomaly Δg_{nc} (mgals)
Column # 2 : Mean variance of observed anomalies (mgals)
Column # 3 : Mean height of observed anomalies (metres)
Column # 4 : Height from the Digital Elevation Model, DEM (metres)
Column # 5 : Free air gravity anomaly corrected for the DEM Δg_c (mgals)
Column # 6 : Mean distance from the centre of the cell to the observations (km)
Column # 7 : Number of observations used to estimate the free air anomaly
Column # 8 : Number of quadrants where the observed anomalies are distributed

FORMAT : F8.2, F8.2, F8.1, F8.1, F8.2, F6.2, I5, I2

N.B.: You must use column # 5,4 and 2 (Do not use 1 and 3). Columns 1, 3, 6, 7 and 8 are for quality and statistical analysis only.

The grid was computed in the following way. If a cell has 5 or more observations, the arithmetic mean of these observations, once corrected to the DEM mean height of the cell, gives the representative value of the cell (column #5). In the other cases, we were averaging the 5 closest observations to the centre of the cell inside a 30 km radius. Again, the estimated mean free-air had to be corrected for the difference between the mean height of the observations and the DEM.

$$\overline{\Delta g_c} = \frac{\sum_{i=1}^{n} \Delta g_i}{n} + 0.1119\,(H_{DEM} - \bar{H})$$

where

$$\bar{H} = \frac{\sum_{i=1}^{n} H_i}{n}$$

N.B.: We found several blunders in the DEM which was used in the previous grid that you used in 1986. We corrected a large number of those errors but some still exist. The file has been considerably improved, to our satisfaction, until a better file is made available from some institution.

The 5' by 5' free-air anomaly grid contains 871,200 cells (660 x 1320). Of those cells 651,766 (74.8%) were computed, the remaining cells are undefined (empty) and are indicated by a record 9999. 9999. 9999. H_{DEM} 9999. 9999. 0 0. One way to reject the edges is to use the distance in column #6. Caution must be taken to use the elevation values (e.g. when computing topographic effect or indirect effect). The H_{DEM} value in column #4 is negative over the oceans (which is the depth of the ocean), however the elevation of the free air in column #5 should in this case be column #3 instead of column #4. Column #3 will often give an elevation zero over ocean but will give a value different than zero near the coast which must be used.

The 1° by 1° Free Air Anomaly Grid

FILE 4 : RECORD = 61 / BLOCK = 6100 / DENSITY = 6250, ASCII.

Column # 1 : Mean free air gravity anomaly Δg_{nc} (1° by 1°, mgals)
Column # 2 : Mean variance of the 5' by 5' (mgals)
Column # 3 : Mean height of observed anomalies (metres)
Column # 4 : Mean height from the DEM (metres)
Column # 5 : Free air gravity anomaly corrected for the DEM Δg_c (mgals)
Column # 6 : Mean distance from the center of the cell to the observations (km)
Column # 7 : Number of 5' by 5' used to estimate free air anomalies
Column # 8 : Number of observed anomalies

FORMAT F8.2, F8.2, F9.2, F9.2, F8.2, F8.2, I4, I7

N.B.: Again, you must use column # 5,4 and 2 (Do not use 1 and 3). Columns 1, 3, 6, 7 and 8 are for quality and statistical analysis only.

The 1° by 1° cells were computed from the 5' by 5' grid. We computed the arithmetic mean on the sphere of the 5' cells inside the 1° cell. The weight is equal to the cosine of the latitude of the 5' cell.

$$\overline{\Delta g}_{C1} = \frac{\sum_{i=1}^{n} \overline{\Delta g}_{C5} \cos \phi}{\sum_{i=1}^{n} \cos \phi}$$

APPENDIX C

"Terrain effects on geoid undulation computations."

Yan Ming Wang
Richard H. Rapp

Terrain Effects On Geoid Undulation Computations

Yan Ming Wang

and

Richard H. Rapp

Department of Geodetic Science and Surveying
The Ohio State University
Columbus, Ohio 43210

January 1989
(revised May 1989)

Abstract

This paper examines two methods for considering the effect of terrain in the calculation of geoid undulations. One method is that associated with the usual terrain correction as documented by Moritz (1968). The second method, developed by Vanicek and Kleusberg (1987), uses a correction term depending on topography, but in a much different way from Moritz. A theoretical discussion shows that the Moritz interpretation results in a free-air anomaly on the surface of a co-geoid defined by the Helmert second condensation procedure. The Vanicek-Kleusberg reduction implies a free-air anomaly that refers to a surface of varying elevation. We feel that the use of such an anomaly in the Stokes' equation is inappropriate.

Numerical calculations were carried out with a 30" digital terrain model in rugged 1° x 1° areas of California and Colorado. The range of the Vanicek-Kleusberg anomaly correction term (-510 to 310 mgal) was substantially larger than the range of the terrain correction (2 to 65 mgals) in California. The range of the undulation correction terms was -110 cm to -9 cm for the Vanicek-Kleusberg method and 115 cm to 174 cm for the terrain correction technique, again in California. The root mean square undulation difference, due to the different anomaly correction terms, was 182 cm in California and 139 cm in Colorado.

We also calculated the undulation differences, due just to the correction terms, for 4 lines 10 km in length, and one line 30 km in length. We found that the difference between the undulation differences implied by the anomaly correction terms reached 32 cm (11ppm) for one 30 km line, and 34 cm (34 ppm) for one 10 km line. Such differences are significant as we go to precise geoid determinations for GPS applications so that the correct procedure, which we believe is that due to Moritz, must be used.

1. Introduction

The effect of topography on the calculation of precise geoid undulations or height anomalies has been considered by many authors in the past. Recognizing that a valid solution to geoid determination would occur only if there were no masses external to the geoid, Helmert suggested that such masses be condensed as a surface layer on the geoid. This condensation implies certain corrections to the gravity anomaly and the introduction of the indirect effect which occurs due to potential changes caused by the condensation process. A discussion of some attributes of the Helmert's second method of condensation may be found in Heiskanen and Moritz (1967, p. 145), Wichiencharoen (1982), etc.

The importance of the terrain (and the terrain correction) was emphasized by Pellinen (1962) in the solution of the Molodensky boundary value problem. Moritz (1968, 1980) examined the role of the terrain to show a relationship between Helmert's condensation reduction and the Pellinen approximate solution of the Molodensky boundary-value problem.

Vanicek and Kleusberg (1987) discussed the effect of the terrain through the "squashing" of all the topographical masses onto the geoid. They calculated the "topographical attraction effect" and an indirect effect. These correction terms were used in the determination of geoid undulations in Canada.

Comparing the anomaly correction terms developed by Moritz and by Vanicek and Kleusberg one finds a difference. The anomaly correction term of Moritz (and others) is the classical terrain correction which is always positive. The correction term used by Vanicek and Kleusberg can be positive or negative. The indirect effect terms are essentially the same.

This paper considers the two different methods for considering the terrain and forms conclusions on the most appropriate technique.

2. Outline of Moritz's and Vanicek-Kleusberg's Results

It is first noted that both Moritz and Vanicek-Kleusberg use the Helmert's second method of condensation. For our discussion we will not distinguish between geoid

undulations and height anomalies which are the emphasis in the Moritz discussion. Both methods require a correction to the free-air anomaly that is used in the Stokes' equation, and the indirect effect term that is caused by the potential change due to the mass condensation. We ignore the secondary indirect effect correction that should be used to refer the anomalies to the co-geoid of the condensation reduction (Wichiencharoen, 1982).

The formula for the geoid undulation is given in Moritz (1980, eq. 48-32):

$$N_T = \frac{R}{4\pi\gamma} \iint_A (\Delta g + C)\, S(\psi)\, dA + t \qquad (1)$$

where C is the terrain correction, which in a linear, planar approximation is:

$$C_p = \frac{1}{2} G\sigma \int_{-\infty}^{\infty} \int \frac{(h-h_p)^2}{d^3}\, dx\, dy \qquad (2)$$

In (2) d is the distance between h and h_p:

$$d^2 = (x-x_p)^2 + (y-y_p)^2 \qquad (3)$$

The t term (Moritz, ibid, eq. 48-30) is neglected in Moritz but should be retained for high precision calculations. In addition we have (in (1) and (2)):

- γ -- average Earth gravity;
- R -- mean Earth radius;
- A -- unit sphere;
- $S(\psi)$ -- Stokes' function;
- G -- gravitational constant;
- σ -- uniform topographic density;
- h_p -- elevation at the point at which C is being computed.

The indirect effect for the Helmert second condensation method has been discussed by Wichiencharoen (1982). The effect can be expressed as the sum of a plane plate effect and the effect of the irregular topography. The latter term is expressed in an infinite series related to odd powers of elevation differences. Retaining only one series term the indirect effect could be written (ibid, eq. (44), (49) and (56)) as:

$$\delta N_I = -\frac{\pi G\sigma h_p^2}{\gamma} - \frac{1}{6}\frac{G\sigma}{\gamma} \int_{-\infty}^{\infty} \int \frac{h^3 - h_p^3}{d^3}\, dx\, dy \qquad (4)$$

where h_p is the elevation at the geoid undulation computation point.

Vanicek and Kleusberg (ibid) have derived a different correction term to be applied to the free-air anomaly due to the topography. This term is given (ibid, eq. (14)) as:

$$\delta g_T \approx \frac{1}{2} G\sigma \int_{-\infty}^{\infty} \int \frac{h^2 - h_p^2}{d^3}\, dx\, dy \qquad (5)$$

Comparison of the terrain correction term (eq. (2)) used by Moritz with equation (5) indicates a difference in the numerators: $(h-h_p)^2$ vs $(h^2-h_p^2)$. Vanicek and Kleusberg (ibid) also derive the indirect effect term which is identical to equation (4).

The correction to the geoid undulation caused by using (2) or (5) may be significantly different. One notes that C_p is always a positive quantity while (5) may be plus or negative. In the next section the theoretical reason for the differences is examined followed by numerical evaluations of the differences in undulation estimates to be expected.

3. A Comparison of the Moritz and Vanicek-Kleusberg Derivations

First consider the potential of the topography at point P, on the topographic surface, as shown in Figure 1. Following Vanicek and Kleusberg (ibid, eq. (11)) the topographic potential can be written as:

$$W = G\sigma \int_{-\infty}^{\infty} \int \int_0^h \frac{1}{L}\, dz\, dx\, dy \qquad (6)$$

with

$$L^2 = d^2 + (h_p - z)^2 \qquad (7)$$

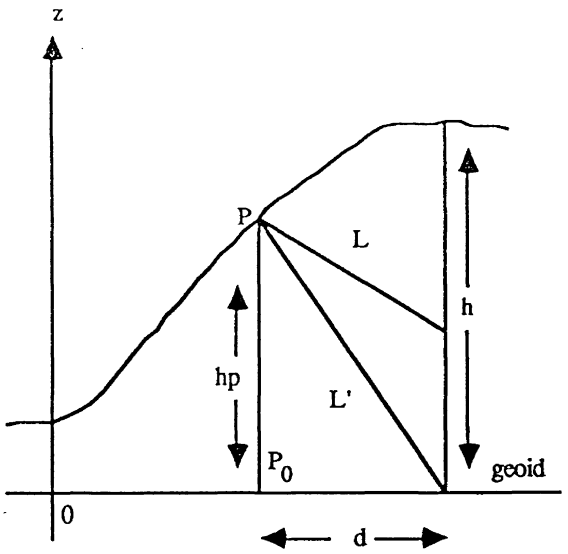


Figure 1. Location of Point P on the Terrain Relative to a Point on the Geoid.

