

**HIGH PRECISION GRAVITY ANOMALY AND SEA SURFACE  
HEIGHT ESTIMATION FROM GEOS-3/SEASAT SATELLITE  
ALTIMETER DATA**

by

**Cheinway Hwang**

**Report No. 399**

**Department of Geodetic Science and Surveying  
The Ohio State University  
Columbus, Ohio 43210-1247**

**July, 1989**

10

Blank

## ABSTRACT

The global prediction of gravity anomalies and sea surface heights and their rigorous accuracy estimates from satellite altimeter data were carried out in the ocean area covering  $\phi = 72^\circ$  to  $-72^\circ$  and  $\lambda = 0^\circ$  to  $360^\circ$ . The data used is the OSU existing combined Geos-3/Seasat data, which was updated by introducing sea surface topography correction, equatorial radius correction, height bias correction and permanent tidal correction and OSU86E reference field to degree 180 for this study. The use of the supercomputer, incorporating with the vectorization technique, has significantly reduced the computer time and a total of 12 CPU hours (on CRAY) was consumed by the production work. The global prediction was carried out in such a way that the sea surface heights were first adjusted by the cross-over arcs method and then the predictions were made by the least squares collocation method. From the tests focusing on various issues, we conclude that the local arc adjustment can remove the biases of arcs with wavelength of 400–500 km and the optimal altimeter data density is 400 points within a prediction cell of  $0.5^\circ$  with  $0.25^\circ$  border.

A total of 2,322,080 point values with  $1/8^\circ$  grid interval were predicted. The mean standard deviations are 12.11 mgals for anomaly and 8 cm for sea surface height. The comparison of the predicted and ship measured anomalies at the 2011 points in the Bermuda area has an RMS difference of 15.9 mgals. The plots of anomalies show that many undetected gravity signatures in Rapp's (1985) work now have been recovered.

## FOREWORD

This report was prepared by Mr. Cheinway Hwang, Graduate Research Associate, Department of Geodetic Science and Surveying, The Ohio State University. This study was supported by Exxon Co., International through The Ohio State University Research Foundation Project Number 720809. This project is under the direction of Richard H. Rapp, Professor, Department of Geodetic Science and Surveying.

Some computer funds for this study were supplied through the Instruction and Research Computer Center. The funds for the supercomputer resources were provided by the Ohio Supercomputer Center through grant pas 160.

This report is a slightly modified version of a thesis submitted in partial fulfillment of the requirements for the degree Master of Science in the Graduate School of The Ohio State University.

The reproduction and distribution of this report were carried out with funds supplied, in part, by the Department of Geodetic Science and Surveying.

## TABLE OF CONTENTS

ABSTRACT.....	ii
FOREWORD.....	iii
LIST OF FIGURES.....	vi
LIST OF TABLES .....	v
CHAPTER	PAGE
I. INTRODUCTION.....	1
II. THE UPDATE OF THE EXISTING COMBINED GEOS-3/SEASAT DATA BASE.....	2
2.1 The existing OSU Geos-3/Seasat data base.....	2
2.2 The update of the current Geos-3/Seasat data base .....	3
III. METHOD OF GRAVITY ANOMALY AND SEA SURFACE HEIGHT ..... RECOVERY .....	12
3.1 Least squares collocation.....	12
3.2 Global covariance models.....	14
3.3 Local covariance model .....	19
3.4 The local adjustment of observed sea surface height .....	21
3.5 Atmospheric correction.....	29
IV. USE OF SUPERCOMPUTER IN THE GLOBAL PREDICTION .....	30
4.1 Computational advance from the supercomputer.....	30
4.2 Vectorization of FORTRAN program .....	32
4.3 Special analysis of some computational kernels for geodetic problems .....	35
4.4 Performance analysis for the production program .....	43
V. THE OPTIMAL PARAMETER FOR PRODUCTION MODE BASED ON VARIOUS TESTS .....	48
5.1 Introduction.....	48
5.2 Bias removal from local arcs .....	48
5.3 Adjustment blocksize and prediction cellsize determination in the Bermuda area.....	67
5.4 Consistency between the observed SSHs and reference geoid heights and the determination of maximum acceptable residual SSH in a prediction cell.....	78

5.5	Matrix singularity problem and its solution in least squares collocation for the global production work.....	82
5.6	Other considerations for altimeter data .....	93
5.7	Summary of tests in the Bermuda area and the final decision for the production procedure.....	93
VI.	GLOBAL PREDICTION OF ANOMALIES AND SEA SURFACE HEIGHTS.....	99
6.1	The initial prediction.....	99
6.2	Revised prediction runs for selected cells.....	106
VII.	STATISTICAL ANALYSIS AND COMPARISON OF THE CURRENT PREDICTION RESULTS WITH THOSE OF RAPP'S (1985).....	114
7.1	The eighth degree data.....	114
7.2	Correlation of the eighth degree data and determination of 30'x 30' mean anomalies and sea surface heights .....	138
VIII.	CONCLUSIONS AND RECOMMENDATIONS.....	145
	LIST OF REFERENCES.....	149

## LIST OF FIGURES

FIGURE	PAGE
2.1 Distribution of OSU adjusted SEASAT data base.....	7
2.2 Distribution of OSU adjusted Geos-3 data base.....	8
2.3 Spatial distribution of Set 3 (Engelis, 1987) 1° x 1° sea surface topography .....	9
2.4 Sea surface topography corrections based on Set 3 of Engelis (1987) with Interpolation for Unknown 1° x 1° cells. Contour interval is 20 cm.....	10
2.5 The geometry of various corrections for the existing Geos-3/Seasat data.....	11
3.1 Covariance functions of undulations with respect to degree 180 field and Tscherning/Rapp anomaly degree variance model.....	17
3.2 Cross-over difference $d_{jk}^{ob}$ between two arcs.....	23
3.3 Poor geometry of cross-overs and an isolated arc.....	26
3.4 Peak SSH observations along arc .....	29
4.1 The structure of computing environment in scalar mode .....	30
4.2 The structure of the supercomputing environment .....	31
4.3 Performance of various matrix inversion routines on IBM and CRAY; N is the dimension of the inverted matrix .....	36
4.4 Performance of matrix product routine .....	39
4.5 Performane of matrix-vector product routine .....	42
4.6 CPU time of collocation prediction part for a 4° x 4° area with standard global production mode, as a function of the number of data points used .....	47
5.1 (a) Step function of bias for large blocksize.....	49
(b) Step function of bias for small blocksize .....	49
5.2 (a) 4° adjustment block.....	50

(b) 2° adjustment block.....	50
5.3 Block with two independent groups of arcs.....	52
5.4 Effect of increasing blocksize of bias adjustment region .....	53
5.5 Example of a bias adjustment block with two independent groups of arcs (Block A) .....	54
5.6 Predicted anomalies from Rapp's (1985) production results.....	58
5.7 Predicted anomalies from 5° adjustment .....	59
5.8 Predicted anomalies from 2° adjustment .....	60
5.9 Predicted anomalies from 4° adjustment .....	67
5.10 Differences in anomalies between 2° and 5° adjustments. Circle indicating the largest difference .....	62
5.11 Difference in anomalies between 2° adjustment and Rapp's (1985) result .....	63
5.12 Difference in anomalies between 5° adjustment and Rapp's (1985) result .....	64
5.13 Distribution of ship measured anomalies in the Bermuda area.....	70
5.14 Bathymetry in the Bermuda area. CI = 100 meters.....	71
5.15 Profiles of anomalies for choice 1 .....	75
5.16 Profiles of anomalies for choice 2.....	76
5.17 Profiles of anomalies for choice 3.....	77
5.18 Profiles of observed SSH and reference geoid heights. ....	80
5.19 Profiles of observed SSH and reference geoid heights. ....	81
5.20 Point distribution in the cell where matrix singularity occurred.....	82
5.21 Altimeter data distribution near the Kuril Trench .....	86
5.22 Predicted anomalies in case of gaps of point distribution. Points in the map indicate altimeter data points. CI = 10 mgals .....	87



5.23	Predicted anomalies in case of multiplying 3 and 4 for Seasat and Geos-3's noises. CI = 10 mgals.....	88
5.24	Predicted anomalies in the Bermuda area using noise scaling factors: 1 for Seasat, 1 for Geos-3.....	90
5.25	Predicted anomalies in the Bermuda area using noise scaling factors: 5 for Seasat, 7 for Geos-3.....	90
5.26	Predicted anomalies in the Bermuda area using noise scaling factors: 10 for Seasat, 15 for Geos-3.....	91
5.27	Predicted anomalies in the Bermuda area from the standard production mode (see text), highest anomaly is 371.83 mgals. CI = 10 mgals .....	97
5.28	Profiles of anomalies for the standard production mode in the Bermuda area.....	98
6.1	Data processing in various systems for the production work .....	100
6.2	File numbers, block number and numbers of altimeter data points for the global production work runs.....	101
6.3	Sketch of file locations for a prediction block.....	103
6.4	The locations of partitioned altimeter files (on CRAY discs).....	104
6.5	The required altimeter data sets (on CRAY) for prediction area at $0^\circ < \phi < 40^\circ$ , $160^\circ < \lambda < 200^\circ$ .....	104
6.6	The predicted anomalies from initial production work. Track errors remained in the adjusted data. CI = 10 mgals .....	108
6.7	The four suspected arcs responsible for linear features in Figure 6.6. The track numbers 1=861 2=416 3=172 4=5008 The intersections between the line crossing the 4 tracks and the 4 tracks are starting points of the profiles in Figures 6.8 (a) and (b).....	109
6.8	The SSH along tracks 861, 416, 172, and 5008 before and after adjustment.....	110
6.9	The predicted anomalies after track 172 was removed. CI=10 mgals.....	111
7.1	Anomaly map from the current production result in the Mariana Trench and Magellan Seamounts. CI = 15 mgals .....	121

7.2	Anomaly map from Rapp's (1985) result in the Mariana Trench and Magellan Seamounts. CI = 15 mgals.....	122
7.3	Sea surface height map from the current production result in the Mariana Trench and Magellan Seamounts. CI = 1 meter .....	123
7.4	Sea surface height map from Rapp's (1985) result in the Mariana Trench and Magellan Seamounts. CI = 1 meter.....	124
7.5	Anomaly map from the current result in the Hawaiian Islands. CI = 15 mgals.....	125
7.6	Anomaly map from Rapp's (1985) result in the Hawaiian Islands. CI = 15 mgals.....	126
7.7	Sea surface height map from the current production result in the Hawaiian Islands. CI = 1 meter.....	127
7.8	Sea surface height from Rapp's (1985) result in the Hawaiian Islands. CI = 1 meter.....	128
7.9	Anomaly map from the current production result in the New England Seamount area. CI = 10 mgals .....	129
7.10	Anomaly mpa from Rapp's (1985) result in the New England Seamount area. CI = 10 mgals.....	130
7.11	Sea surface height map from the current result in the New England Seamount area. CI = 1 meter.....	131
7.12	Sea surface height map from Rapp's (1985) result in the New England Seamount area. CI = 1 meter .....	132
7.13	Bathymetry map in the New England Seamount area. CI = 200 meters .....	133
7.14	Potential degree variances from various sources in selected areas.....	137
7.15	Distribution of 30' x 30' cells where the absolute differences between the 30' x 30' mean anomalies from the current result and Rapp's (1985) result are greater than 15 mgals.....	138

## LIST OF TABLES

TABLE	PAGE
3.1 Correlation length with respect to 180 field and Tscherning/Rapp's anomaly degree variance model.....	18
4.1 CPU time comparison between a vectorized code and a non-vectorized code .....	34
4.2 Performance analysis for LINDS.....	37
4.3 (a) CPU time ratio in the cross-over adjustment part.....	45
(b) CPU time ratio in the collocation prediction part .....	45
5.1 (a) Biases corresponding to Figure 5.2 (a) .....	51
(b) Biases corresponding to Figure 5.2 (b).....	51
5.2 Quantities of interest for the different prediction choices A, B and C.....	56
5.3 Comparison of anomalies from various adjustment block sizes.....	57
5.4 Variations of anomalies for 5° and 2° adjustment .....	65
5.5 The results of 4° adjustment at $-16^{\circ} < \phi < -20^{\circ}$ , $256^{\circ} < \lambda < 260^{\circ}$ , predicted anomalies corresponding to Figure 5.9.....	67
5.6 Comparison between ship measurement and predicted anomalies.....	72
5.7 Standard deviations of predictions for choices 1, 2 and 3.....	72
5.8 Status of matrix inversion at $\phi = 55^{\circ}0$ , $\lambda = 166^{\circ}0$ .....	83
5.9 Status of matrix inversion at $\phi = 53^{\circ}0$ , $\lambda = 159^{\circ}0$ .....	84
5.10 Differences between ship measured and predicted anomalies based on various choices of minimum spacings (at $30^{\circ} < \phi < 34^{\circ}$ , $293^{\circ} < \lambda < 297^{\circ}$ ).....	85
5.11 Sample listing of anomalies based on various noise's scaling factors .....	89
5.12 Comparison of predicted anomalies and ship measurements based on various noise's scaling factors.....	92

5.13 (a) Summary of tests in the Bermuda area (cases).....	95
(b) Results (comparison between predicted anomalies and ship measurements).....	96
6.1 Cpu time for the initial block prediction .....	106
6.2 Summary of revised prediction areas .....	113
7.1 Number of points with $\Delta g \geq 300$ mgals that will be deleted according to depth criteria .....	116
7.2 The statistics of the global eighth degree data based on the point value acceptance criteria .....	117
7.3 Frequency distribution of the accepted anomalies .....	117
7.4 Standard deviation distribution of the accepted anomalies.....	118
7.5 Three areas where potential degree variances were computed .....	135
7.6 Average correlations of predicted anomalies and sea surface heights based on various numbers of points .....	140

## **CHAPTER I**

### **Introduction**

The role of satellite altimeter data in the recovery of the gravity anomalies and sea surface heights has been clearly demonstrated by Rapp (1979), Rapp (1983) and Rapp (1985). The capability of the recovery is, however, limited by the number of altimeter observations, the data accuracy, the computer resource and software's efficiency. Other important factors include the model of the gravity field, the consideration of sea surface topography and the satellite altimeter's instrument calibrations.

Rapp (1985) first used the combined Geos-3/Seasat altimeter data to perform the global recovery of gravity anomalies and sea surface heights. Due to computer resource and other considerations, he reduced the original 5.9 million observations to 1.1 million observations. Accuracy estimates for the predicted quantities were not rigorously computed. It is clear that the detailed information could be lost due to such simplification. With the advent of supercomputers, the above considerations become unnecessary. Thus, our goal now is to use the original combined Geos-3/Seasat data to exploit the full potential of the recovery of gravity anomalies and sea surface heights through the use of optimal procedures.

In addition to improved data handling, several correction terms, not considered by Rapp (ibid.) will be considered. These terms will be illustrated in the following chapter.

## CHAPTER II

### **The update of the existing combined Geos-3/Seasat data base**

#### **2.1 The existing OSU Geos-3/Seasat data base**

The development of the existing OSU Geos-3/Seasat data base can be summarized chronologically as follows:

- (i) Rowlands (1981) performed crossing-arc adjustments of Seasat data provided by the Jet Propulsion Laboratory. The adjustments were split into two parts, i.e., the primary adjustment and local adjustment. Approximately 2.2 million points were obtained from the adjustments. Figure 2.1 shows the distribution of Seasat arcs.
- (ii) In 1982, Cruz improved Rowland's local adjustment but kept Rowland's primary adjustment unchanged (Rapp, 1985).
- (iii) Liang (1983) carried out the adjustment and combination of Seasat data from (ii) and the 3.5 year Geos-3 data from National Geodetic Survey. The adjustment is implemented in such a way that the primary arcs of Seasat were held fixed and Geos-3 data were treated as local observations which were "forced" to fit the Seasat data. After the adjustment, the two data sets were based on the same system and can be regarded as one data set.

At this point, 3.7 million Geos-3 observations were obtained. Therefore, a total of approximately 5.9 million combined Seasat/Geos-3 observations were available. The Seasat data and Geos-3 data are stored in two different tapes. The distribution of the adjusted Geos-3 data is given in Figure 2.2.

- (iv) Liang (1984) readjusted some of Cruz's primary arcs (item (i) and (ii)) since 53 Seasat's primary arcs were found to have cross-over discrepancies greater than 40 cm (Rapp, 1985, appendix A). Only the observations along the 53 arcs were

corrected with the new biases associated with these particular arcs. Liang then merged the Seasat and Geos-3 data so that the observations in a particular geographical region will reside in the same file. The result is the current OSU combined Geos-3/Seasat data base.

## 2.2 The update of the current Geos-3/Seasat data base

Considering the expected accuracies of this study, several corrections for altimeter data should be applied. The use of new reference undulations (to be described later) is also necessary. These updates are summarized as follows:

### (i) Sea Surface Topography correction

In the method of recovering gravity anomalies from satellite altimeter, the primary data quantity of interest is the geoid undulation. The geoid is an equipotential surface which deviates from the sea surface by a term called sea surface topography (SST). SST can be divided into a time invariant (stationary) part and a time dependent part. For the purpose of this study, we assume that the time dependent part is removed by averaging data at different time intervals. The stationary SST has been estimated by a number of authors (Litvin (1974), Levitus (1982)) using oceanographic data.

For this study we use the modified SST of Levitus described as SET3 by Engelis (1987). This data set consists of 30922 estimated  $1^\circ \times 1^\circ$  mean values, including data in the Mediterranean Sea. The mean value of 2.01 meters has been removed from the data and the data has a RMS value of 62.4 cm. The spatial distribution of Engelis' SET3 data is shown in Figure 2.3.

Using Levitus' SST data, Rapp (1985) found that the effect of SST on gravity anomalies was on the order of  $\pm 2$  mgals which is far below the accuracy of the predicted anomaly. However, since SST creates a long wavelength effect on the predicted anomalies, it should be removed from the observed sea surface height (SSH).

For oceanwide corrections of SST to altimeter data, two steps were carried out:

- a. Interpolate the mean SST in the blocks where SST are not available from the 5 closest mean values surrounding the block of interest. The program used for interpolation is GEOGRID written by Forsberg (1982). The method of interpolation is based on

$$SST_p = \frac{\sum_{i=1}^5 \frac{SST_i}{r_i^2}}{\sum \frac{1}{r_i^2}} \quad (2-1)$$

where  $SST_p$  is the interpolated value,  $SST_i$  is the value used and  $r_i$  is the distance between two blocks (center to center). This step leads to a complete SST data set in the ocean. A contour map of such data is given in Figure 2.4.

- b. Interpolate the point SST to altimeter data point from the complete  $1^\circ \times 1^\circ$  data set using a bilinear interpolation procedure.

(ii) Equatorial radius correction

The current Geos-3/Seasat data were referenced to the ellipsoid of the Geodetic Reference System 1980 ( $a = 6378137.0$  m,  $f = 1/298.257222101$ ). However, a more accurate equatorial radius (Rapp, 1987) of 6378136.2 m will be used in this study. We designate  $a_{old} - a_{new} = da$ .

(iii) Seasat's bias correction

Based on the data analysis in the Bermuda area, Kolenkiewicz and Martin (1982) found that the estimated height bias of Seasat altimeter was  $0.0 \pm 0.07$  meter, which should be added to Seasat's SSH. For the existing Geos-3/Seasat data, 11 cm had been assumed to be Seasat's height bias and this quantity apparently is not consistent with Kolenkiewicz and Martin's result. Therefore, 11 cm should be subtracted from the existing Geos/Seasat's SSH. (Note: Due to the way of combining Geos-3 and Seasat data, Geos-3 data should also receive such correction). We designate such bias as  $b$ .



(iv) Permanent tidal correction

The tidal effects on the satellite altimeter measurement consist of two parts: the ocean tides and the solid Earth tides. The ocean tides can be estimated through some models such as the Schwiderski model. The solid Earth tides are associated with the elastic deformation of the earth and are induced from the sun and moon. Let  $\overline{\Delta h_s}$  be the average (constant) sun induced earth tide and  $\overline{\Delta h_m}$  be the average (constant) moon induced earth tide, we wish to have a correction  $\Delta h_c$  to  $\overline{\Delta h_s}$  and  $\overline{\Delta h_m}$  so that

$$\overline{\Delta h_s} + \overline{\Delta h_m} + \Delta h_c = 0 \quad (2-2)$$

A sea surface with the  $\Delta h_c$  correction refers to a mean surface (i.e., in the presence of the constant effect of the sun and moon).  $\Delta h_c$  can be found in Rowlands (1981):

$$\Delta h_c = 0.124 \left( \frac{3}{2} \sin^2 \phi - \frac{1}{2} \right) \quad (\text{meter}) \quad (2-3)$$

Note that a sign error exists in equation (12) of Rowlands (ibid.).

Our goal now is to represent the 'geoid' in the absence of the sun and moon, thus the constant effect  $\Delta h_c$  has to be removed from the satellite altimeter measurement. Let  $S_I$  be the sea surface height corrected for the ocean tides, the various heights during the development of the existing Geos-3/Seasat data can be summarized as follows:

- Rowlands (1981):  $S_{II} = S_I - (\Delta h_s + \Delta h_m + \Delta h_c) + \Delta h_c$
- Cruz (1982):  $S_{III} = S_I - (\Delta h_s + \Delta h_m + \Delta h_c) = S_{II} - \Delta h_c$
- Liang (1983):  $S_{IV} = S_{III}$
- Liang (1984):  $S_V = S_{IV}$

Apparently,  $S_{II}$  represents the height in the absence of the sun and moon while  $S_{III}$ ,  $S_{IV}$ ,  $S_V$  have included the constant effect  $\Delta h_c$ . For the purpose of this study, we wish to have the  $S_{II}$  system, thus  $\Delta h_c$  must be added to the sea surface heights of the existing altimeter data (i.e.,  $S_V$ ).

(v) Use of new reference undulations

The existing Geos-3/Seasat data has reference undulations based on OSU81 to degree 180. The new reference undulations based on OSU86E (Rapp and Cruz, 1986) to degree 180 will be used instead. To do this, program F388 in Rapp's program library was used to generate a global 1/4 gridded undulations from OSU86E to degree 180 (The equation used for this is (3-5)). Then, interpolation of undulation was made for each altimeter data point using a bicubic spline interpolation procedure with 16 closest grid points.

As a summary for these updates, let SSH be the sea surface height before corrections and N be the undulation implied by the sea surface height after corrections; we have:

$$N = SSH - SST + da - b + \Delta h_c \quad (2-4)$$

Now we can treat the corrected sea surface heights as undulations which will be used for our global gravity anomaly and undulation recovery. The geometry and sign conventions of the corrections are shown in Figure 2.5

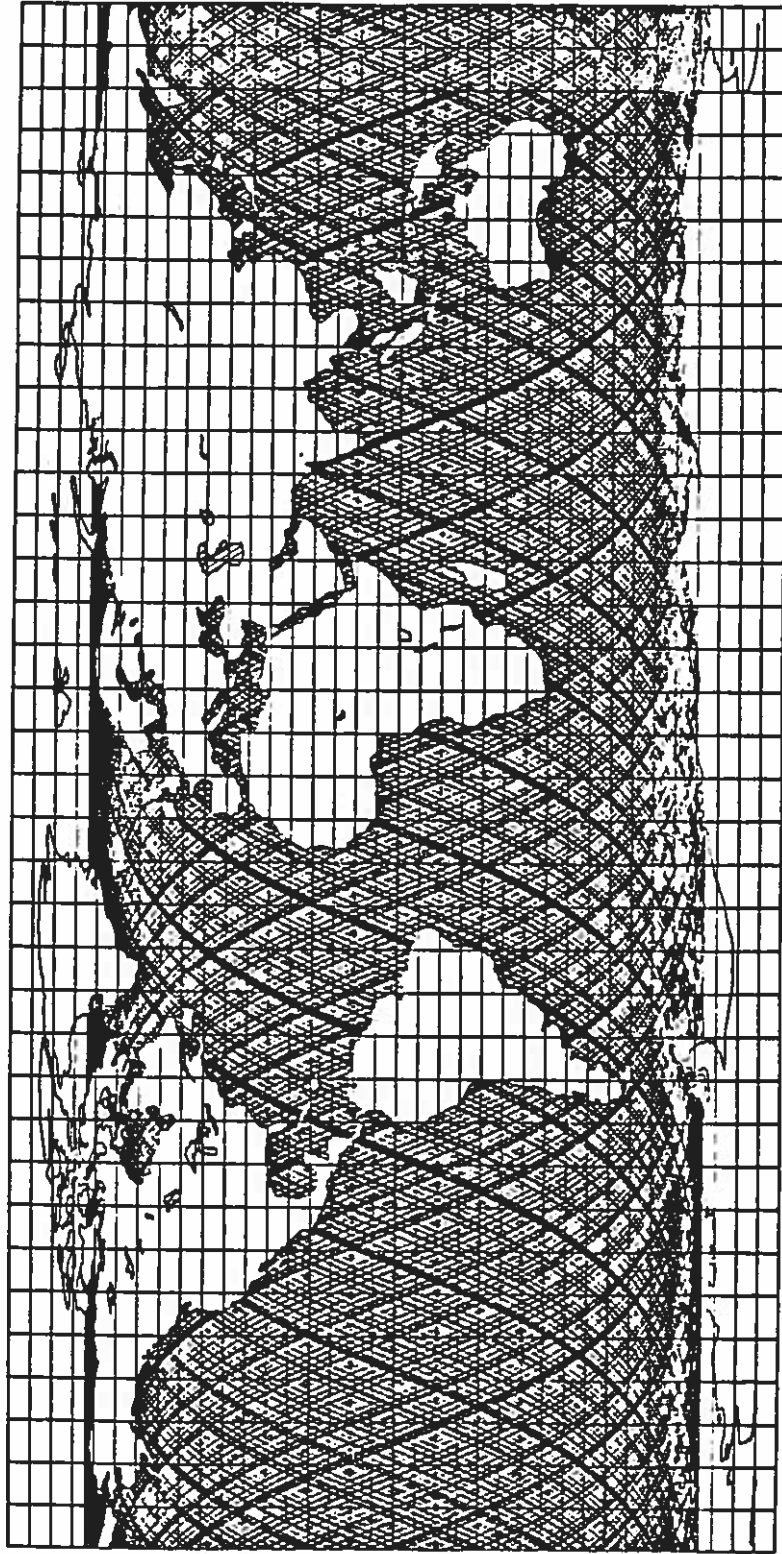


Figure 2.1 Distribution of OSU adjusted SEASAT data base

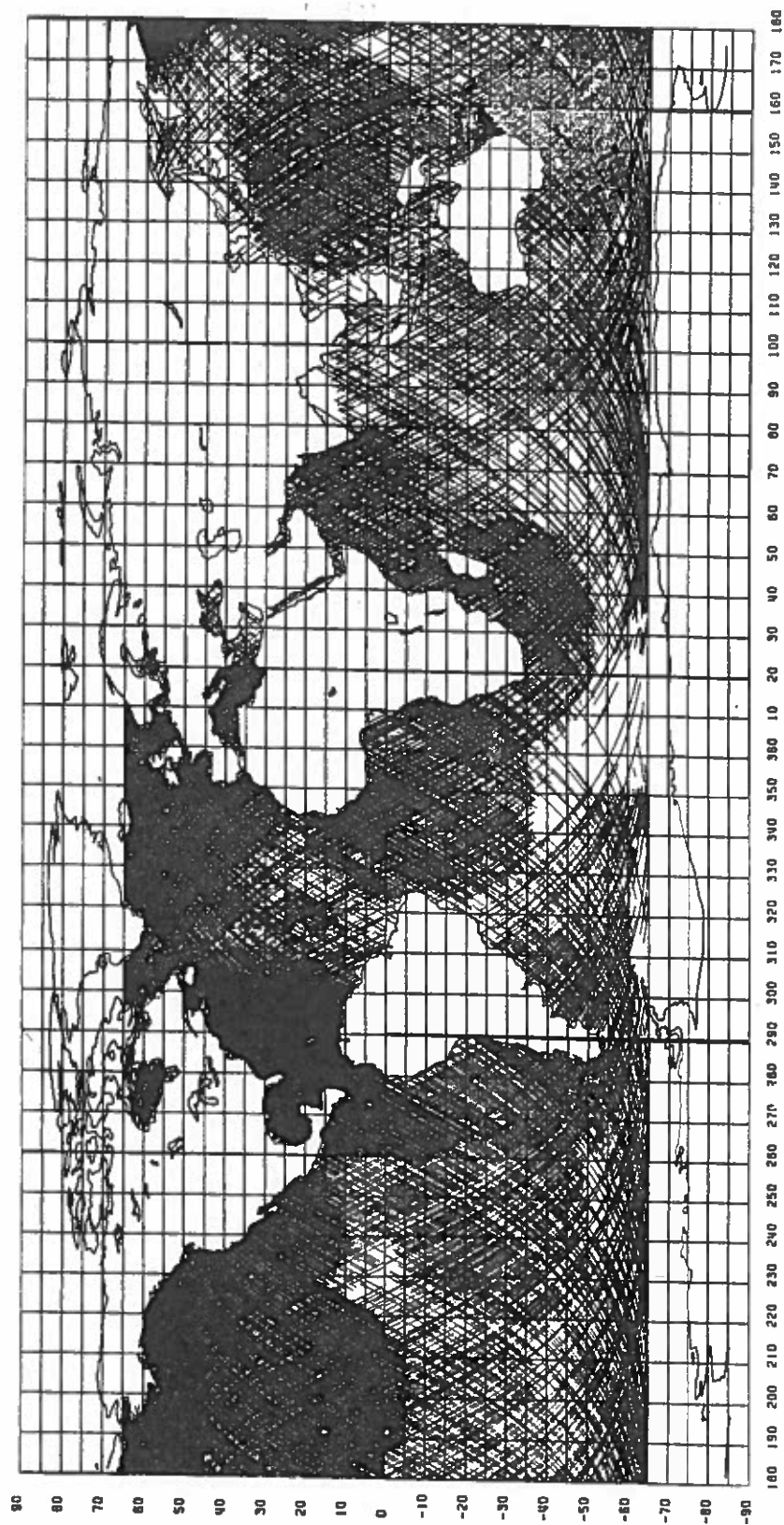


Figure 2.2 Distribution of OSU adjusted GEOS-3 data base

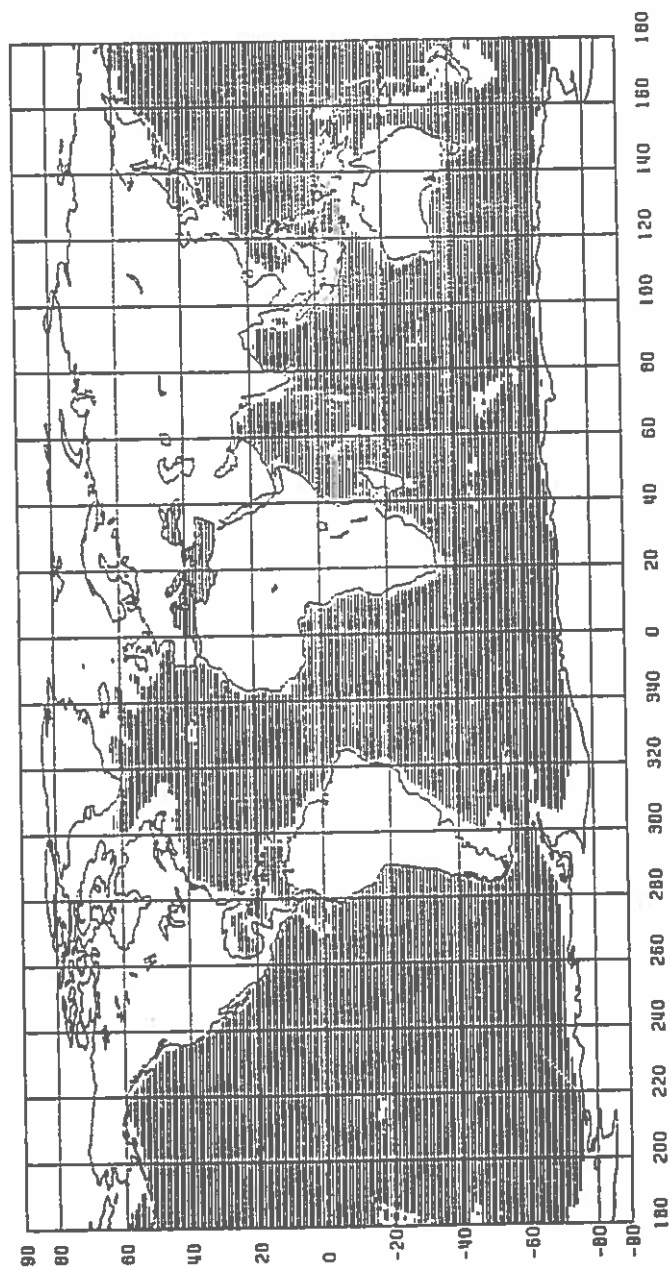


Figure 2.3 Spatial distribution of Set 3 (Engelis, 1987)  $1^\circ \times 1^\circ$  sea surface topography

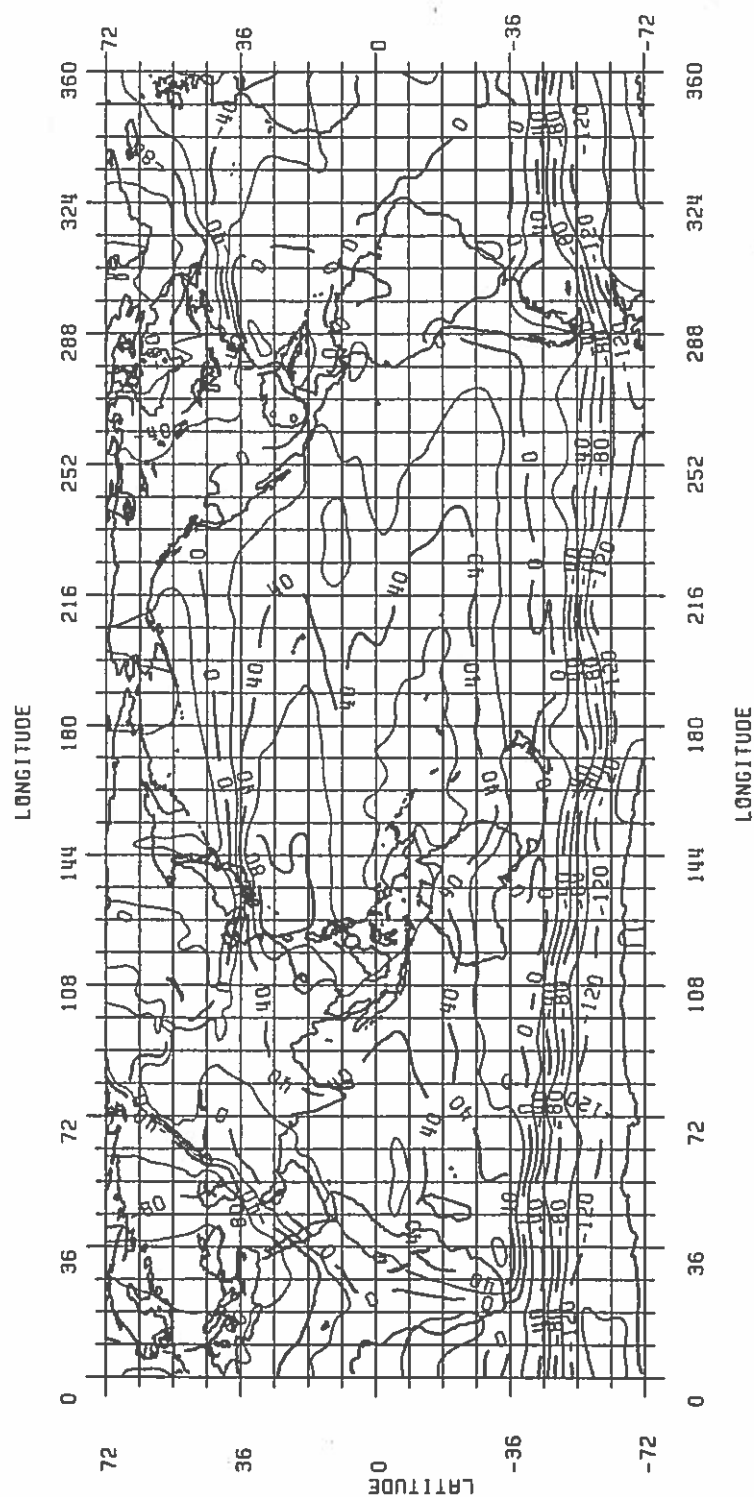


Figure 2.4 Sea surface topography corrections based on Set 3 of Engelis (1989) with Interpolation for Unknown  $1^\circ \times 1^\circ$  cells. Contour interval is 20 cm.