The attraction of the topography at point P is:

$$A = -\frac{\partial W}{\partial h_p} = G\sigma \int_{-\infty}^{\infty}\int \left.\frac{1}{L}\right|_0^h dx\, dy$$

$$= G\sigma \int_{-\infty}^{\infty}\int \left[\frac{1}{\sqrt{d^2+(h-h_p)^2}} - \frac{1}{\sqrt{d^2+h_p^2}}\right] dx\, dy$$

$$= G\sigma \int_{-\infty}^{\infty}\int \left[\frac{1}{d} - \frac{1}{\sqrt{d^2+h_p^2}}\right] dx\, dy -$$

$$-\frac{1}{2} G\sigma \int_{-\infty}^{\infty}\int \frac{(h-h_p)^2}{d^3} dx\, dy + \ldots$$

$$\approx 2\pi G\sigma\, h_p - C \qquad (8)$$

Equation (8) indicates that the attraction of the topography consists of the attraction Bouguer plate and the terrain correction. This equation is the same as equation (44) in Moritz (1968) after a plane approximation.

The potential of the condensed layer at the point P is (Vanicek and Kleusberg (ibid) eq. (12)):

$$W' = G\sigma \int_{-\infty}^{\infty}\int \frac{h}{L'} dx\, dy \qquad (9)$$

with

$$L'^2 = d^2 + h_p^2 \qquad (10)$$

The attraction of the condensed layer at the point P is given by:

$$A_s = -\left.\frac{\partial W'}{\partial h_p}\right|_{h_p} = G\sigma \int_{-\infty}^{\infty}\int \frac{h\, h_p}{L'^3} dx\, dy$$

$$= 2\pi G\sigma\, h_p + G\sigma\, h_p \int_{-\infty}^{\infty}\int \frac{h-h_p}{\left(d^2+h_p^2\right)^{3/2}} dx\, dy$$

$$\approx 2\pi G\sigma\, h_p - g_1 \qquad (11)$$

where

$$g_1 = -G\sigma\, h_p \int_{-\infty}^{\infty}\int \frac{h-h_p}{d^3} dx\, dy \qquad (12)$$

The correction, to the free-air anomaly at point P, is given by:

$$\delta g_T = -A + A_s = C - g_1 \qquad (13)$$

Eq. (13) is identical to eq. (5). The C - g_1 term was given in (Wang, 1988, p. 3). Here it is emphasized that the reduced gravity anomaly $\Delta g + \delta g_T$ is at the point P, not at a point on the geoid!

If one chooses the computation point P_0 on the condensed layer as Moritz did, then (Heiskanen and Moritz, 1967, eq. (1-17a)):

$$A_s^M = -\left.\frac{\partial W'}{\partial h_p}\right|_{h_p=0} = 2\pi G\sigma h_p -$$

$$-\left[\frac{\partial}{\partial h_p}\int_{-\infty}^{\infty}\int \frac{G\sigma h}{L'} dx dy\right]_{h_p=0}$$

$$= 2\pi G\sigma h_p \qquad (14)$$

Combining eq. (8) and eq. (14), one has:

$$\delta g_M = - A + A_s^M = C \qquad (15)$$

as shown by Moritz (1968, eq. (63)). Note now the computation point P_0 is on the condensed layer and the correction term δg_M is referenced to the geoid.

This discussion is now related to the classical free-air anomaly, which should refer to the geoid, and the surface free-air anomaly of the Molodensky theory.

The classical free-air anomaly is defined as (Heiskanen and Moritz, 1967, p. 146):

$$\Delta g_G = g - \frac{\partial g}{\partial h} H - \gamma_0$$

$$\approx g - \gamma_0 - \frac{\partial \gamma}{\partial h} H \qquad (16)$$

where g is the gravity on the earth's surface, H is the orthometric height and γ_0 is the normal gravity on the ellipsoid. The gravity anomaly, Δg_G, is referenced to the geoid. Adding the terrain correction, to take into account, the topography yields the Faye anomaly (Moritz, 1980, p. 419). This anomaly is a terrain corrected anomaly referenced to the geoid and is given as:

$$\Delta g_M = \Delta g + C \qquad (17)$$

The free-air anomaly has a different meaning in Molodensky's problem. The surface free-air anomaly is defined as (Moritz, 1980, p. 293):

$$\Delta g_S = g - \left(\gamma_0 + \frac{\partial \gamma}{\partial h} H^* + \frac{1}{2} \frac{\partial^2 \gamma}{\partial h^2} H^{*2} + \ldots \right) \qquad (18)$$

where H^* is the normal height. We can write:

$$\Delta g_S \approx g - \left(\gamma_0 + \frac{\partial \gamma}{\partial h} H^* \right) \qquad (19)$$

Since $H^* \approx H$ the two anomalies (eq. (16) and (19)) are essentially the same. The δg_T value (from Vanicek and Kleusberg (ibid)) has been calculated at the topographic surface and should be added to Δg_S from (19). This anomaly is designated as:

$$\Delta g_V = \Delta g + \delta g_T \qquad (20)$$

Since this anomaly refers to a surface that is continually varying in height, it cannot be used in the Stokes' equation. If, Δg_V, is reduced to the geoid using the linear gradient correction terms (g_1), (see Wang (1988, eq. 12)) the anomaly (eq. (17)) on the geoid is obtained.

4. The Evaluation of the Effect of δg_T on the Geoid Undulation Computation

Consider next an expression for the contribution of Vanicek-Kleusberg's anomaly correction term to the geoid undulation. The Stokes' operator (Moritz, 1980, p. 393-394) in planar form, is:

$$S(f) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int \frac{f}{d} dx dy \qquad (21)$$

where f can be the gravity anomaly. One can introduce the gradient operator L (f):

$$L(f) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int \frac{f - f_p}{d^3} dx dy \qquad (22)$$

Applying (22) to (21) one has (Moritz, ibid):

$$L\{S(f)\} = S\{L(f)\} = -f \qquad (23)$$

Applying (22) to the Vanicek-Kleusberg correction term, eq. (5), yields:

$$\delta g_T = \pi G \sigma L(h^2) \qquad (24)$$

The contribution of δg_T to the geoid undulation, as computed by Vanicek and Kleusberg, is given by applying (21) to (24) and using (23):

$$\delta N_V = S(\delta g_T) \frac{1}{\gamma} = -\frac{1}{\gamma} \pi G \sigma h_p^2 \qquad (25)$$

Comparing (25) with the formula of the indirect effect, one finds that δN_V is the same order of magnitude as the indirect effect.

In this paper the 2D Fourier transform of h (for example) is defined as follows:

$$G(u, v) = F\{h\} = \int_{-\infty}^{\infty} \int e^{-j2\pi(ux+vy)} h(x, y)\, dx\, dy \qquad (26)$$

where u, v are frequency variables. The inverse 2D Fourier transform would then be:

$$h(x, y) = F^{-1}\{G(u, v)\}$$

$$= \int_{-\infty}^{\infty}\int e^{j2\pi(ux+vy)} G(u, v)\, du\, dv \qquad (27)$$

Applying the Fourier transformation to eq. (5) yields:

$$\delta g_T = F^{-1}\left\{-2\pi^2 G\sigma\omega\, F\left\{h^2\right\}\right\} \qquad (28)$$

where $\omega = (u^2 + v^2)^{1/2}$. In practice the data are given as discrete point values or mean block values. Therefore, the discrete Fourier transformation can be used to evaluate eq. (5). One first constructs a new kernel function:

$$\frac{1}{d_0^3} = \begin{cases} 0 & d=0 \\ \dfrac{1}{d^3} & d\neq 0 \end{cases} \qquad (29)$$

Then eq. (5) can be written as:

$$\delta g_T = \frac{1}{2} G\sigma \left[DF^{-1}\left\{ DF\left\{\frac{1}{d_0^3}\right\} DF\left\{h^2\right\}\right\} - h_p^2\, DF^{-1}\left\{ DF\left\{\frac{1}{d_0^3}\right\} DF\{1\}\right\}\right] \Delta x\, \Delta y \qquad (30)$$

where DF, DF^{-1} denote the discrete Fourier transformation and its inverse; Δx, Δy are the grid interval. The value of the kernel function d_0^{-3} at the origin has no effect on the computation value, because the center block (point) of the integral (5) has no contribution.

For the numerical calculation of the terrain correction, eq. (2), a similar procedure was used.

Applying the planar Stokes' formula to the Vanicek-Kleusberg's correction, one has:

$$\delta N = \frac{1}{2\pi\gamma}\int_{-\infty}^{\infty}\int \frac{\delta g_T}{d}\, dx\, dy \qquad (31)$$

The plane approximation is not accurate enough for the geoid determination by using the full gravity anomaly, but it is sufficient for the computation of a geoid undulation correction.

Applying the discrete Fourier transformation to (31), one obtains:

$$\delta N = \frac{1}{2\pi\gamma} DF^{-1}\left\{ DF\{\delta g\}\, DF\left\{\frac{1}{d}\right\}\right\} \Delta x\, \Delta y \qquad (32)$$

In the computation of (32) there is a singularity at the origin. The method to overcome this problem is described in Wang (1988, p. 13).

Other methods to eliminate the singularity problem are discussed in (Heiskanen and Moritz, (1967) and Schwarz, et al., (1988)).

5. Numerical Calculations and Computations

In order to study the magnitude and differences of the two correction terms, two test areas, 1° x 1° in size, were chosen. Area A ($36° \le \phi \le 37°$, $241° \le \lambda \le 242°$) was in California where the cell contained topography that ranged from 150 m to 4270 m. Area B ($38°\le\phi\le39°$, $253°\le\lambda\le254°$) was in a high rugged area of central Colorado. The elevations were given at 30" interval. In order to avoid leakage, elevation data out to 1° from the test cell borders was used.

Table 1 shows information related to the anomaly correction terms in the two test areas.

Table 1. Magnitudes of the Anomaly Correction Terms (mgals)

	Area A		Area B	
	C	δg_T	C	δg_T
minimum	2	-510	1	-479
maximum	65	310	44	251
mean	12	-2	4	-1
RMS	13	68	6	67
Std. Dev.	±7	68	4	67

From Table 1 it is apparent that Area A could be considered a rougher area through a comparison of the magnitudes of the terrain correction (C). Although the mean value of δg_T is smaller than C, the magnitudes of δg_T are considerably larger than the terrain correction.

The anomaly correction terms were then used in equation (32) to obtain undulation correction terms. These values are shown in Table 2.

Table 2. Undulation Correction Terms (cm)

	Area A		Area B	
	C	δg_T	C	δg_T
minimum	115	-100	78	-100
maximum	174	-9	103	-31
mean	150	-24	85	-53
RMS	151	26	85	54
Std. Dev.	±17	±10	±6	±12

From Table 2 it is seen that the terrain correction has a dominantly systematic correction term. This was also seen by Rapp and Wichiencharoen (1984). The mean undulation corrections are significantly (greater than 1 m) different in both test areas.

The undulation correction terms were next differenced at each of the grid points in the 1° cell. The statistics on these differences are shown in Table 3.

Table 3. Information on Undulation Correction Term Differences (δN_T - δN_V) (cm)

	Area A	Area B
Minimum	121	114
Maximum	265	194
Mean	180	138
RMS	182	139
Std. Dev.	± 28	16

Table 3 shows that the difference between the undulation correction terms can reach 2.6 m in Area A and 1.9 m in Area B. Since most of the difference is a systematic one the standard deviation of the differences is considerably smaller than the root mean square difference.

Of increasing importance for GPS application is the calculation of geoid undulation differences (e.g. Kadir and Rapp, 1988). Consider the role of the anomaly and undulation correction terms in calculating such differences. To do this a base point is picked in the center of the test area and then 4 points at 10 km from the base point in 4 different directions (point 2 was south west, 3 north west, 4 north east and 5 south east). Point 6 was taken 30 km south of the first point. The undulation correction terms from each method were differenced with the results given in each area in Table 4.

Table 4. Undulation Difference for Selected Lines Due to Anomaly Correction Terms (cm)

	Area A		Area B	
Line	C	δg_T	C	δg_T
1-2	-3.8	-11.1	1.0	-6.8
1-3	3.1	15.1	-1.9	4.3
1-4	2.1	12.7	-4.1	30.0
1-5	-1.4	-13.4	0.6	-4.5
1-6	15.1	-16.5	0.4	1.5

From Table 4 it is clear that the correction term effects are generally larger when δg_T is used. In area A the largest difference between the C and δg_T corrections is for the 30 km line (1-6) reaching 32 cm (11 ppm). In area B the largest discrepancy is 34 cm (34 ppm) for line 1-4. Each of these differences is substantially higher than the accuracy (1 ppm) needed. It is therefore important to choose the proper correction term. Note that the results given in Table 4 are sensitive to the points selected. Other test point selections could yield different results. However the results shown in Table 4 indicate the differences between the C and δg_T method, for undulation difference computations, is not negligible.

A final calculation was made to assess the accuracy of the approximation to δN_V given by equation (25). This was done by comparing the result from (25) with the value given from the equation, (32). The results are given in Table 5.

Table 5. Difference of Simple (eq. (25)) and Complex (eq. (32)) Calculation for δN_V (cm)

	Area A	Area B
Minimum	-6	-6
Maximum	6	3
Mean	-1	-2
RMS	2.2	2.7
Std. Dev.	±2.1	±1.2

Recalling from Table 2 that the root mean square undulation correction (RMS) for area A is 26 cm, and area B, 54 cm, it is apparent that the approximate formula works well in calculating δN_V.

5. Summary and Conclusion

Vanicek and Kleusberg (ibid) have used an anomaly correction that differs from the terrain correction term proposed by Moritz (1968). This paper examines the theoretical background for each technique and argues that the Moritz procedure of reduction is the more correct method as it reduces the anomaly to a constant elevation surface, the geoid; the Vanicek-Kleusberg correction leaves the anomaly referenced to a surface of varying height. Such anomalies should not be used in the Stokes' equation. The key point giving rise to the anomaly correction term difference is the selection of the elevation point where the attraction of the condensed topography is evaluated.

Numerical evaluations were carried out in two test areas; 1°x 1° cells in California (area A) and in Colorado (area B), where 30" elevation data was available. The computations showed that the magnitude of δg_T was considerable larger than the magnitude of C (the terrain correction).

The undulation correction terms showed disagreement at the 1 m level. For example, the root mean square undulation corrections in area A were 151 cm (from C) and 26 cm (from δg_T). A great part of this difference is attributed to the systematic contributions of C (150 cm).

Undulation differences associated with the anomaly correction terms were also calculated. Such differences are important when undulation differences are needed in the application of GPS derived ellipsoidal height differences. Four lines of 10 km and one line of 30 km in length were considered. The differences between the correction term effects reached 32 cm (30 km line) in area A and 34 cm (10 km line) in area B. These large differences indicate that it is important to determine the proper method to calculate the anomaly correction terms. This paper argues that the proper method is the one derived by Moritz (1968) and that the Vanicek-Kleusberg method is incorrect. The use of the Vanicek-Kleusberg method in mountainous areas will yield undulation differences that can be substantially wrong in comparison with GPS elevation difference accuracies.

Finally, note the importance of applying the indirect effect term (eq. (4)) when precise undulations (or undulation differences) are being computed. Although not of significance for the computations of this paper since the indirect effect is the same for both methods studied, the magnitude of the correction can be significant.

Acknowledgment

The research described in this paper has been supported by the Air Force Geophysics Laboratory, Contract F19628-86-K-0016, The Ohio State University Research Foundation Project 718188.

References

Forsberg R (1984) A Study of Terrain Reductions, Density Anomalies and Geophysical Inversion Methods in Gravity Field Modelling. Report 355, Dept. of Geodetic Science, The Ohio State University

Heiskanen WH, Moritz H (1967) Physical Geodesy. W. H. Freeman and Co., San Francisco

Kadir M, Rapp R (1988) A Preliminary Geoid for the State of Tennessee, Surveying and Mapping. Vol. 48, No. 4, December

Moritz H (1968) On the Use of the Terrain Correction in Solving Molodensky's Problem. Report No. 108, Department of Geodetic Science, The Ohio State University

Moritz H (1980) Advanced Physical Geodesy. Herbert Wichmann Verlag, Karlsruhe FRG

Pellinen LP (1962) Accounting for topography in the calculation of quasigeoidal heights and plumbline deflections from gravity anomalies. Bulletin Geodesique 63:57-65

Rapp RH, Wichiencharoen C (1984) A Comparison of Satellite Doppler and Gravimetric Geoid Undulations Considering Terrain-Corrected Gravity Data. J. Geophys. Res. 89:1105-1111

Schwarz KP, Sideris MG, Forsberg R (1988) The Use of FFT Techniques in Physical Geodesy. submitted for publication

Sideris MG (1985) A Fast Fourier Transform Method for Computing Terrain Corrections. manuscripta geodaetica Vol. 10, No. 1

Vanicek P, Kleusberg A (1987) The Canadian geoid - Stokesian approach. manuscripta geodaetica, Vol. 12, 2:86-98

Wang Y (1988) Downward Continuation of the Free-Air Gravity Anomalies to the Ellipsoid Using the Gradient Solution, Poisson's Integral and Terrain Correction - Numerical Comparison and the Computations. Report No. 393, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus

Wichiencharoen C (1982) The Indirect Effects on the Computation of Geoid Undulations. Report No. 336, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus

APPENDIX D

"Reformulation of Stokes's theory for higher than second-degree reference field and modification of integration kernels."

Lars E. Sjöberg
Petr Vaníček

Reformulation of Stokes's Theory for Higher than Second-Degree Reference Field and Modification of Integration Kernels

Lars E. Sjöberg

Department of Geodesy, The Royal Institute of Technology, S-10044 Stockholm, Sweden

Petr Vaníček

Department of Surveying Engineering, University of New Brunswick, P.O. Box 4400, Fredericton, N.B., Canada

Short Title: Reformulation of Stokes's Theory

Submitted to *Journal of Geophysical Research*, January 1989

Re-submitted, June 1989
Re-submitted, February 1990

Abstract

An argument is put forward in favour of using a model gravity field of an order higher than 2 as a reference in gravity field studies. Stokes's approach to the evaluation of the geoid from gravity anomalies is then generalized to be applicable to a higher than second-order reference spheroid. The effects of truncating Stokes's integration and of modifying the integration kernels are investigated in the context of the generalized approach. Equations are derived for changing the geoid evaluated by means of the generalized approach with modified kernels in response to change of the reference field. As an example, the effect of exchanging GEM9 for T1 reference fields on the UNB Dec.'86 geoid is shown.

Introduction

After the famous Newton-Cassini argument about the basic shape of the earth (oblate vs. prolate) was settled by the French Academy's mid-eighteenth century expeditions, the oblate biaxial ellipsoid became the reference surface in geodesy. There is, of course, much to be said about the appropriateness of such a simple surface for the purpose of positioning. We wish to argue here, however, that for gravity field studies a higher than second-order surface — a spheroid of degree and order M — may now be used with considerable advantage.

This is not, of course, a new idea. Earth Gravity Models (EGM) expressed in terms of a series of zonal spherical harmonics have become part of the definition of geodetic reference systems — see, for instance, the definition of GRS 80 [Moritz, 1980], or WGS 1984 [Smith, 1988]. Low-degree and -order fields, mostly determined from satellite orbits, have been used by many researchers, e.g., Nagy and Paul [1973], in the past few decades, mostly without acknowledging their reference role explicitly. We wish to point out that there is a definite gain in insight and thus a didactic advantage in the explicit acknowledgment of the reference field role played in effect by the EGMs.

To be sure, there is, and always will be, an error in any EGM to be adopted as a reference field. But this situation is no different from that we now face with the second-degree zonal (Somigliana-Pizzeti) field, a situation we have been living with for at least two centuries. There are ways of dealing with this problem and we shall try to point them out as appropriate.

Throughout this paper, we shall be expressing the gravity field (the geoid in particular) or its components interchangeably in terms of convolution integrals of Green's type and in terms of finite or infinite series of spherical harmonics. We shall speak of those respectively as integral or spectral representations, as has recently become the custom in geodesy.

In addition, it should be emphasized that the residual geoid obtained using a higher-degree and -order reference field may have particular advantages for regional geophysical interpretation, see,

e.g., Sjöberg [1984c] and Christou et al. [1989]. This fact alone should motivate a closer study of Stokes's formula for a higher-degree and -order reference field.

Reformulation of Stokes's Convolution Integral

Let us start by considering Stokes's original formula for geoid undulation N referred to the geodetic reference ellipsoid [Heiskanen and Moritz, 1967]:

$$N \doteq \kappa \iint_{\mathcal{E}} S(\psi)\, \Delta g\, d\mathcal{E} \quad , \tag{1}$$

where $\kappa = R/(4\pi\gamma)$, R and γ are the mean surface radius and mean gravity of the earth, Δg is the gravity anomaly defined as

$$\Delta g = g_{\mathcal{G}} - \gamma_{\mathcal{E}} = g - \gamma_0 \tag{2}$$

($g_{\mathcal{G}} = g$ is the actual gravity on the geoid, and $\gamma_{\mathcal{E}} = \gamma_0$ is the normal gravity on the ellipsoid, both along the same normal to the ellipsoid), and the integration kernel S is called the Stokes function. The integration is carried over the whole reference ellipsoid $\mathcal{E}$ or, equivalently, over a unit sphere. We use the approximate equality sign because the expression is correct only to the order of e^2 (the square of eccentricity of the reference ellipsoid) [Vaníček and Krakiwsky, 1986]; this is known as the "spherical approximation." Stokes's function is usually written as a series of Legendre polynomials P_n,

$$S(\psi) = \sum_{n=2}^{\infty} \frac{2n+1}{n-1} P_n(\cos\psi) \quad , \tag{3}$$

where ψ is the geocentric angle between the point of interest and the dummy point in the integration.

It has been shown by many authors, e.g., Lachapelle [1977], or Vaníček and Krakiwsky [1986], that if a spheroid of degree M given by the first M degree spherical harmonic components (N_i) of the geoid

$$(N)_M = \sum_{i=2}^{M} N_i \quad , \tag{4}$$

where N_i contains (2i+1) spherical harmonics of appropriate orders, is taken as a reference surface, then the geoidal height N^M above that spheroid is given by the following equation, correct to the order of the eccentricity of the reference ellipsoid squared (e^2). We have:

$$N^M \doteq \kappa \iint_{\mathcal{E}} S^M(\psi)\, \Delta g^M\, d\mathcal{E} \ , \tag{5}$$

where

$$\begin{aligned} S^M(\psi) &= \sum_{n=M+1}^{\infty} \frac{2n+1}{n-1} P_n(\cos\psi) \\ &= S(\psi) - \sum_{n=2}^{M} \frac{2n+1}{n-1} P_n(\cos\psi) \ , \end{aligned} \tag{6}$$

$$\Delta g^M = g - \gamma^M \ ; \ \gamma^M = \sum_{n=2}^{M} \gamma_n \ , \tag{7}$$

and γ_n is defined below. In the sequel, we shall call the spheroid of M-th degree and order simply "spheroid of M-th degree."