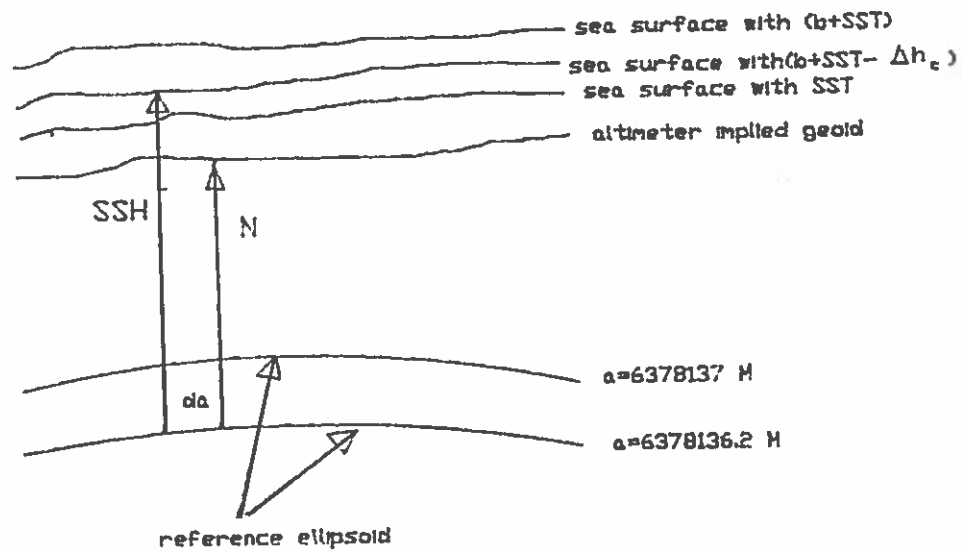


Figure 2.5 The geometry of various corrections for the existing Geos-3/Seasat data

## CHAPTER III

### Method of gravity anomaly and sea surface height recovery

#### 3.1 Least squares collocation

The method used in this study was the well-known least squares collocation which now takes altimeter implied geoid undulations as observables. A remove-restore procedure was introduced using a set of potential coefficients. Our basic equations are (following Rapp (1985)):

$$\hat{\Delta g} = \underline{C}_{gh} (\underline{C}_{hh} + \underline{D})^{-1} (\underline{h} - \underline{h}_R) + \Delta g_R \quad (3-1)$$

$$M_{\hat{\Delta g}}^2 = \underline{C}_{gg} - \underline{C}_{gh} (\underline{C}_{hh} + \underline{D})^{-1} \underline{C}_{hg} \quad (3-2)$$

where

- $\Delta g_R$  - The reference anomaly (at the prediction point) computed from a given set of potential coefficients up to a certain degree.
- $\underline{h}_R$  - The column vector of reference undulations of altimeter implied geoid undulations computed from a given set of potential coefficients up to a certain degree.
- $\underline{h}$  - The column vector of altimeter implied geoid undulations
- $\underline{D}$  - The diagonal matrix containing diagonal elements from the error variances of observables, i.e.  $\underline{h}$ .
- $\underline{C}_{hh}$  - The autocovariance matrix of observables  $\underline{h}$ .
- $\underline{C}_{gh}$  - The cross covariance matrix between predicted quantity (g) and observables  $\underline{h}$ .
- $\hat{\Delta g}$  - The predicted anomaly.
- $M_{\hat{\Delta g}}$  - The accuracy estimate of the predicted anomaly.
- $\underline{C}_{gg}$  - The autocovariance matrix of the predicted anomaly.

Eqns (3-1) and (3-2) are valid for the predicted point anomaly  $\Delta g$  and its accuracy  $M_{\Delta g}$ .



Similar equations for the predicted geoid undulations are the following:

$$\hat{S} = \underline{C}_{sh} (\underline{C}_{hh} + \underline{D})^{-1} (\underline{h} - \underline{h}_R) + S_R \quad (3-3)$$

$$\underline{M}_f^2 = \underline{C}_m - \underline{C}_{sh} (\underline{C}_{hh} + \underline{D})^{-1} \underline{C}_{hs} \quad (3-4)$$

The corresponding meanings of the notations in (3-3) and (3-4) can be immediately obtained by comparing (3-3) and (3-4) with (3-1) and (3-2).

It is necessary to introduce the practical computation of the reference anomaly  $\Delta g_R$  and undulation  $h_R$  and  $S_R$  (Note:  $h_R$  is the reference undulation for SSH and  $S_R$  is the reference undulation at the prediction point). Following Rapp (1979),  $\Delta g_R$  and  $h_R$  can be computed from a given disturbing potential coefficients,  $\bar{C}_{nm}$ ,  $\bar{S}_{nm}$ , obtained by subtracting the reference potential of a specific ellipsoidal system from the total potential, in the following spherical harmonic expansions:

$$\Delta g_R(r, \theta, \lambda) = \frac{KM}{a^2} \sum_{n=2}^N (n-1) \left(\frac{a}{r}\right)^{n+2} \sum_{m=0}^n (\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\cos \theta) \quad (3-5)$$

$$h_R(r, \theta, \lambda) = \frac{KM}{\gamma r} \sum_{n=2}^N \left(\frac{a}{r}\right)^n \sum_{m=0}^n (\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\cos \theta) \quad (3-6)$$

$a$  is a scaling factor which is usually the equatorial radius of a specific reference system,  $N$  is the maximum degree of harmonic expansion,  $KM$  is the product of gravitational constant and the mass of the Earth,  $r$  is the distance from geocenter to the point where the computation is made. For this study, we use the following parameters for the computation:

$$\begin{aligned} N &= 180 \\ KM &= 398600.5 \text{ Km}^3 \text{ s}^{-2} \\ f &= 298.257222101 \\ a &= 6378137 \text{ meters} \end{aligned} \quad (3-7)$$

For the global computation of reference gravity anomalies and geoid undulations, program F388TAPE in Rapp's program library was used. The computation was made at  $1/8^\circ$  grid interval from latitude  $90^\circ$  to  $-90^\circ$  and longitude from  $0^\circ$  to  $360^\circ$ . The CPU time consumed by calculating  $\Delta g_R$  is 27.0 minutes and the CPU time spent by calculating  $N_R$  is 27.3 minutes on the IBM 3081.

### 3.2 Global covariance models

Due to the use of remove-restore procedure, it is necessary to consider the coefficient errors of the reference field. The total covariance used in eqns (3-1) and (3-3) consists of two parts, i.e., a signal part and an error part:

$$\text{cov}(\psi) = \text{cov}_R(\psi) + \text{cov}_E(\psi) \quad (3-8)$$

Note that the covariance function is assumed to be isotropic and the spherical separation  $\psi$  is the unique variable in the covariance function.

For the error part, the error degree variance related to the reference field can be computed by (Rapp, classnotes, GS871)

$$\delta C_n = \left( \frac{KM}{a^2} \right)^2 (n-1)^2 \left( \frac{a}{R} \right)^{2(n+2)} \sum_{m=0}^n \left( \epsilon_{c_{nm}}^2 + \epsilon_{s_{nm}}^2 \right) \quad (3-9)$$

where  $\bar{\epsilon}_{c_{nm}}$  and  $\bar{\epsilon}_{s_{nm}}$  represent the errors of the coefficients  $\bar{C}_{nm}$  and  $\bar{S}_{nm}$  in the potential set, and  $n$  is the harmonic degree. Using OSU86C noise model (Rapp and Cruz, 1986), we put

$$a = R_E = 6371 \text{ km}$$

$$R = R_B = 6369 \text{ km} \quad (3-10)$$

where  $R_B$  is the Bjerhammer's sphere (Moritz, 1980).

For the signal part, the anomaly degree variance  $C_n$  can be determined from a specific model. For this study, we adopt Tscherning/Rapp's model (Tscherning and Rapp, 1974) having the form:

$$C_n = \frac{A(n-1)}{(n-2)(n+B)} \quad (3-11)$$

where  $A = 425.28 \text{ mgal}^2$ ,  $B = 24$

Another important parameter of the covariance function is  $s$  defined as

$$s = \frac{R_B^2}{R_E^2} \quad (3-12)$$

The value of  $s$  is specified to be 0.999617.

Having the error degree variance for the error part and the adopted anomaly degree variance for signal part, the three fundamental covariance functions between two arbitrary points can be expanded into a series of spherical harmonics as follows:

$$\text{cov}(\Delta g_p, \Delta g_q) = \sum_{n=2}^N \delta C_n s^{n+2} P_n(\cos \psi_{pQ}) + \sum_{N+1}^{\infty} C_n s^{n+2} P_n(\cos \psi_{pQ}) \quad (3-13)$$

$$\text{cov}(\Delta g_p, h_Q) = \frac{R_B^2}{r_p \gamma_Q} \sum_{n=2}^N \frac{\delta C_n}{n-1} s^{n+1} P_n(\cos \psi_{pQ}) + \frac{R_B^2}{r_p \gamma_Q} \sum_{N+1}^{\infty} \frac{C_n}{(n-1)} s^{n+1} P_n(\cos \psi_{pQ}) \quad (3-14)$$

$$\text{cov}(h_p, h_Q) = \frac{R_B^2}{\gamma_p \gamma_Q} \sum_{n=2}^N \frac{\delta C_n}{(n-1)^2} s^{n+1} P_n(\cos \psi_{pQ}) + \frac{R_B^2}{\gamma_p \gamma_Q} \sum_{N+1}^{\infty} \frac{C_n}{(n-1)^2} s^{n+1} P_n(\cos \psi_{pQ}) \quad (3-15)$$

In (3-13) through (3-15) the first part on the right hand side of the equation is the error part, second part being the signal part.

In the previous chapter, we mentioned that OSU86E field with expansion to degree 180 will be used for the reference field. Therefore, the noise model adopted for this study will be based on OSU86C which closely resembles OSU86E.

In addition to the parameters given by Tscherning and Rapp (ibid), Jekeli (1978) also recommended the parameters

$$\begin{aligned} A &= 343.3408 \text{ mgals}^2 \\ B &= 24 \\ s &= 0.9988961 \end{aligned} \quad (3-16)$$

Parameters in (3-16) associated with Tscherning/Rapp's anomaly degree variance model were extensively used by Kadir (1987) in his tests on the recovery of gravity anomalies and sea surface heights using fairly dense altimeter data.

Figure 3.1 illustrates the behavior of the covariance functions of undulations with respect to a degree 180 field and the anomaly degree variance model of (3-11). It can be seen that the covariance function based on Tscherning/Rapp's parameters decays faster than the covariance function based on Jekeli's parameters. When very dense data is used, the relative smoothness of the later covariance function can easily cause the matrix singularity during the collocation process. Therefore, it will be suggested that Tscherning/Rapp's parameters be used if one has to use dense data, or, if the working area is restricted to being small.

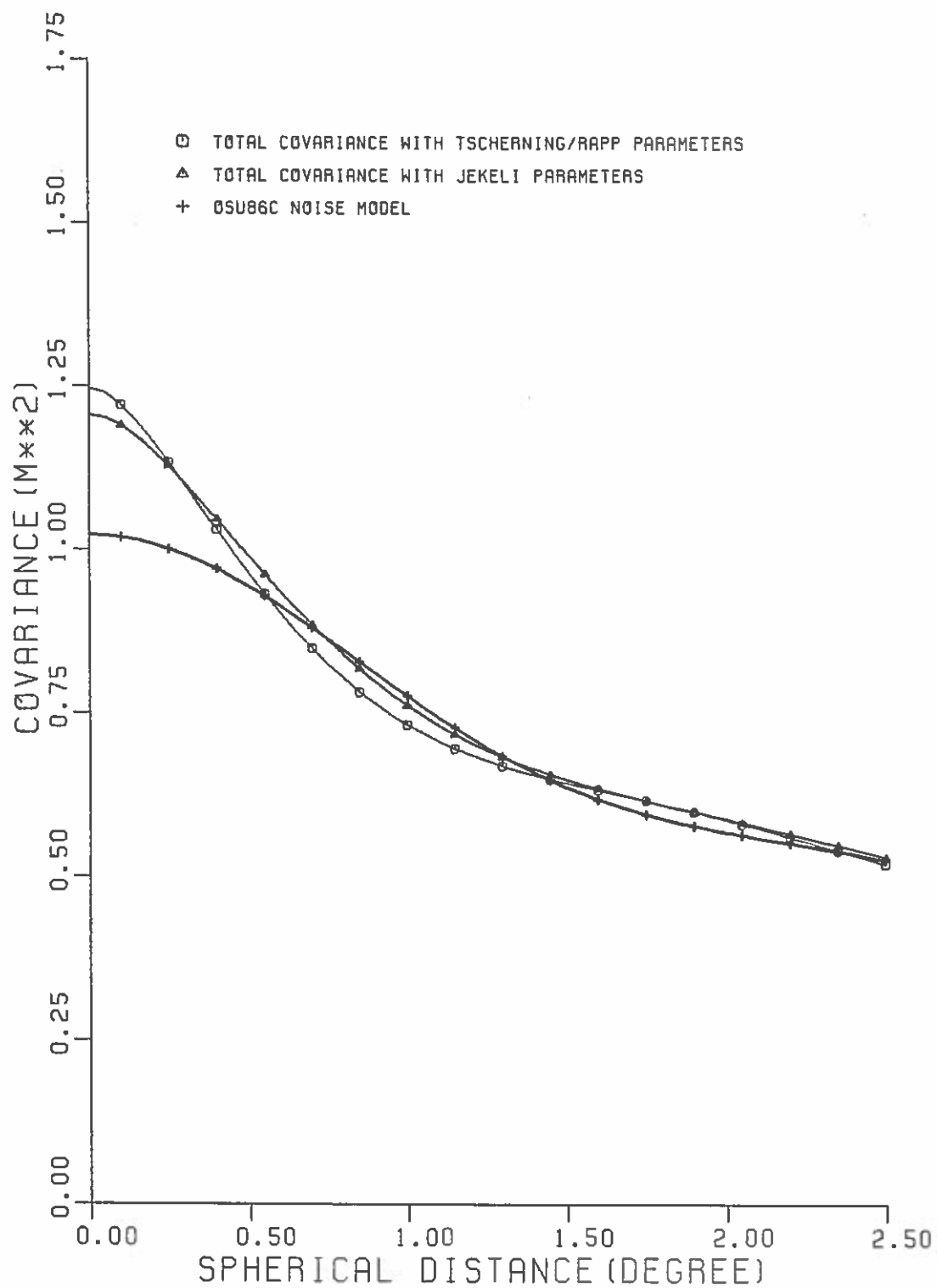


Figure 3.1 Covariance functions of undulations with respect to degree 180 field and Tscherning/Rapp anomaly degree variance model

Recalling the idea of recovering gravity anomalies using geoid undulations, clearly we are clearly dealing with heterogeneous gravimetric quantities. Therefore, it is not sufficient for the determination of prediction choice if the covariance function of undulation is discussed alone. As pointed out by Moritz (1980, p. 177), the prediction is accurate enough if the station distances are well below the correlation length, which is the spherical distance  $\xi$  having the feature:

$$C(\xi) = \frac{1}{2} C_0 \quad (3-17)$$

where  $C_0$  is the variance and  $C(\xi)$  is the covariance at a distance  $\xi$ . The discussion then will be concentrated on the correlation lengths of the various covariance functions. Table 2.1 shows the correlation lengths related to  $\text{cov}(\Delta g_p, \Delta g_Q)$ ,  $\text{cov}(\Delta g_p, h_Q)$  and  $\text{cov}(h_p, h_Q)$  in (3-13), (3-14) and (3-15), respectively.

Table 2.1 Correlation length with respect to degree 180 field<sup>+</sup> and Tscherning/Rapp's anomaly degree variance model

	model A <sup>*</sup>	model B <sup>**</sup>
$\text{cov}(\Delta g_p, \Delta g_Q)$	0.125°	0.208°
$\text{cov}(\Delta g_p, h_Q)$	0.339°	0.441°
$\text{cov}(h_p, h_Q)$	1.695°	1.877°

+ OSU86C noise model

\* Tscherning/Rapp's parameters

\*\* Jekeli's parameters

Again, we compare the correlation lengths from the two sets of parameters for the purpose of selecting a prediction procedure. Both sets indicate that the covariance function of undulations undergoes a relatively smooth change while the covariance function of anomalies has the fastest decay. Although Moritz (ibid.) observed the importance of correlation length to the prediction accuracy, he did not specifically discuss case where two kinds of covariance functions are used. For example, the prediction of anomalies from geoid undulations (eq. (3-1)) is exactly such a case.

Nevertheless, it is appropriate to assume that the covariance function having the shortest correlation length could dominate the accuracy. Thus, it is expected that  $\text{cov}(\Delta g_p, h_Q)$  will have an influence on the selection of a prediction cell (the cell where one individual collocation prediction is performed). This aspect will be discussed in Chapter 5.

### 3.3 Local covariance model

For a local prediction of gravity anomalies using satellite altimeter data, the global covariance function should be scaled to the local covariance function in the sense that the residual undulations should appropriately reflect the accuracies of the predicted quantities. A practical application of such concept to altimeter-related prediction has been given by Rapp (1985).

It is now necessary to explain the scaling procedure for this study: In a prediction cell, we first compute the variance of the residual undulations by

$$\sigma_h^2 = \sum \frac{(h_i^{\text{res}} - \bar{h})^2}{n} \quad (3-18)$$

where  $n$  is total number of points used for the prediction and  $\bar{h}$  is the mean of the residual undulations defined by

$$h_i^{\text{res}} = h_i - h_{Ri} \quad (3-19)$$

where  $h_i$  and  $h_{Ri}$  are the observed and reference geoid undulations, at point  $i$ , respectively.

Let  $\text{CNN}(0)$  be the variance of undulations based on (3-15). We have to apply a scaling factor  $\alpha$  to this quantity so that the resultant quantity will be identical to the variance of residual undulations. Thus  $\alpha$  is:

$$\alpha = \frac{\sigma_h^2}{\text{CNN}(0)} \quad (3-20)$$

The local covariance functions then can be constructed by multiplying (3-13), (3-14) and (3-15) by the scaling factor  $\alpha$ .

To ensure a realistic scaling factor, we assume that the predicted anomaly has a minimum anomaly variance of 400 mgal<sup>2</sup>. Thus a limiting scaling factor  $\alpha'$  will be the lowest value for actually scaling (3-13), (3-14) and (3-15).  $\alpha'$  is computed by

$$\alpha' = \frac{400}{\text{CGG}(0)} = 0.425 \quad (3-21)$$

for Tscherning/Rapp's covariance model (with their parameters), CGG (0) is the variance of the anomaly based on (3-13).

Having  $\alpha$ , we can re-write (3-1) and (3-2) as follows:

$$\hat{\Delta g} = (\alpha \underline{C}_{gh}) [(\alpha \underline{C}_{hh}) + \underline{D}]^{-1} (\underline{h} - \underline{h}_R) + \Delta g_R \quad (3-22)$$

$$M_{\Delta g}^2 = (\alpha \underline{C}_{gg}) - (\alpha \underline{C}_{gh}) [(\alpha \underline{C}_{hh}) + \underline{D}]^{-1} (\alpha \underline{C}_{hg}) \quad (3-23)$$

They can be reduced to:

$$\hat{\Delta g} = \underline{C}_{gh} \left( \underline{C}_{hh} + \frac{1}{\alpha} \underline{D} \right)^{-1} (\underline{h} - \underline{h}_R) + \Delta g_R \quad (3-24)$$

$$M_{\Delta g}^2 = \alpha \left[ \underline{C}_{gg} - \underline{C}_{gh} \left( \underline{C}_{hh} + \frac{1}{\alpha} \underline{D} \right)^{-1} \underline{C}_{hg} \right] \quad (3-25)$$

Forms of (3-24) and (3-25) indicate the problem related to matrix singularity. First of all, we recall the definition of  $\underline{D}$  matrix: The elements of  $\underline{D}$  are essentially the variances of the observations, which break into two categories:



a constant of 20 cm for Geos-3 observations  
 5 cm - 20 cm or higher for Seasat observations

A scaling factor of  $\alpha$  which is much larger than 1 then can reduce the variances of observations so that the resulting matrix  $1/\alpha \underline{D}$  has no effect on the covariance matrix  $\underline{C}_{hh}$ . Then, we investigate the case when a small prediction cell and dense data are used. Since some of the points are so close, the corresponding row vectors of  $\underline{C}_{hh}$  nearly become dependent. If  $1/\alpha \underline{D}$  has relatively large elements, the dependence can be reduced through the effect of  $(\underline{C}_{hh} + 1/\alpha \underline{D})$ . If  $\alpha$  is large, i.e.  $1/\alpha \underline{D}$  becomes small, then  $1/\alpha \underline{D}$  is no help for resolving the dependence problem. These statements will be verified by the tests performed in area of higher latitude where only Seasat data is available; see Chapter 5 for more discussion.

#### 3.4 The local adjustment of observed sea surface heights

The adjustment of altimeter data performed by Rowlands (1981) and Liang (1983) were on a global basis. The global adjustments have theoretically treated the major track problems in a global sense. However, the track problem can remain in a local area where we thought a further adjustment of the arcs is necessary. The goal of the local adjustment thus is to remove the possible track biases that were not removed by the global adjustments.

The arc adjustment can be theoretically broken into two categories:

- (i) Adjustment using cross-over differences of arcs
- (ii) Adjustment using least squares collocation with biases of track as parameters

Category (i) again can be broken into two parts, according to different models for cross-over differences.

Based on categories (i) and (ii), we present 3 models as follows:

Model A:

In this model, we assume that each arc has a constant bias; the adjusted sea surface height can be obtained by adding bias correction to the observed sea surface height as follows:

$$\hat{h}_{ij} = h_{ij}^{ob} + b_j \quad (3-26)$$

where  $\hat{h}_{ij}$  and  $h_{ij}^{ob}$  are the adjusted and observed sea surface heights, respectively, pertaining to point  $i$  along arc  $j$ . Superscript "ob" denotes observed quantities. If two arcs  $j, k$  have a cross-over difference at point  $i$ , then we form a cross-over difference by:

$$d_{jk}^{ob} = h_{ij}^{ob} - h_{ik}^{ob} = b_j - b_k \quad (3-27)$$

(3-27) clearly is due to the assumption that:

$$\hat{h}_{ij} = \hat{h}_{ik} \quad (3-28)$$

which means the two sea surface heights at the same point should be the same. A sketch of the cross-over between two arcs is shown in Figure 3.2. Note that there is no actual observation at point  $i$ . Observations at point  $i$  can be obtained by linear interpolation from points A and B, C and D, separately.

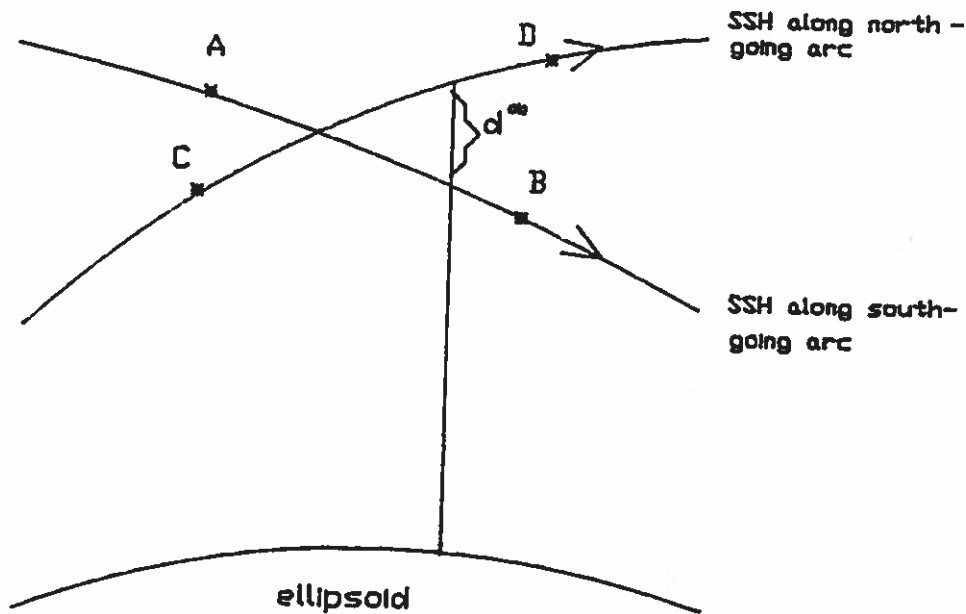


Figure 3.2 Cross-over difference  $d_{jk}^{ob}$  between two arcs

(3-27) is fundamental to the observation equation and the cross-over differences are the observables. Assuming  $n$  cross-over observations and  $q$  bias parameters, the observation equation can be set up by (Uotila, 1986):

$$V = A\hat{X} + L \quad (3-29)$$

where

$V$  = vector of  $n$ -elements containing the residuals of cross-over differences

$A$  =  $n \times q$  matrix containing elements 1, -1 and 0

$\hat{X}$  = vector of  $q$ -elements containing the bias parameters associated with arcs

$L$  = vector of  $n$ -elements containing the cross-over differences

The system in (3-29) has rank deficiency of one because all arcs are not fixed in the working area. A supplementary condition can be added to avoid the problem, for example, we assume the sum of all biases is zero (Knudsen, 1987), hence

$$\sum_{j=1}^q b_j = 0 \quad (3-30)$$

which implies the average of the corrections is zero. (3-30) yields the constraint that has to be added to (3-29):

$$C^T \hat{X} = 0 \quad (3-31)$$

Since 1 and -1 will simultaneously appear in any row of matrix A (the nature of cross-over difference) and C is a vector with elements "1" alone, we immediately see that  $AC=0$ . Therefore A and C are orthogonal. With the orthogonality property, the least squares solution of (3-29) with (3-31) is found by (Caspary, eqn (3-51), 1987):

$$\hat{X} = - \left( A^T P A + C C^T \right)^{-1} A^T P L \quad (3-32)$$

$$\Sigma_{\hat{X}} = \left( A^T P A + C C^T \right)^{-1} A^T P A \left( A^T P A + C C^T \right)^{-1} \quad (3-33)$$

where P is the weight matrix of cross-over differences obtained from error variances of SSH using error propagation. (The propagation is based on (3-27)). Since  $C C^T$  is  $q \times q$  matrix of elements "1", it has no substantial influence on the solution of (3-33) in terms of programming efficiency (since the process of adding  $C C^T$  to  $A^T P A$  is simple).

A crossover adjustment has been used by Rowlands (1981) and Liang (1983) who also augmented the model by adding "tilt" parameters and used different methods to handle the rank deficiency problem. cf: Rowlands (ibid.) and Liang (ibid.) and Knudsen (ibid.).

Model B:

This model again uses cross-over differences as observables. However, it is assumed that the bias is a function of time instead of a constant and the covariance

function of biases can be known through some model. The observable  $d_{jk}$  at time  $t_j$  and  $t_k$ , at tracks  $j$  and  $k$ , is

$$d_{jk}^{ob} = b_j - b_k \quad (3-34)$$

Since bias is a time varying quantity, for  $n$  cross-over differences, we actually have  $2n$  estimates for biases. Thus the biases can be obtained through (Jackson, 1979 and Knudsen, 1987)

$$\hat{X} = -C_x A^T (A C_x A^T + P^{-1})^{-1} L \quad (3-35)$$

$$\hat{\Sigma X} = C_x - C_x A^T (A C_x A^T + P^{-1})^{-1} A C_x \quad (3-36)$$

where  $C_x$  is the covariance matrix of biases with dimension  $2n \times 2n$ . Note that now  $A$  is  $n \times 2n$  matrix. The major difference between model A and B is the use of the covariance function (matrix) of biases, i.e.,  $C_x$ , which is known as "a priori" information. Knudsen (1987) suggested a model for evaluating such a function using Gaussian functions:

$$\text{cov}(S_a, S_n) = \sum_i A_i^2 \frac{1}{e} \left( \left( \frac{S_a}{\xi_{a,i}} \right)^2 + \left( \frac{S_n}{\xi_{n,i}} \right)^2 \right) \quad (3-37)$$

where  $S_a$  is along track distance and  $S_n$  is cross track distance (the distance can be computed from the time tag or the sequential point numbers along a track) at the point where the cross-over is formed.  $A_i$  is the amplitude of the error associated with the  $i^{\text{th}}$  phenomenon.  $\xi_{a,i}$  and  $\xi_{n,i}$  are along and cross track correlation lengths of errors associated with the  $i^{\text{th}}$  phenomenon.

A practical calculation using model B has been carried out by Knudsen (ibid.) who found that the RMS value of adjusted cross-over differences was less than that calculated from model A. cf. Knudsen (ibid., p. 624).

### Model C:

The cross-over adjustment based on model A or B fails if no cross-overs between arcs can be obtained. Or, even though the crossovers can be found the geometry of the arcs could be unstable, cf. Figure 3.3

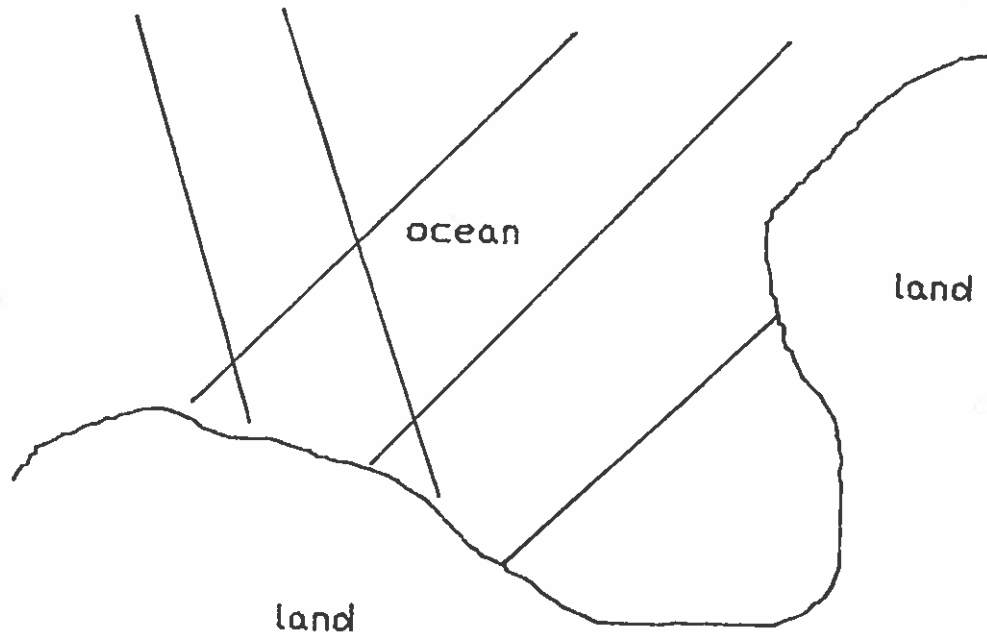


Figure 3.3 Poor geometry of cross-overs and an isolated arc

A solution to this problem is to incorporate the biases into the least squares collocation (lsc) solution. For the simplest case, we still assume one constant bias associated with one arc. Then, we have the lsc with parameters (Moritz, eq. (16-1), 1980).

$$l = A \hat{X} + BT + n \quad (3-38)$$

where

- $l$  - a vector containing  $m$  elements of observed sea surface heights
- $\hat{X}$  - a vector containing  $q$  bias parameters

- $A$  - a  $m \times q$  matrix containing only 0 and 1  
 $BT$  - a vector containing signal part of vector  $l$   
 $n$  - a vector containing noises of the observations

The solution of (3-38) is given by (ibid., p. 116)

$$\hat{X} = \left( A^T \bar{C}^{-1} A \right)^{-1} A^T \bar{C}^{-1} l \quad (3-39)$$

$$\Sigma_{\hat{X}} = \left( A^T \bar{C}^{-1} A \right)^{-1} \quad (3-40)$$

where  $\bar{C}$  is the total covariance of undulations, together with the variance of observations, i.e.

$$\bar{C} = C_{hh} + D \quad (3-41)$$

cf. eq. (3-1). Note that the definition  $A$  in (3-39) is different from that in (3-29), due to the different types of observables.

An a-priori weight matrix can be added to the bias parameters, as suggested by Tscherning et al (1985): The solution then becomes:

$$\hat{X} = \left( A^T \bar{C}^{-1} A + P_x \right)^{-1} A^T \bar{C}^{-1} l \quad (3-42)$$

$$\Sigma_{\hat{X}} = \left( A^T \bar{C}^{-1} A + P_x \right)^{-1} \quad (3-43)$$

where  $P_x$  is the weight matrix for the bias parameters.

Tscherning et al (ibid) has carried out bias determination using model A and model C in a test area located at  $38^{\circ}.25 < \phi < 39^{\circ}.75$  and  $295^{\circ}.25 < \lambda < 296^{\circ}.75$ . It was found that

the two results have an RMS difference of  $\pm 0.02$  m for the estimated biases except for a constant of 0.22 m.

At this point, three models have been discussed. However, only model A will be used here in the global prediction due to an extremely large data set. Model A features simplicity, efficient programming and easy debugging. Using Model A, Knudsen (1988) has developed a sequence of programs that can determine the cross-over differences, solve for the bias parameters and apply the bias corrections to the original altimeter data. With some modification in input/output and dimension statements, these programs can process a fairly large data set in an extremely efficient manner.

For the practical calculation of bias parameters based on Model A, two editing criteria have been suggested by Knudsen (1988, program documentations):

- (i) The mean distance between two successive points along a track is first calculated. Then the maximum distance within which a cross-over can be formed is 2.5 times this mean distance.
- (ii) The standard deviations of "peak" and "outlier" SSH observations along one particular arc are multiplied by a factor of 10.0. A "peak" is an SSH observation that satisfies the following sequence of equations:

$$\begin{aligned} & \left| 2H_i - (H_{i-1} + H_{i+1}) \right| > (2\text{err}_i + \text{err}_{i-1} + \text{err}_{i+1}) * 2.0 \\ & |H_i - H_{i-1}| > (\text{err}_i + \text{err}_{i-1}) * 2.0 \\ & |H_i - H_{i+1}| > (\text{err}_i + \text{err}_{i+1}) * 2.0 \end{aligned} \tag{3-44}$$

Distance ( $H_{i-1}, H_{i+1}$ ) < 15.0 km

$H_{i-1}, H_i, H_{i+1}$  are 3 consecutive SSH observations along an arc. cf. Figure 3.4.  $\text{err}$  denotes the standard deviation associated with a SSH observation.



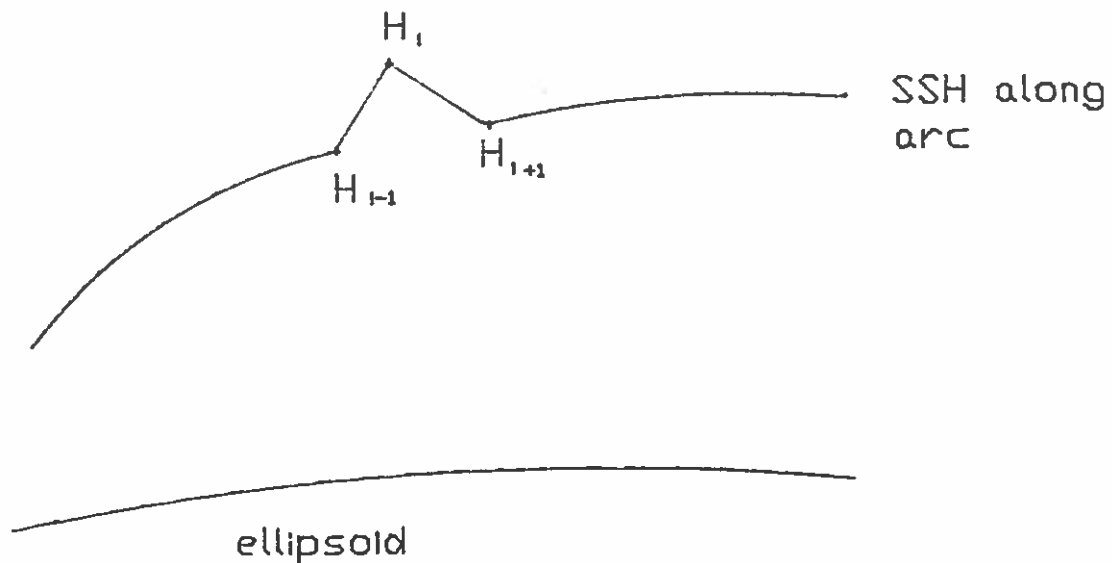


Figure 3.4 Peak SSH observations along arc.

An "outlier" is an SSH observation satisfying:

$$|DH_i - \overline{DH}| > STD * 2.5 + err_i \quad (3-45)$$

where  $\overline{DH}$  is the mean value of the residual SSHs and STD is the standard deviation of the residual SSHs in the area where the adjustment is performed.  $DH_i$  is the  $i^{\text{th}}$  residual SSH.

Such an enlargement of standard deviations will not propagate to lsc process, namely, the lsc still uses the original standard deviations of SSH.

### 3.5 Atmospheric correction

In order to compare the altimeter predicted anomalies with terrestrial measurements, an atmospheric correction must be taken into account (Rapp, 1985). Namely, 0.87 mgal must be subtracted from the altimeter predicted anomalies. However, all the values in the figures and tables in this study do not apply such corrections.

## CHAPTER IV

### Use of the supercomputer in the global prediction

#### 4.1 Computational advance from the supercomputer

A revolutionary advance in computational techniques has been achieved by the supercomputer. An extensive application of the supercomputer to industrial engineering, chemistry, physics and molecular dynamics, etc. has brought science into a new era. For geodetic problems, especially those related to matrix operations, such an advance has a profound influence.

In contrast to the traditional computer, the supercomputer uses its exclusive vector functional units to perform calculations. The application of vector hardware, including the famous tools such as pipeline and chaining, can actually linearize the computational time which could be in the exponential or quadratic form originally in a traditional scalar-operated machine. It is an important concept that the full application of supercomputer in essence needs a conceptual update of the roles of hardware, algorithm designer and application programmer (BCS, 1984). In the scalar mode, the architecture, algorithm and application may work almost independently. Such a case is indicated in Figure 4.1 (ibid.)

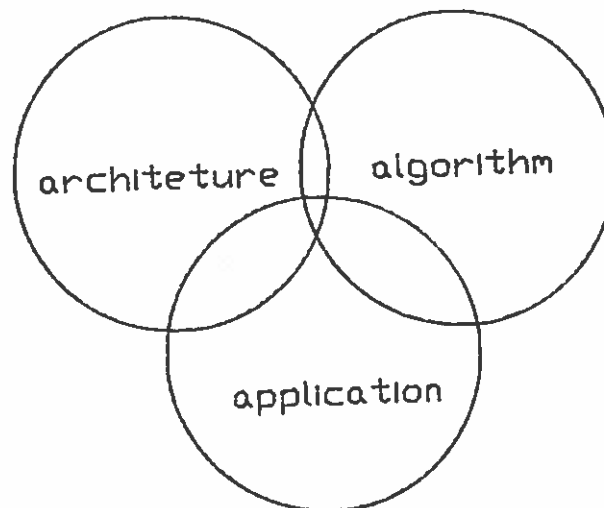


Figure 4.1 The structure of computing environment in the scalar mode

With the advent of the supercomputer, such a traditional structure must be updated. Namely, a designer can no longer ignore the hardware structure of the machine and the algorithms have to be re-investigated so that they would work efficiently in a supercomputer environment. Such an update can be illustrated in Figure 4.2 (ibid.)

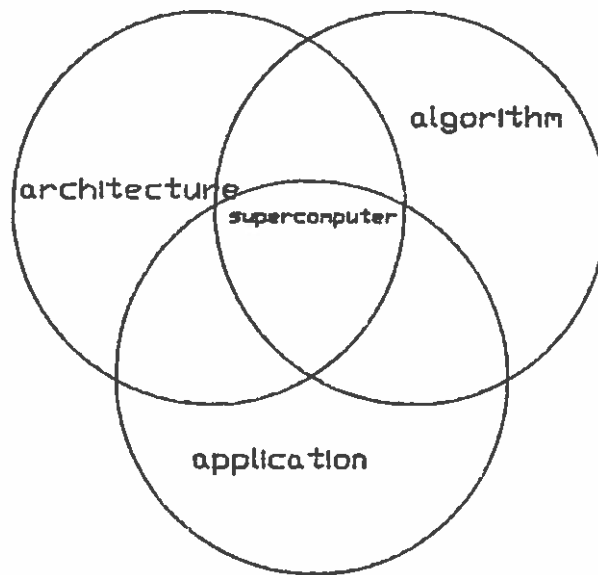


Figure 4.2 The structure of the supercomputing environment

Let us now give an example of such an update. Suppose a programmer has to design a specific code for completing a certain calculation which involves some treatments for an irregular problem. If he is working under a traditional scalar computing environment he may use some "clever" ways to avoid such irregularity and his code might be fairly short. He also disregards the system-provided routines and generates equivalent routines by himself. He, however, will not or should not do so if he is working under supercomputing environment. He first should notice that the "clever" ways might not perform efficiently for a vector machine, say the CRAY X-MP. Instead of short code, a longer code should be generated so that the irregular problem can be treated in a vectorized mode, though some duplicated calculations may be made. He also

has to realize that the system has provided some fundamental computational kernels written in, say CRAY Assembly Language, which are not equivalent to his own code. Such system kernels, unlike his code, can be executed much faster than what he expected.

Another example is the compiler directive (command that instructs the compilation of the program) which can be inserted in the FORTRAN codes. A user must understand the function of the directive and then the directive will become meaningful to his code. The understanding in turn means the interaction between user and machine and further indicates the dependence between each other. Thus, it is full communication between user and supercomputing facility that makes the computation reach its boundary.

In terms of optimization, the items such as vector hardware, compiler directives and computational kernels can be regarded as global optimization that can be thought to be built-in facilities. The important effect of local optimization of a user's code, however, cannot be neglected. The optimization will mostly concentrate on the vectorization of the code, as will be illustrated in the following section.

#### 4.2 Vectorization of a FORTRAN program

One of the major features of a supercomputer is the use of vector hardware, which has been mentioned in the previous section. The meaning of vectorization thus becomes straightforward. Any manual of CRAY FORTRAN or any supercomputer workshop material will define vectorization. Generally speaking, vectorization is the implementation of the code so that some particular parts of the program can take advantage of the machine's vector hardware and execute efficiently.

In terms of FORTRAN language, two types of loops can be vectorized:

Innermost DO - loop

IF - loop with backward branch

However, not all such loops can be vectorized. For example, the loops will not be vectorized if the loops contain:

CALL statement  
 I/O statement  
 Branch out of loop  
 Dependency of array  
 Bound checking for an array (compiler option)  
 RETURN statement

Understanding these prohibitions, the vectorization then can be carried out. In other words, the work now will concentrate on removing these prohibitions or changing the algorithm so that the code can be executed under vector environment. A sophisticated skill can be obtained through only practical experience with the basic knowledge of vectorization.

The achievement of vectorization can be remarkable. Consider the FORTRAN code for calculating the  $(\underline{C}_{hh} + \underline{D})$  matrix in eq. (3-3). The non-vectorized code is:

```

K=0
DO 40 I = 1, N
DO 40 J = 1, I
K=K+1
IF (I.EQ.J) GO TO 42
DIST=DSQRT ((X(I)-X(J))*(X(I)-X(J))+(Y(I)-Y(J))*(Y(I)-Y(J)))
ID=DIST/FINT+1.01
CSTAR(K)=CNN(ID)+(CNN(ID+1)-CNN(ID))/FINT*(DIST-(ID-1)*FINT)
GO TO 40
42  CSTAR(K)=VARU+AAC2(I)
40  CONTINUE
  
```

which can be regarded as terse code. CSTAR is a one-dimensional array that stores elements of  $(\underline{C}_{hh} + \underline{D})$ , AAC2 is an array containing variances of SSH observations. As indicated by the CFT77 compiler, this code is not qualified for vectorization. Notice that the calculation of distances and covariances is expensive. A modified code which

introduces an additional array and duplicates the calculation of diagonal elements can be vectorized. The modified code is:

```

      K=0
      DO 40 I = 1, N
      DO 40 J = 1, I
      K=K+1
      DIST=DSQRT ((X(I)-X(J))*(X(I)-X(J))+(Y(I)-Y(J))*(Y(I)-Y(J)))
      ID=DIST/FINT+1.01
      CSTAR(K)=CNN(ID)+(CNN(ID+1)-CNN(ID))/FINT*(DIST-(ID-1)*FINT)
40  CONTINUE
      DO 60 I=1, N
      INDX (I)=I*(I+1)/2
60  CONTINUE
      DO 61 I=1, N
      CSTAR (INDX(I))=VARU+AAC2(I)
61  CONTINUE

```

A timing comparison for the execution of the two codes is given in Table 4.1.

Table 4.1

CPU time comparison between vectorized code and non-vectorized code  
(Time in  $10^{-3}$  seconds)

N	non-vector	vector	non-vect./vect.
20	0.024	0.013	1.85
50	0.059	0.026	2.27
100	0.163	0.054	3.02
200	0.498	0.122	4.08
300	1.025	0.211	4.86
400	1.728	0.320	5.40
500	2.616	0.448	5.84

The code was executed on the CRAY X-MP/28 supercomputer (2 stands for 2 processors, or CPUs, 8 stands for 8 mgawords). It is obvious that the vectorization becomes more important as the size of N increases. We also note that a terse code does not guarantee a faster execution and the "redundant" calculation might help the program's speedup.

From this example, we can see the significant impact of vector calculation on the traditional algorithm design. If the problem is related to a mammoth calculation, which is exactly the case in this study, the contribution of vectorization can definitely not be neglected.

#### 4.3 Special analysis of some computational kernels for geodetic problems

Geodetic problems are frequently related to matrix operations, e.g., the lsc problem in this study. Therefore, it is necessary to investigate the corresponding computational kernels before they can be introduced to the production mode. Three such kernels (routines) are analyzed as follows:

##### (i) Matrix inversion routines

The CRAY Library provides the vectorized LINPACK routines that can compute the inverse of a positive definite matrix. They are:

- SPPCO:** computes Cholesky factorization of A (a PD matrix) i.e.,  $A = R^T R$ , where R is an upper triangular matrix. It also estimates the condition of the inversion and provides the working accuracy in the proceeding computation.
- SPPDI:** Inverts A by using R computed by SPPCO. It can optionally compute the determinant of A.

IMSL routines are also available in our supercomputer system. The routine LINDS is equivalent to the combination of SPPCO and SPPDI and it inverts a PD matrix without breaking the procedure into two parts, namely

$$\text{LINDS} = \text{SPPCO} \rightarrow \text{SPPDI}$$

The performances of LINDS and SPPDI (including SPPCO) are presented in Figure 4.3. In this performance analysis, we first created a covariance matrix ( $\underline{C}_{hh} + \underline{D}$ ) of dimension 500 x 500 using the altimeter data in the Bermuda area; then routines LINDS and SSPDI (with SPPCO) inverted the covariance matrices with dimensions varying from 10 to 500 (a total of 50 matrices were inverted).

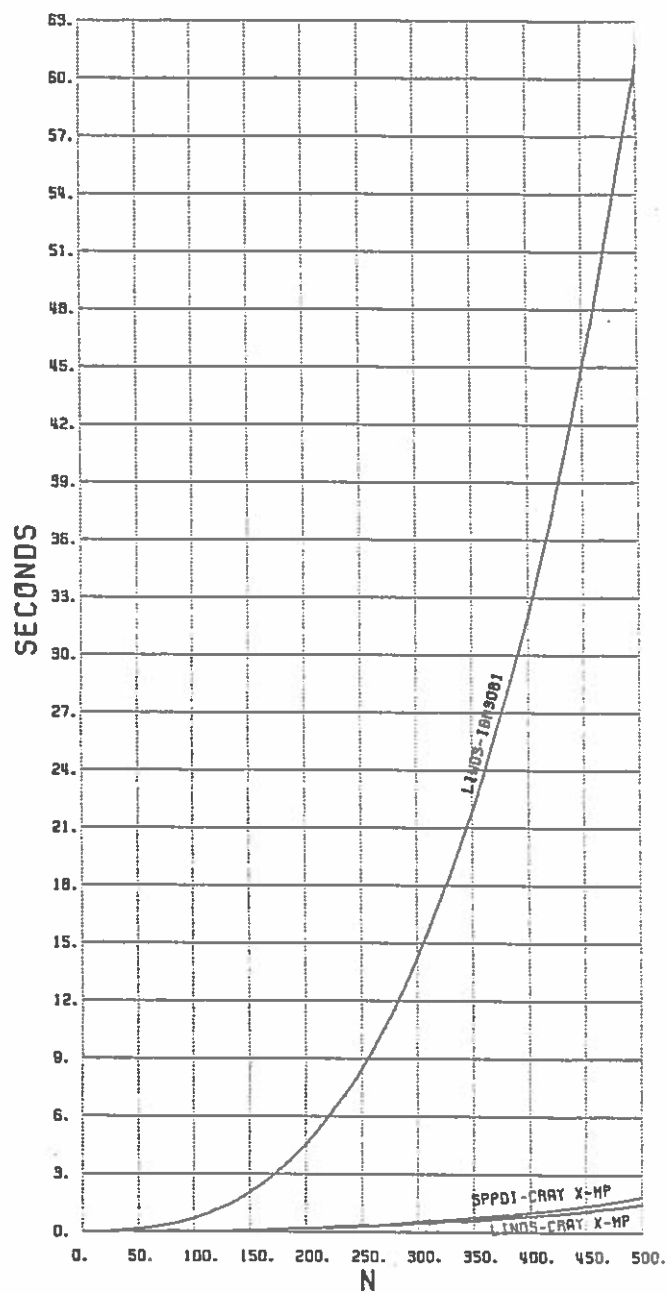


Figure 4.3 Performance of various matrix inversion routines on IBM and CRAY;  $N$  is the dimension of the inverted matrix



Since the routine LINDS is also available on the IBM 3081 system, such a performance analysis was also carried out in such a scalar machine. Figure 4.3 clearly demonstrates the linearization of CPU time in a vector machine for a routine whose execution time undergoes a non-linear increase in a scalar machine. To have a feeling of the efficiency of the supercomputer, Table 4.2 also provides the CPU time saving factors for routine LINDS under IBM3081 and CRAY X-MP environments.

Table 4.2 Performance analysis for LINDS

(Time in seconds)

Dimension	CPU in IBM3081	CPU in CRAY X-MP	gain factor (IBM/X-MP)
20	0.0203	0.0035	5.80
50	0.1323	0.0129	10.26
100	0.7093	0.0431	16.46
200	4.5748	0.1740	26.29
300	14.1970	0.4295	33.05
400	32.0196	0.8427	38.00
500	60.8148	1.4490	41.97

Apparently, the gain factors increase as dimensions increase. This is due to the operational characteristic of vector hardware: the longer the vector is, the more efficient the process is. Thus, it is preferred that the problem involves an inversion of a large matrix so that vectorization can be fully utilized.

For our global prediction purpose, we decided to choose LINPACK's routines, i.e., SPPCO and SPPDI, since the SPPCO routine can provide the working accuracy for collocation prediction. More specifically, the parameter RCOND in

SPPCO (A, N, RCOND, WORK, INFO)

is an estimate of the reciprocal condition,  $1/\kappa(A)$ , which is a condition number that measures the sensitivity to the inverse solution. If the logical expression

(1.0 + RCOND). EQ.1.0

is true, then A can be regarded as singular to working accuracy (Dongorra et al, 1979) (In the practical prediction of anomalies, the predicted anomalies have extremely abnormal values if the above logical expression is true). Knowing such conditions, we can easily handle the flow of the program and the program will not be interrupted. The IMSL routine, LINDS, though can execute slightly faster, will completely stop the program's execution if the matrix to be inverted is nearly singular, thus it is not adopted as our computational kernel.

(ii) Matrix product routine

Our collocation solution inevitably uses the matrix product operation. Such matrix product operations can be easily written with in-line FORTRAN codes. However, the supercomputer's system usually provides such basic routines which are written in CRAY Assembly Language (CAL) and callable by a FORTRAN program. Written in CAL, the routine will execute faster than the equivalent in-line FORTRAN codes.

Consider the following in-line FORTRAN codes:

```

DIMENSION C (100, 500), A (100, 500), B (500, 500)
DO 1 I=1,100
DO 1 J=1, N
DO 1 K=1,500
C(I,J)=C(I,J)+A(I,K)*B(K,J)
1  CONTINUE

```

which is identical to

$$C = AB$$

A CRAY routine MXM will do the same job. We invoke

```
CALL MXM (A, 100, B, 500, C, N)
```

The output C will be exactly the same as that from in-line FORTRAN codes.

The performances of MXM and its equivalent codes are given in Figure 4.4. It is evident that the routine written in CAL indeed provides a faster solution. A wise choice will be using such system-provided routine instead of writing the codes if the problem

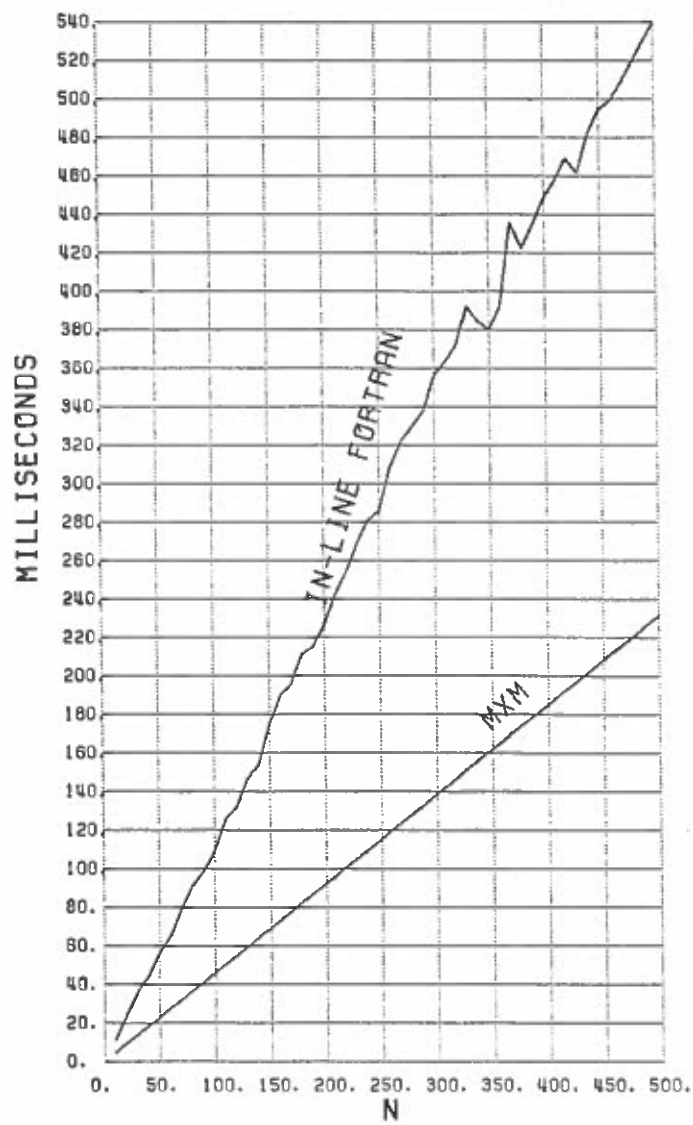


Figure 4.4 Performance of matrix product routine

involves a huge number of matrix products, e.g., a total of 345,600, in our global production work.

(iii) Matrix-vector product routine

Our collocation process involves an operation such as:

$$(\underline{C}_{hh} + \underline{D})^{-1} (\underline{h} - \underline{h}_R) \quad (4-1)$$

which is a product of a matrix and a vector. In a collocation solution without accuracy estimates, we actually do not need to invert  $(\underline{C}_{hh} + \underline{D})$  and then perform the product of the inverted matrix and the vector,  $(\underline{h} - \underline{h}_R)$ . Any numerical analysis book, e.g., Gerald et al (1983), will suggest such a solution

$$(\underline{C}_{hh} + \underline{D})^{-1} (\underline{h} - \underline{h}_R) = X \quad (4-2)$$

Then

$$(\underline{C}_{hh} + \underline{D}) X = (\underline{h} - \underline{h}_R) \quad (4-3)$$

$(\underline{C}_{hh} + \underline{D})$  can be factorized, thus

$$\underline{C}_{hh} + \underline{D} = \underline{R}^T \underline{R} \quad (4-4)$$

Therefore

$$\underline{R}^T (\underline{R} X) = (\underline{h} - \underline{h}_R) \quad (4-5)$$

The advantage of eq. (4-5) is that  $(\underline{R} X)$  is a vector and  $\underline{R}^T$  is a lower triangular matrix and  $(\underline{h} - \underline{h}_R)$  is also a vector. The solution will be found in a forward substitution phase without inverting  $\underline{R}^T$ . However, a collocation solution with an accuracy estimate

requires a matrix inversion. Thus, a product of matrix and vector cannot be avoided in such a case. This work can be easily accomplished by using the FORTRAN codes:

```
DIMENSION C (100), A (100, 500), B (500)
DO 1 I=1,100
DO 1 J=1, N
C (I) = C (I) + A (I, J) * B (J)
1  CONTINUE
```

Again, we recall another useful routine written in CAL

```
MXV (A, 100, B, N, C)
```

which is equivalent to the above in-line codes. The performances are presented in Figure 4.5. Again, it is remarkable the routine MXV can save much more computer time than the in-line codes do.

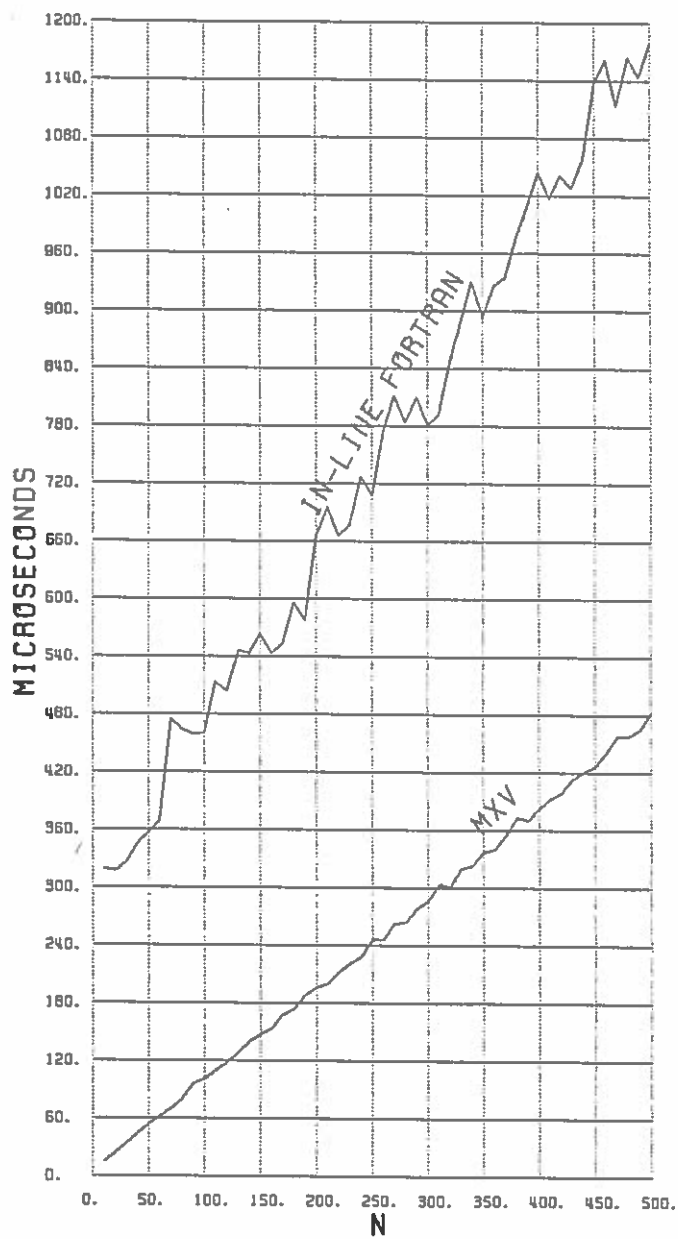


Figure 4.5 Performance of matrix-vector product routine

#### 4.4 Performance analysis for the production program

Having realized the characteristics of the supercomputing environment, a fully vectorized computer program was designed for our global production work. This program is especially suitable for accuracy computation.

As pointed out by Rapp (1985), if the accuracy calculation is performed in a collocation solution on a scalar machine, the time needed is eight times more than the time needed for solution without accuracy calculation. The key factor is the computation of

$$\hat{\Delta g} = B(h - h_R) + \Delta g_R \quad (4-6)$$

$$M_{\Delta g}^2 = C_{gg} - BC_{hg} \quad (4-7)$$

where

$$B = C_{gh}(C_{hh} + D)^{-1} \quad (4-8)$$

must be computed for each predicted point although the matrix needs to be inverted once (ibid.). In essence, the following sections of the production program have been optimized (or vectorized):

- (i) The selection of altimeter points from the adjusted data (i.e., output from the adjustment program)
- (ii) Loops containing I/O
- (iii) The computations of covariance matrices
- (iv) The inefficient part of computing B using eq. (4-8)
- (v) The formations of total  $\Delta g$  and SSH and their accuracy estimates using (4-6) and (4-7)

In addition to these considerations, the matrix product operations were replaced by SCILIB routine MXM and the matrix-vector product operations were replaced by SCILIB

routine MXV. In essence, to accomplish the vectorization of the original IBM version of the OSU altimeter analysis, i.e., F459IBM3, the entire algorithm and the codes were completely changed.

Tests also showed that a matrix instability can take place if data is too dense, especially for the case in which the original combined Geos-3/Seasat data was used. Therefore, we had to eliminate some data points which were close to each other. The original points thus are processed so that two arbitrary points will be separated from each other by a specified distance. Such a process was also vectorized and it took a small amount of time as compared to the time required for the entire collocation prediction. In the test area located at  $\varphi = 70^\circ - 66^\circ$ ,  $\lambda = 10^\circ - 14^\circ$ , it was found that, by specifying 1 km as the minimum spacing between two arbitrary points, a distance check process can reduce the number of altimeter points by 15 to 20 points in a  $0.5^\circ \times 0.5^\circ$  prediction cell (with  $0.5^\circ$  border) with 300 points as the maximum number of points used, and hence reduce the CPU time by 8% in the collocation part. The predicted quantities were not substantially changed from a case where the check was made.

Another effort made to speed up the calculation is the simplification of the determination of the reference values at the predicted points. Since our goal is to predict a set of regularly gridded point values, their reference values can be calculated beforehand and stored in a matrix that will generally cover the entire earth or 1/4 of the earth. The interpolation then becomes a simple process which only involves the identification of the corresponding elements in such a particular matrix. Furthermore, the accuracy of interpolation method needs not be concerned since we did not perform interpolation.

At this point, we have tried all the effort that will accelerate the calculation of point anomalies and sea surface heights on an oceanwide basis. The completion of the optimizations yields a program called F459PRD. Since we incorporated the cross-over adjustment program into it, the steps of the computation carried out by this program is then split into two parts:

- (i) Cross-over adjustment of arcs
- (ii) Prediction of point anomalies and sea surface heights using lsc.



Let's denote F459IBM3 as the original program that was executed in an IBM3081 machine, F459NOV being the CRAY version of F459IBM3 without any further modification. The comparison among F459IBM3 (on IBM3081), F459NOV (on CRAY X-MP/28) and F459PRD (on CRAY X-MP/28) is shown in Table 4.3 (a) and (b).

Table 4.3  
(a) CPU time ratio in the cross-over adjustment part  
(unitless)

program	F459IBM3	F459NOV	F459PRD
ratio	6.0	1.5	1.0

(b) CPU time ratio in the collocation prediction part  
(unitless)

program	F459IBM3	F459NOV	F459PRD
ratio	136.0	18.0	1.0

The ratios in Table 4.3 (a) and (b) are subject to change, depending on the choice of prediction mode, but the variations are small. Note that not too much effort has been made in optimizing the cross-over adjustment program and hence the speedup factor will be approximately 5 which is the fundamental saving factor on transferring a scalar-processed machine to a vector-processed machine. The CPU time comparison between cross-over adjustment part and collocation part is hard to evaluate, due to the considerations such as number of cross-overs, number of arcs, number of points used for each prediction cell, etc. To have a feeling of CPU times consumed by the two parts, an example is given as follows:

Case:

Test area:  $\varphi = 34^\circ - 30^\circ$ ,  $\lambda = 293^\circ - 297^\circ$

Adjustment blocksize =  $4^\circ \times 4^\circ$

No. of crossovers = 23,273

No. of arcs = 466

No. of prediction cells = 64 ( $0.5^\circ \times 0.5^\circ$ , plus  $0.25^\circ$  border)

Grid interval =  $0.125^\circ$

Minimum spacing between two altimeter data points = 4 km

Average number of altimeter points used in a prediction cell = 280

CPU time:

Cross-over adjustment = 15.95 seconds

Collocation prediction = 35.95 seconds

This particular run involved dense data that is not representative of the global case. Typically, the cross-over adjustment will consume only 1/10 of the time consumed by the collocation prediction, or even less.

Finally, it is necessary to present the statistics of cpu time taken by the collocation prediction part in a regular production mode, which is just the case mentioned above and will be comprehensively discussed in Chapter 5. The statistics are shown in Figure 4.6. By CPU time we mean the time needed for processing 64 prediction cells which individually involves one matrix inversion, matrix products, data point manipulations, etc. A total of 1024 gravity anomalies and 1024 sea surface heights were predicted in such a case. Figure 4.6 is very useful for the CPU time estimate for a global production work.

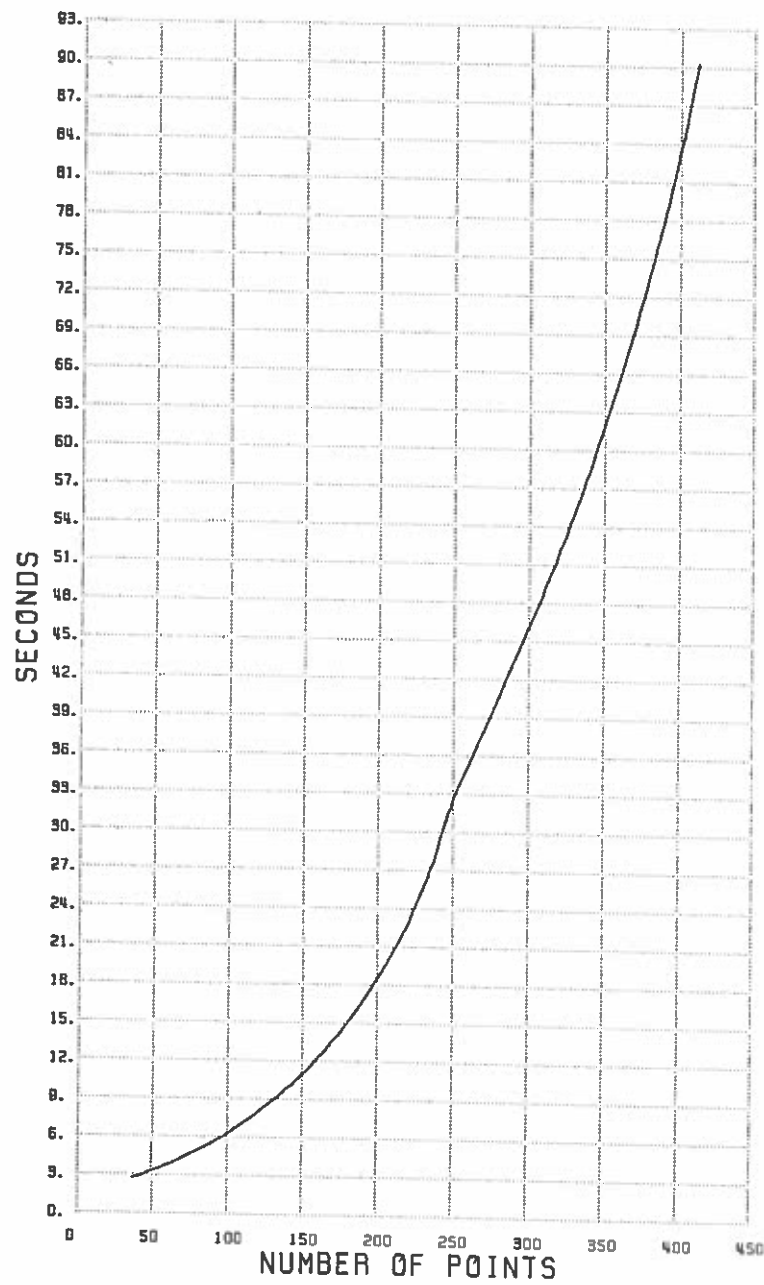


Figure 4.6 CPU time on the CRAY X-MP/28 for a collocation prediction part in a  $4^\circ \times 4^\circ$  area with the standard global production mode, as a function of the number of data points used.