The following two notes are required. First, writing eqn. (5), we have neglected a term

$$\varepsilon_1 = \kappa \iint_{\mathcal{E}} (S - S^M)\, \Delta g^M\, d\mathcal{E} \ , \tag{8}$$

which equals to zero when $\forall n \leq M$: $\Delta g_n^M = g_n - \gamma_n = 0$, i.e., in the absence of errors in the low-degree harmonic components g_n and γ_n of the observed gravity g and the model gravity γ. This condition will clearly not be satisfied generally — as the analogous condition for g_o and γ_o is not satisfied in the original Stokes formulation — and we will bring the neglected term into the picture later when we start discussing observational errors. Second, we note the non-standard use of the symbols γ_n and γ^M. By γ_n we denote the spherical harmonic components of gravity generated by the EGM, i.e., the "model gravity" on the spheroid. The model gravity γ^M plays exactly the same role here as the normal gravity γ_o does in the original Stokes development. (It should be noted that γ_o does not figure in any of the new expressions as it should not. In practical computations, however, Δg around the world would be available rather than g in which case γ_o obviously has to be taken into account. But this step is irrelevant to the theory presented here.) The only place the reference ellipsoid is implied is in eqn. (4); by eqn. (4) the reference spheroid is presumed to be referred to the reference ellipsoid. Clearly $(N)_M + N^M = N$ as required. As expected, as the degree M grows, Δg^M tends to zero. Moreover, for $n > M$: $\Delta g_n^M = g_n$, since, by definition, for $n > M$: $\gamma_n = 0$.

Equation (5) is an exact counterpart of the original Stokes formula (1) derived for the second-degree reference spheroid (i.e., the reference ellipsoid) and we shall be calling it the "generalized

Stokes formula." It is accurate, like its original second-degree counterpart, to terms of the order of e^2, i.e., to the order of 0.3%. Since for a reasonable choice of M, say M = 20, $|N^M|$ is about one order of magnitude smaller than |N|, the effect of this inaccuracy is also one order of magnitude smaller, which amounts to a few centimetres.

We note that in eqn. (5) Δg^M can be replaced by Δg without any effect on the resultant N^M because S^M is "blind" to the first M harmonic components of Δg. This result follows from the orthogonality of sperical harmonics on the sphere. Nevertheless we shall systematically use Δg^M, because we will want to perform certain operations on the generalized Stokes function S^M which may destroy its "blindness," and the retention of the low-order part of Δg in the convolution integral would give rise to unjustifiable terms.

We wish to point out that the adoption of the higher-degree reference spheroid in the Stokes theory as discussed above is responsible for the change in the shape of the Stokes kernel from S (really S^2) to S^M. The latter tapers off more rapidly than the former, see, e.g., Jekeli [1980], i.e., the influence of distant gravity anomalies on local geoidal height is reduced. The reduction is proportional to the degree M of the reference spheroid. Thus, to evaluate the geoidal height N^M above the M-th degree reference spheroid, distant gravity anomalies may be treated in a more cavalier way than in the standard second-order Stokes theory.

As with the standard Stokes formula, the generalized formula is oblivious to the scale of the reference spheroid and to the geocentricity of the reference field. The question of "forbidden harmonics" [Heiskanen and Moritz, 1967] does not arise since we do not need to transform Δg_1 to N_1. We leave the topographical, indirect, and atmospheric effects in the generalized Stokes approach out of the discussion here. These effects were discussed exhaustively by Vaníček and Kleusberg [1987]. However, again we wish to point out that most errors, linear and non-linear, coming from various sources, are reduced when transferring to a higher-degree reference field (see Heck [1989]).

Numerical Evaluation of the Generalized Stokes Convolution Integral

Theoretically, the integration implied by eqn. (5) has to be carried out over the whole earth. This is a nuisance, because gravity coverage of the earth's surface is irregular and incomplete. Also the numerical effort involved would be huge. This is where the fast convergence of S^M to zero with growing ψ becomes very helpful. The integration does not have to be carried out all the way to $\psi = \pi$, because the contributions to N^M from distances ψ larger than a certain value ψ_0 become manageably small. For practical evaluation of the convolution integral, we would welcome the critical distance ψ_0 to be as small as possible. This would imply that as high a degree

of M should be used for the reference spheroid as possible. On the other hand, the error in the reference field grows with growing M. Normally, therefore, a compromise value of M is used.

Writing now eqn. (5) as

$$N^M \doteq \kappa \iint_{\mathcal{C}_o} S^M(\psi)\, \Delta g^M\, d\mathcal{E} + \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^M(\psi)\, \Delta g^M\, d\mathcal{E}\,, \tag{9}$$

where $\mathcal{C}_o$ denotes a spherical cap of radius ψ_o, we can study the effect of M and ψ_o on the geoidal height N^M. The first term denotes the "truncated" convolution integral (for $\psi \le \psi_o$), while the second term describes the "truncation correction," or the negative "truncation error" δN^M committed when the truncated integral is taken instead of the complete integral over the whole earth. Note that we still use the approximate equality because of the spherical approximation.

It was Molodenskij et al. [1960], who first introduced the idea of reducing the value of ψ_o further by allowing the kernel to be modified in such a way as to minimize the second term in eqn. (9), i.e., the truncation correction or truncation error. This idea has since been explored and developed in different directions by scores of researchers.

To explain the similarities and differences between the various possibilities, let us first explain how the Molodenskij type modification works within the framework of generalized Stokes theory. Molodenskij's idea is to change (modify) the integration kernel by subtracting a modifying function M_S from it. We then get the modified kernel S^{M*} in the following form:

$$S^{M*}(\psi) = S^M(\psi) - M_S(\psi)\quad . \tag{10}$$

Substituting this into eqn. (9) we obtain

$$N^M \doteq \kappa \iint_{\mathcal{C}_o} S^{M*}\, \Delta g^M\, d\mathcal{E} + \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^{M*}\, \Delta g^M\, d\mathcal{E} + \kappa \oiint_{\mathcal{E}} M_S\, \Delta g^M\, d\mathcal{E}\,, \tag{11}$$

where the first term on the right-hand side represents a new approximation of N^M, and the last two terms are the new truncation correction to be minimized.

To make things easier, the modifying function M_S is now chosen so that the last term disappears. Disregarding, once again, any errors in the EGM and in the low-degree components

of Δg, all components Δg_n^M of degree lower than or equal to M disappear, and any polynomial in P_n of degree lower than or equal to M will satisfy the above requirement. (The effect of this disregarded term in the presence of long wavelength errors will be treated later.) We then choose

$$M_S(\psi) = \sum_{n=0}^{M} \frac{2n+1}{2}\, t_n\, P_n(\cos\psi) \ , \tag{12}$$

where the factors (2n+1)/2 are introduced for computational convenience, and t_n, called "Molodenskij's modification coefficients," are to be determined so that the new truncation error

$$\delta N^{M*} = -\ \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^{M*}\, \Delta g^M\, d\mathcal{E} \tag{13}$$

is minimized in one sense or another.

If the generalized Stokes formula is applied to a properly scaled EGM expressed in a geocentric coordinate system, then the summation (12) may begin with n=2. This is what we shall assume for simplicity from now on, so that we write

$$S^{M*}(\psi) = S^M(\psi) - \sum_{n=2}^{M} \frac{2n+1}{2}\, t_n\, P_n(\cos\psi) \ . \tag{14}$$

From Schwarz's inequality applied to eqn. (13) it follows that

$$(\delta N^{M*})^2 \le \kappa^2\ \|S^{M*}\|^2\ \|\Delta g^M\|^2 \ , \tag{15}$$

where

$$\|\cdot\|^2 = \iint_{\mathcal{E}-\mathcal{C}_o} (\cdot)^2\, d\mathcal{E} \ . \tag{16}$$

Now, Molodenskij required that the upper bound of $|\delta N^{M*}|$ (cf. eqn. (15)) be the minimum. For a given Δg^M (fixed reference field and location), the norm $\|\Delta g^M\|$ is constant, while $\|S^{M*}\|$ varies with the choice of t_n (n = 2,3,...,M). Minimizing the latter norm leads to the following system of normal equations:

$$\forall n \le M: \ \frac{\partial}{\partial t_n} \iint_{\mathcal{E}-\mathcal{C}_o} (S^{M*})^2\, d\mathcal{E}$$

$$= \frac{\partial}{\partial t_n} \int_{\psi=\psi_o}^{\pi} (S^{M*})^2 \sin\psi\, d\psi = 0 \ ,$$

or

$$\forall n \le M: \quad \int_{\psi=\psi_o}^{\pi} S^{M*} \frac{\partial M_S^*}{\partial t_n} \sin\psi \, d\psi = 0 \, . \tag{17}$$

Carrying out the differentiation we obtain

$$\forall n \le M: \quad \int_{\psi=\psi_o}^{\pi} (S^M - M_S) \, P_n(\cos\psi) \sin\psi \, d\psi = 0 \, . \tag{18}$$

Employing the usual notation,

$$\int_{\psi=\psi_o}^{\pi} P_i(\cos\psi) \, P_j(\cos\psi) \sin\psi \, d\psi = e_{ij}(\psi_o) \, , \tag{19}$$

$$\int_{\psi=\psi_o}^{\pi} S(\psi) \, P_i(\cos\psi) \sin\psi \, d\psi = Q_i(\psi_o) \, , \tag{20}$$

$$\int_{\psi=\psi_o}^{\pi} S^M(\psi) \, P_i(\cos\psi) \sin\psi \, d\psi = Q_i^M(\psi_o) \, , \tag{21}$$

we get finally:

$$\forall n \le M: \quad \sum_{k=2}^{M} \frac{2k+1}{2} \, e_{nk} \, t_k = Q_n^M = Q_n - \sum_{k=2}^{M} \frac{2k+1}{2} \, e_{nk} \, . \tag{22}$$

This represents a system of M-1 linear equations for t_k which can be solved for any given ψ_o. These (Molodenskij-like) coefficients t_k are then substituted into eqn. (14) to give the Molodenskij-type modified kernel for the generalized Stokes formula. This approach was used in producing the "UNB Dec. '86" Canadian geoid [Vaníček et al., 1988].

Equations (22) are slightly different from the original Molodenskij equations and the resulting parameters t_n are also different from Molodenskij's. But it turns out that the generalized Stokes function S^M modified *à la* Molodenskij (S_{Mol}^{M*}) is exactly the same as Molodenskij's modified original Stokes's function (S_{Mol}^*). This is understandable because in both approaches we seek a function so modified as to have the minimal L_2-norm, and in both cases only the low frequencies (up to wave number M) are allowed to change. There is however a significant difference between the geoidal height N_{Mol}^* obtained by applying Molodenskij's modified kernel in the original Stokes theory and applying it in the generalized Stokes theory: the upper bound of the truncation error is,

for the same radius of integration ψ_0, significantly smaller for the generalized theory. To show this, let us write

$$\left(\delta N_{Mol}^{M*}\right)^2 \le \left(\|S_{Mol}^{M*}\|\,\|\Delta g^M\|\right)^2 = \left(\|S_{Mol}^{*}\|\,\|\Delta g^M\|\right)^2 \tag{23}$$

and, similarly, for the original truncation error:

$$\left(\delta N_{Mol}^{*}\right)^2 \le \left(\|S_{Mol}^{*}\|\,\|\Delta g\|\right)^2 \ . \tag{24}$$

The expected value of $\|\Delta g^M\|$ is significantly smaller than the expected value of $\|\Delta g\|$, which proves the point.

Spectral Representation of Different Kinds of Geoidal Heights

Before we discuss the modification issue further, let us derive the spectrum (harmonic series) representation of the individual kinds of geoidal heights. The simplest expression is obtained for the "exact" geoidal height N^M given by eqn. (5). Expressing both S^M and Δg^M in Legendre's polynomials, we get

$$N^M \doteq c \sum_{n=M+1}^{\infty} \frac{2}{n-1} \Delta g_n^M \ , \tag{25}$$

where $c = R/(2\gamma)$, and Δg_n^M can be replaced by Δg_n because for $n>M$ the two are identical. We use again the approximate equality symbol because the accuracy is only to the order of e^2. Bringing into the discussion also the term ε_1, given by eqn. (8) which was neglected originally, we get an additional term

$$\varepsilon_1 = c \sum_{n=2}^{M} \frac{2}{n-1} \Delta g_n^M \ . \tag{26}$$

The sum of eqns. (25) and (26) gives the complete expression:

$$N^M \doteq c \sum_{n=2}^{\infty} \frac{2}{n-1} \Delta g_n^M \ , \tag{27}$$

in the expected and correct form.

Let us now turn to the spectral representation of the geoidal height obtained from the truncated integration, i.e., from eqn. (9). First, to simplify the forthcoming equations, we introduce a new kernel $\bar{S}^M$ by the following expression,

$$\bar{S}^M = \begin{cases} S^M & \text{for} \quad \psi \le \psi_o \\ 0 & \text{for} \quad \psi > \psi_o \end{cases} \tag{28}$$

and write it in a Legendre's polynomial series form as

$$\bar{S}^M = \sum_{n=2}^{\infty} \frac{2n+1}{2} s_n P_n(\cos\psi) , \tag{29}$$

where s_n are some coefficients to be determined. Disregarding the truncation error we have

$$\begin{aligned} \bar{N}^M &\doteq \kappa \iint_{\mathcal{C}_o} S^M \Delta g^M \, d\mathcal{E} \\ &= \kappa \oiint_{\mathcal{E}} \bar{S}^M \Delta g^M \, d\mathcal{E} \ . \end{aligned} \tag{30}$$

Taking eqn. (29) into account, we can transfer this convolution integral into its spectral form as follows:

$$\bar{N}^M \doteq c \sum_{n=2}^{\infty} s_n \Delta g_n^M \ ,$$

or, equivalently,

$$\bar{N}^M \doteq c \sum_{n=2}^{M} s_n \Delta g_n^M + c \sum_{n=M+1}^{\infty} s_n \Delta g_n \ . \tag{31}$$

Here we recognize the first term to be again caused by long wavelength errors in γ^M and g; it would disappear if the EGM were errorless and if g were not contaminated by long wavelength errors.

The coefficients s_n can be easily determined from

$$\forall n: \quad s_n = \int_{\psi=0}^{\pi} \bar{S}^M(\psi) P_n(\cos\psi) \sin\psi \, d\psi \ ,$$

or, equivalently, from

$$\forall n: \quad s_n = \int_{\psi=0}^{\pi} S^M P_n \sin\psi \, d\psi - \int_{\psi=\psi_0}^{\pi} S^M P_n \sin\psi \, d\psi \, . \tag{32}$$

Clearly,

$$s_n = \begin{cases} - Q_n^M & \text{for} \quad n \le M \\ \frac{2}{n-1} - Q_n^M & \text{for} \quad n > M \end{cases} , \tag{33}$$

and eqn. (31) becomes

$$\bar{N}^M \doteq - c \sum_{n=2}^{M} Q_n^M \Delta g_n^M + c \sum_{n=M+1}^{\infty} \left(\frac{2}{n-1} - Q_n^M \right) \Delta g_n \, . \tag{34}$$

Here again, the first term will have a non-zero value only because of long wavelength errors in the EGM and in g.

It is interesting now to have a look also at the spectrum of the truncation error. The complete truncation error can be obtained from eqn. (9) and by considering the originally neglected term ε_1. We obtain

$$\delta\bar{N}^M = - \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^M \Delta g^M \, d\mathcal{E} - \kappa \oiint_{\mathcal{E}} (S - S^M) \Delta g^M \, d\mathcal{E} \, . \tag{35}$$

This can be written in a spectral form as

$$\delta\bar{N}^M = - c \sum_{n=2}^{\infty} Q_n^M \Delta g_n^M - c \sum_{n=2}^{M} \frac{2}{n-1} \Delta g_n^M \, . \tag{36}$$

Subtracting $\delta\bar{N}^M$ from $\bar{N}^M$ (given by eqn. (34)) we obtain N^M (eqn. (27)) as we should.

Similarly, the geoidal height N^{M*} computed by means of a modified generalized Stokes function S^{M*} (eqn. (11)) has the following spectrum:

$$N^{M*} \doteq - c \sum_{n=2}^{M} (t_n + Q_n^{M*}) \Delta g_n^M + c \sum_{n=M+1}^{\infty} \left(\frac{2}{n-1} - Q_n^{M*} \right) \Delta g_n , \tag{37}$$

where we use the symbol Q_n^{M*} to denote

$$\forall n: Q_n^{M*} = Q_n^M - \sum_{i=2}^{M} \frac{2i+1}{2} e_{in} t_i . \tag{38}$$

Both the long and short wavelengths are affected by the modification. The difference between the geoidal heights computed from the truncated integration using the generalized Stokes kernel and the modified generalized Stokes kernel is given by

$$N^{M*} - \bar{N}^M = - c \sum_{n=2}^{M} \left(t_n - \sum_{i=2}^{M} \frac{2i+1}{2} e_{in} t_i \right) \Delta g_n^M + c \sum_{n=M+1}^{\infty} \sum_{i=2}^{M} \frac{2i+1}{2} e_{in} t_i \Delta g_n . \tag{39}$$

The critical role played by the Molodenskij modification coefficients t_n is clearly demonstrated in this equation.

The complete truncation error of modified geoidal height increased by the term ε_1 is

$$\delta N^{M*} = - \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^{M*} \Delta g^M \, d\mathcal{E} - \kappa \oiint_{\mathcal{E}} M_S \Delta g^M \, d\mathcal{E} - \varepsilon_1 . \tag{40}$$

In a spectral form it can be written as

$$\delta N^{M*} = - c \sum_{n=2}^{\infty} Q_n^{M*} \Delta g_n^M - c \sum_{n=2}^{M} \left(\frac{2}{n-1} + t_n \right) \Delta g_n^M , \tag{41}$$

which, once again, subtracted from N^{M*} (eqn. (37)) gives N^{M}.

We note that Molodenskij's modification (eqn. (22)) implies that

$$\forall n \le M: \quad Q_n^{M*} = 0$$

rendering the following result

$$\delta N_{Mol}^{M*} \doteq - c \sum_{n=2}^{M} \left(\frac{2}{n-1} + t_n \right) \Delta g_n^M - c \sum_{n=M+1}^{\infty} Q_n^{M*} \Delta g_n \, . \tag{42}$$

The Molodenskij modified geoidal height spectrum is then

$$N_{Mol}^{M*} \doteq - c \sum_{n=2}^{M} t_n \Delta g_n^M + c \sum_{n=M+1}^{\infty} \left(\frac{2}{n-1} - Q_n^{M*} \right) \Delta g_n \, . \tag{43}$$

Comparing now eqns. (37) and (43), we see that the Molodenskij modification affects not only the long wavelength geoidal heights — in the presence of long wavelength errors in Δg^M — but also the short wavelengths. The Q_n^{M*} factors in the two equations are different, as is apparent from eqn. (38): in the former case the modification coefficients are unspecified, in the latter case they are given by eqns. (22). This is true, of course, only for the modification which uses Legendre's polynomials up to the M-th degree. Other modification schemes can be used, and it will be interesting to investigate, for instance, modification schemes which use Legendre's polynomials up to a degree L higher than the degree M of the reference spheroid.

To do so, we rewrite eqn. (5) in the following rather general way:

$$N^M = \kappa \oiint_{\mathcal{E}} S^L(\psi) \, \Delta g^M \, d\mathcal{E} + c \sum_{n=M+1}^{L} s_n \Delta g_n \, , \tag{44}$$

or equivalently,

$$\begin{aligned} N^M = {} & \kappa \iint_{\mathcal{C}_o} S^L(\psi) \, \Delta g^M \, d\mathcal{E} \\ & + c \left\{ \sum_{n=2}^{L} \left[Q_n^L(\psi_o) + s_n \right] \Delta g_n^M + \sum_{n=L+1}^{\infty} Q_n^L(\psi_o) \, \Delta g_n \right\} \, , \end{aligned} \tag{45}$$

where

$$\forall L \ge M: \quad S^L(\psi) = S(\psi) - \sum_{k=2}^{L} \frac{2k+1}{2} \, s_n \, P_n(\cos\psi) \, , \tag{46}$$

$$\forall\, n \le L:\quad Q_n^L(\psi_o) = \int_{\psi_o}^{\pi} S^L(\psi)\, P_n(\cos\psi)\, \sin\psi\, d\psi$$

$$= Q_n(\psi_o) - \sum_{k=2}^{L} \frac{2k+1}{2}\, e_{nk}(\psi_o)\, s_k\,. \tag{47}$$

In eqn. (45) we have divided the integration area into a cap $\mathcal{C}_o$ of spherical angle ψ_o around the computation point and a remote zone area $\mathcal{E}$-$\mathcal{C}_o$. The latter integral becomes

$$\kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^L(\psi)\, \Delta g^M\, d\mathcal{E} = c \sum_{n=2}^{\infty} Q_n^L(\psi_o)\, \Delta g_n^M\,, \tag{48}$$

which when inserted into eqn. (44) yields eqn. (45). We have assumed that the maximum degree of modification (L) is at least as high as the degree of the reference field, i.e., L≥M, and we will use this assumption throughout the rest of this paper. We also assume that potential coefficients, typically determined from satellite orbit analyses, are available to degree and order L. However, only the first M degrees and orders are used to define the reference spheroid.

Referring now to eqn. (45), it follows that $\forall n>M$: $\Delta g_n^M = \Delta g_n$. Furthermore, due to the orthogonality of the Laplace harmonics, it follows that eqn. (44) is equivalent to eqn. (5) for any choice of the parameters s_n (k=2,3,...,L). (For clarification, S^L is denoted by S^{L*}, when $s_n = (2/(n-1) + t_n$.) Equation (44) is essential in the error estimates to be derived below.

One thing is clear, however: compared to the standard Stokes approach, truncation of the convolution integral does much less damage in the generalized Stokes approach. The price one has to pay for this is the introduction of unmitigated errors in the EGM (really in $(N)_M$). We shall investigate the effect of these errors together with the effect of errors in Δg later.

The other source of errors, the discretization error in the numerical evaluation of the convolution integral for N^{M*}, is considered outside the scope of this paper. It represents a problem from the domain of numerical analysis and as such calls for development of techniques from that mathematical domain.

Two Generic Estimators

To be able to treat the case of different degrees (and orders) of the reference spheroid (M) and the modification (L), we shall directly introduce two kinds of generic estimators of N^M rather than deriving them from some desired properties. Later, we shall show that specific selections of free parameters s_n lead to specific properties of these estimators.