## CHAPTER V

### The optimal parameter for production mode based on various tests

#### 5.1 Introduction

Rapp (1985) has carried out the global prediction of  $0.125^\circ$ -gridded anomalies and SSH using the combined Geos-3/Seasat altimeter data. He used the following specifications:

Reference Field = OSU 180 x 180 (1981)

Covariance Function = Scaled Tscherning/Rapp

Blocksize =  $3^\circ \times 3^\circ$  (For one matrix inversion)

Data Border =  $0.5^\circ$  about block border

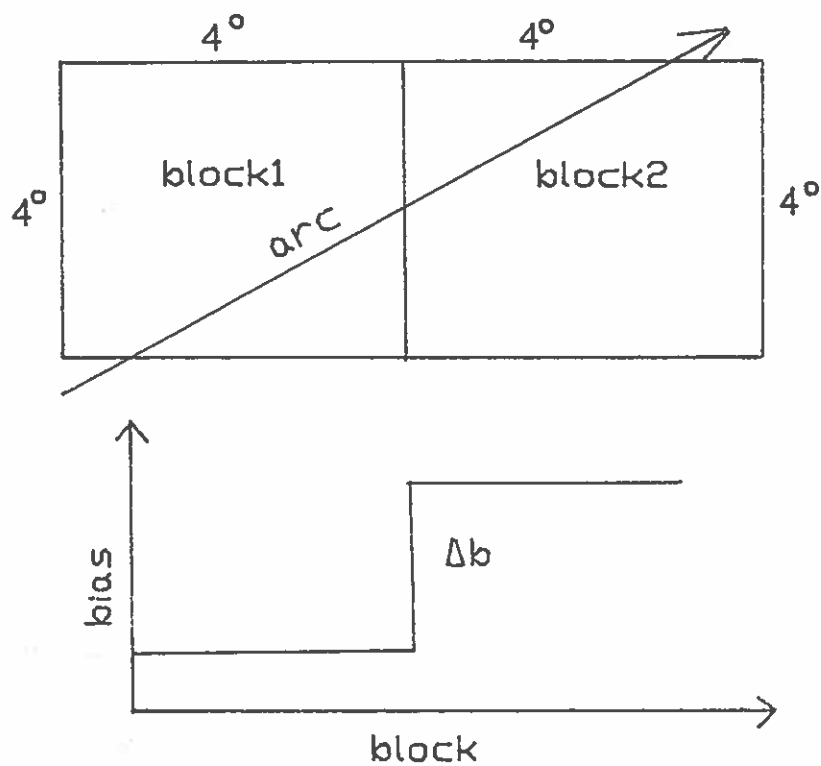
Grid interval =  $0.125^\circ$

As mentioned in Chapter I, he virtually used a thinned data set which contained only 1/6 of the original data points. Our job, however, is to recover the anomalies and SSH using the original, unthinned altimeter data. Furthermore, we are also concerned about the track errors associated with local areas that were not removed in Liang's adjustment and combination process (Liang, 1983), thus many tests will be dedicated to this part.

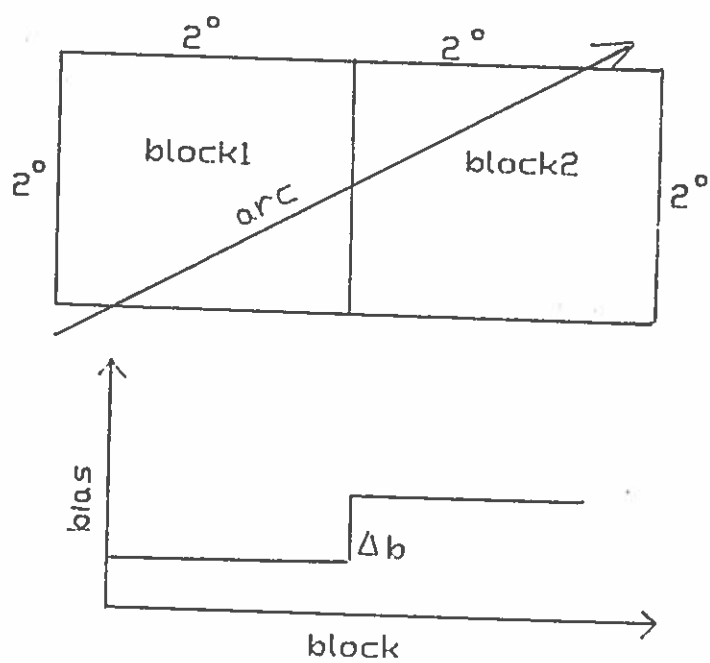
Since our prediction procedure is carried out by first adjusting the local arcs and secondly predicting the desired quantities in the local areas, we have to keep the other part unchanged when one part is undergoing tests. In the following sections, our discussion will concentrate on these two parts separately, though sometimes the two aspects are of the same importance for interpreting the resulting quantities.

#### 5.2 Bias removal from local arcs

In section 3.4 we mentioned that Model A will be the adopted model for the local arc adjustment. In this model many factors should be taken into account. First of all, let's illustrate the effects of the adjustment blocksize when an arc passes through adjacent adjustment blocks and separate biases are found. Figure 5.1 (a) and (b) shows such a case.



(a) Step function of bias for large blocksize



(b) Step function of bias for small blocksize

Figure 5. 1

Considering Figures 5.1 (a) and (b), we somewhat loosely define that a blocksize of  $4^\circ \times 4^\circ$  is large and  $2^\circ \times 2^\circ$  is small. The bias determined by the adopted model is a step function whose variable is the location of the block. When an arc passes through a specific block, it will receive a bias from the adjustment in this particular block. The bias will be mostly influenced by the local phenomena which will vary from block to block. It is such variation of local phenomena that makes the bias a step function. From this point of view, it is expected that the bias function of a specific track will be relatively smooth if such a function is determined from some continuous "small" blocks. On the contrary, the function will become rough if some "large" blocks are used for determining such a function. More specifically, we have  $|\Delta b_1| > |\Delta b_2|$  for Figures 5.1 (a) and (b).

An example of the biases determined from the adjustment is given in Figures 5.2 (a) and 5.2 (b) and Tables 5.1 (a) and 5.1 (b). (Using the updated Geos-3/Seasat data, see Chapter 2).

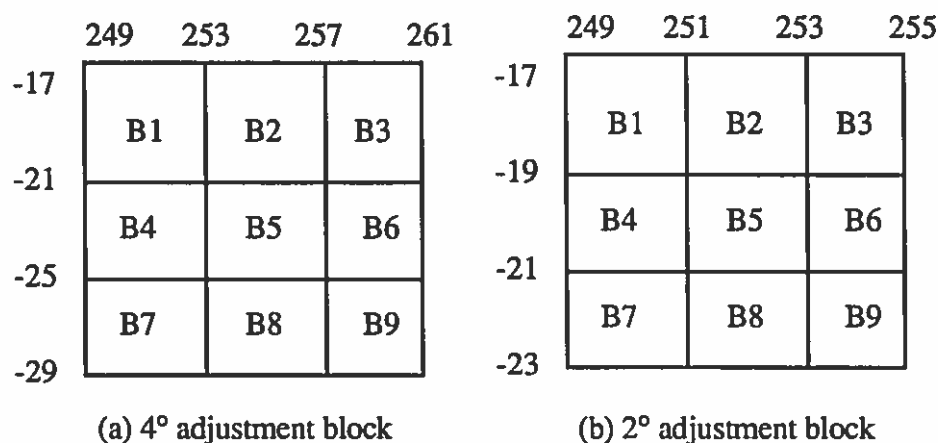


Figure 5.2

Table 5.1

(a) Biases corresponding to Figure 5.2 (a)

Revolution number	B1	B2	B3	B4	B5	B6
N** 4100	-0.051	-0.093			-0.154	
N 3972	0.111	0.131				
N 3958	0.243				0.289	
S 4064	0.381	0.126			0.162	0.183
S 16489				0.591		-0.056
S 3808					0.289	-0.097
S+ 90544					-0.231	

(b) Biases corresponding to Figure 5.2 (b)

Revolution number	B1	B2	B3	B4	B5	B6	B7	B8
N4100	0.053	0.022		-0.278	-0.130	-0.178		-0.161
N3972		0.129	0.09		0.083	0.157		
N3958				0.238			0.212	
S4064					0.367	0.327		-0.029
S90544		-0.102	-0.111	-0.154	-0.149		-0.198	-0.202

\* 0.5° border about the block for both 4° and 2° blocks

\*\* N denotes ascending arc and S descending arc

+ Revolution number starting from 9 is Seasat track; otherwise it is Geos-3 track.

From this example, no substantial difference was found among biases determined from different locations of blocks and different block sizes. From the adjustment's result, we also knew that, in the above two areas, the RMS cross-over differences before adjustment are on the order of 20 to 30 cm, and approximately 5 to 10 cm less after adjustment. For the production work, it will be appropriate to say that the choice between 2° and 4° is not important since the biases estimated from such two cases have no significant differences as compared to their accuracy estimates. However, an extremely large block size, say 10 degree, will lose the meaning of adjusting short wavelength error or it cannot be handled by the current computer program due to large number of bias parameters. On the other hand, an extremely small block size, say 0.5 degrees, will yield meaningless results due to the instability of the geometry and the lack of cross-overs. Therefore, a block size varying from 2° to 5° will be acceptable.

Another discussion of block size will be related to the density of the altimeter data distribution. Basically, the cross-overs form the connections between arcs. If the data is

sparse or the distribution of the arcs is not even in a block, it is possible that the connection will be lost between two or more independent groups of arcs. This case is shown in Figure 5.3.

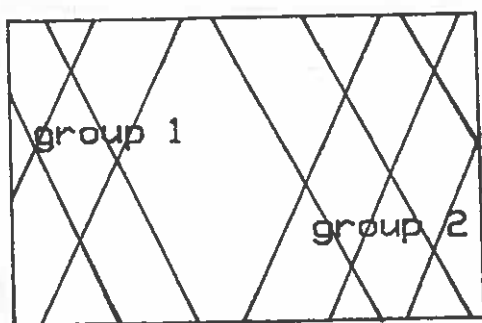


Figure 5.3 Block with two independent groups of arcs

The case in Figure 5.3 will yield dependency of the parameters in the adjustment solution and the normal equation becomes singular. To overcome this situation, two separate constraints given by eq. (3-31) must be added to the observation equations, or generally,  $n$  constraints are required for  $n$  groups of arcs. Although the additional constraints can resolve the singularity problem, the biases could be inconsistent from group to group in this block. Therefore, a better solution is recommended by increasing the blocksize so that the connection between two or more groups of arcs become possible. Figure 5.4 shows the effect of increasing the blocksize.



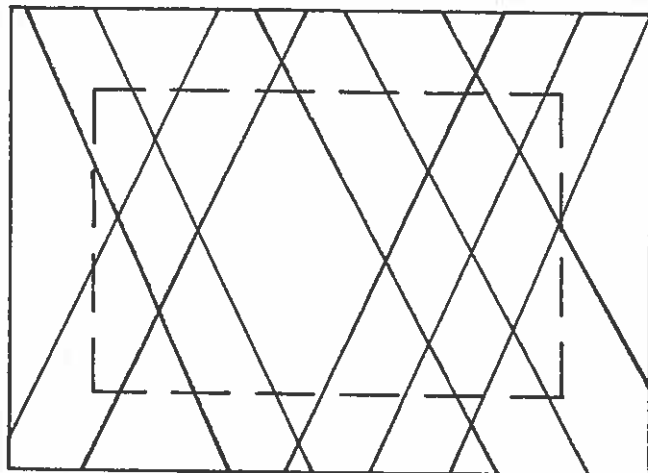


Figure 5.4 Effect of increasing blocksize of bias adjustment region

Generally speaking, two types of areas will create independent groups: Area where only Seasat or Geos-3 data is available and areas near a coastline. An example of the second type is given in Figure 5.5. This area is located in Northern Australia. Block A cannot be adjusted without adding an additional constraint, while Block B can be adjusted without adding such a constraint.

At this point, we have discussed the non-gravity field related features of the adjustment of the altimeter data in which the geometry of the altimeter arcs play an important role. We should now proceed to the prediction of anomalies and SSHs which will be our final product.

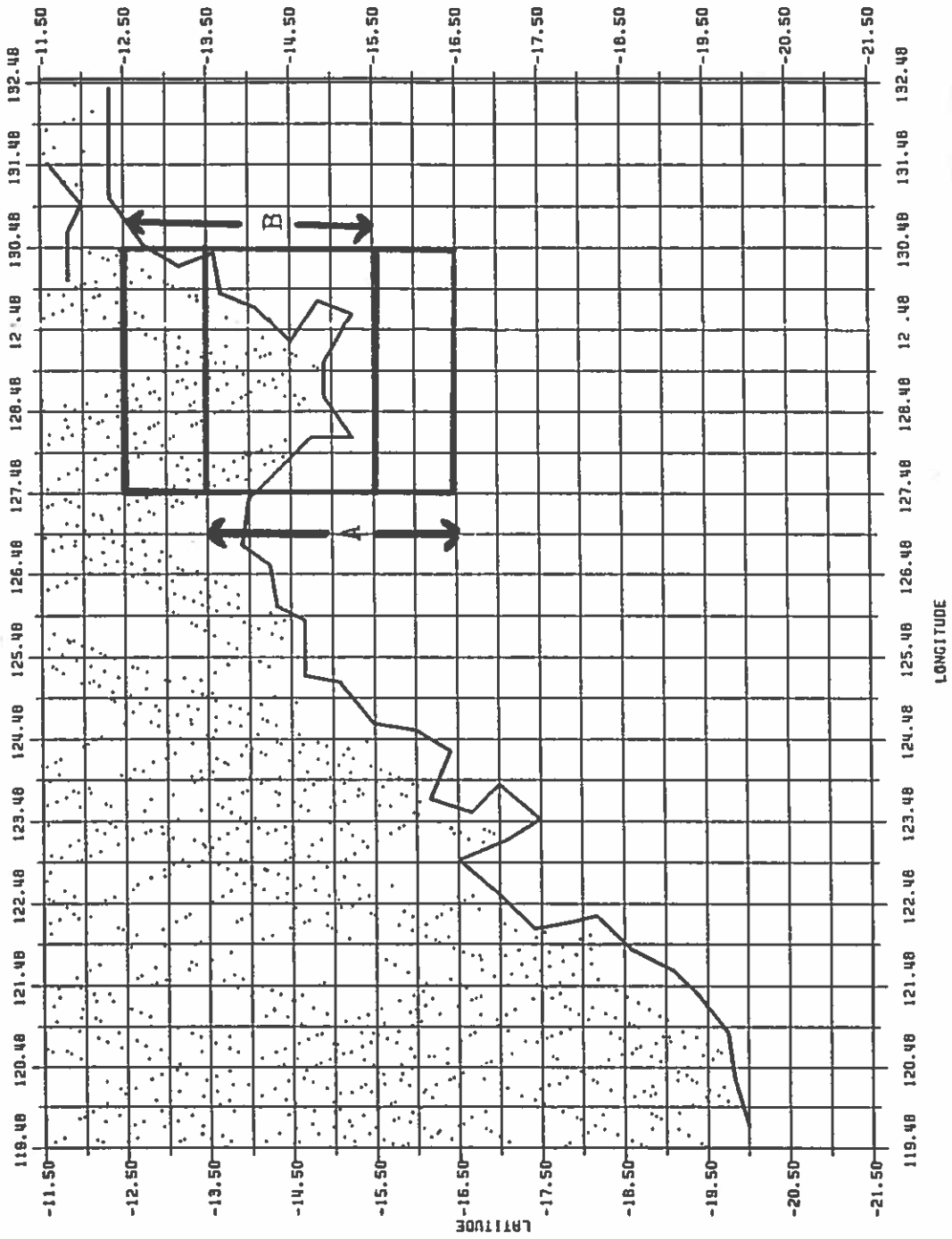


Figure 5.5 Example of a bias adjustment block with two independent groups of arcs  
(Block A)

Rapp (1985) has given an example of track problem occurring in the area at  $-33.0^\circ < \phi < 0.0^\circ$  and  $220.0^\circ < \lambda < 259.75^\circ$ . Thus we decide to choose part of this area for testing the bias adjustment program. In order to select the appropriate blocksize for the bias adjustment, three predictions were carried out. The three predictions were performed in an  $10^\circ \times 10^\circ$  area so that the continuity of the biases and predicted anomalies and SSH can be investigated from block to block. The three predictions had the following specifications which were designated as choice A, choice B and choice C.

(i) Location: (for A, B and C\*)

$$-27^\circ < \phi < -17^\circ$$

$$249^\circ < \lambda < 259^\circ$$

- \* For choice C, a  $12^\circ \times 12^\circ$  area was used; then the result for comparison was extracted from this area (since  $10^\circ$  is not an integer multiple of  $4^\circ$ , we have to do so, see (iii))

(ii) Common choice: (for A, B and C)

Reference Field: OSU86E  $180 \times 180$

Covariance Function: Scaled Tscherning/Rapp with Jekeli's parameters

Prediction cell size:  $1^\circ \times 1^\circ$  (For one matrix inversion)

Data border:  $0.5^\circ$  about block border

Grid interval:  $0.125^\circ$

arcs without bias parameters: excluded\*

- \* The isolated arcs, which cannot be adjusted by using cross-over techniques, are not used in the collocation solution.

(iii) Adjustment blocksize:

Choice A:  $5^\circ \times 5^\circ$  (with  $0.5^\circ$  border)

Choice B:  $2^\circ \times 2^\circ$  (with  $0.5^\circ$  border)

Choice C:  $4^\circ \times 4^\circ$  (with  $0.5^\circ$  border)

Quantities associated with choices A, B, and C are given in Table 5.2.

Table 5.2 Quantities of interest for the different prediction choices A, B and C

	No. of Adjust. block	No. of matrix inversions	No.* of $\Delta g$	No.* of SSH
Choice A	4	100	6400	6400
Choice B	25	100	6400	6400
Choice C	9	144	9216	9216

\* predicted

In order to demonstrate the ability of the bias removal by the adjustment program, we first made a contour map for the predicted anomalies from Rapp's (1985) results. This contour map is given in Figure 5.6 (a). Using the same data a color contour plot was generated in a Tektronix 4115B machine with the same interval (color interval for this plot) as used in Figure 5.6 (a). Figure 5.6 (b) was prepared by using the image of the color contour plot in a regular Xerox machine. The advantage of the original color plot or the image of the color plot (as given in Figure 5.6 (b)) was its ability to show some track patterns that cannot be detected in a regular contour plot. Such features were clearly revealed by Figure 5.6 (b) in which some linear features caused by track errors were found.

We next used the adjusted geoid heights obtained from the 3 adjustments described in Table 5.2 to perform the prediction of anomalies and SSH. The predicted anomalies were given in Figures 5.7 (a) and (b), Figures 5.8 (a) and (b) and Figures 5.9 (a) and (b) which correspond to choice A, choice B and choice C, respectively. It is obvious that the linear features in Figures 5.6 (a) and (b) were removed from the predicted anomalies and the adjusted data yields a smooth surface of anomalies that are free of the most obvious features.

We are further interested in the differences between the different sets of predicted anomalies. Table 5.3 gives the differences between the anomalies predicted from the three choices being considered for production work.

Table 5.3 Comparison of anomalies from various adjustment blocksizes

Difference	plot	max	RMS	mean	Location of max diff.	
					$\phi$	$\lambda$
2° vs 5°	5.10 (a) & (b)	27.37	3.69	-0.01	-24.000	258.875
2° vs 4°		20.82	3.67	0.01	-17.000	255.250
2° vs Rapp*	5.11 (a) & (b)	49.36	6.53	-0.77	-26.375	254.375
4° vs Rapp*		46.96	6.27	-0.76	-25.875	258.625
5° vs Rapp*	5.12 (a) & (b)	42.97	6.21	-0.78	-26.375	253.500

\* Result from Rapp (1985) (1/8° point anomaly and SSH data base)

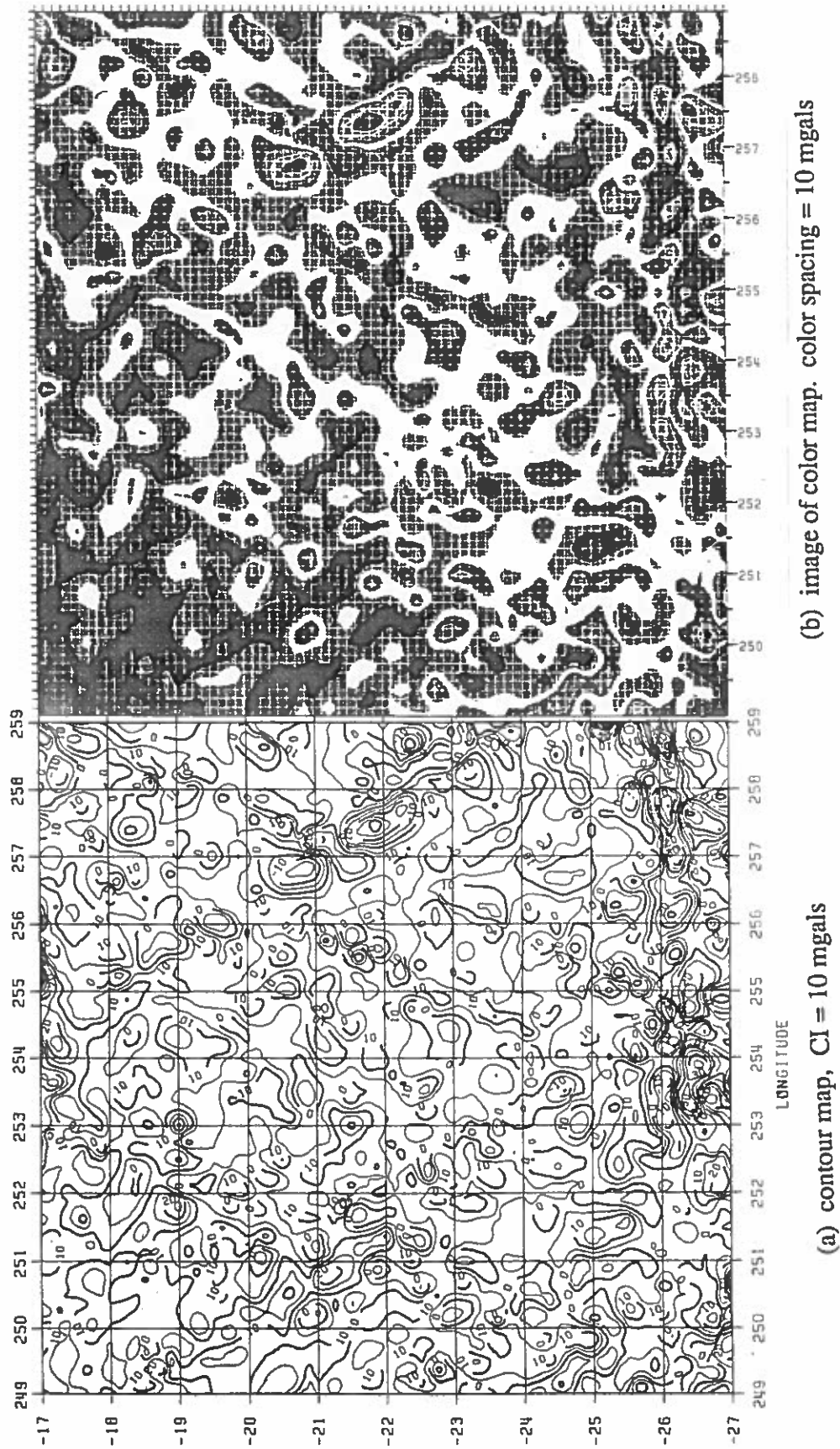
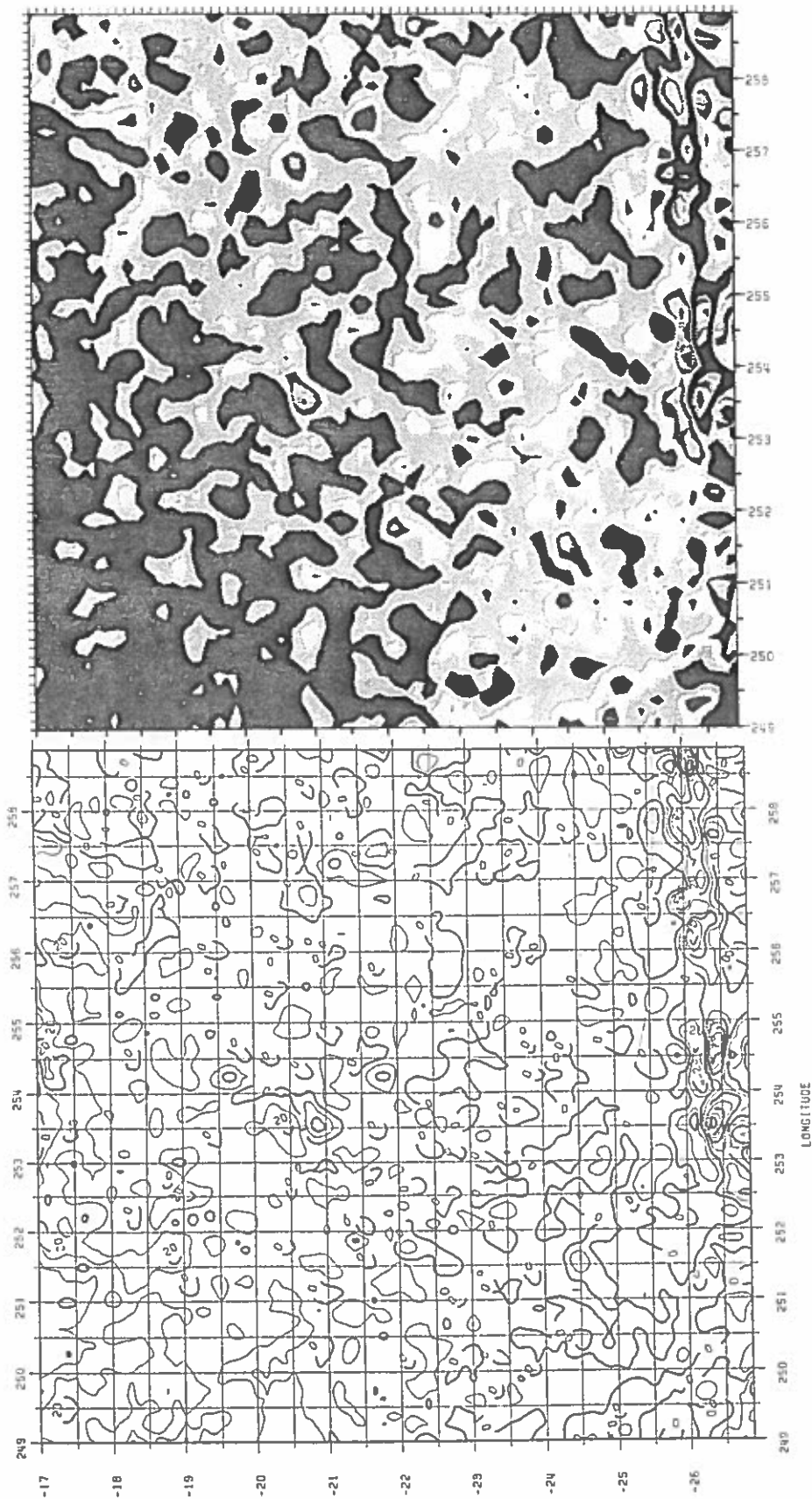


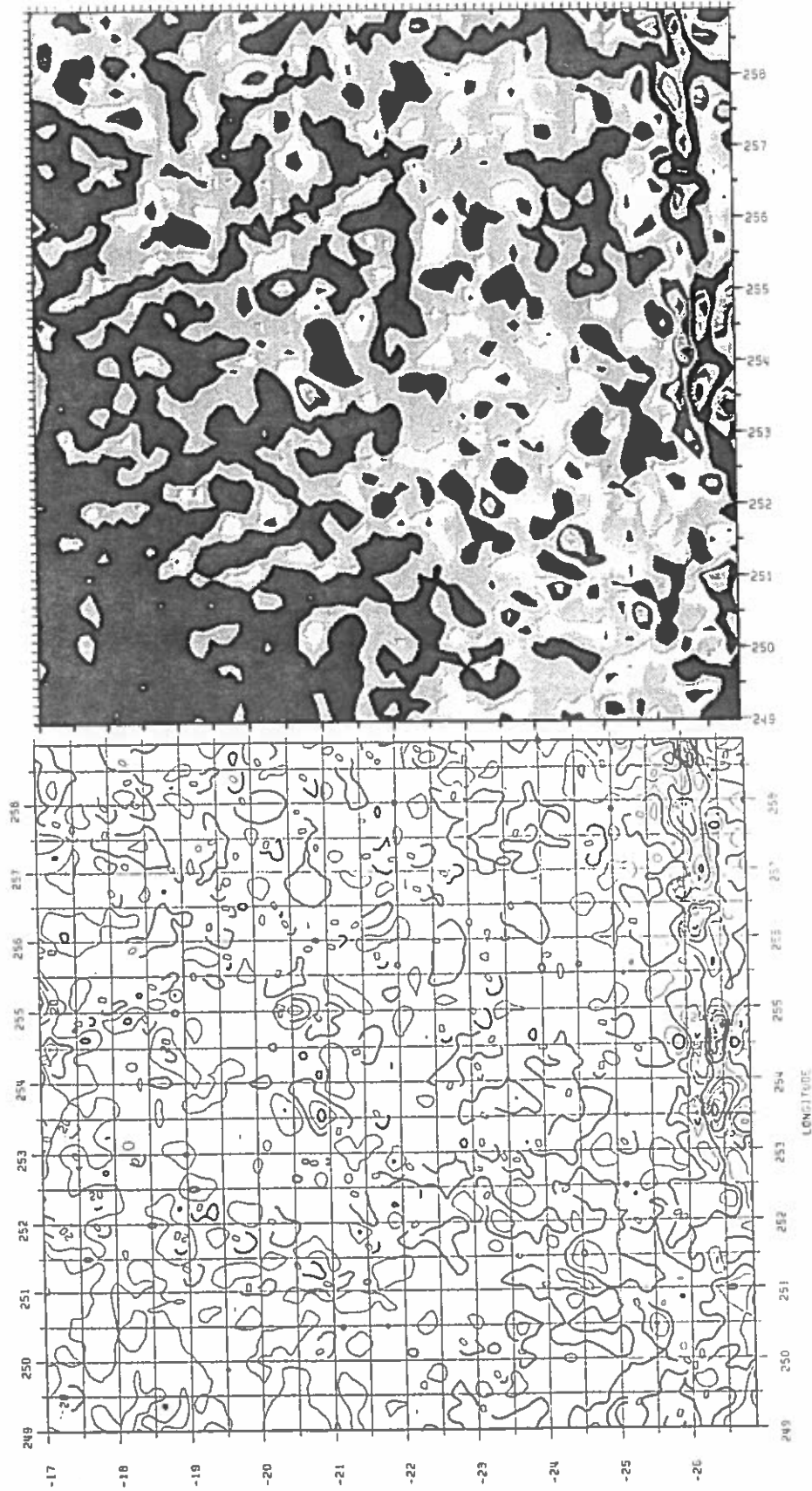
Figure 5.6 Predicted anomalies from Rapp's (1985) production results