Consider first the following general estimator of N^M (cf. Sjöberg [1987])

$$\tilde{N}^{M'} = \kappa \iint_{\mathcal{C}_o} S^L(\psi)\, \hat{\Delta g}^M\, d\mathcal{E} + c \sum_{n=M+1}^{L} (Q_n^L + s_n)\, \hat{\Delta g}_n\,, \tag{49}$$

where $\hat{\Delta g}^M$ and $\hat{\Delta g}_n$ are observed values (estimates) of Δg^M and Δg_n, respectively. $\tilde{N}^{M'}$ suffers from errors in $\hat{\Delta g}^M$ and $\hat{\Delta g}_n$ as well as from the truncation error arising from the limited integration area $\mathcal{C}_o$.

A slightly different general estimator is given by (cf. Sjöberg [1984a,b; 1987])

$$\tilde{N}^{M''} = \kappa \iint_{\mathcal{C}_o} S^L(\psi)\, \hat{\Delta g}^M\, d\mathcal{E} + c \sum_{n=M+1}^{L} s_n\, \hat{\Delta g}_n\,. \tag{50}$$

Note that the summation terms on the right-hand side of eqns. (49) and (50) reflect the assumptions that $\forall n \le M$: $\hat{\Delta g}_n^M = 0$, i.e., that the reference field represents the first M degrees of the actual field perfectly, and, by definition, $\forall n > M$: $\hat{\Delta g}_n^M = \hat{\Delta g}_n$ (cf. discussion following eqn. (8)).

The errors of the two estimators (49) and (50) will be discussed in the next section.

Error Estimation

Let us denote the errors in terrestrial gravity and the reference field harmonics of anomalies by ε^T and ε^S, respectively. The n-th Laplace harmonics of these errors are denoted by ε_n^T and ε_n^S. Then the estimator $\tilde{N}^{M'}$ can be rewritten as

$$\begin{aligned} \tilde{N}^{M'} = \kappa \iint_{\mathcal{C}_o} & S^L(\Delta g^M + \varepsilon^T - \varepsilon^S)\, d\mathcal{E} \\ & + c \sum_{n=M+1}^{L} (Q_n^L + s_n)(\Delta g_n + \varepsilon_n^S)\,, \end{aligned} \tag{51}$$

where we have taken, following eqn. (7):

$$\hat{\Delta g}^M = \hat{\Delta g}^T - \hat{\Delta g}^S = \Delta g + \varepsilon^T - (\Delta g_M + \varepsilon^S) = \Delta g^M + \varepsilon^T - \varepsilon^S\,, \tag{52}$$

and

$$\varepsilon^S = \sum_{n=2}^{M} \varepsilon_n^S\,. \tag{53}$$

By adding the stipulated (estimated) reference spheroid

$$\tilde{N}_M = c \sum_{n=2}^{M} \frac{2}{n-1} (\Delta g_n + \varepsilon_n^S) , \tag{54}$$

one obtains an estimator ($\tilde{N}'$) for the total undulation above the reference ellipsoid. Equations (51), (54), (4), and (45) yield the expression for the total geoid undulation error which includes both the truncation error and the errors in the satellite and terrestrial gravity anomalies:

$$\begin{aligned} \delta\tilde{N}' = \tilde{N}' - N = &- \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^L \Delta g^M \, d\mathcal{E} \\ &+ \kappa \iint_{\mathcal{C}_o} S^L(\varepsilon^T - \varepsilon^S) \, d\mathcal{E} + c \sum_{n=2}^{M} \frac{2}{n-1} \varepsilon_n^S \\ &+ c \sum_{n=M+1}^{L} (Q_n^L + s_n)(\Delta g_n + \varepsilon_n^S) . \end{aligned} \tag{55}$$

Note that eqn. (52) yields

$$\forall n \le M: \quad \Delta\hat{g}_n^M = \varepsilon_n^T - \varepsilon_n^S \tag{56}$$

and

$$\forall n > M: \quad \Delta\hat{g}_n^M = \Delta\hat{g}_n = \Delta g_n + \varepsilon_n^T . \tag{57}$$

Using the following relations between integral convolutions and Laplace expansions:

$$\kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^L \Delta g^M \, d\mathcal{E} = \sum_{n=M+1}^{\infty} Q_n^L \Delta g_n \tag{58}$$

and

$$\kappa \iint_{\mathcal{C}_o} S^L(\varepsilon^T - \varepsilon^S) \, d\mathcal{E} = c \sum_{n=2}^{\infty} \left(\frac{2}{n-1} - Q_n^L - s_n^*\right)(\varepsilon_n^T - \varepsilon_n^S), \tag{59}$$

where

$$\forall n > L: \quad \varepsilon_n^S = 0 \tag{60}$$

and

$$s_n^* = \begin{cases} s_n & \text{for} \quad 2 \le n \le L \\ 0 & \text{for} \quad n > L \end{cases} , \tag{61}$$

one obtains, finally, the spectral form of the total error:

$$\delta\tilde{N}' = c \sum_{n=2}^{\infty} \left(\frac{2}{n-1} - s_n^* - Q_n^L \right) \varepsilon_n^T + c \sum_{n=2}^{L} \left(Q_n^L + s_n \right) \varepsilon_n^S - c \sum_{n=L+1}^{\infty} Q_n^L \Delta g_n \ . \tag{62}$$

Assuming that the observation errors have zero expectation, i.e.,

$$\forall n: \ E(\varepsilon_n^T) = E(\varepsilon_n^S) = 0, \text{ and thus also } E(\varepsilon^T) = E(\varepsilon^S) = 0 \ , \tag{63}$$

it follows that the expected value of the total error is

$$\overline{\delta N}' = E(\delta\tilde{N}') = - c \sum_{n=L+1}^{\infty} Q_n^L \Delta g_n \ . \tag{64}$$

Introducing the global average operator

$$\Gamma(\cdot) = \frac{1}{\mathcal{E}} \iint_{\mathcal{E}} \cdot \, d\mathcal{E} \ , \tag{65}$$

we note that

$$\Gamma(\overline{\delta N}') = - c \sum_{n=L+1}^{\infty} Q_n^L \ \Gamma(\Delta g_n) = 0 \ , \tag{66}$$

i.e., the global average of $\delta\tilde{N}'$ is unbiased. However, $\delta\tilde{N}'$ itself is locally biased, because for any given locality, the expected value $\overline{\delta N}'$ is not equal to zero. To study the local bias, we introduce:

$$\Gamma[(\overline{\delta N}')^2] = c^2 \sum_{n=L+1}^{\infty} (Q_n^L)^2 \, c_n > 0 \ , \tag{67}$$

where c_n are the so-called "anomaly degree variances"

$$\forall n: \ c_n = \Gamma(\Delta g_n^2) \geq 0 \ , \tag{68}$$

which reflect the global behaviour of the gravity field and have nothing to do with observing errors. Note also that the anomaly degree covariances vanish:

$$\forall n \neq k: \quad \Gamma(\Delta g_n, \Delta g_k) = 0 \ . \tag{69}$$

This shows that the estimator $\tilde{N}'$ is locally biased from degree and order L+1 up. This is caused by the truncation of Stokes's integration to a cap $\mathcal{C}_o$ and by the use of a harmonic series to degree and order L in eqn. (49) (see also Sjöberg [1987]).

In a similar way, the error of the second general estimator $\tilde{N}''$ becomes (cf. eqns. (49) and (50))

$$\delta\tilde{N}'' = \tilde{N}'' - N = \delta\tilde{N}' - c \sum_{n=M+1}^{L} Q_n^L \, \hat{\Delta g}_n \tag{70}$$

with the expectation

$$\overline{\delta N}'' = E(\delta\tilde{N}'') = - c \sum_{n=M+1}^{\infty} Q_n^L \, \Delta g_n \ . \tag{71}$$

It follows that the estimator $\tilde{N}''$ is again globally unbiased but is locally biased from degree and order M+1 up. Clearly, if L is selected to equal to M, both estimators will have the same local bias.

Next we will determine the variances of the estimators $\tilde{N}'$ and $\tilde{N}''$. We will use the following notations for the gravity anomaly error covariances

$$\forall n,k: \quad E(\varepsilon_n^T \, \varepsilon_k^T) = \Lambda_{nk} \quad \text{and} \quad \forall n,k \leq L: \quad E(\varepsilon_n^S \, \varepsilon_k^S) = \Omega_{nk} \ . \tag{72}$$

Note that Λ_{nk} and Ω_{nk} are position dependent! Furthermore, we assume that errors in $\hat{\Delta g}^T$ and $\hat{\Delta g}_n^S$ are uncorrelated, i.e.,

$$\forall n \neq k: \quad E(\varepsilon_n^T \, \varepsilon_k^S) = 0 \ . \tag{73}$$

This implies that terrestrial gravity should not have been used in the computation of the potential coefficients defining either the reference spheroid or the potential coefficients of degrees between M and L! The case of correlated ε^T and ε^S was treated by Sjöberg [1987]. Finally, we note that potential coefficient derived anomaly errors (ε_n^S) will contribute to the variances of $\tilde{N}'$ and $\tilde{N}''$ both through the reference field (see eqn. (54)) ($2 \leq n \leq M$) and through the potential coefficient representation for degrees between M and L.

Then the (expected) "local mean square error" of $\tilde{N}'$ becomes

$$MSE(\tilde{N}') = E[(\tilde{N}' - N)^2] = Var(\tilde{N}') + Bias^2(\tilde{N}') \ , \tag{74}$$

where

$$Var(\tilde{N}') = E\{[\tilde{N}' - E(\tilde{N}')]^2\}$$

$$= c^2\left[\sum_{n=2}^{\infty}\left(\frac{2}{n-1} - s_n^* - Q_n^L\right)\sum_{k=2}^{\infty}\left(\frac{2}{k-1} - s_k^* - Q_k^L\right)\Lambda_{nk}\right.$$

$$\left. + \sum_{n=2}^{L}(Q_n^L + s_n)\sum_{k=2}^{L}(Q_k^L + s_k)\,\Omega_{nk}\right] \tag{75}$$

and

$$Bias^2(\tilde{N}') = [E(\tilde{N}') - N]^2 = c^2\left(\sum_{n=L+1}^{\infty} Q_n^L\,\Delta g_n\right)^2 \ . \tag{76}$$

Similarly, one obtains the (expected) local mean square error of the estimator $\tilde{N}''$:

$$MSE(\tilde{N}'') = Var(\tilde{N}'') + Bias^2(\tilde{N}'') \ , \tag{77}$$

where

$$Var(\tilde{N}'') = c^2\left[\sum_{n=2}^{\infty}\left(\frac{2}{n-1} - s_n^* - Q_n^L\right)\sum_{k=2}^{\infty}\left(\frac{2}{k-1} - s_k^* - Q_k^L\right)\Lambda_{nk}\right.$$

$$\left. + \sum_{n=2}^{L}(P_n^L + s_n)\sum_{k=2}^{L}(P_k^L + s_k)\,\Omega_{nk}\right], \tag{78}$$

where

$$P_n^L = \begin{cases} Q_n^L & \text{for} \quad 2 \le n \le M \\ 0 & \text{for} \quad M < n \le L \end{cases} \tag{79}$$

and

$$Bias^2(\tilde{N}'') = c^2\left(\sum_{n=M+1}^{\infty} Q_n^L\,\Delta g_n\right)^2 \ . \tag{80}$$

Note that the above local mean square errors, variances, and biases are position dependent!