(a) contour map, CI = 10 mgals

(b) image of color map. color spacing = 10 mgals

Figure 5.7 Predicted anomalies from 5° adjustment

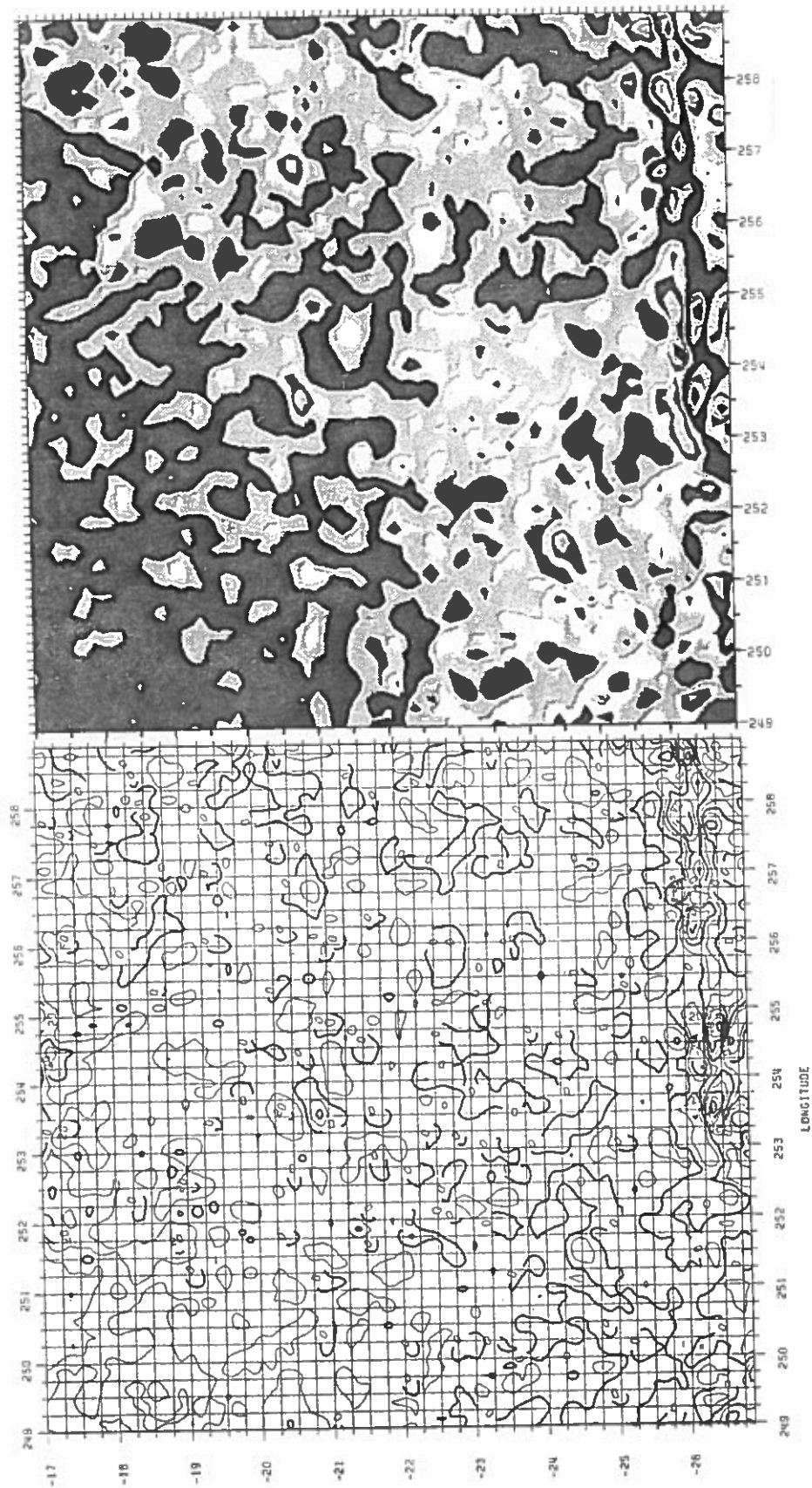


(a) contour map, CI = 10 mgals

(b) image of color map, color spacing = 10 mgals

Figure 5.8 Predicted anomalies from 2° adjustment

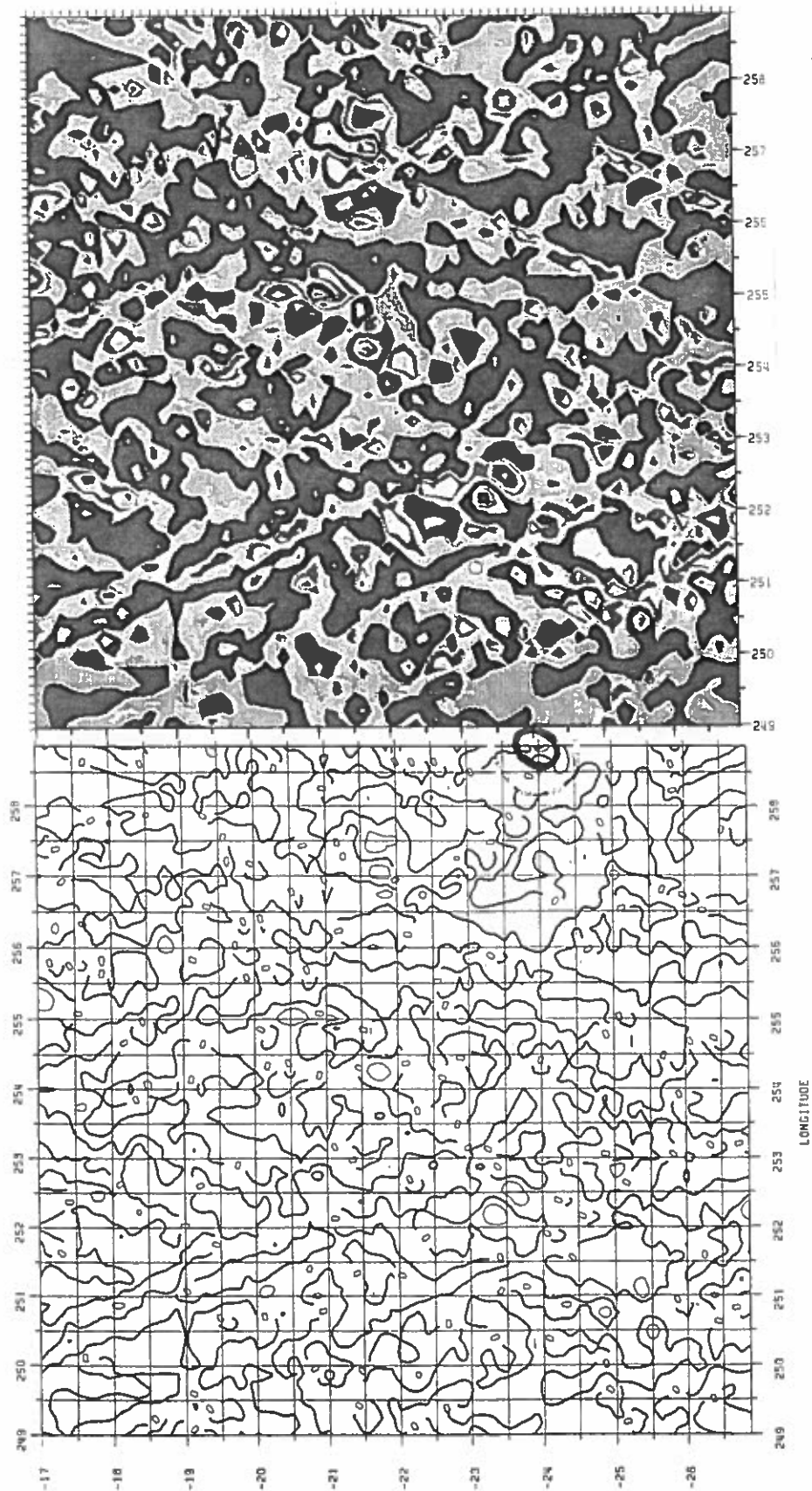




(b) image of color map. color spacing = 10 mgals

(a) contour map, CI = 10 mgals

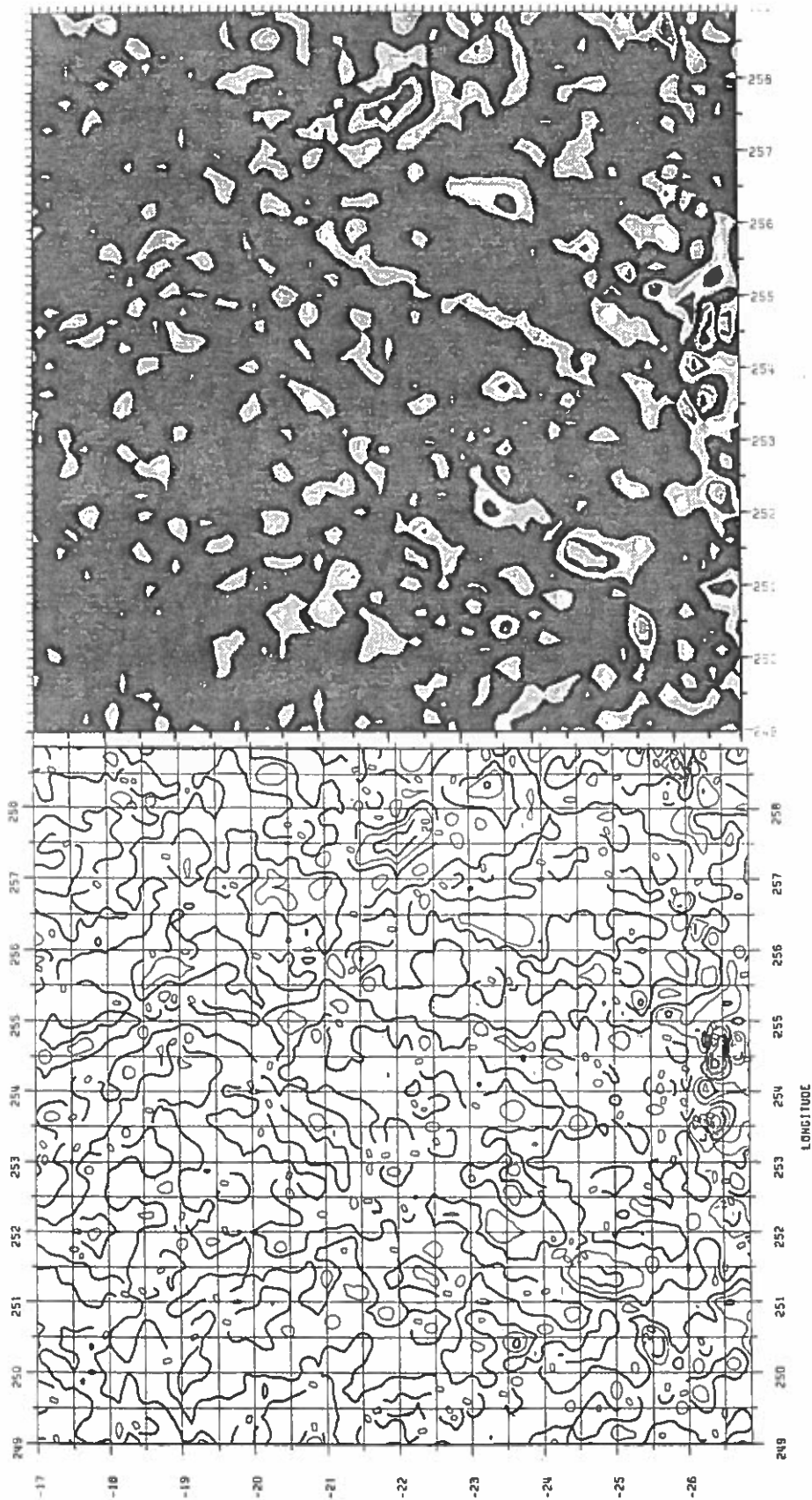
Figure 5.9 Predicted anomalies from 4° adjustment



(a) contour map, CI = 10 mgals

(b) image of color map, color spacing = 10 mgals

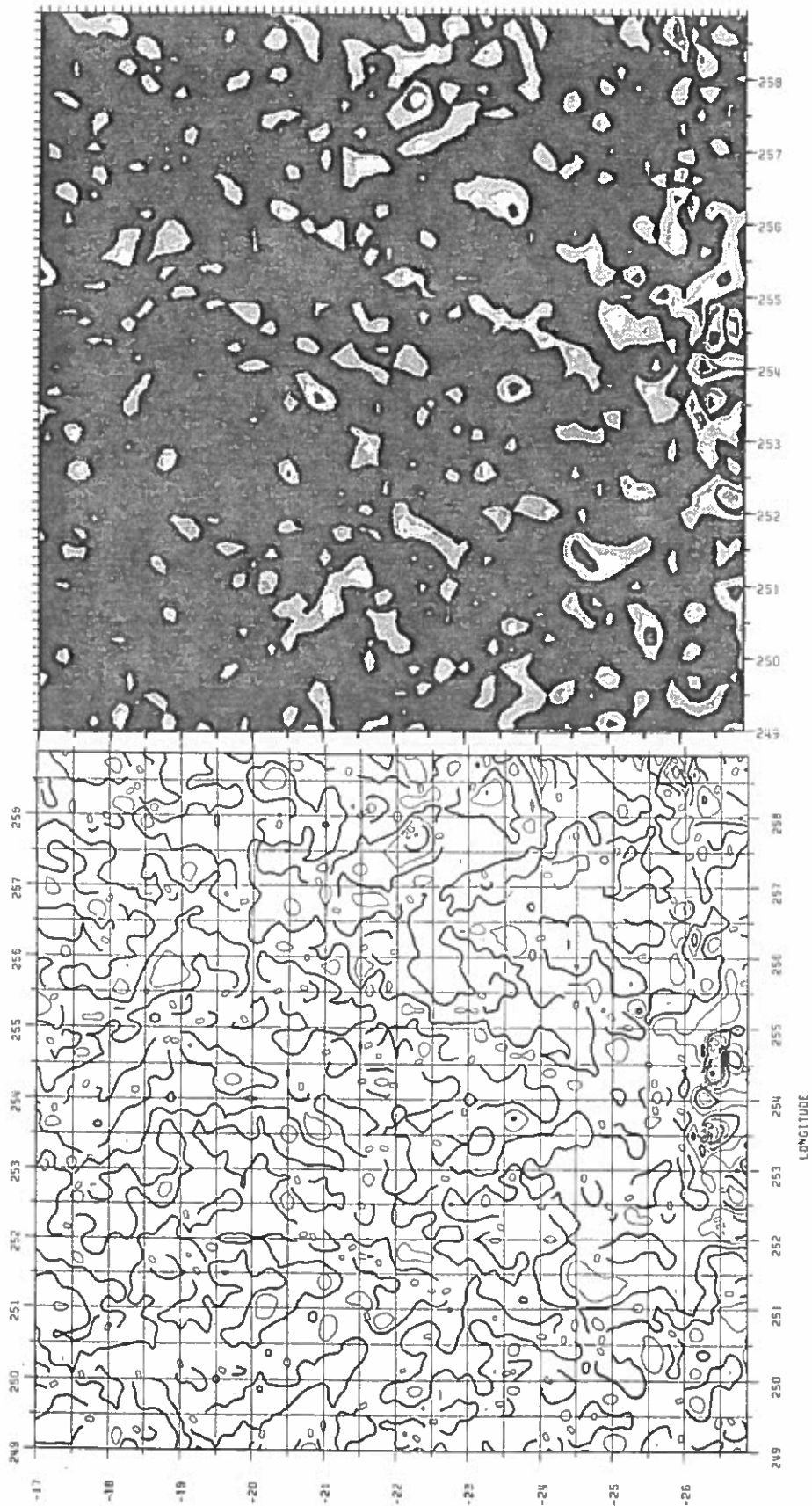
Figure 5.10 Difference in anomalies between 2° and 5° adjustments.  
Circle indicating the largest difference.



(a) contour map, CI = 10 mgals

(b) image of color map. color spacing = 10 mgals

Figure 5.11 Difference in anomalies between 2° adjustment and Rapp's (1985) production result



(a) contour map, CI = 10 mgals

(b) image of color map. color spacing = 10 mgals

Figure 5.12 Difference in anomalies between 5° adjustment and Rapp's (1985) production result

As shown in Figures 5.11 and 5.12, the large discrepancies of anomalies between the predictions from the unadjusted (and reduced data file) and adjusted arcs took place in the rough area (located at  $-27^\circ < \phi < -26^\circ$ ,  $253^\circ < \lambda < 255^\circ$ ) and areas where significant track errors occurred (located at  $-22^\circ < \phi < -21^\circ$ ,  $256^\circ < \lambda < 259^\circ$ ). The large differences of anomalies caused by using different adjustment block sizes, as shown in Figures 5.10 (a) and (b), however, took place near the border of the prediction block (located at  $-24.5^\circ < \phi < -23.5^\circ$ ,  $258.5^\circ < \lambda < 259^\circ$ ). A copy of the color plot shown in Figure 5.10 (b) reveals linear features which correspond to zero contour lines. The reason is apparently that the same arcs receive different biases from different choices of the blocksize and hence yield different anomalies along the arcs. The smoothness of this area is also responsible for the visibility of such features.

It is interesting to study the variations of anomalies around the point where maximum discrepancy of anomaly occurred. Table 5.4 shows such variations.

Table 5.4 Variations of anomalies for  $5^\circ$  and  $2^\circ$  adjustments  
unit = mgals

$\phi \backslash \lambda$	$258^\circ.625$	$258^\circ.750$	$258^\circ.875$
$-24^\circ.000$	20.5 7.7	8.3 8.5	-15.4 12.0
$-24^\circ.125$	14.7 9.2	11.6 14.0	2.9 16.5
$-24^\circ.250$	2.2 5.2	5.3 8.4	9.3 11.5

The first number in Table 5.4 is the predicted anomaly from the  $5^\circ$  adjustment while the second number is the predicted anomaly from the  $2^\circ$  adjustment. It is seen that the variation of anomalies of the first set is more drastic than that of a second set. At the point  $\phi = -24^\circ.00$ ,  $\lambda = 258^\circ.875$ , the anomaly differs from the adjacent anomaly by 23.4 mgals for the  $5^\circ$  case while only 3.5 mgals of difference was found for the  $2^\circ$  case. The unreasonable change of sign in anomalies at this point also indicates the instability of the estimated biases which could be improperly assumed to be constants along the arcs in a  $5^\circ \times 5^\circ$  block.

From Table 5.3, we also know that  $5^\circ$  and  $4^\circ$  cases produce similar results since the corresponding RMS differences with  $2^\circ$  case are close. One remarkable achievement of the adjustment process is that the adjusted arcs are almost continuous from block to block since we have found systematic differences which are also continuous from block to block in Figure 5.10. We thus conclude that the adjustment process can remove the track errors if the cross-overs between arcs can be properly formed. The choice of adjustment blocksize can vary from  $2^\circ$  to  $5^\circ$  and the result from these will be similar.

Another issue will be the accuracies of the bias estimation. The adjustment basically tries to remove the biases of the arcs so that better agreement between observations at the same locations can be achieved. Therefore, the RMS value of cross-overs of those unadjusted arcs will be always reduced after adjustment. However, the standard deviations of the estimated biases primarily depends on two factors. One factor is the number of cross-overs along a particular arc. If the data distribution is such that few cross-overs were formed along the arc, then the standard deviations are expected to be high. From this viewpoint, a larger blocksize of adjustment is preferred since the number of cross-overs could be increased in this case. Another factor is the RMS cross-over differences along a particular arc. A large RMS value usually implies a great fluctuation of the observations which in turn reveals the large inconsistencies between the observations along this arc and the observations along the other arcs. A large standard deviation of the estimated bias can be expected for a large RMS value.

To show the relationship between the standard deviations and the above two factors, Table 5.5 presents part of the result of the bias adjustment in the adjustment block  $-16^\circ < \phi < -20^\circ$ ,  $256^\circ < \lambda < 260^\circ$ . This is a  $4^\circ$  adjustment whose result corresponds to Figure 5.9 (the anomalies were only given up to  $\lambda = 259^\circ$  in Figure 5.9, see Table 5.2 for explanation). We select this block because we already found the track errors in this block (see Figure 5.6).

Table 5.5 The result of 4° adjustment + at  $-16^\circ < \phi < -20^\circ$ ,  $256^\circ < \lambda < 260^\circ$ , predicted anomalies corresponding to figure 5.9 (unit for mean, RMS, bias, std. dev.: meters)

Rev. No.	No. CRS.	Mean *		RMS		Bias	Std. dev.
		before	after	before	after		
3716	19	-0.600	-0.011	0.622	0.202	-0.572	0.046
90264	30	0.030	-0.001	0.131	0.083	0.046	0.012
90752	30	-0.023	-0.008	0.121	0.061	0.000	0.014
3936	3	-0.654	0.033	0.736	0.405	-0.701	0.180
90271	44	0.023	0.019	0.174	0.086	-0.020	0.013

\* Mean is the mean value of cross-over differences along an arc, the mean values were computed before and after the adjustment.

+ Additional information:

No. of tracks: 41

No. of cross-overs: 396

RMS cross-over difference before adjustment: 0.227 meter

RMS cross-over difference after adjustment: 0.144 meter

From Table 5.5, we also found that the standard deviations of estimated biases can vary from 0.012 to 0.180 meters and the estimated biases also have a large range of variations. According to the analysis of the path of arcs, arc 3936 could be one of the arcs that are responsible for the linear features in Figure 5.6. The small number of cross-overs along this arc also accounts for large standard deviation of the bias.

### 5.3 Adjustment blocksize and prediction cell size determination in the Bermuda area

In the previous section, we have demonstrated the capability of bias removal through the cross-over adjustment model. A further analysis on this problem and other related problems was then carried out in the Bermuda area where we have an opportunity to compare the predicted anomalies with the ship measured anomalies.

We recall our goal of the study: Recovery of high precision gravity anomalies and SSH through using dense altimeter data. In order to accomplish this goal a thorough understanding of the relationship between precision of the predicted quantities and the selection of altimeter data in the prediction area is needed (Covariance function is not our issue now, but it is still a very important factor for the prediction process). Many authors, e.g. Rapp (1985), and Kadir (1988), have shown the capability of detecting

seamounts through the use of dense altimeter data. An example given by Rapp (ibid.) illustrated the effect of prediction cell sizes in New England Seamount Area. In this example, the prediction cell size varied from  $3^\circ$  to  $1^\circ$  and the strongest signature of seamount was for the  $1^\circ$  case. In essence, two factors are involved in the choice: prediction cell size and number of points used in a prediction cell.

In order to investigate these two factors, together with the adjustment block size factor, six tests were performed with the following parameter selections:

Choice 1:

Block size =  $1^\circ \times 1^\circ$  (for one matrix inversion)  
 Data number = approximately 300 points  
 Data border =  $0.5^\circ$  about block border  
 Adjustment block size =  $2^\circ \times 2^\circ$   
 Adjustment border =  $0.5^\circ$   
 Grid interval =  $0.125^\circ$   
 Arcs without parameters = excluded

Choice 2:

Block size =  $0.5^\circ \times 0.5^\circ$  (for one matrix inversion)  
 Data number = approximately 300 points  
 Data border =  $0.5^\circ$  about block border  
 Adjustment block size =  $2^\circ \times 2^\circ$   
 Adjustment border =  $0.5^\circ$   
 Grid interval =  $0.125^\circ$   
 Arcs without parameters = excluded

Choice 3:

Block size =  $0.5^\circ \times 0.5^\circ$  (for one matrix inversion)  
 Data number = approximately 400 points  
 Data border =  $0.5^\circ$  about block border  
 Adjustment block size =  $2^\circ \times 2^\circ$   
 Adjustment border =  $0.5^\circ$   
 Grid interval =  $0.125^\circ$   
 Arcs without parameters = excluded

Choice 4:

Block size =  $1^\circ \times 1^\circ$  (for one matrix inversion)  
 Data number = approximately 300 points  
 Data border =  $0.5^\circ$  about block border  
 Adjustment block size =  $2^\circ \times 2^\circ$   
 Adjustment border =  $0.5^\circ$   
 Grid interval =  $0.125^\circ$   
 Arcs without parameters = included



Choice 5: Scaling factor\* standard deviation of 1.5 for Seasat and 2.0 for Geos-3  
 Block size =  $1^\circ \times 1^\circ$  (for one matrix inversion)  
 Data number = approximately 300 points  
 Data border =  $0.5^\circ$  about block border  
 Adjustment block size =  $2^\circ \times 2^\circ$   
 Adjustment border =  $0.5^\circ$   
 Grid interval =  $0.125^\circ$   
 Arcs without parameters = excluded

Choice 6:  
 Block size =  $1^\circ \times 1^\circ$  (for one matrix inversion)  
 Data number = approximately 300 points  
 Data border =  $0.5^\circ$  about block border  
 Adjustment block size =  $5^\circ \times 5^\circ$   
 Adjustment border =  $0.5^\circ$   
 Grid interval =  $0.125^\circ$   
 Arcs without parameters = excluded

\* will be discussed in the following section

A set of gravity measurements made at sea containing 2011 points was selected for the comparison with predicted anomalies. This data was on a tape provided by the National Geodetic Survey. The distribution of the data is shown in Figure 5.13. The bathymetry surrounding the island is given in Figure 5.14. The ship data is based on the GRS80 ellipsoid and has a maximum anomaly of 358.10 mgals and a minimum of -46.40 mgals with an RMS value of 70.15 mgals. For the purpose of comparison, the predicted anomaly was interpolated at each ship measurement point from the predicted (also gridded) data using a bicubic-spline interpolation procedure with 16 points. The result of comparison is given in Tables 5.6 (a) & (b). In Table 5.6 (a), we only accept the differences less than 70 mgals; thus the number of points compared reduced to approximately 500 points. However, in Table 5.6 (b), we compared 2011 points.

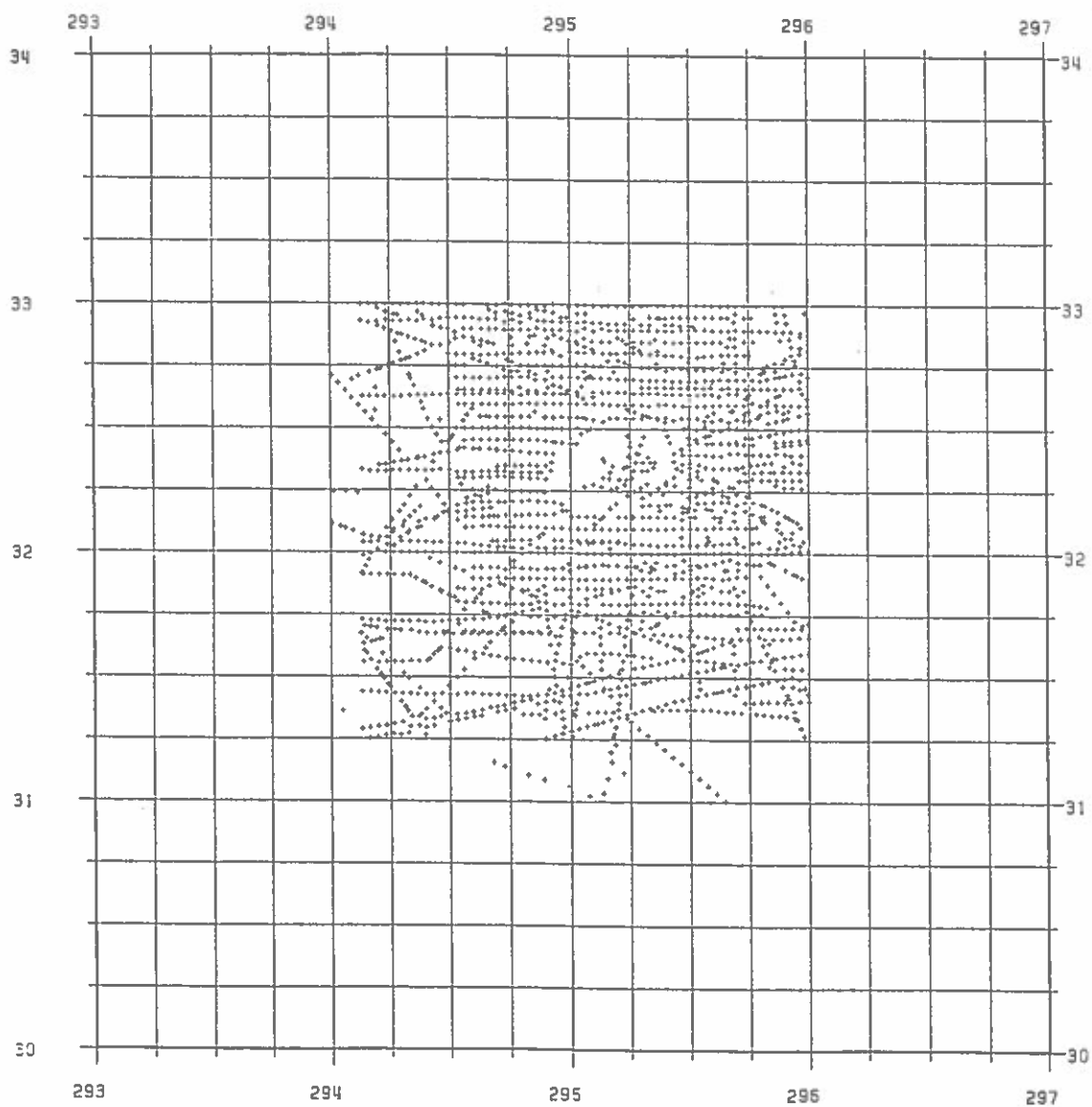


Figure 5.13 Distribution of ship determined anomalies in the Bermuda area (from NGS data)

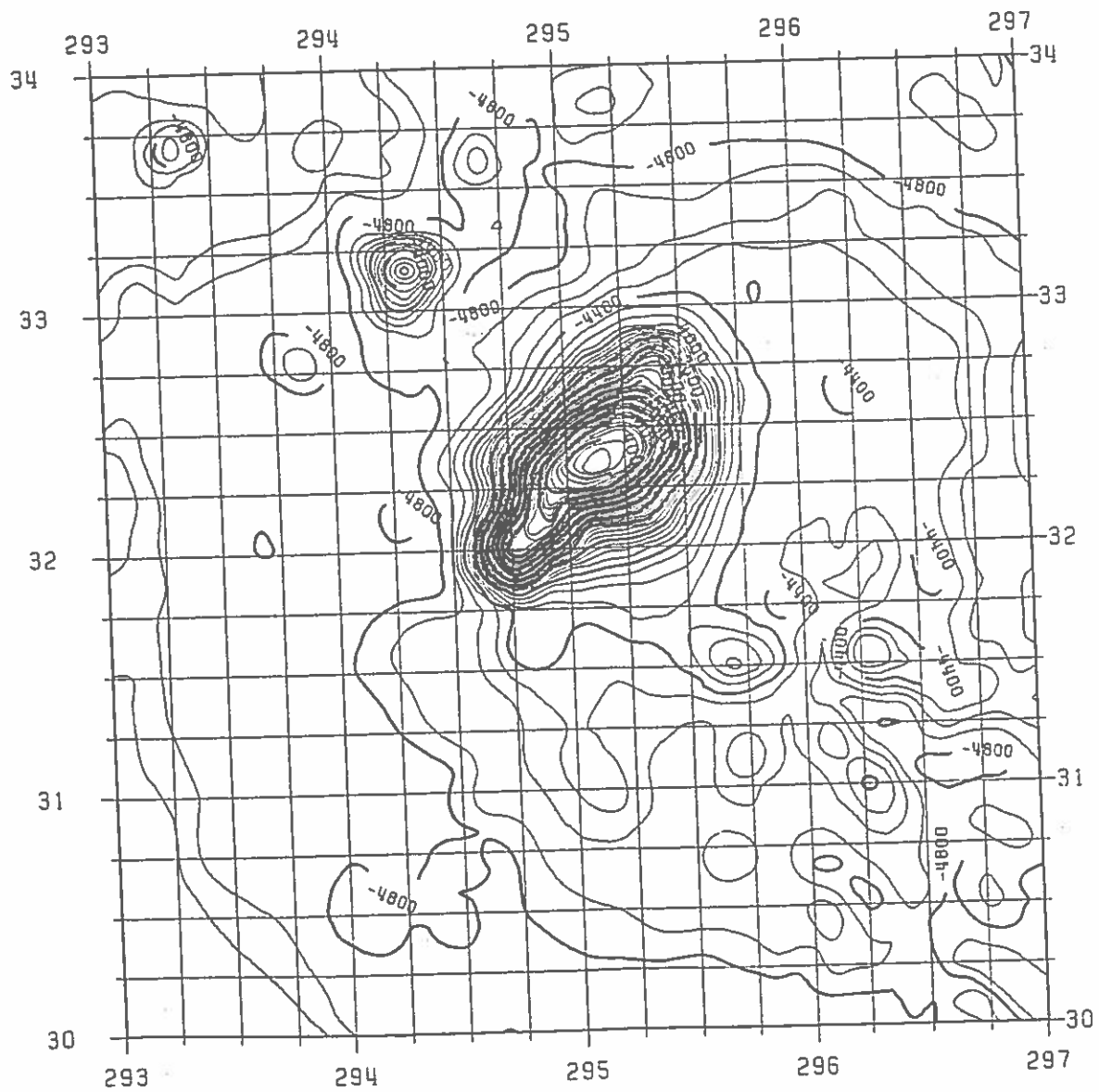


Figure 5.14 Bathymetry in the Bermuda area, CI = 100 meters

Table 5.6 Comparison between ship measurement and predicted anomalies

- (a) Location =  $30^\circ < \phi < 33^\circ$ ,  $294^\circ < \lambda < 296^\circ$   
max discrepancy = 70 mgals

Choice	max	mean	RMS
1	69.3	-7.5	17.7
2	69.6	-3.4	15.1
3	69.4	-3.6	15.7
4	69.8	-6.6	16.8
5	69.8	-9.3	18.5
6	69.8	-7.9	17.7

- (b) Location =  $30^\circ < \phi < 33^\circ$ ,  $294^\circ < \lambda < 296^\circ$   
max discrepancy = 400 mgals

Choice	max	mean	RMS
1	183.6	-4.0	28.4
2	152.3	-1.5	21.6
3	138.2	-1.9	20.5
4	170.9	-3.1	26.6
5	216.1	-4.1	34.3
6	185.9	-3.5	30.1

The accuracies of the prediction depend primarily on the number of points used in the prediction and whether the variation of residual SSH is smooth or rapid. In an area with dense altimeter data (e.g., the Bermuda area), the standard deviations of predicted result are expected to be low, even though the gravity field is not smooth. To investigate the accuracies of predictions, we present some statistics of standard deviations for choices 1, 2, 3 in Table 5.7.

Table 5.7 Predicted standard deviations\* for choices 1, 2 and 3

Choice	max std		min std		mean std	
	$\Delta g$	N	$\Delta g$	N	$\Delta g$	N
1	18.04	0.26	6.79	0.04	10.38	0.09
2	16.16	0.16	5.68	0.03	9.32	0.07
3	14.30	0.13	5.60	0.03	8.89	0.07

\* unit for anomalies = mgals, unit for undulations = meters

From Table 5.7, it is clear that increasing number of points (i.e., from choice 1 to choice 2) in a prediction cell has improved the accuracy of prediction. As compared with the RMS differences that were shown in Table 5.6, the standard deviations in Table 5.7 could be too low. These relatively low standard deviations should be due to using Jekeli's parameters. If T/R's parameters were used, the standard deviations of the predictions could be more realistic. This will be discussed later in this chapter.

In order to detect the capability of recovering detailed anomalies, profile plots along selected latitude belts were made for choices 1, 2 and 3. These profiles are shown in Figures 5.15, 5.16 and 5.17, respectively. Generally speaking, good agreement between ship and predicted anomalies took place in the low amplitude domain. Nevertheless, the high amplitude part has been seriously degraded in all of the three choices; this effect leads to the large differences in anomalies reflected in Table 5.4. However, it is obvious that the predicted anomalies can be improved by reducing the prediction cellsize or by increasing the number of altimeter points in the prediction cell. For example, at the latitude belt of  $32^{\circ}0$ , the result from choice 1 generated a smooth profile, but the result from choice 2 made the profile closer to the ship measurement's, the best agreement then was obtained by choice 3. For the other latitude belts, similar phenomena can be observed.

The possible explanation of the loss of the detailed anomalies in the altimeter predictions could be:

- No altimeter observations were made at the points where large anomalies occurred.
- The density of altimeter data is not sufficient
- The observed SSH were corrected by improper tide models
- Covariance function is not proper.

Regarding item 2, several tests were made by Kadir (1988) by using very dense data in a small prediction cell (e.g., 300 points in  $0^{\circ}25$  cell, see p. 21, *ibid.*). In these tests, he was able to recover the detailed anomalies without losing the highest signature such as seen in Table 5.4, cf. Kadir (*ibid.*). Unfortunately, the current Geos-3/Seasat data base does not allow such predictions to be made everywhere in the ocean area. For example, in the South Pacific area where we just finished the bias removal tests, the

average number of altimeter points in a  $1^\circ \times 1^\circ$  prediction cell (with  $0.5^\circ$  border) varied from 100 to 200, so that the test carried out by Kadir (ibid.) can only be made in limited regions. Another concern is the large computer resources needed for the prediction with the very dense data.

Although the current density of altimeter data is sparse in some regions it will be helpful to understand what sort of density of altimeter data that will make the high precision anomaly recovery possible from the experiments performed in the area such as Bermuda, where the distribution of altimeter points can be considered to be very dense. For example, in a  $1.5^\circ \times 1.5^\circ$  area, the number of altimeter points can be as high as 2000, and in a  $1^\circ \times 1^\circ$  area, the maximum number can reach 1430. Such dense data provides an opportunity for flexible selection of data points and also an opportunity to recover more detailed anomalies and SSHs.