We now proceed to the global averages of the local mean square errors which we shall call "global mean square errors." For the estimator $\tilde{N}'$ we obtain:

$$\Gamma[MSE(\tilde{N}')] = \Gamma\,[Var(\tilde{N}')] + \Gamma\,[Bias^2(\tilde{N}')]\ , \tag{81}$$

where

$$\Gamma\,[Var(\tilde{N}')] = c^2\left[\sum_{n=2}^{\infty}\left(\frac{2}{n-1}\,s_n^* - Q_n^L\right)^2 (\sigma_n^T)^2 + \sum_{n=2}^{L}(Q_n^L + s_n)^2(\sigma_n^S)^2\right], \tag{82}$$

and

$$\Gamma\,[Bias^2(\tilde{N}')] = c^2 \sum_{n=L+1}^{\infty} (Q_n^L)^2\, c_n\ . \tag{83}$$

In these derivations, we have employed the following notation:

$$\Gamma(\Lambda_{nk}) = \begin{cases} (\sigma_n^T)^2 & \text{for } n=k \\ 0 & \text{for } n\neq k \end{cases} \tag{84}$$

and

$$\Gamma(\Omega_{nk}) = \begin{cases} (\sigma_n^S)^2 & \text{for } n=k \\ 0 & \text{for } n\neq k \end{cases}\ , \tag{85}$$

where $(\sigma_n^T)^2$ and $(\sigma_n^S)^2$ are called the "error degree variances" of the terrestrial and of the potential-coefficient-generated anomalies, respectively. In addition, equations (68) and (69) were used to derive the global averages of squared biases.

In the same way, the global mean square error of $\tilde{N}''$ is derived:

$$\Gamma\,[MSE(\tilde{N}'')] = \Gamma\,[Var(\tilde{N}'')] + \Gamma\,[Bias^2(\tilde{N}'')]\ , \tag{86}$$

where

$$\Gamma\,[Var(\tilde{N}'')] = \Gamma\,[Var(\tilde{N}')] - c^2 \sum_{n=M+1}^{L} Q_n^L(Q_n^L + 2s_n)(\sigma_n^S)^2 \tag{87}$$

and

$$\Gamma\,[\mathrm{Bias}^2(\tilde{N}'')] = c^2 \sum_{n=M+1}^{\infty} (Q_n^L)^2\, c_n \ . \tag{88}$$

It is interesting to note that the mean square errors of $\tilde{N}'$ are independent of the choice of degree $M(\leq L)$ of the reference field, while for $\tilde{N}''$ they depend on M (cf. eqns. (74) to (87)).

Some Special Modifications

We will now consider some special cases of modifying Stokes's formula, namely the "Molodenskij modification," the "strict separation modification," and the "least-squares modification." The first method, limited to L=M, was formulated already above and will now be only restated for $L \geq M$.

Molodenskij's Modification

To begin with, we rewrite eqn. (45) as

$$N^M = \kappa\left[\iint_{\mathcal{C}_o} S^L(\psi)\Delta g^M d\mathcal{E} + \iint_{\mathcal{E}-\mathcal{C}_o} S^L(\psi)\Delta g^M d\mathcal{E}\right] + c \sum_{n=M+1}^{L} s_n\, \Delta g_n \ . \tag{89}$$

The truncation error is described by the second term (which is not contained in $\tilde{N}^{M''}$):

$$\delta N^L = \kappa \iint_{\mathcal{E}-\mathcal{C}_o} S^L(\psi)\ \Delta g^M\, d\mathcal{E} \ . \tag{90}$$

In view of Molodenskij's choice of parameters s_n, this solution implies that

$$\forall n \leq L: \quad Q_n^L = 0 \ . \tag{91}$$

Hence the two estimators (49) and (50) become identical, with

$$\hat{N}^M_{\mathrm{Mol}} = \kappa \iint_{\mathcal{C}_o} S^L\, \Delta\hat{g}^M\, d\mathcal{E} + c \sum_{n=2}^{L} s_n\, \Delta\hat{g}_n \tag{92}$$

with the local mean square error:

$$\begin{aligned} \mathrm{MSE}(\hat{N}^M_{\mathrm{Mol}}) = c^2 \sum_{n=2}^{\infty}\left(\frac{2}{n-1} - s_n^* - Q_n^L\right)\sum_{k=2}^{\infty}\left(\frac{2}{k-1} - s_k^* - Q_k^L\right)\Lambda_{nk} \\ + c^2 \sum_{n=2}^{L}\sum_{k=2}^{L} s_n\, s_k\, \Omega_{nk} + c^2\left(\sum_{n=L+1}^{\infty} Q_n^L\, \Delta g_n\right)^2 . \end{aligned} \tag{93}$$

The global mean square error becomes:

$$\Gamma\,[\mathrm{MSE}(\hat{N}^{M}_{\mathrm{Mol}})] = c^2 \sum_{n=2}^{L}\left[\left(\frac{2}{n-1} - s_n\right)^2 (\sigma^T_n)^2 + s_n^2(\sigma^S_n)^2\right] + c^2 \sum_{n=L+1}^{\infty}\left[\left(\frac{2}{n-1} - Q^L_n\right)^2 (\sigma^T_n)^2 + (Q^L_n)^2\, c_n\right]. \tag{94}$$

Clearly, the errors of Molodenskij's modification of Stokes's formula are independent of the degree (M) of the reference field; rather they depend on the degree L of modification. Due to eqn. (69), the resulting truncation error becomes

$$\delta N^L_{\mathrm{Mol}} = - c \sum_{n=L+1}^{\infty} Q^L_n\, \Delta g_n\,, \tag{95}$$

independent of the choice of degree M. This implies that the smaller truncation error bound for a higher-degree (M) reference field discussed earlier must be taken as applying rather to the degree L of modification for which M is the lowest bound.

The Strict Separation Modification

Frequently the combined solution of the truncated (and modified) Stokes's formula and the reference field of degree M is said to be a merger of long-wavelength features given by the latter and short-wavelength features given by the former. This separation is generally not rigorous, and it is highly dependent on the type of modification of Stokes's function, i.e., the choice of the parameters s_n. For some applications, a strict wavelength separation might be advantageous. The derivation of such a solution, which we have not found in the open literature, is the intention of this section.

Reconsider the total error (62) of the general estimator $\tilde{N}^{M'}$. If s_n is selected in such a way that

$$\forall\ 2 \le n \le M \le L:\quad Q^L_n + s_n = \frac{2}{n-1} \tag{96}$$

then the error becomes

$$\delta\hat{N}' = c \sum_{n=2}^{L} \frac{2}{n-1}\, \varepsilon^S_n + c \sum_{n=L+1}^{\infty}\left[\left(\frac{2}{n-1} - Q^L_n\right)\varepsilon^T_n - Q^L_n\, \Delta g_n\right]. \tag{97}$$

We have thus shown that for this choice of s_n, the estimator

$$\hat{N}' = \kappa \iint_{\sigma_0} S^L \hat{\Delta g}\, d\sigma + c \sum_{n=2}^{L} \frac{2}{n-1} \hat{\Delta g}_n \tag{98}$$

is strictly determined from estimated Laplace harmonics of Δg up to degree L and from observed terrestrial gravity anomalies above degree L. Here it is completely irrelevant whether $\hat{\Delta g}^M$ or $\hat{\Delta g}$ is the argument under Stokes's integral of eqn. (98). The local and global mean square errors become

$$\mathrm{MSE}(\hat{N}') = c^2 \left[\sum_{n=2}^{L} \sum_{k=2}^{L} \frac{2}{n-1} \frac{2}{k-1} \Omega_{nk} + \sum_{n=L+1}^{\infty} \sum_{k=L+1}^{\infty} \left(\frac{2}{n-1} - Q_n^L \right) \left(\frac{2}{k-1} - Q_k^L \right) \Lambda_{nk} + \left(\sum_{n=L+1}^{\infty} Q_n^L \, \Delta g_n \right)^2 \right] \tag{99}$$

and

$$\Gamma\,[\mathrm{MSE}(\hat{N}')] = c^2 \left\{ \sum_{n=2}^{L} \left(\frac{2}{n-1} \right)^2 (\sigma_n^S)^2 + \sum_{n=L+1}^{\infty} \left[\left(\frac{2}{n-1} - Q_n^L \right)^2 (\sigma_n^T)^2 + (Q_n^L)^2 \, c_n \right] \right\} . \tag{100}$$

Similarly, we could get the estimator $\hat{N}''$ and its mean square errors from eqns. (96), (50), (74) to (76), and (77) to (80). $\hat{N}''$ and its errors are dependent on the degree (M) of the reference field, while this is not the case with $\hat{N}'$.

The Least-Squares Modifications

The least-squares modification minimizes the mean square errors of the solution with respect to the choice of the parameters s_n (k=2,3,...,L). For each of the general estimators $\tilde{N}^{M'}$ and $\tilde{N}^{M''}$, one obtains two least-squares solutions: one locally best estimator, minimizing the (local) mean square error (eqns. (74) to (76)), and one globally best estimator, minimizing the global mean square error (eqns. (77) to (80)). In each case, the MSEs of the general solutions, i.e., equations (74) to (76) and (77) to (80), can be written in the following general form [Sjöberg, 1987]

$$\mathrm{MSE} = a + s^T A s - 2 s^T h \,, \tag{101}$$

where a is the MSE without modification, i.e., for all s_n set to zero, A is a symmetrical matrix, and h is a vector. For instance

$$a = a_1 = c^2\left[\sum_{n=2}^{\infty}\left(\frac{2}{n-1} - Q_n\right)^2 (\sigma_n^T)^2 + \sum_{n=2}^{L} Q_n^2 (\sigma_n^S)^2 + \sum_{n=L+1}^{\infty} Q_n^2\, c_n\right] \tag{102}$$

and

$$a = a_2 = a_1 + c^2 \sum_{n=M+1}^{L} Q_n^2 [c_n - (\sigma_n^S)^2] \ , \tag{103}$$

where a_1 and a_2 refer to the Γ(MSE) of $\tilde{N}'$ and $\tilde{N}''$, respectively.

The minimum variance is obtained for

$$\frac{\partial MSE}{\partial s} = As - h = 0 \ . \tag{104}$$

Thus the optimum set of parameters $\hat{s}$ are the solution of the system

$$A\hat{s} = h \ , \tag{105}$$

yielding the MSE:

$$MSE = a - \hat{s}^T A\hat{s} = a - \hat{s}^T h \ . \tag{106}$$

The elements of A and h are given as follows:

(a) Least-squares estimator unbiased to degree and order L.