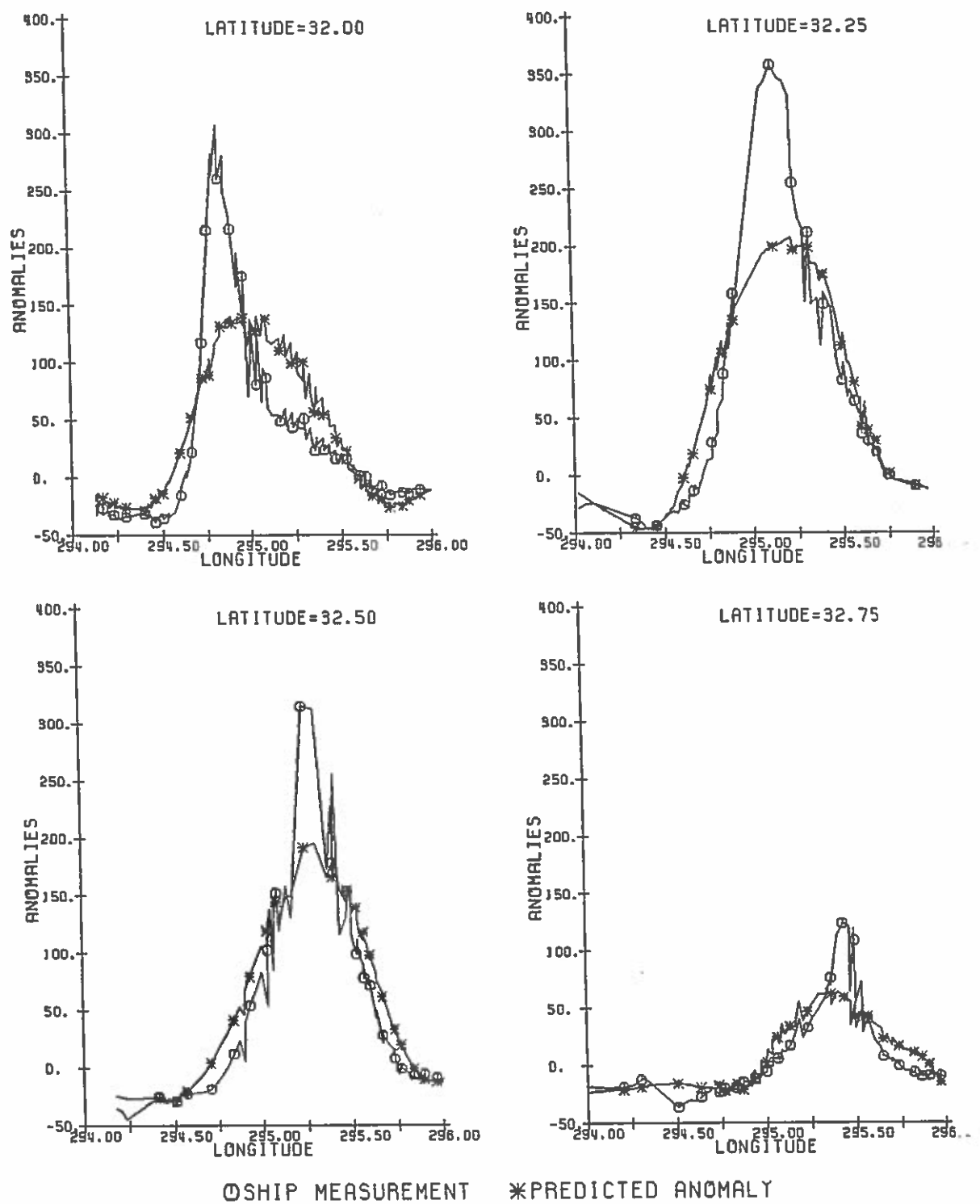


Figure 5.15 Profiles of anomalies for choice 1

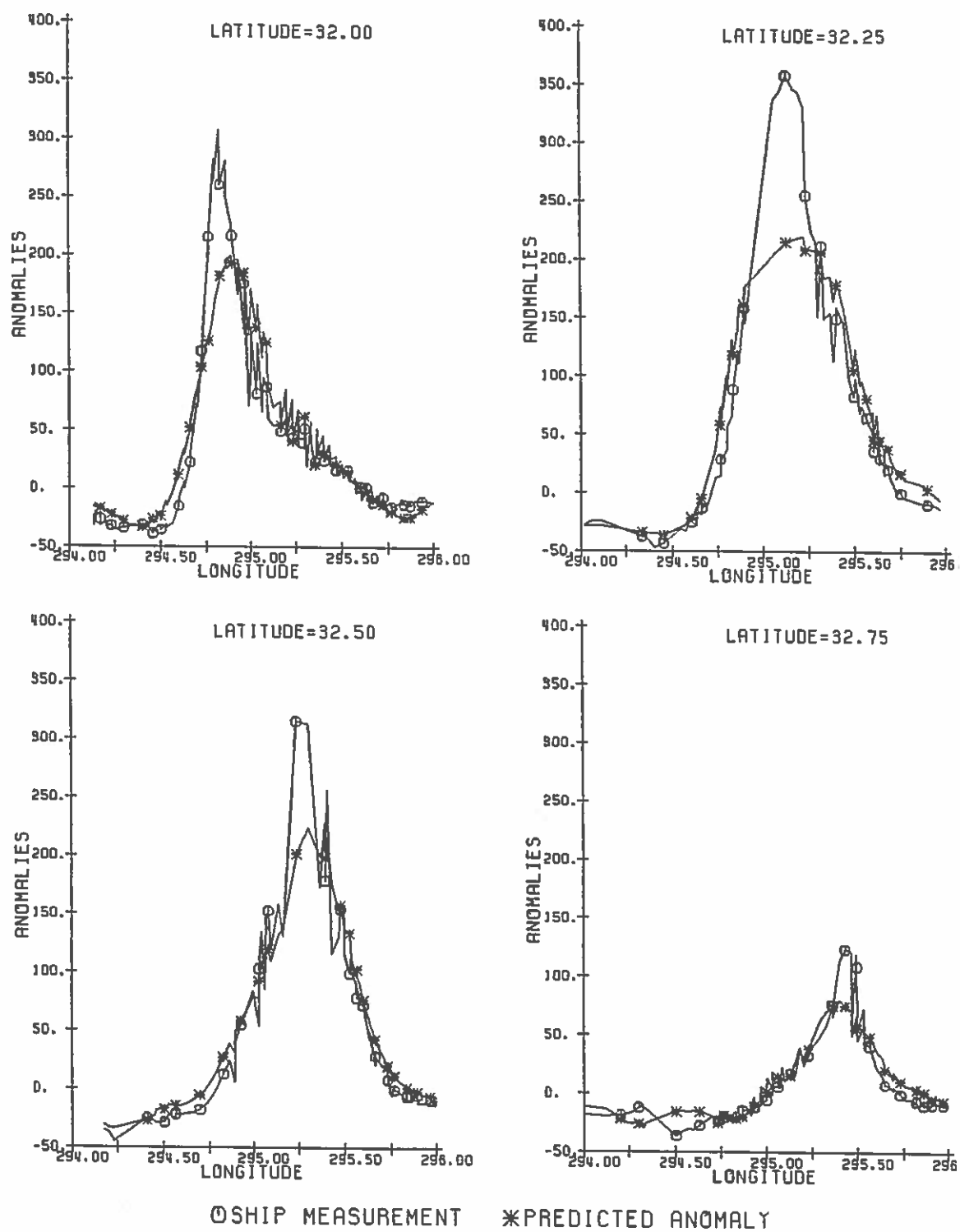


Figure 5.16 Profiles of anomalies for choice 2



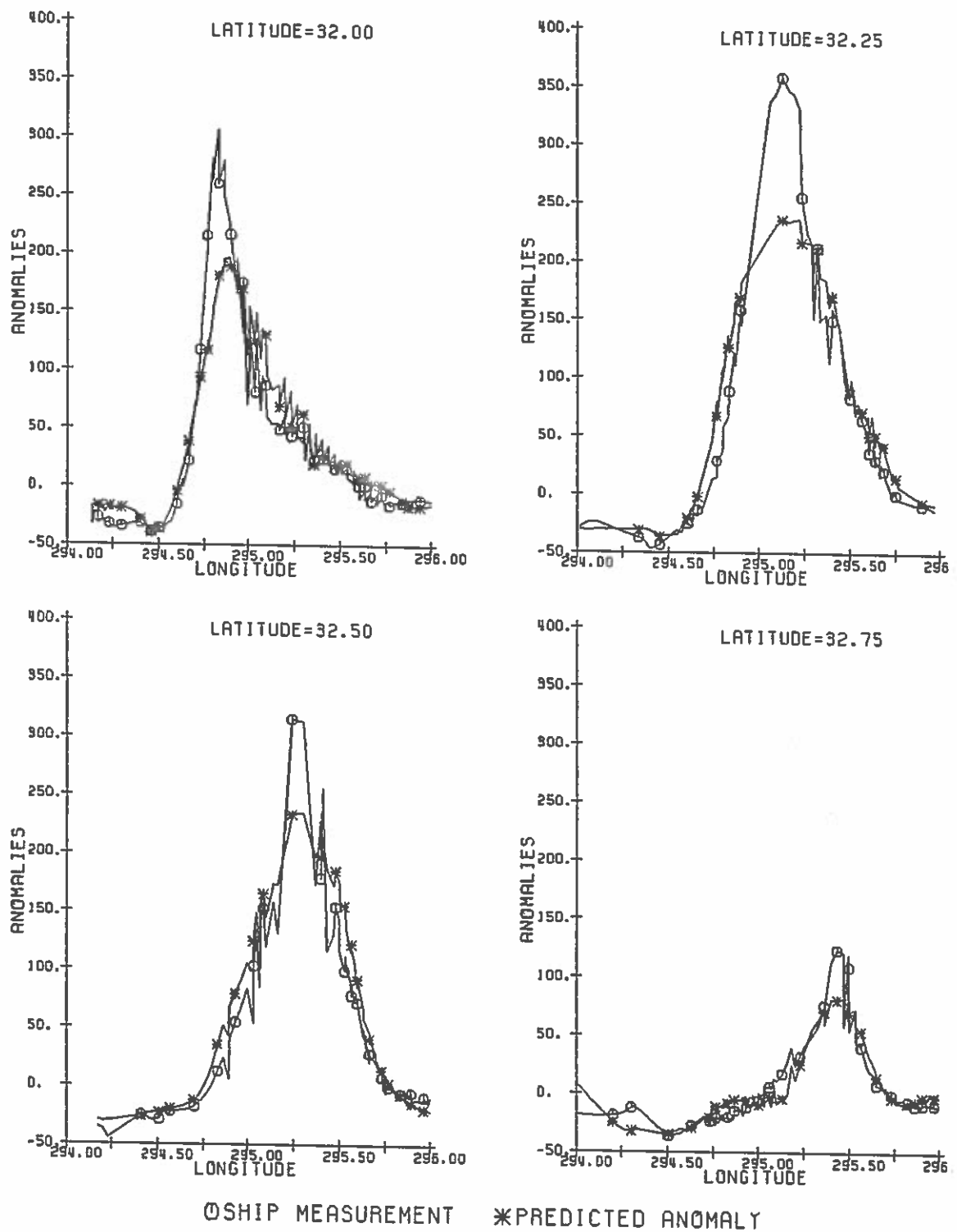


Figure 5.17 Profiles of anomalies for choice 3

Finally, we will draw some conclusions from choices 1 through 6:

- From choices 1 and 6:  $2^\circ$  and  $5^\circ$  adjustment block sizes yield the similar result.
- From choices 1 and 2:  $0.5^\circ$  prediction cell is better than  $1^\circ$  prediction cell.
- From choices 2 and 3: increasing the altimeter points can yield better agreement with ship data.
- From choices 1 and 4: including arcs which have no bias parameters can produce better agreement.
- From choices 1 and 5: Increasing the standard deviations of the SSHs will degrade the result.

#### 5.4 Consistency between the observed SSHs and reference geoid heights and the determination of maximum acceptable residual SSH in a prediction cell

One of the editing criteria in selecting altimeter points is related to the residual SSH within the prediction cell. The procedure of such editing begins with computing the standard deviation of the residual SSH for all the points falling into the prediction cell. Then, the maximum acceptable difference between the residual SSH and the mean residual SSH is 2.5 times of the standard deviation of the residual SSH or a specified extremum if such maximum value exceeds the specified extremum. The extremum used by all the tests in the above sections was 3.6 meters with which very few points were rejected, even in the rough area such as Bermuda.

However, Rapp (1985, p. 92) has observed that the maximum sea surface horizontal gradient can reach 0.28 m/km in the Kuril trench. This means that for a  $0.5^\circ$  prediction cell with a  $0.5^\circ$  border, the variation of the observed SSH could be quite large in an area of  $1.5^\circ \times 1.5^\circ$ . This information indicates the need of investigating the behavior of observed SSH, especially in the rough area such as trenches and seamounts. In the early stage of production work, many large standard deviations of residual SSHs were found in the Kamchatka Trench (roughly at  $50^\circ < \phi < 60^\circ$ ,  $160^\circ < \lambda < 170^\circ$ ) and hence many points were deleted due to the 3.6 meters limit. Therefore, this area provided a good opportunity for investigating the agreement between SSHs and reference geoid heights.

Two sets of profiles of SSHs (combined Geos-3/Seasat data) and reference geoid undulations are examined in the two areas (the profiles were interpolated by GSPP (Süinkel, 1980) using inverse distance of power 2 as weights):

- $54.0 < \phi < 55.50$ ,  $165.5 < \lambda < 168^\circ$
- $52.0 < \phi < 53.50$ ,  $158.5 < \lambda < 161^\circ$

The corresponding plots are given in Figure 5.18 and 5.19 respectively.

These two sets of profiles show the significant inconsistency between the observed SSH and reference geoid undulations. The inconsistency is partly due to the resolution of  $180 \times 180$  field and partly due to the imperfection of OSU86E geopotential model in these areas. However, it is clear that the variation of observed SSHs itself is continuous and no evidence of track errors can be found. The numerical listing also showed that the observations are quite consistent. In particular, both observed SSH and reference geoids generate the similar sea surfaces along the six profiles in Figure 5.19 and this implies that the observed SSHs can properly reflect the fluctuations of the sea surface. Furthermore, a change of 12 meters in one degree was found along the profile at  $\phi = 54.60$  in Figure 5.18. The largest difference for all the Geos-3/Seasat between SSH and reference geoid heights was 12.50 meters, which took place along the profile at  $\phi = 53.20$  in Figure 5.19.

With such possible drastic variations of sea surface within a small area, the limited extreme of 3.6 meters thus becomes unpractical since some detailed information could be lost due to the deletion of the points where differences are greater than 3.6 meters. Based on the altimeter data in the Kuril trench and the Kamchatka Trench, the average RMS difference between SSH and reference geoid heights is on the order of 2-3 meter, thus it is suggested that 6 meters (roughly 2.5 times of the RMS residual) be the maximum acceptable difference.

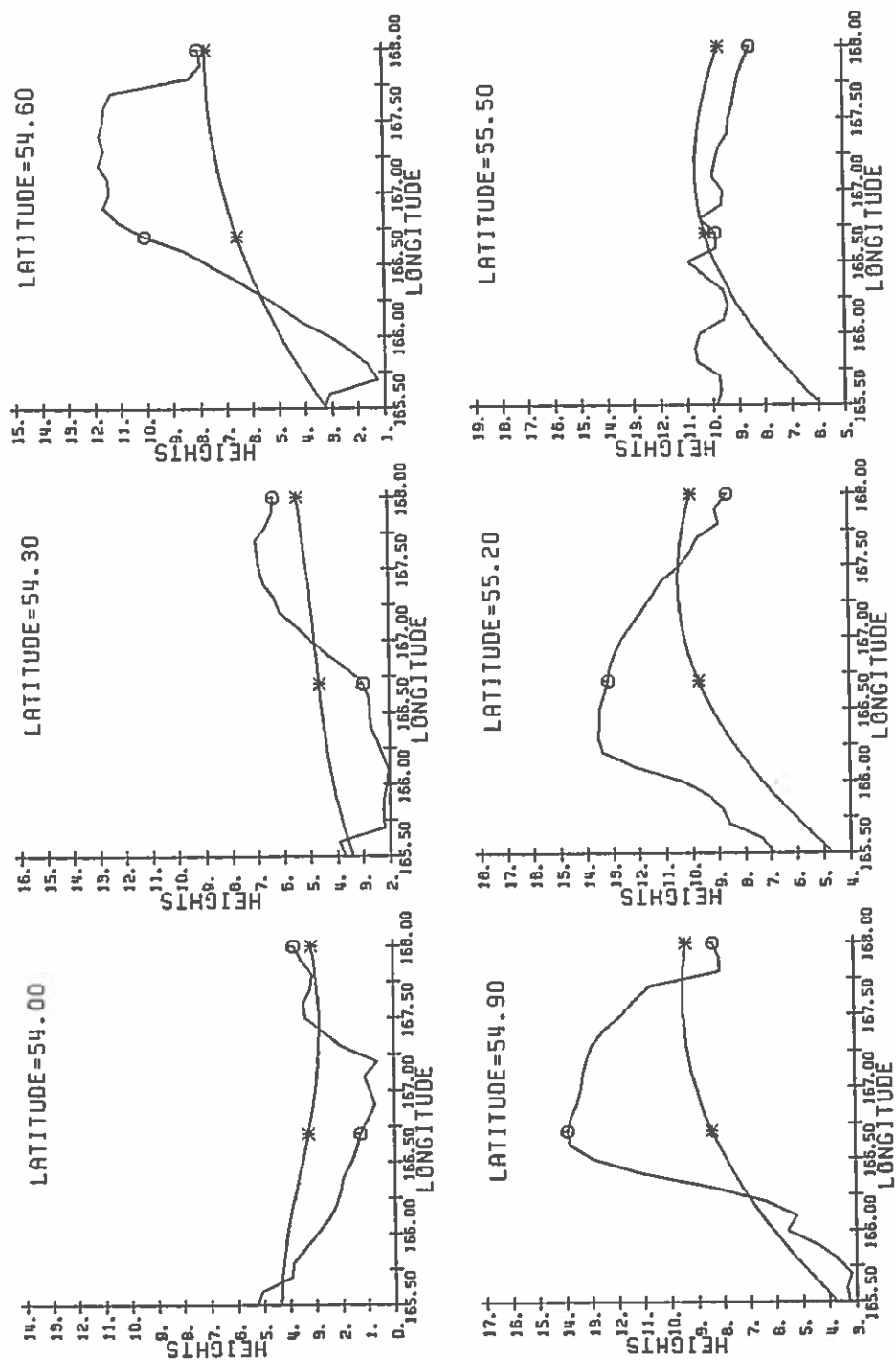


Figure 5.18 Profiles of observed SSH and reference geoid heights.

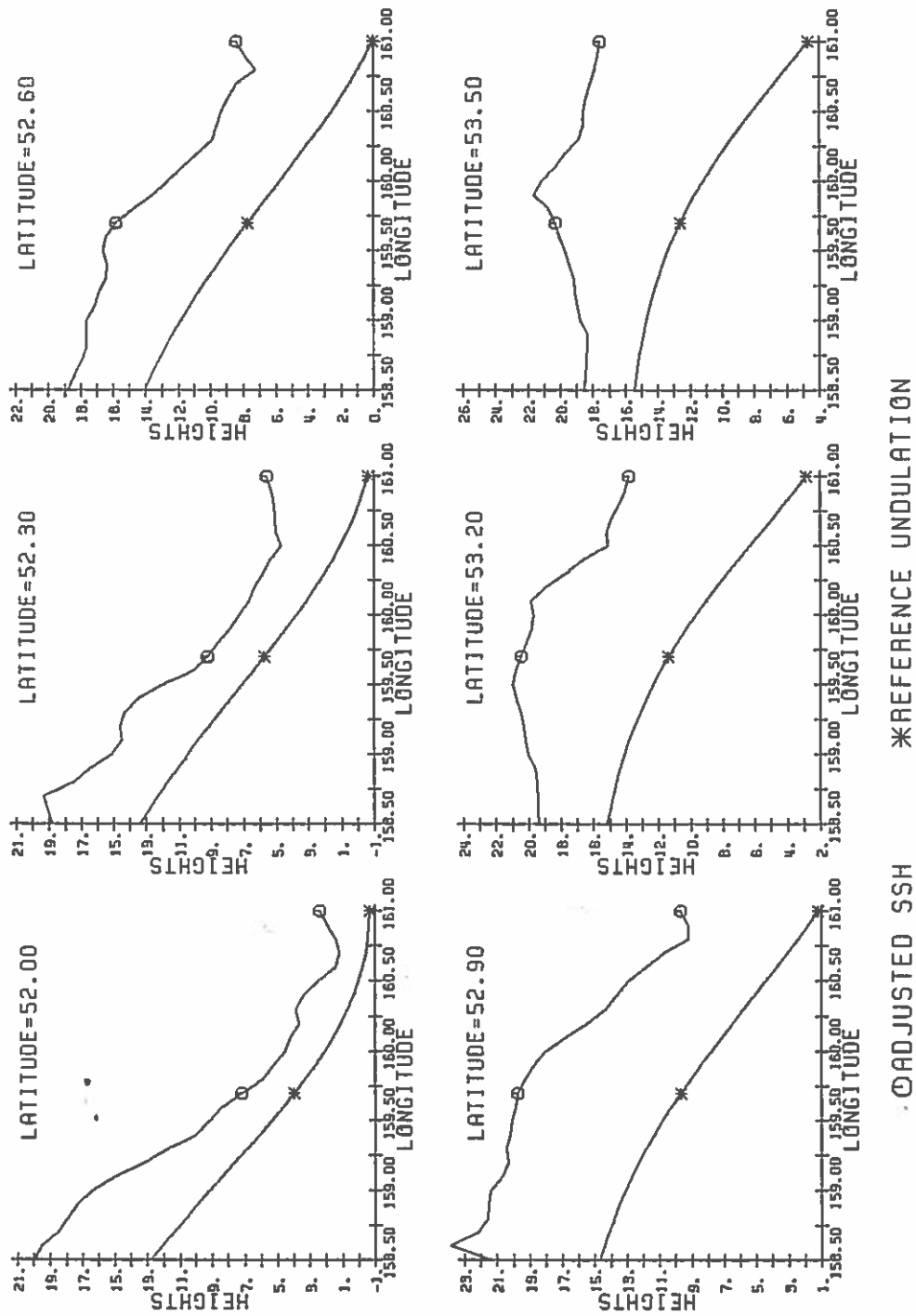


Figure 5.19 Profiles of observed SSH and reference geoid heights.

### 5.5 Matrix singularity problem and its solution in least squares collocation for the global production work

When it came to the stage of first production run at the first block, i.e., at  $40^\circ < \phi < 72^\circ$ ,  $0^\circ < \lambda < 40^\circ$ , the matrix

$$C_{hh} + \frac{1}{\alpha} C_{nn} \quad (5-1)$$

became singular in some prediction cell with the choice of  $4^\circ$  adjustment blocksize,  $0.5^\circ$  prediction cell ( $0.5^\circ$  border) and Tscherning/Rapp's model with Jekeli's parameter. The phenomenon is the high scaling factor  $\alpha$  for those cells having singularity problem.

The first singularity was found in the cell

$$70.0^\circ < \phi < 70.5^\circ, \quad 28.5^\circ < \lambda < 29.0^\circ \quad (\text{not including border})$$

within which only Seasat data is available. Since the latitude is high, the points are close to each other, as indicated in Figure 5.20.

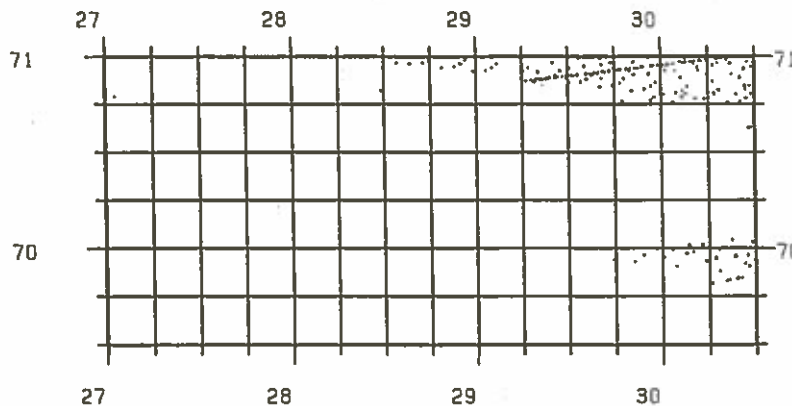


Figure 5.20 Point distribution in the cell where a matrix singularity occurred

The reason for singularity is obviously that two or more rows in matrix  $(C_{hh} + 1/\alpha C_{nn})$  are effectively linearly dependent, which in turn means that either the covariance function of the undulation is so smooth that the regularizing matrix  $1/\alpha C_{nn}$  becomes invalid or many points are so close that the matrix is still singular regardless of the behavior of covariance function. Therefore, our solution to the singularity problem will be based on the information derived from the distribution of the data, covariance function characteristic, and the regularizing matrix  $1/\alpha C_{nn}$ .

It is obvious that the singularity can be avoided by specifying a minimum spacing in the data selection procedure, but care must be exercised since the selection of points might cause the loss of detailed information. To give an example of minimum spacing between altimeter points that will make the matrix invertible, the status of matrix inversion is given in Table 5.8 at the prediction cell whose north-west corner is  $\phi=55^\circ 0'$ ,  $\lambda=166^\circ 0'$ .

Table 5.8 Status of matrix inversion at  $\phi=55^\circ 0'$ ,  $\lambda=166^\circ 0'$

max points used	min spacing	inver. cond.	no. of pts rejected*
315	1 km	error	2
315	1.5 km	error	3
315	2.0 km	no error	8

\* Due to residual SSH criterion

In Table 5.8, the covariance model was T/R's with Jekeli's parameters, and the limiting value for residual SSH is 3.6 meters. The choice was  $4^\circ$  adjustment blocksize and  $0.5^\circ$  prediction cell with  $0.5^\circ$  border.

Although the singularity problem was solved by specifying 2 km as the minimum spacing, this value does not guarantee no inversion error in other places. Another important consideration is the stability of the inversion even though the matrix is successfully inverted. For example, in the first example of matrix singularity problem ( $70^\circ 0' < \phi < 70^\circ 5'$ ,  $28^\circ < \lambda < 29^\circ 0'$ ), the minimum distance between any two points was found to be 317 meters in the cell, so that a specified minimum spacing of 500 meters was able to make the matrix positive definite. However, the predicted quantities became unstable since many unreasonably large quantities (residual anomalies and standard deviations)

were found. Such numerical instability was then overcome by specifying a minimum spacing of 1000 meters. This example also shows the importance of numerical stability in the interpretation of the predicted quantities.

Another concern is the limiting value for the residual SSH. Based on the analysis in section 5.4, 6 meters was suggested to be such a limit. Furthermore, the covariance function is also an important factor in the matrix singularity problem. In the following example, we will combine these considerations and analyze the inversion status.

Table 5.9 Status of matrix inversion at  $\varphi = 53^{\circ}0$ ,  $\lambda = 159^{\circ}0$

max points used	residual limit	min spacing	model	inver. cond.
315	3.6	2 km	Jekeli	error
315	3.6	3 km	Jekeli	no error
330	6.0	3 km	Jekeli	error
330	6.0	1 km	T/R	no error

In Table 5.9, the choice for adjustment blocksize and prediction cell are the same as those in Table 5.8. The model of Jekeli (in Table 5.9) is the T/R's model with  $A=343.408 \text{ mgals}^2$ ,  $B = 24$ ,  $s = 0.9988961$  while the model of T/R uses:  $A=425.28 \text{ mgal}^2$ ,  $B=24$ ,  $s=0.999617$ . It is noted that the T/R's model gives a more stable inversion than Jekeli's model. Thus, for future work the T/R's model with the original numerical parameters will be preferred.

A concern arises when the minimum spacing is getting larger and too many points are deleted. The concern is primarily on the quality of the predicted quantities. In order to investigate this, several tests were performed in the Bermuda area. The predicted anomalies were compared with ship measurements in the same way as described in Section 5.3. Table 5.10 shows the results.



Table 5.10 Differences between ship measured and predicted anomalies based on various choices of minimum spacings (at  $30^\circ < \phi < 34^\circ$ ,  $293^\circ < \lambda < 297^\circ$ )

Model	min spacing	max	mean	RMS	no. of pts used
T/R	3 km	117.4 mgal	-1.43 mgal	17.8 mgal	~270
T/R	1 km	117.2	-0.90	17.7	~320
Jekeli	3 km	118.6	-1.64	18.5	~270
Jekeli*	0	151.8	-2.47	18.5	~400
Jekeli**	0	138.3	-2.00	20.5	~400

unit for anomalies: mgals

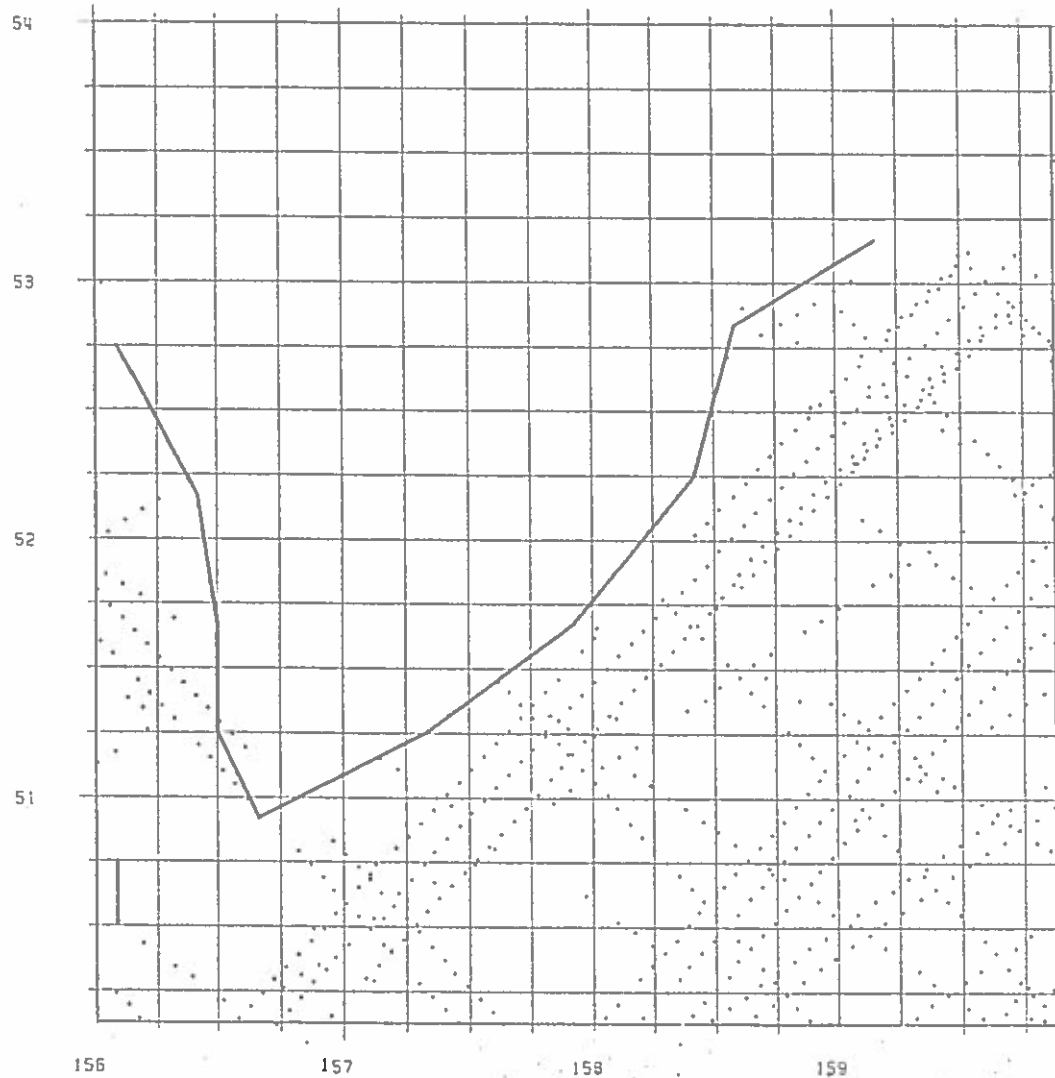
\* specified max no. of points = 400

\*\* limit residual SSH = 3.6 meters, adjustment blocksize = 2°0, specified max no. of points = 400

In Table 5.10, except for cases marked, the adjustment blocksize is 4°0 and prediction cellsize is 0°5, the maximum limiting residual SSH is 6 meters and the specified maximum number of points for the initial point selection is 330. It was found that a minimum spacing of 1 km will delete approximately 10 points and 3 km will delete 60 points. It is remarkable that the deletion of points did not degrade the prediction's quality. The use of additional points, such as in the last two cases, even yields worse results (It could be due to matrix singularity). Another success of the minimum spacing is the saving of computer time. For example, in Table 5.10, the case with the 3 km spacing took about 70 percent of the time spent by the case with 0 km minimum spacing and about 80 percent of time spent by the case with 1 km minimum spacing.

#### Scaling factors for standard deviation of SSH

From previous discussions, we realize that the matrix  $(C_{hh} + 1/\alpha C_{nn})$  can be positive definite if the distance between two arbitrary points exceeds a minimum spacing. However, in the area where the distribution of altimeter data is not even and the gravity field is relatively rough, the predicted anomalies again became unreliable even though a large minimum spacing was used. Example of such a case can be shown in Figure 5.21. In this figure, several gaps of altimeter points can be found. (A data gap means a location where data distribution is not continuous and regular as compared to the data distribution of the adjacent areas).



**Figure 5.21** Altimeter data distribution near the Kuril Trench

Using a choice of  $4^\circ$  adjustment blocksize,  $0.5^\circ$  prediction cellsize with  $0.25^\circ$  border and 4 km minimum spacing (it is not necessary to specify the maximum number of points, since the total number of points in a prediction cell can hardly exceed 100), a set of anomalies was predicted in the area given in Figure 5.21. As indicated by Figure 5.22 two circular areas of rapid anomaly change are seen at the location where data gaps occurred. The numerical listing showed no significant errors in the observed SSHs in the area. Observe that the sign of anomalies changed from positive to negative right around the loops. These unrealistic signatures of anomalies cannot be reduced even by using a 10 km minimum spacing.

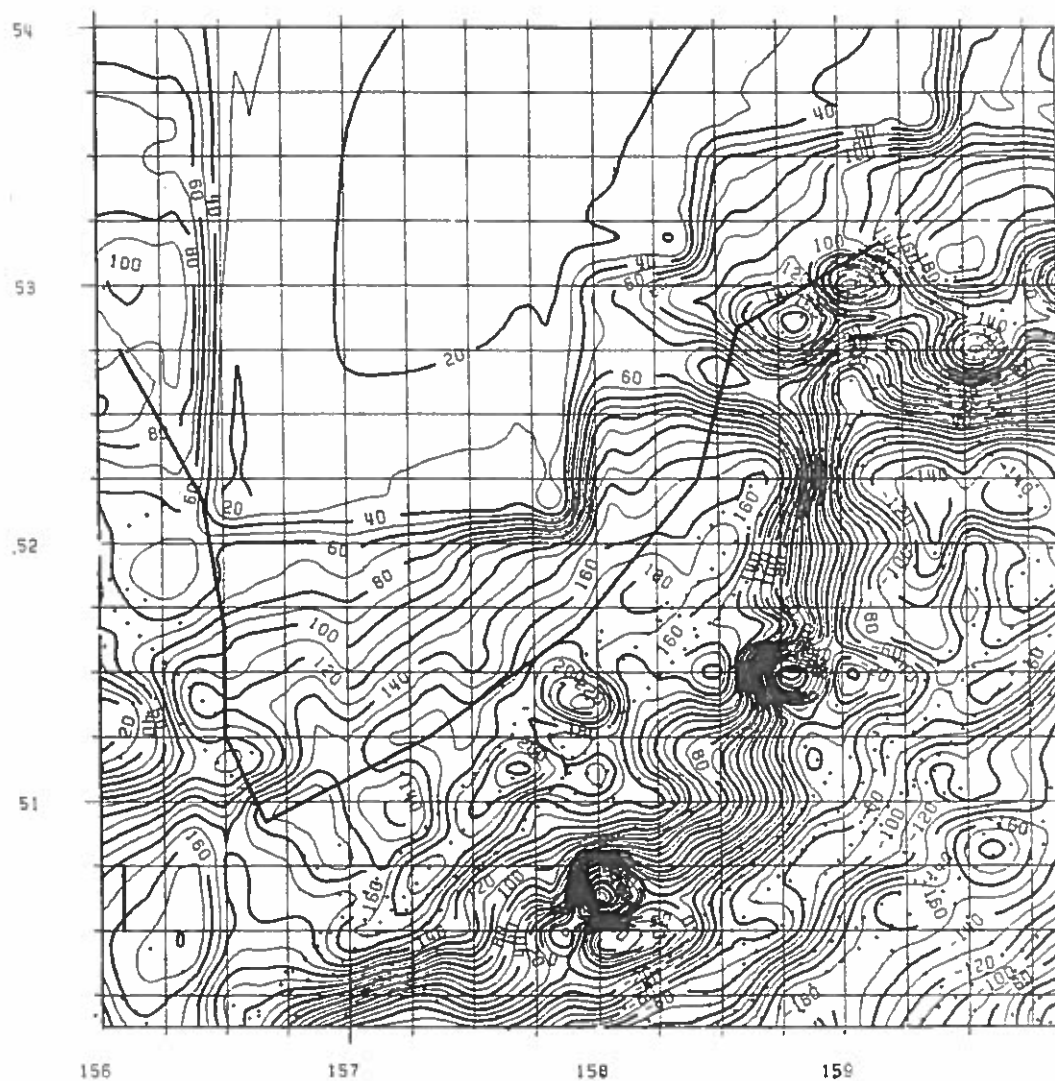


Figure 5.22 Predicted anomalies in case of gaps of point distribution. Points in the map indicate altimeter data points. CI = 10 mgals

Using the same prediction structure, but multiplying the Seasat noise by 3.0 and the Geos-3 noise by 4.0, the loops in Figure 5.22 then disappeared. Such a case is shown in Figure 5.23.

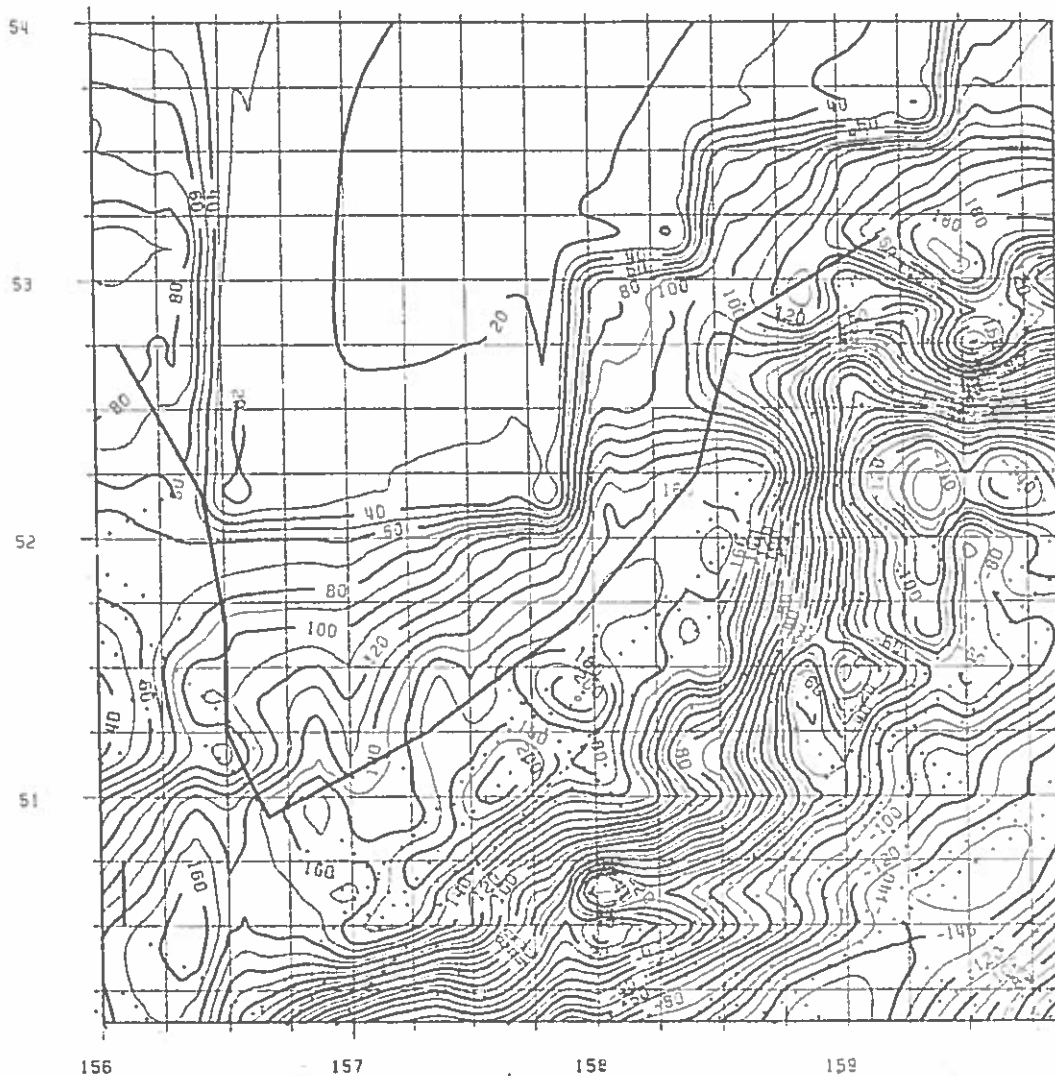


Figure 5.23 Predicted anomalies in case of multiplying 3 and 4 for Seasat and Geos-3's noises. CI = 10 mgals

This example shows the smoothing effect of the noise's scaling factor. In terms of matrix singularity problem, the noise's scaling factors become fairly important when a large  $\alpha$  value occurs, since for such a case the regularizing matrix  $1/\alpha C_{nn}$  has no significant effect to the matrix's stability. A sample listing of anomalies corresponding to Figures 5.22 and 5.23 is also given in Table 5.11. Clearly, the change of anomalies from  $\lambda = 157.875$  to  $\lambda=158.00$  is not acceptable for the first case.

Table 5.11 Sample listing of anomalies based on various noise's scaling factors

noise factors*:	(1, 1)	(3, 4)
$\varphi = 50.625$	(anomalies)	
$\lambda = 157.500$	$119 \pm 28$	$127 \pm 32$
157.625	$94 \pm 39$	$99 \pm 41$
157.750	$64 \pm 44$	$20 \pm 46$
157.875	$36 \pm 45$	$42 \pm 48$
158.000	$-165 \pm 50$	$-54 \pm 61$
158.125	$-52 \pm 58$	$-32 \pm 61$

\*First factor is for Seasat, second for Geos-3.

Although the noise's scaling factors might help the prediction's stability, it is not recommended in the area where the altimeter data has regular and even distribution, and the observed SSH do reflect the roughness of gravity field. To illustrate this, three successive anomaly plots, namely, Figures 5.24, 5.25 and 5.26, were prepared for the Bermuda area. These figures correspond to the predicted anomalies based on the scaling factors: (choice:  $4^\circ$  adjustment blocksize,  $0.5^\circ$  prediction cell size, maximum number of points used = 315, T/R's model, 2 km minimum spacing, 6 meters limit residual)

Figure 5.24 1 for Seasat, 1 for Geos-3

Figure 5.25 5 for Seasat, 7 for Geos-3

Figure 5.26 10 for Seasat, 15 for Geos-3

Enormous loss of details is found in Figure 5.26, and the signature in Figure 5.25 has reduced to such a level that it was thought to be unacceptable. Table 5.12 also shows the comparison between the predicted results and the ship measurements.

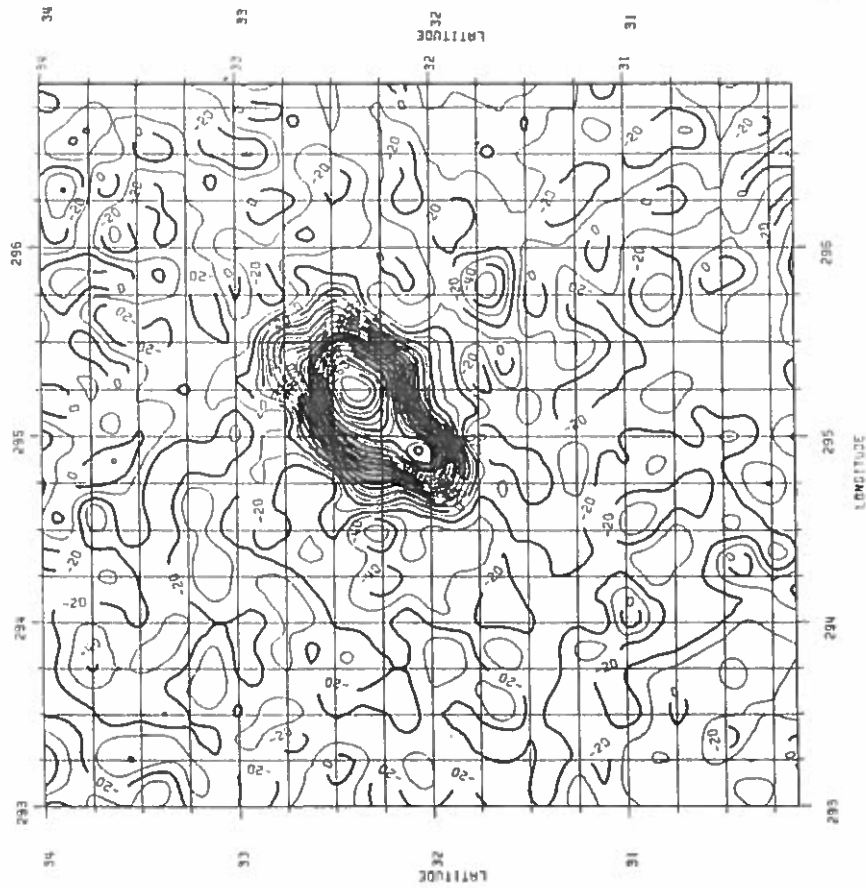


Figure 5.24 Predicted anomalies in the Bermuda area using noise scaling factors: 1 for Seasat, 1 for Geos-3

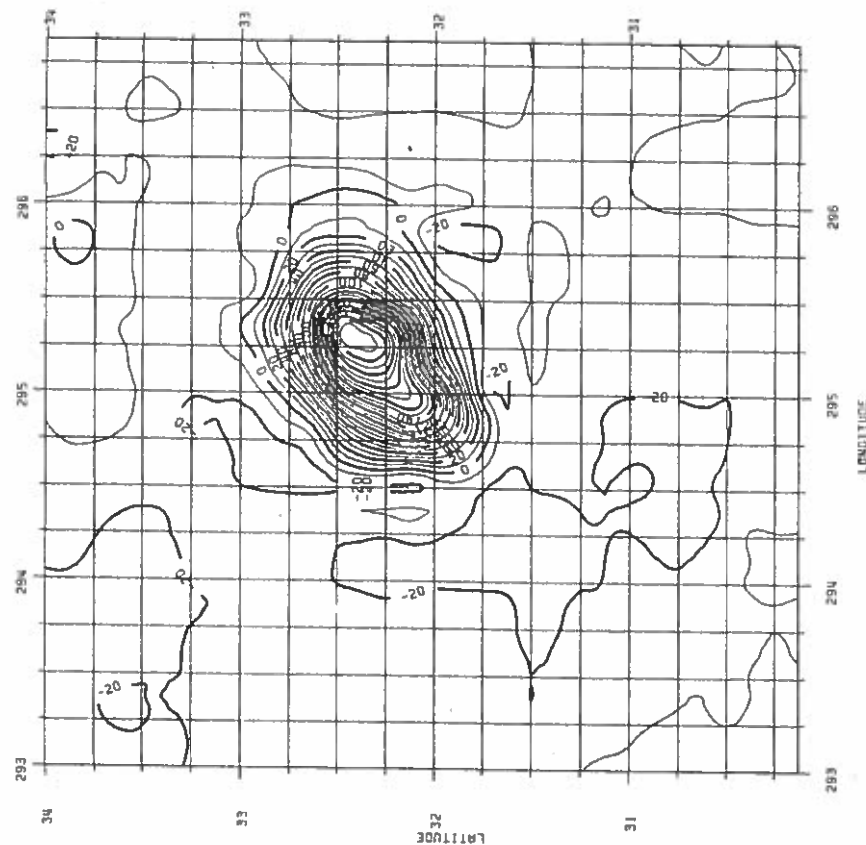


Figure 5.25 Predicted anomalies in the Bermuda area using noise scaling factors: 5 for Seasat, 7 for Geos-3

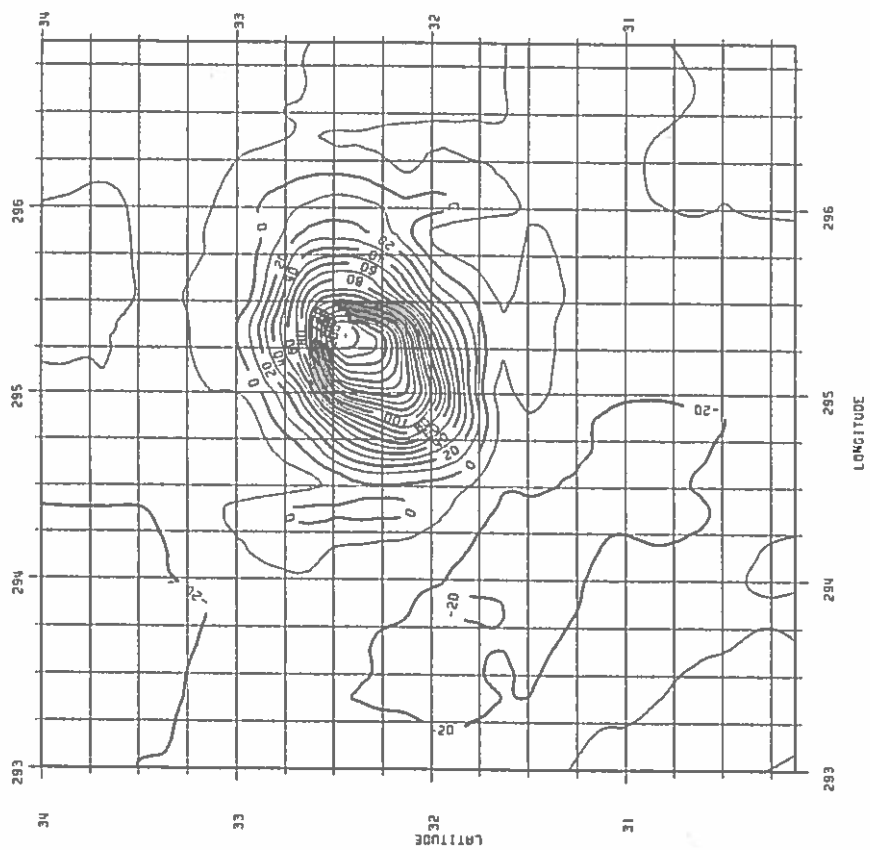


Figure 5.26 Predicted anomalies in the Bermuda area using noise scaling factors: 10 for Seasat, 15 for Geos-3

**Table 5.12** Comparison of predicted anomalies and ship measurements based on various noise's scaling factors  
unit = mgals

Figure	noise factors*	max	mean	RMS
5.24	1, 1	126.7	-1.3	19.3
5.25	5, 7	197.1	-5.2	26.3
5.26	10, 15	237.0	-7.7	33.7

\* First factor for Seasat, second for Geos-3

This test, on the other hand, gives an excellent example of the importance of the accuracy estimate of the altimeter data. The implication is: If the observed SSH are sufficiently accurate but the accuracy estimates are improperly high, the predicted anomalies are then degraded. Needless to say, if the observed SSH are seriously affected by orbit error, tidal error and other factors and large noises are inherent in observations, then the contaminated signals cannot be recovered by the lsc method and hence no recovery of highly detailed anomaly is possible.

Based on the above analysis, advantage and disadvantage of scaling factors can be found. For our global production work, we will suggest not to use the scaling factors since the case shown in Figure 5.22 can rarely occur and our tests for scaling factors in Bermuda area have shown degraded anomalies if such factors are used. Although the plot corresponding to the predicted anomalies based on scaling factors 3 for Seasat data and 4 for Geos-3 data (these are factors corresponding to the resulting anomalies in Figure 5.23) was not prepared, such anomalies have been compared with the ship data (the same ship data used in Table 5.11) and an RMS difference of 22.9 mgals was found.

In case that no scaling factors are used and erratic anomalies are predicted (such as those in Figure 5.22), an alternative way to eliminate (or detect) such anomalies is to apply some point value acceptance criteria (e.g., standard deviation, depth, etc.) after the prediction has been made. This will be discussed in Chapter 7.



### 5.6 Other considerations for altimeter data

Although the adjustment process can produce consistent Seasat and Geos-3 data in most of areas, there are large discrepancies between the mean SSH values of Seasat and Geos-3 data within some prediction cells where the data distributions are not even or the observations are nearly on land. For such a case, we will only retain the Seasat data in the cells since the formation of the combined Seasat/Geos-3 data was performed in such a way that the Geos-3 arcs were adjusted to the Seasat primary arcs. In the practical computations, the mean SSH values of the Seasat and Geos-3 were first calculated at the initial data selection stage within the prediction cell. Then, all the Geos-3 data would be deleted if the difference between the mean values was greater than a specified limit. Based on some tests, 1.5 meters was suggested to be the maximum absolute difference.

In some areas, it was found that the Seasat data with noise greater than 15 cm has abnormal SSH values as compared with the adjacent points. Therefore, it was decided that such data will not be adopted in the adjustment and prediction parts.

### 5.7 Summary of tests in the Bermuda area and the final decision for the production procedure

In addition to the tests described in the above sections, some other similar tests focusing on various issues were also carried out in the Bermuda area. Among these issues, the most important one is the determination of the data border about the prediction cell. In Table 5.13, we have used a  $0.5^\circ$  data border for cases 1 through 13 and  $0.25^\circ$  for cases 14 through 17. As we can see, better results have been achieved by using  $0.25^\circ$  data border. This is due to two reasons:

- (i) In Table 2.1, we have listed the correlation lengths of the covariance function  $\text{COV}(\Delta g_p, h_Q)$ . The correlation of  $\text{COV}(\Delta g_p, h_Q)$  is  $0.339^\circ$  for T/R's model (with T/R's parameters). With such a short correlation length, a smaller data border ( $0.25^\circ$ ) would be preferred since the distances between the predicted point and most of altimeter data points can be shorter than the correlation length.
- (ii) A larger data density has been used by the cases with a  $0.25^\circ$  data border. For example, in case 3 ( $0.5^\circ$  border), 400 points were used in a  $1.5^\circ \times 1.5^\circ$  area, the

corresponding data density is 178 points per square degree. In case 16 ( $0.25^\circ$  border), 198 points were used in a  $1^\circ \times 1^\circ$  area (see Table 5.14), the corresponding data density is 198 points per square degree. In addition to the fact that case 16 has larger data density, we also performed the selection of altimeter data points so that any two points will be at least 4 km apart in case 16. This larger data density not only improved the quality of the prediction but also reduced the computer time enormously.

Based on Table 5.13 and Table 5.14 and all the tests carried out in the current chapter, we finally decided to use the following parameters for the global production mode:

Reference Field = OSU86E to degree 180  
 Covariance function = scaled Tscherning/Rapp with the T/R parameters  
 Adjustment block size =  $4^\circ \times 4^\circ$   
 Adjustment border =  $0.5^\circ$  about block border  
 Prediction cell size =  $0.5^\circ$  (for one matrix inversion)  
 Prediction border =  $0.25^\circ$  about cell border  
 Data number = approximately 400 points in one prediction cell  
 Prediction grid interval =  $0.125^\circ$   
 Arcs without parameters = included  
 Minimum spacing = 4 km  
 Scaling factors for noises = none (i.e., 1, 1)  
 Residual limit for SSH = 6.0 meters  
 Minimum no. of points in one prediction cell = 7  
 Applying point acceptance criteria in sec. 5.6

Such a choice corresponds to case 16 in Table 5.13. The predicted anomalies in the Bermuda area using the production mode are shown in Figure 5.27. The profiles of anomalies at selected latitude belts are also given in Figure 5.28. From Figure 5.28, it is found that the excellent agreement between ship measurements and predicted anomalies has been achieved by using our production mode. In particular, the highest anomalies along profile  $\phi = 32.25$  and  $\phi = 32.50$  have been successfully recovered. Comparing the anomalies in Figures 5.24 and 5.27, two observations can be made: First, in the flat area the two sets of anomalies have good agreements; Second, in the rough area (near the

Bermuda Islands) the two sets of anomalies have similar trends of signatures but the magnitudes along the peak of the signature are substantially increased. (If one draws a line from the point at  $\phi=32^{\circ}0$ ,  $\lambda=294^{\circ}875$ , to another point at  $\phi=32^{\circ}50$ ,  $\lambda=295^{\circ}375$ , this line will correspond to the peak of the signature). Therefore, the signatures resulting from the use of the production mode have decreased the large discrepancies between ship determined and predicted anomalies along the profiles  $\phi = 32^{\circ}25$  and  $\phi = 32^{\circ}50$ .

Table 5.13  
Summary of tests in the Bermuda area

I Cases

case	model	adjust. block	border	prediction cell	border	min spacing	factors	max no.pt.	residual limit
1	Jekeli	2°x2°	0°5	1°x1°	0°5	0	1,1	300	3.6 m
2	"	"	"	0°5x0°5	"	0	1,1	300	"
3	"	"	"	"	"	0	1,1	400	"
4	"	"	"	1°x1°	"	0	1,1	300	"
5	"	"	"	"	"	0	1,5,2	300	"
6	"	5°x5°	"	"	"	0	1,1	300	"
7	T/R	4°x4°	"	0°5x0°5	"	2km	1,1	315	6 m
8	"	"	"	"	"	2km	5,7	315	"
9	"	"	"	"	"	2km	10,15	315	"
10	"	"	"	"	"	6km	1,1	450	"
11	"	"	"	"	"	10km	1,1	450	"
12	"	"	"	"	"	4km	3,4	315	"
13	"	"	"	"	"	4km	1,1	315	"
14	"	"	"	"	0.25	4km	3,4	315	"
14A	"	"	"	"	"	4km	3,4	400	"
15	"	"	"	"	"	4km	1,1	315	"
15A	"	"	"	"	"	4km	1,1	400	"
16*	"	"	"	"	"	4km	1,1	400	"
17	"	"	"	"	"	3km	1,1	400	"

\* Applying point acceptance criteria (see sec. 5.6)

- For cases 1, 2, 3, 5, and 6, the arcs without parameters are not used in the prediction. For the remaining cases, such arcs are used.

Table 5.14

Summary of the results of the tests in the Bermuda area

case	max diff.	mean diff.	RMS diff.	ave. no. of pts used
1	183.6	-4.0	28.4	300
2	152.3	-1.5	21.6	300
3	138.2	-1.9	20.5	400
4	170.9	-3.1	26.6	300
5	216.1	-4.1	34.3	300
6	185.9	-3.5	30.1	300
7	126.7	-1.3	19.3	280
8	197.1	-5.2	26.3	280
9	237.0	-7.7	33.7	280
10	184.5	-1.6	22.6	200
11	160.1	-3.3	23.4	100
12	158.3	-4.0	23.3	227
13	122.7	-1.4	19.5	227
14	224.9	-3.5	22.9	177
14A	200.4	-2.8	20.8	198
15	145.1	-2.1	16.4	177
15A	113.4	-1.5	15.9	198
16	113.4	-1.5	15.9	198
17	113.9	-1.7	15.3	235

\* Case 16 is the production choice.

In the Bermuda area, case 15A is equivalent to case 16

As mentioned in section 5.3, the standard deviations of predictions from use of the Jekeli's parameters (in covariance function) were relatively low as compared to RMS differences of ship measurement comparison (see Table 5.6 and 5.7). However, the standard deviations of the results from use of the T/R's parameters, have better reflected the accuracies of the predictions in terms of ship measurement comparison. For example, the anomalies shown in Figure 5.26 have a maximum standard deviation of 26.2 mgals and a minimum of 7.6 mgals. The mean standard deviation is 12.3 mgals, which should be more appropriate for such an area with rough gravity field.

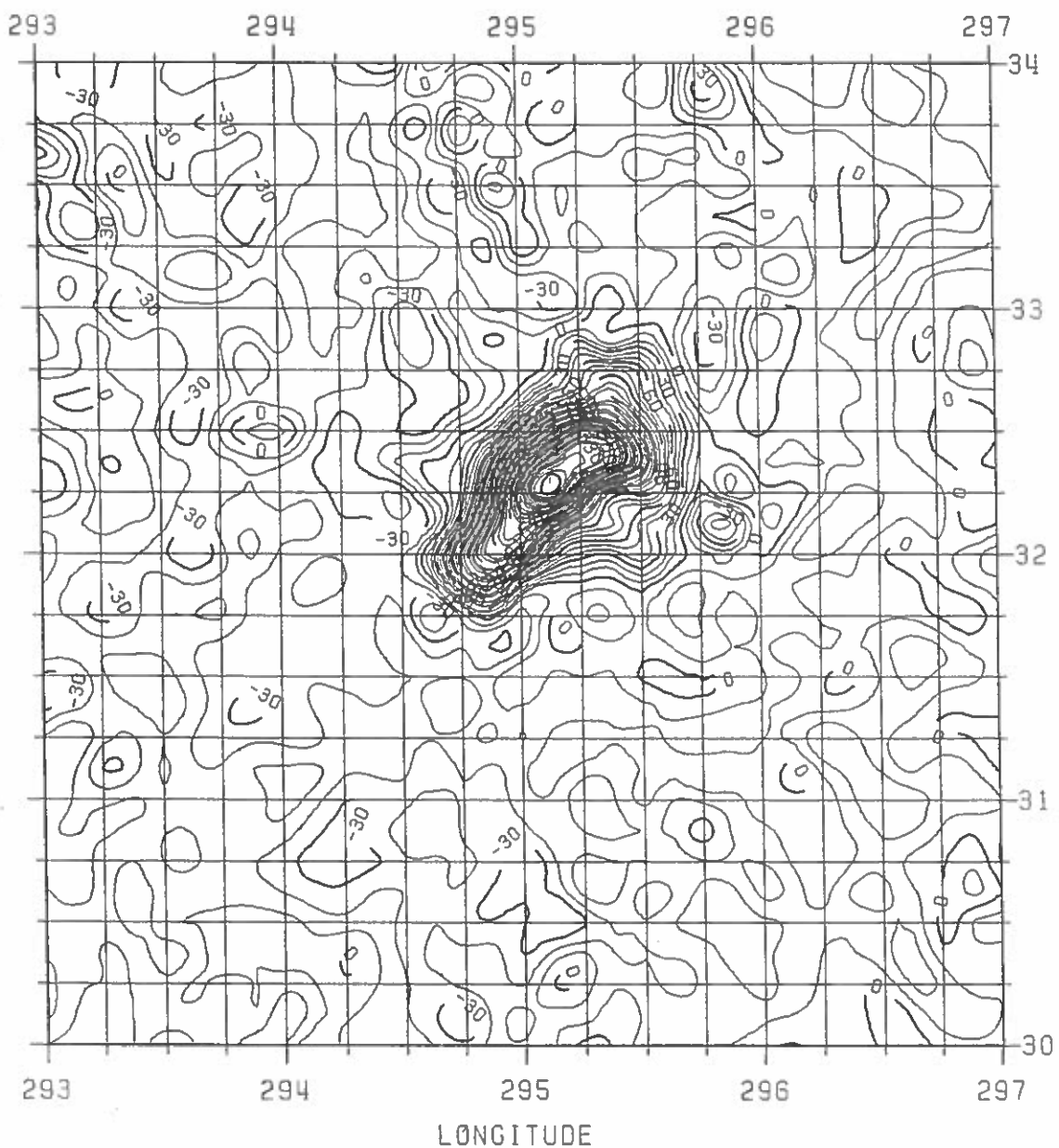


Figure 5.27 Predicted anomalies in the Bermuda area from the standard production mode (see text), highest anomaly is 371.83 mgals. CI = 10 mgls.

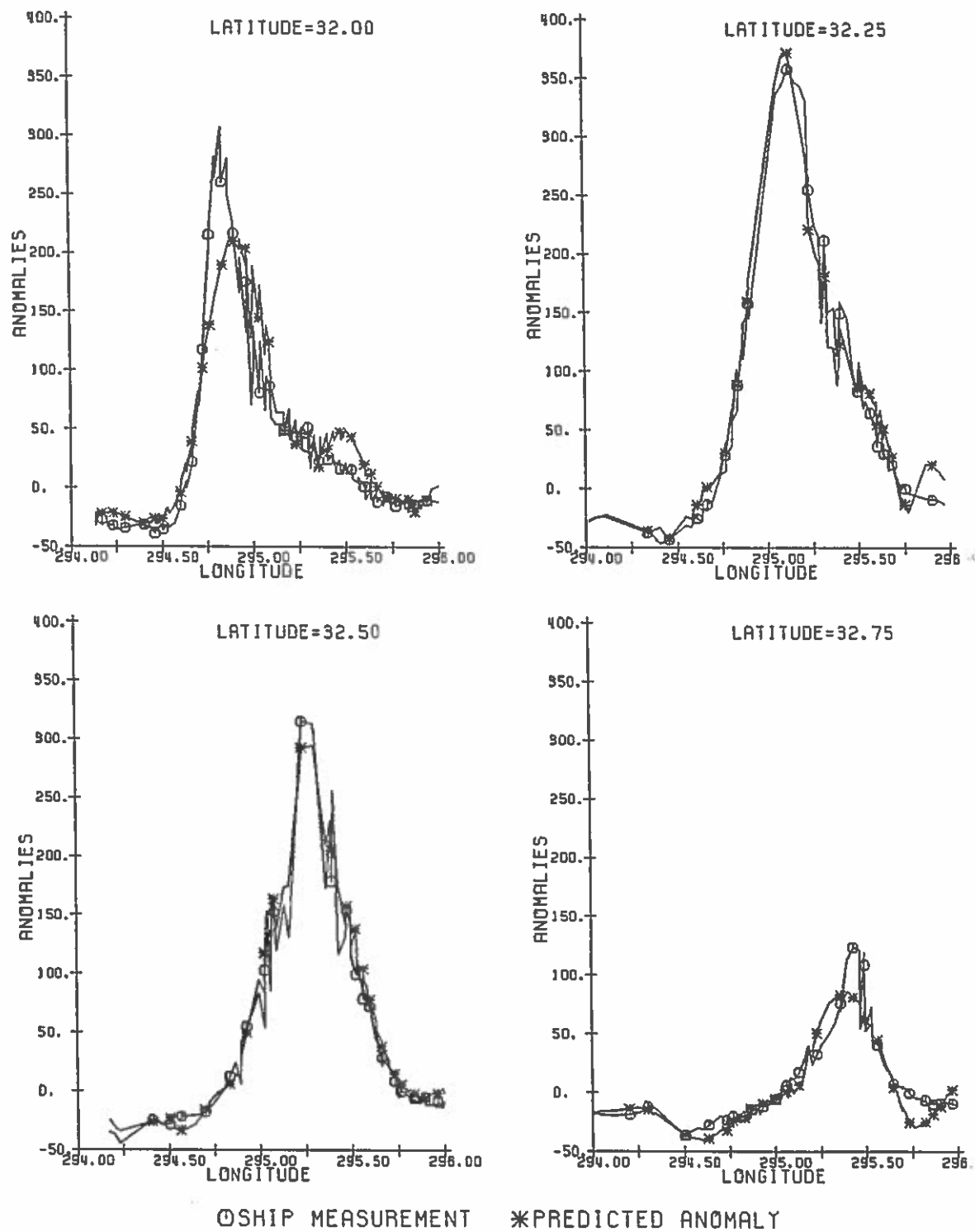


Figure 5.28 Profiles of anomalies for the standard production mode in the Bermuda area

## CHAPTER VI

### Global prediction of anomalies and sea surface heights

#### 6.1 The initial prediction

The initial prediction contained the global production of gravity anomalies and SSH using the combined Geos-3/Seasat data base. The predictions were carried out on the CRAY X-MP/28 machine residing at the Ohio Supercomputer Center while the data have been stored on the tapes on the IBM system. In addition, the required data sets such as covariance tables,  $1/8^\circ$  reference values also resided on the IBM system. As the prediction run was completed on the CRAY, all the output data had to be transferred back to the IBM system for further processing. The steps and data involved in the production work are indicated in Figure 6.1.

With the decisions made in Section 5.7, we decided to perform the global prediction by following the altimeter data storage scheme on the two master tapes. The storage scheme provided the systematic organization of data according to the geographical areas of altimeter data. Typically, a file on the tapes contains the altimeter data in a  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$  area with sequential file numbers running from north to south and west to east. More specifically, tape GS386 contains data in the Northern Hemisphere and tape GS387 contains data in the Southern Hemisphere. The file numbers on the two tapes, the prediction block numbers and the numbers of altimeter data points in the files are shown in Figure 6.2.

By (prediction) block number we mean that it is the sequence number associated with one prediction run. After some preliminary tests, it was confirmed that the prediction of gravity anomalies and SSHs in one  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$  block can be completed in one execution of the production program, F459PRD. A prediction run thus means the execution of program F459PRD with prediction coverage of one  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$  block. However, due to the limited temporary storage space of CRAY discs (more specifically, the Solid State Device) and maximum accessible memory, we had to read the altimeter data within the prediction block several times. This factor had a significant influence on the production work.