(i) The locally best estimator:

$$\forall n,k=2,\ldots,L: \quad A_{nk} = \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} (\delta_{ni} - E_{ni})(\delta_{kj} - E_{kj})(\Omega_{ij} + \Lambda_{ij}) + \sum_{i=L+1}^{\infty}\sum_{j=L+1}^{\infty} E_{ni} E_{kj} \Delta g_i \Delta g_j \ , \tag{107}$$

$$\forall n=2,\ldots,L: \quad h_n = \sum_{i=2}^{\infty}\sum_{j=2}^{\infty} (\delta_{ni} - E_{ni})\left[\Lambda_{ij}\left(\frac{2}{j-1} - Q_j\right) - \Omega_{ij} Q_j\right] + \sum_{i=L+1}^{\infty} E_{ni} \Delta g_i \sum_{j=L+1}^{\infty} Q_j \Delta g_j \ , \tag{108}$$

and

$$a = c^2\Bigg[\sum_{i=2}^{\infty} \sum_{j=2}^{\infty} \left(\frac{2}{i-1} - Q_i\right)\left(\frac{2}{j-1} - Q_j\right)\Lambda_{ij} + \sum_{i=2}^{L} \sum_{j=2}^{L} Q_i\, Q_j\, \Omega_{ij} + \left(\sum_{i=L+1}^{\infty} Q_i\, \Delta g_i\right)^2 \Bigg], \tag{109}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{for} \quad i=j \\ 0 & \text{for} \quad i\neq j \end{cases} \tag{110}$$

is the Kronecker symbol, and

$$E_{ni} = \frac{2n+1}{2}\, e_{ni}\ . \tag{111}$$

(ii) The globally best estimator (see also Sjöberg [1984b; 1987]):

$$\forall k,r = 2,\ldots,L: \quad A_{kr} = \chi_k\, \delta_{kr} - \frac{2r+1}{2}\, \chi_k\, e_{kr} - \frac{2k+1}{2}\, \chi_r\, e_{rk} + \frac{2k+1}{2}\, \frac{2r+1}{2} \sum_{n=2}^{\infty} e_{nk}\, e_{nr}\, \chi_n\ , \tag{112}$$

$$\forall k = 2,\ldots,L: \quad h_k = \frac{2(\sigma_k^T)^2}{k-1} - Q_k\, \chi_k + \frac{2k+1}{2} \sum_{n=2}^{\infty}\left[Q_n\, e_{nk}\, \chi_n - \frac{2}{n-1}\, e_{nk}\, (\sigma_n^T)^2\right]\ , \tag{113}$$

and

$$a = c^2\left[\sum_{i=2}^{\infty} \left(\frac{2}{i-1} - Q_i\right)^2 (\sigma_i^T)^2 + \sum_{i=2}^{L} Q_i^2\, (\sigma_i^S)^2 + \left(\sum_{i=L+1}^{\infty} Q_i\, \Delta g_i\right)^2\right] \tag{114}$$

where

$$\chi_n = (\sigma_n^T)^2 + \begin{cases} (\sigma_n^S)^2 & \text{for} \quad 2 \le n \le L \\ c_n & \text{for} \quad n > L \end{cases} \tag{115}$$

(b) Least-squares estimator unbiased to degree M (degree and order of the reference field.)

(i) The locally best estimator:

$$\forall n,k=2,\ldots,M:\ A_{nk} = \sum_{i=2}^{\infty}\sum_{j=2}^{\infty}(\delta_{ni} - E_{ni})(\delta_{kj} - E_{kj})(\Lambda_{ij} + \Omega^{*}_{ij}) + \sum_{i=M+1}^{\infty}\sum_{j=M+1}^{\infty} E_{ni}\,E_{kj}\,\Delta g_i\,\Delta g_j\ , \tag{116}$$

$$\forall n=2,\ldots,M:\ h_n = \sum_{i=2}^{\infty}\sum_{j=2}^{\infty}(\delta_{ni}-E_{ni})\left[\left(\frac{2}{j-1} - Q_j\right)\Lambda_{ij} - \Omega^{*}_{ij}\right] + \sum_{i+M+1}^{\infty} E_{ni}\,\Delta g_i \sum_{j=M+1}^{\infty} Q_j\,\Delta g_j \tag{117}$$

and

$$a = c^2\left\{\sum_{i=2}^{\infty}\sum_{j=2}^{\infty}\left[\left(\frac{2}{i-1} - Q_i\right)\left(\frac{2}{j-1} - Q_j\right)\Lambda_{ij} + Q_i\,Q_j\,\Omega^{*}_{ij}\right] + \left(\sum_{i=M+1}^{\infty} Q_i\,\Delta g_i\right)^2\right\}, \tag{118}$$

where

$$\Omega^{*}_{ij} = \begin{cases} \Omega_{ij} & \text{for} \quad i \vee j \le M \\ 0 & \text{otherwise} \end{cases}\ . \tag{119}$$

(ii) The globally best estimator:

$$\forall k,r=2,\ldots,M:\ A_{kr} = \delta_{kr}\,\chi_r - E_{rk}\chi^{*}_{k} - E_{kr}\chi^{*}_{r} + \sum_{n=2}^{\infty} E_{kn}\,E_{rn}\,\chi^{**}_{n} \tag{120}$$

$$\forall k=2,\ldots,M:\ h_k = \left(\frac{2}{k-1} - Q_k\right)(\sigma^{T}_{k})^2 + \sum_{n=2}^{M}(E_{kn} - \delta_{kn})\,(\sigma^{S}_{n})^2 + \sum_{n=M+1}^{\infty} E_{kn}\,Q_n\,c_n\ , \tag{121}$$

and

$$a = c^2\left[\sum_{i=2}^{\infty}\left(\frac{2}{i-1} - Q_i\right)^2 (\sigma^{T}_{i})^2 + \sum_{i=2}^{\infty} Q_i^2\,\chi^{**}_{i}\right]\ , \tag{122}$$

where

$$\chi_n^* = (\sigma_n^T)^2 + \begin{cases} (\sigma_n^S)^2 & \text{for} \quad 2 \le n \le M \\ 0 & \text{for} \quad n > M \end{cases} , \tag{123}$$

and

$$\chi_n^{**} = \begin{cases} (\sigma_n^S)^2 & \text{for} \quad 2 \le n \le M \\ c_n & \text{for} \quad n > M \end{cases} . \tag{124}$$

Final Remarks and Conclusions

The generalized Stokes approach parallels the classical Stokes approach when the reference ellipsoid is replaced by an M-th degree spheroid, normal gravity γ_0 on the reference ellipsoid is replaced by a model gravity γ^M on the spheroid, and the Stokes integration kernel S is replaced by the spheroidal kernel S^M. In standard geodetic practice, gravity values g are reduced to gravity anomalies Δg by subtracting normal gravity γ_0. In the context of the generalized Stokes approach, gravity values are reduced even more by subtracting the model gravity γ^M to obtain generalized gravity anomalies Δg^M. For growing M, the Δg^M tends to zero, and for $n > M$ the harmonic components Δg_n^M become identically equal to the harmonic components g_n of the gravity itself.

In both the classical Stokes approach and the generalized Stokes approach, the spherical approximation causes an error in the evaluated geoidal height but for growing M the error diminishes. For M = 20, the error is estimated to be within a few centimetres. In both approaches, the error in the reference field (second degree and M-th degree, respectively) has to be treated separately, and both approaches are oblivious to the scale of their respective reference surfaces and their geocentricity.

The main advantage of the generalized Stokes approach is that the integration kernel S^M converges rapidly to zero for the growing integration distance. Consequently, the effect of individual gravity anomalies vanishes more rapidly with distance from the point of interest, and the numerical evaluation of the convolution integral may thus be truncated much closer to the point of interest to achieve the same accuracy as with the classical Stokes approach. Thus the evaluation of the generalized Stokes convolution integral requires less computational effort.

Various modification schemes may also be used for practical evaluation of the generalized Stokes convolution integral. It turns out that if the classical Molodenskij modification is used with modification of degree M, the modified S^M has exactly the same shape as the original Molodenskij-modified S. However, the upper bound of the truncation error for the same radius of integration ψ_0 is significantly smaller.

We have considered two generic models ($\tilde{N}^{M'}$ and $\tilde{N}^{M''}$) for the modification of Stokes's formula for a higher-degree reference field. For unbiased data, the first model is locally unbiased to the degree of modification (L) which equals to the maximum degree of harmonic coefficients, while the second model is locally unbiased to the degree ($M \leq L$) of the reference field. The error of $\tilde{N}^{M'}$ is therefore independent of the choice of M, while the error of $\tilde{N}^{M''}$ is generally dependent on M. The choice between the two models, and the choice of degree M for $\tilde{N}^{M''}$ are still open questions. However, for Molodenskij's modification, the two estimators coincide. The Examples 2 and 4 in Sjöberg [1987] show that the biased least-squares estimator is superior to the unbiased estimator in the limiting case of a vanishing cap size. These examples also indicate that a low-degree reference field is to be preferred in this particular case.

We have derived a new type of least-squares estimator, namely the locally best one. Its application is restricted by the limited knowledge of the local error covariance functions of terrestrial gravity, the correlations among the potential coefficients, and the high degree spectrum of gravity. These limitations are considerably relaxed in the more modest global least-squares modifications.

Acknowledgement

We wish to acknowledge the support of the Natural Sciences and Engineering Research Council (NSERC) of Canada for this research. The work was partly conducted while Sjöberg was a visiting scientist at the University of New Brunswick. We are indebted to Mr. Bernard Chovitz for his many thoughtful comments and to Ms. Wendy Wells for her flawless word processing of the many iterations of this paper.

References

Christou, N., P. Vaníček, and C. Ware, Geoid and density anomalies, *EOS, Transactions of the AGU, 70*(22), 625-631, 1989.

Heck, B., On the non-linear geodetic boundary value problem for a fixed boundary surface, *Bulletin Géodésique, 63* (1), 57-67, 1989.

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

Jekeli, Ch., Reducing the error of geoid undulation computations by modifying Stokes' function, Department of Geodetic Science, The Ohio State University Report No. 301, Columbus, Ohio, 1980.

Lachapelle, G., Estimation of disturbing potential components using a confined integral formula and collocation, Second Int. Summer School in the Mountains, Ramsau, 1977.

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

Moritz, H., Geodetic Reference System 1980, *Bulletin Geodesique,57* (3), 395-405, 1980.

Nagy, D. and M. Paul, Gravimetric geoid of Canada, *Proceedings of the Symposium on the Earth's Gravity Field and Secular Variations in Position,* Eds. R.S. Mather and P.V. Angus-

Leppan, Australian Academy of Science, IAG, Sydney, Australia, 20-30 November, 188-201, 1973.

Sjöberg, L.E., Least squares modification of Stokes' and Vening-Meinesz' formulas by accounting for truncation and potential coefficient error, *Manuscripta Geodaetica, 9*, 209-229, 1984a.

Sjöberg, L.E., Least squares modification of Stokes' and Vening-Meinesz' formulas by accounting for errors of truncation, potential coefficients and gravity data, Department of Geodesy Report No. 27, The University of Uppsala, Uppsala, 1984b.

Sjöberg, L.E., The Fennoscandian land uplift spectrum and its correlation with gravity, Proceedings of the IAG Symposia, IUGG, Hamburg, F.R.G., 1983, Department of Geodetic Science and Surveying, The Ohio State University, Columbus Oh, Vol. 1, 166-179, 1984c.

Sjöberg, L.E., Comparison of some methods of modifying Stokes' formula, *Boll. Geod. Sci. Aff., 45* (3) 229-248, 1986.

Sjöberg, L.E., Refined least squares modification of Stokes' formula, proceedings of the XIX General Assembly of the IUGG, the IAG Scientific Meeting GSM3I "The Challenge of the cm-Geoid — Strategies and State of the Art," Vancouver, B.C., 9-22 August, 1987.

Smith, R.W., Department of Defense World Geodetic System 1984 - Its definition and relationship with local geodetic systems, The Defense Mapping Agency Technical Report 8350.2, 1988.

Vaníček, P. and E.J. Krakiwsky, *Geodesy: the Concepts*, 2nd edition, North Holland, Amsterdam, 1986.

Vaníček, P. and A. Kleusberg, The Canadian geoid — Stokesian approach, *Manuscripta Geodaetica, 12* (2), 86-98, 1987.

Vaníček, P., A. Kleusberg, R.G. Chang, H. Fashir, N. Christou, M. Hofman, T. Kling, and T. Arsenault, The Canadian Geoid, Department of Surveying Engineering Technical Report No. 129, University of New Brunswick, Fredericton, N.B., 1986.