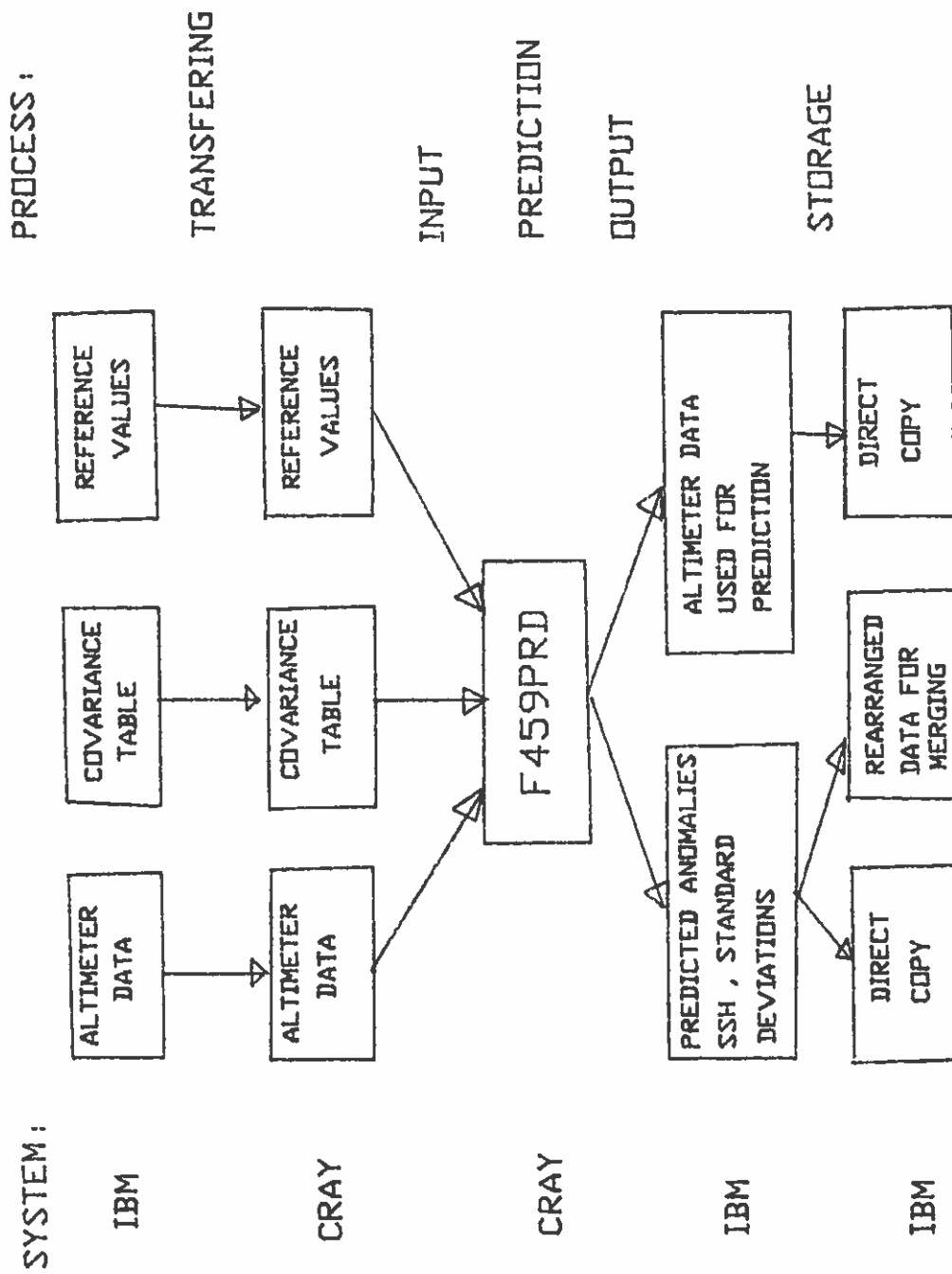


Figure 6.1 Data processing in various systems for the production work



		LONGITUDE													
LATITUDE		0	40	80	120	160	200	240	280	320	360				
		72													
	1	2	3	4	5	6	7	8	9						
	1	2	3	4	5	6	7	8	9						
	77583	13293	7	36824	219092	270262	12857	202850	315057						
	10	11	12	13	14	15	16	17	18						
	10	11	12	13	14	15	16	17	18						
	36010	95958	54743	229538	285717	339951	300870	727722	193147						
	1	2	3	4	5	6	7	8	9						
	19	20	21	22	23	24	25	26	27						
	135861	205338	155179	126758	184370	185215	245511	56898	228458						
	10	11	12	13	14	15	16	17	18						
	28	29	30	31	32	33	34	35	36						
	93335	102890	98276	117769	130470	117506	141994	102424	80707						

Numbers in one block:

First row = File number on the tape

Second row = Block number for global production runs

Third row = Number of altimeter data points in one file

Figure 6.2 File numbers, block numbers and numbers of altimeter data points for the global production work

At this point, we should clarify some terms in order to better illustrate the production work. We have the following definitions:

- Prediction block: the block, as shown in Figure 6.2, that covers a  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$  geographical region. See also the above description.
- Adjustment block: a  $4^\circ \times 4^\circ$  region within which one cross-over adjustment is performed.
- Prediction cell: a  $0.5^\circ \times 0.5^\circ$  cell within which 16 point anomalies and 16 sea surface heights are predicted.
- Altimeter file: a file containing altimeter data in a  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$  area.

These definitions are only valid for the production mode as specified in Section 5.7. For a prediction block, the entire prediction process involves the following steps:

- (1) Select altimeter data points for four  $4^\circ \times 4^\circ$  adjacent adjustment blocks (a total of  $8^\circ \times 8^\circ$  area) and write the data points to 4 separate temporary disc units (FORTRAN I/O units).
- (2) Perform a cross-over adjustment for the first adjustment block
- (3) Use the adjusted altimeter data from (2) to perform the predictions of gravity anomalies and SSH in the 64  $0.5^\circ \times 0.5^\circ$  prediction cells individually.
- (4) Repeat (2) and (3) for the second, third and fourth adjustment blocks.
- (5) Select the altimeter data points for the next 4 adjustment blocks and repeat (2) - (4).

Based on the above procedure, a total of 20 data selections are required for a  $32^\circ \times 40^\circ$  block and 25 data selections for a  $40^\circ \times 40^\circ$  block. This means that we have to read the altimeter data files 20 times or 25 times in order to complete the prediction in a prediction block. In terms of the predicted result, a set of predicted gravity anomalies and SSHs is defined as the predicted data located in the 4 adjustment blocks and derived from steps (1) to (4). This explanation of "set" will be useful for checking the computer's printouts.

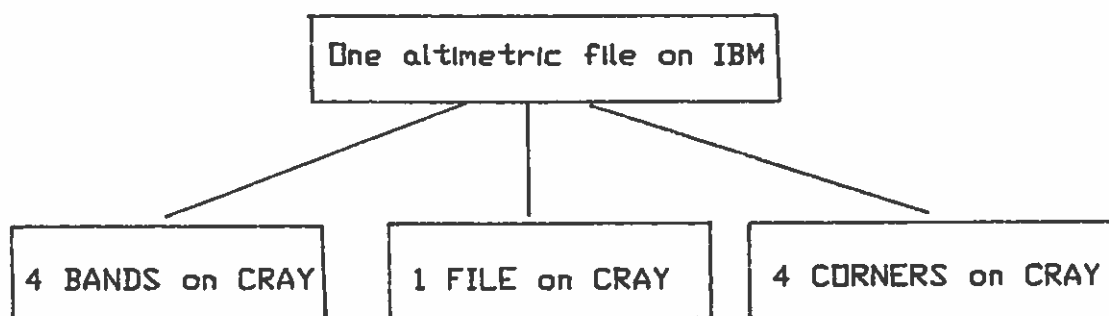
With such a procedure, theoretically we have to read a maximum of 9 altimeter files which have coverage of an area of either  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$ . The CPU time consumed by reading an entire altimeter file could be significant if the total number of points in the file is large. For example, a total of 12.6 CPU seconds is needed for reading 163,888

records (on CRAY X-MP/28). However, only a portion of data in the files surrounding the "central file" are needed in the prediction. By central file we mean the file whose corresponding geographical location is identical to the location of the prediction block. cf. Figure 6.3.

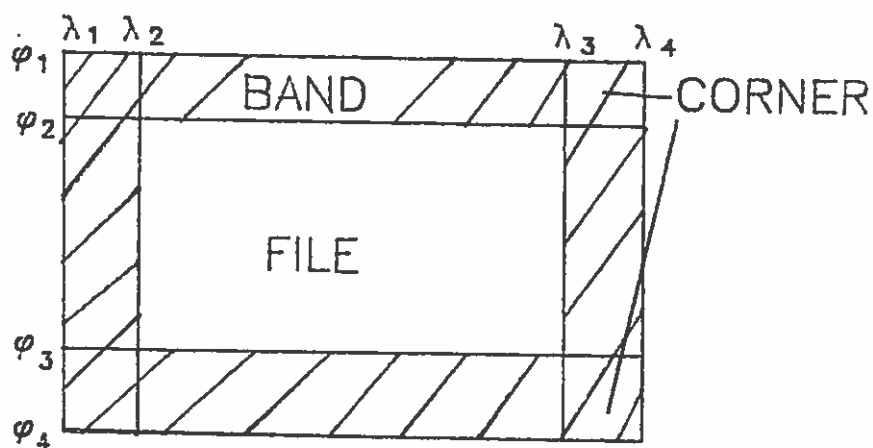
file 8	file 1	file 2
file 7	prediction block (central file)	file 3
file 6	file 5	file 4

Figure 6.3 Sketch of file locations for a prediction block

In order to reduce CPU time for reading the altimeter data, an altimeter file was partitioned into several sub-files, according to the desired borders of the adjustment blocks and prediction cells. The partition was completed when an altimeter file was transferred from the IBM system (in ASCII) to CRAY (in binary). Specifically, we have the following scheme:



A "FILE" contains data in a  $32^\circ \times 40^\circ$  or  $40^\circ \times 40^\circ$  block, while BAND and CORNER cover the data at the boundary areas of this particular block. Figure 6.4 shows the locations of FILE, BAND and CORNER that are obtained from an altimeter file (on IBM).



FILE:  $\phi_4 < \phi < \phi_1, \lambda_1 < \lambda < \lambda_4$  (total 1)

BAND:  $\phi_4 < \phi < \phi_1, \lambda_1 < \lambda < \lambda_2$ , etc (total 4)

CORNER:  $\phi_2 < \phi < \phi_1, \lambda_1 < \lambda < \lambda_2$ , etc (total 4)

Figure 6.4 The locations of partitioned altimeter files (on CRAY discs)

The width of the band will typically be  $0.5^\circ$ . Comparing Figure 6.4 with 6.3, we immediately see that the central file is FILE. To avoid any mistake, the names of these data sets have been systematically organized on CRAY discs. Using this manipulation of data, the cpu time for reading data has been reduced enormously.

To show an example of what data sets should be used in a  $40^\circ \times 40^\circ$  block prediction, we show the data set names associated with the regions in and surrounding the prediction area in Figure 6.5.

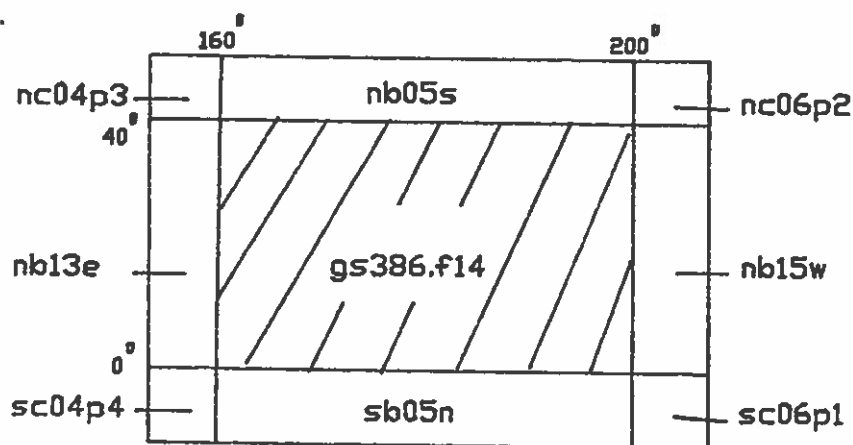


Figure 6.5 The required altimeter data sets (on CRAY) for prediction area at  $0^\circ < \phi < 40^\circ$ ,  $160^\circ < \lambda < 200^\circ$

When one prediction block was completed, two formatted data sets were generated and transferred back to the IBM system. One data set contained the anomalies, SSHs and their accuracy estimates. It was then reorganized and stored on a tape. The other data set contained the altimeter data used in the prediction cells. It was then copied to a file on another tape. The space on a tape needed for the first data set is small while the space needed for the second one is large. For the global production runs, it was estimated that 3 2400-foot tapes are required for the first data set and 4 to 5 2400-foot tapes for the second data set.

To estimate the CPU time for global production work, only some empirical timing can be obtained. Based on the tests in the S. Pacific Ocean and the Bermuda area and the performance analysis for program F459PRD, an average time of 915 seconds was needed for a production run which completed a prediction in a  $40^{\circ} \times 40^{\circ}$  block. Approximately 10 CPU hours will be consumed for the entire production work. However, this estimate does not count the repeated runs, so a factor of 1.2 should be considered and the final timing will be close to 12 CPU hours on CRAY. As far as the CPU time on IBM is concerned, it was expected that 1 CPU hour will be consumed by generating reference values and covariance table and by some other manipulations. If the predicted anomalies and SSH are plotted, then additional CPU time should be included.

When the initial global production runs were completed in January, 1989, the CPU time (on CRAY) spent by each block was tabulated. Table 6.1 shows the result.

Table 6.1 CPU time (on CRAY) for the initial block prediction

Block*	time	Block	time	Block	time
1	460	13	1202	25	1265
2	102	14	1582	26	340
3	44	15	1574	27	1280
4	212	16	1492	28	518
5	1230	17	1913	29	589
6	1354	18	1071	30	567
7	106	19	791	31	672
8	1016	20	1135	32	753
9	1704	21	799	33	676
10	245	22	691	34	805
11	546	23	996	35	585
12	299	24	992	36	444

unit = cpu second

\* The locations of the blocks are shown in Figure 6.2

Based on Table 6.1, a total of 8.34 hours was used in the 36 predictions. The average CPU time for a prediction run is 834 seconds which is close to the original estimate.

## 6.2 Revised Prediction Runs for Selected Cells

The necessity of revised prediction runs was due to unreasonable results such as track patterns and abnormal signatures that were found in the anomaly plots associated with the production results. Such phenomena, as discussed before, are mostly caused by erroneous observations and uneven point distribution. The sea surface heights usually generate a smooth surface and such unreasonable phenomena cannot be seen from plots of this data.

Although we have adjusted the altimeter tracks before the predictions were performed, the track biases really cannot be detected and adjusted by the current adjustment algorithm if the biases turn out to be errors within a very short satellite orbit segment. In an area with a smooth gravity field, the systematic patterns of anomalies caused by these errors will be enormously enhanced. This effect will be demonstrated in the following example.

The first problem area found is located in the South Indian Ocean. As shown in Figure 6.6, linear features still showed up even though the adjustment of tracks for bias removal has been carried out. Apparently, these features are associated with some tracks that pass through the features. Another plot, as given in Figure 6.7 shows 4 suspected tracks, i.e., tracks 861, 416, 172 and 5008 are responsible for the features. Note that at this point we disregard the signatures at the right-lower corner of Figure 6.6. These four tracks are Seasat's tracks and they are almost parallel as presented in Figure 6.4. Also note that the major "error" signatures occur at the center of the areas where the data distributions are not even (for example,  $-42.25 < \phi < -41.25$ ,  $57^\circ < \lambda < 58.25$ ). In order to study the behavior of SSH, the SSH along the four tracks are plotted in Figure 6.8 (a) and (b) where (a) contains the unadjusted tracks and (b) contains adjusted tracks. In Figure 6.8 (a) and (b), the spherical distances of the tracks started from a common cross section as indicated in Figure 6.7.

The four tracks are approximately 12.5 km apart. The change of sea surface height within such short distance is expected to be small. However, both Figure 6.8 (a) and (b) indicate that the SSHs along track 172 have abnormal fluctuations as compared to the

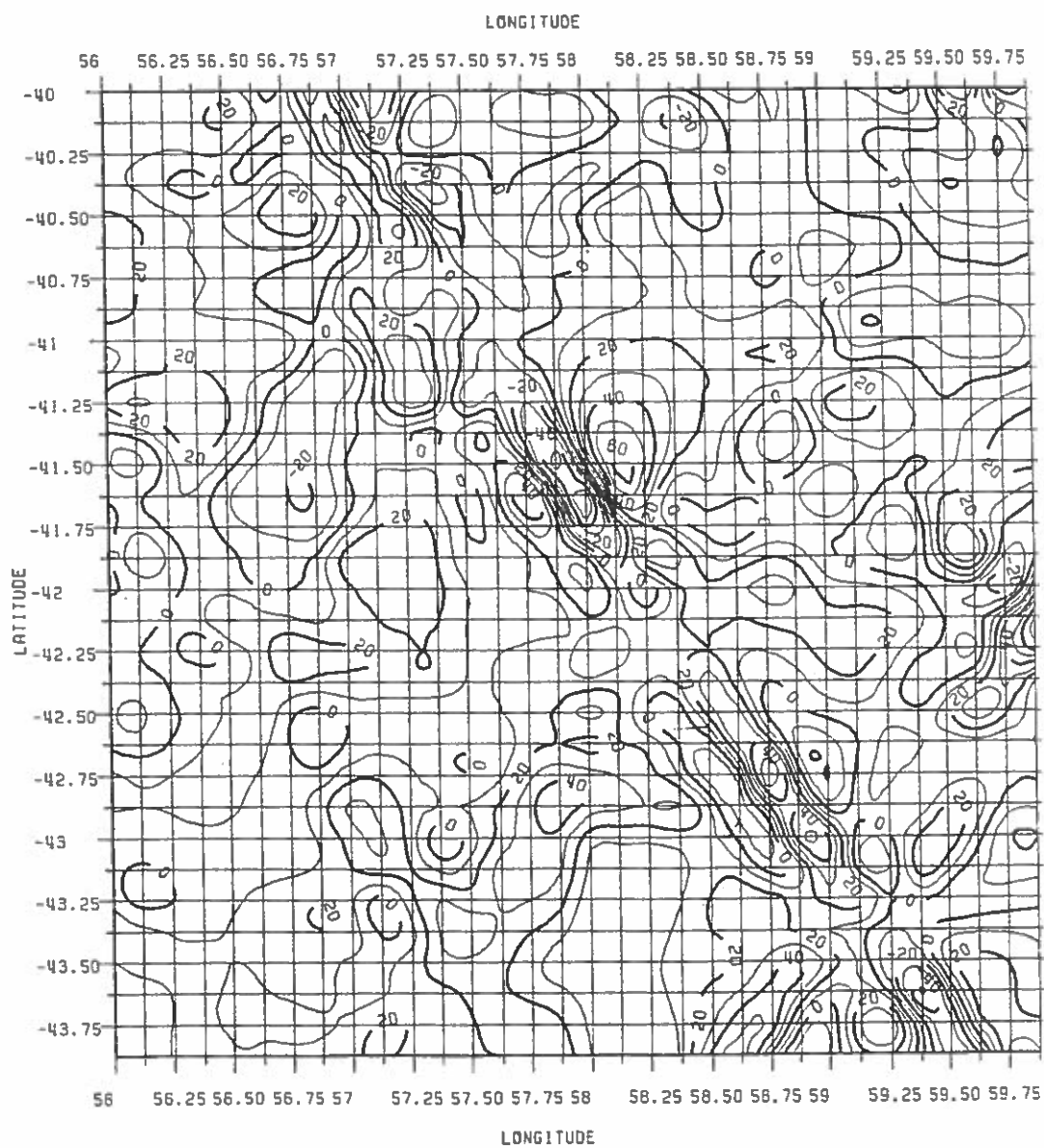


Figure 6.6 The predicted anomalies from initial production work. Track errors remained in the adjusted altimeter data. CI = 10 mgals.



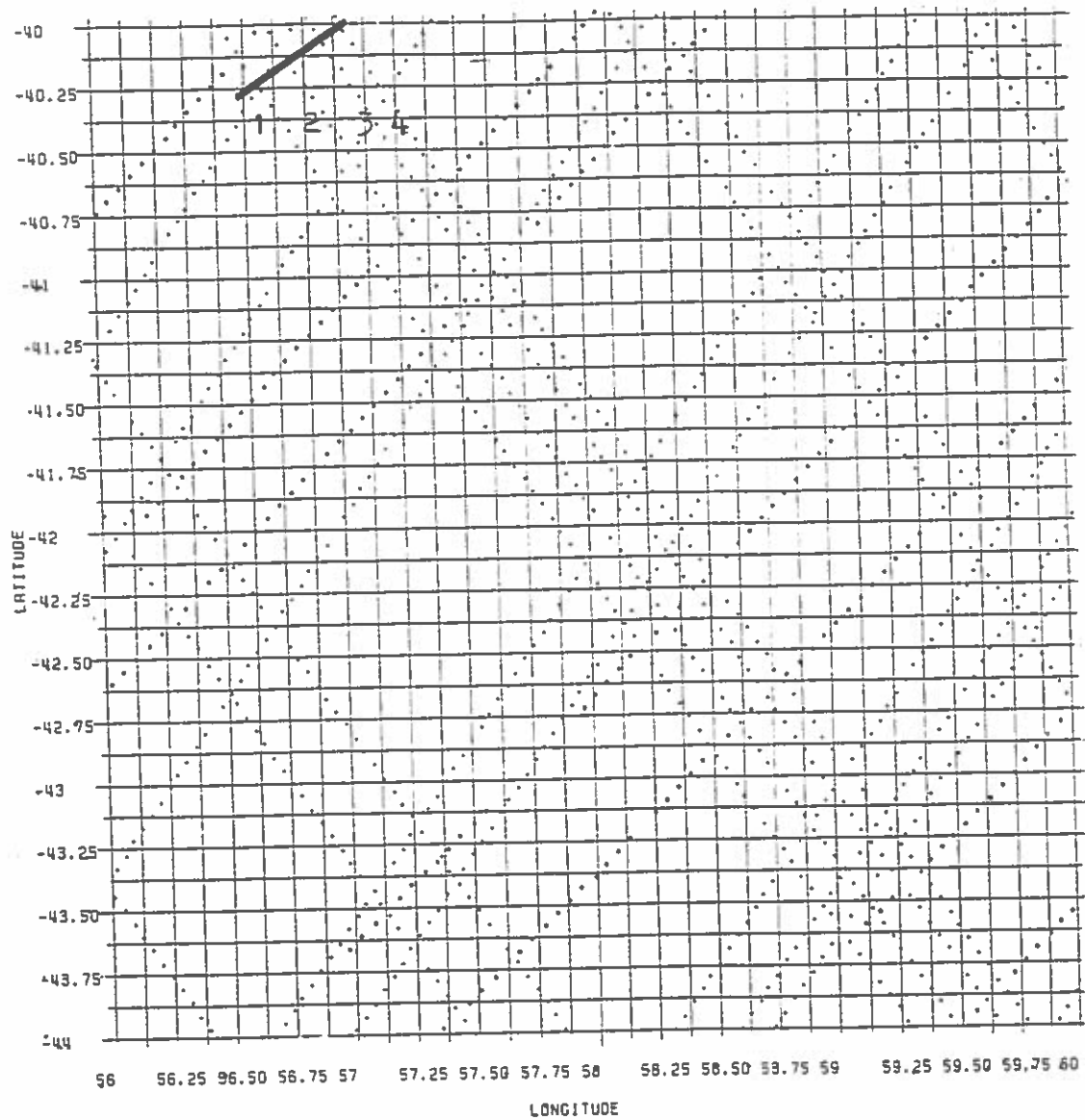


Figure 6.7 The four suspected tracks responsible for linear features in Figure 6.6. The track numbers

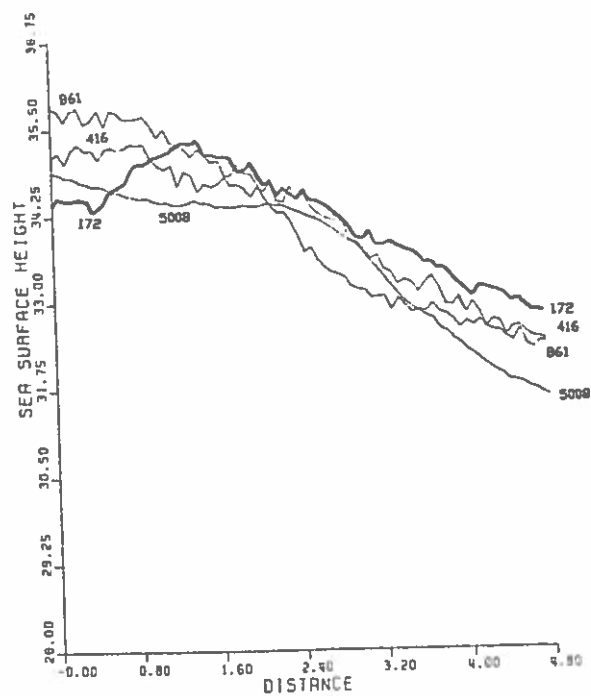
1=861

2=416

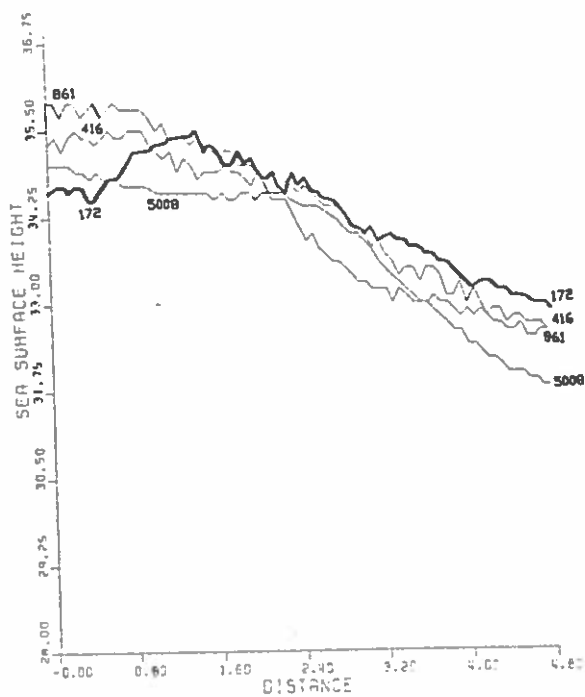
3=172

4=5008

The intersections between the line crossing the 4 tracks and the 4 tracks are the starting points of the profiles in Figures 6.8 (a) and (b).



(a) The SSH before adjustment



(b) The SSH after adjustment

Figure 6.8 The SSH along tracks 861, 416, 172, and 5008 before and after adjustment.

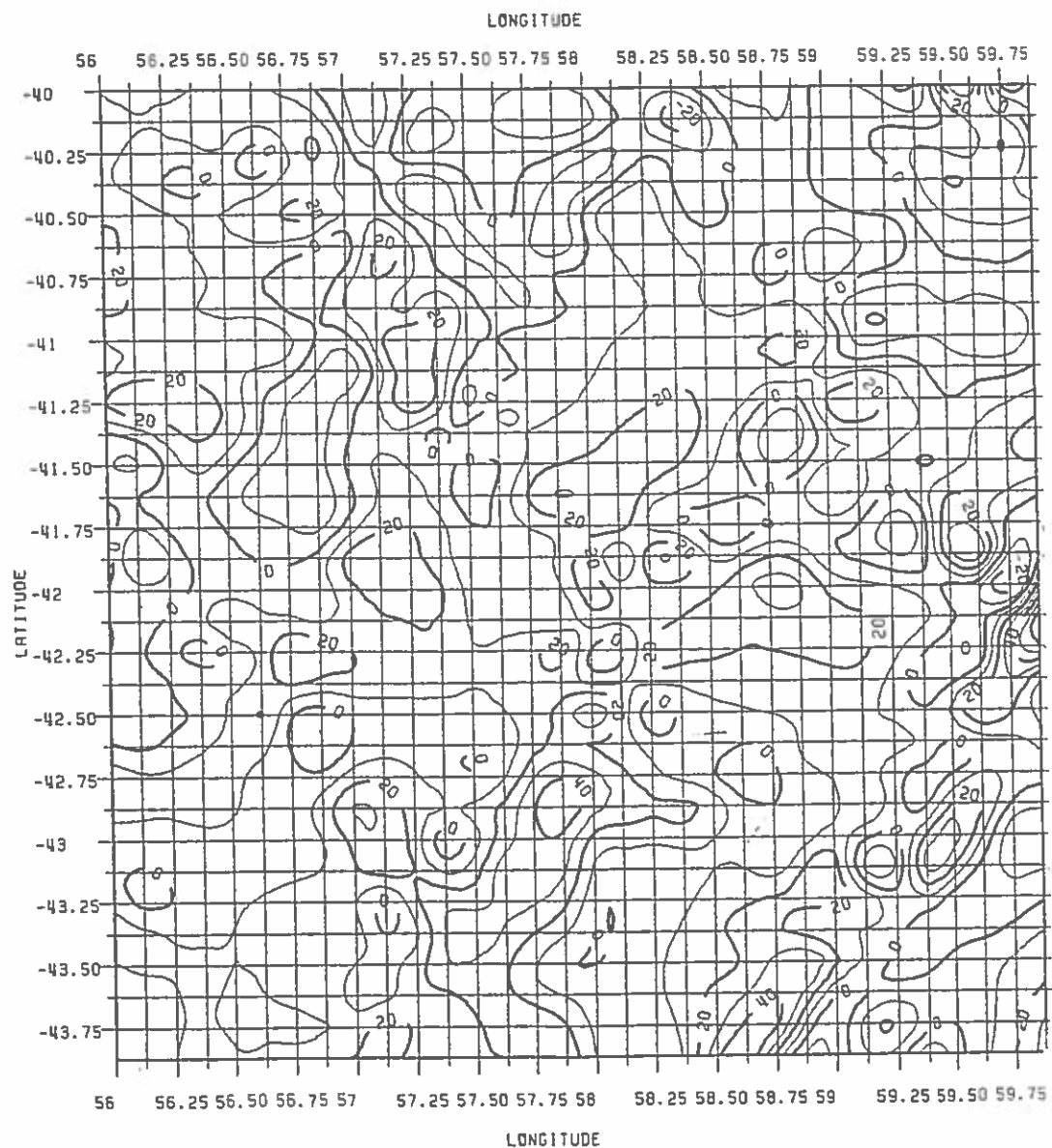


Figure 6.9 The predicted anomalies after track 172 was removed. CI = 10 mgals

adjacent SSH. The adjustment obviously cannot eliminate the errors along this track since the wavelengths of the errors are too short. Before a better model of adjustment is developed, the only way to resolve the problem is to identify the "bad" track and remove it from the prediction. Figure 6.9 shows the predicted anomalies after track 172 has been removed. It can be seen that the linear features have disappeared. (The features in right-lower corner need another investigation).

Unfortunately, this particular block was not the only one that has the track error problems. Based on the global  $10^{\circ} \times 10^{\circ}$  anomaly plots of the production results made by K. Hong, the problems still exist in some other areas. The solution to the problems for all the areas is the same: First plot the track numbers, remove the suspected tracks and repeat the prediction. Then, plot the anomalies and repeat the process until the signatures associated with track errors completely disappear. Tremendous effort has been given to identifying and fixing the track problem. The amount of the work is almost equivalent to the initial prediction work.

As a result, Table 6.2 lists the areas that receive revisions and the tracks that cause the abnormal gravity anomalies. The information in Table 6.2 could be very important for future altimeter work since the "bad" tracks have been recorded and these tracks can be simply ignored in future work.

In Table 6.2, we find that a large amount of tracks have been deleted in some revised areas. For example, in block 17 of Table 6.2, a total of 13 tracks were removed. However, it is expected that the gravity information will not be substantially affected by the removal of tracks even if some of the observations along the deleted tracks could provide reliable information. In essence, as shown in Table 6.2, the number of tracks deleted is proportional to the data density (or the "track density"), the removal of tracks will be equivalent to eliminating the altimeter data points to fulfill the minimum spacing criterion. The data density information is shown in Figure 6.2 (third number in a block).

Table 6.2 Summary of revised prediction areas

ID	location*	block <sup>+</sup>	tracks removed
1	40-32.125 144-151.875	13	656, 857, 463, 420, 168, 613, 707, 10415
2	40-32.125 300-307.875	17	7120, 1604, 463, 1346, 1387, 1432, 543, 14015, 245, 7666, 10077, 1732, 291
3	-10- -17.875 124.0-131.875	22	664, 865, 1467, 4166, 1381
4	-32.0- -39.875 308-315.875	26	771, 283, 291, 168, 455, 211
5	-40- -47.875 56-63.875	29	172, 422, 465, 502, 221
6	-40- -47.875 64-71.875	29	422, 465, 502, 680, 393, 717, 723
7	-40- -47.875 80-87.875	30	286, 220, 421, 5009, 177, 4923, 1767
8	-40- -47.875 304-311.875	35	283, 211, 205, 277,
9	32-16.125 32-39.875	10	12481
10	-60.0 -63.875 0.0-3.875	28	674
11	-52.0- -55.875 156.0-159.875	31	642

\* Latitude and longitude in degree

+ See figure 6.2 for block numbers

Having completed the revised predictions, the original predictions were replaced by the revised result according to the block it resided on (as indicated in Table 6.2). Based on Table 6.2, a total of 10 such revised files were created.

After the initial and revised predictions were completed, a total of 36 files were created. Each of the 36 file stores the predicted anomalies, SSH and their accuracy estimates in a prediction block. For future application, these 36 files were merged in such a way that each file occupies a 72° x 40° area. The result is another tape containing such 18 files. Also, 18 additional files were created so that each grid point has an elevation with it. The elevations were based on 5 minute mean elevations of TUG87 digital terrain model (Wieser, 1987).

## CHAPTER VII

### **Statistical analysis and comparison of the current prediction results with those of Rapp's (1985)**

#### **7.1 The eighth degree data**

The global production work yielded a total of 2,322,080 point anomalies and sea surface heights and their rigorous accuracy estimates. The point or grid interval is  $0^{\circ}125$ . Since one individual  $40^{\circ} \times 40^{\circ}$  or  $32^{\circ} \times 40^{\circ}$  block prediction was completed in one production run without interruption, theoretically the prediction will take place in any of the 6400 or 5120 prediction cells provided that the number of altimeter in a cell is equal to or greater than 7. However, for many obvious reasons, some altimeter observed SSHs have significant deviations from the geoid heights at the corresponding locations. This type of observed SSHs, for example, includes the SSHs in in-land seas, the SSHs affected by ice, rain and clouds, the SSHs without valid ocean tidal correction. If these undesired SSHs are used for the prediction, the predicted result will apparently be unreliable.

Our global prediction of gravity anomalies and SSHs described in chapter 6 has been carried out without considering these undesired SSHs, though local arc adjustments were performed before predictions were made (we have interpreted the meaning of local arc adjustment). Based on the global anomaly plots and numerical investigation of the predicted results, the following phenomena were observed:

- Two groups of large anomalies were found in the Caspian Sea and Black sea, respectively. The depths of the seas (with respect to their sea surfaces) are around 10 to 30 meters.
- Abnormal anomalies were detected in Hudson Bay and Strait.
- Predictions were made in some isolated prediction cells on land. The anomaly can be as high as 600 mgal.
- Most of the predicted quantities near (about 50 km) continental coastal area have erratic behaviors. The accuracy estimates are large. (the largest 78.40 mgals for anomaly)

- Abnormal behavior of anomalies was found in areas below  $\phi = -63^{\circ}0$ . This could be due to the surface ice when the observations were made.

Additionally, large standard deviations of the predictions are found in areas with a rough gravity anomaly field, especially in trench areas. The associated phenomena with these large standard deviations are the large scaling factors of the covariance functions. Based on the above analysis of the predicted result, it is not appropriate to adopt all the 2,322,080 point values as our final prediction result. Several criteria could be applied if one has to judge if a predicted value is reliable.

The decision for selecting the point acceptance criteria is hard to be made since we want to keep the reliable information as much as possible under the consideration of the criteria. However, a point elevation at each gridded point could be always one of the criteria since the predictions made on land will never be adopted. To conveniently apply the elevation criterion, the 5' x 5' elevation block was searched for each point and the elevation of the point was assumed to be the mean elevation of this 5' x 5' block. As mentioned in Chapter 6, the 5' x 5' mean elevations for this study is TUG87 (Wieser, 1987). By doing this, a point in the ocean receives a negative elevation and a point on land receives a positive elevation. The question is then what the minimum depth should be if the point value can be acceptable. Obviously, we wish to choose a minimum depth for which the predicted result in the Caspian Sea, Hudson Bay, Hudson strait, isolated cells on land and coastal area will be rejected. Unfortunately, many locations surrounding the isolated islands in the open ocean could have shallow water while the predictions are still very reliable. A good example can be found in the vicinity of Bermuda. Table 7.1 illustrates the effect of deleting point values if the depths at the points are not greater than a specified minimum depth.

Table 7.1

Number of points with  $\Delta g \geq 300$  mgals that will be deleted according to the depth criteria

area	minimum acceptable depth (meters)			
	100	250	350	500
Bonin Trench <sup>+</sup>	0	3	5	6
Alleutian Is. <sup>+</sup>	3	4	4	4
Hawaii Is. <sup>+</sup>	6	9	14	18
Bermuda Is. <sup>+</sup>	0	0	2	2
Bering Strait *	All			
Hudson Bay *			All	
Caspian Sea *	All			
Black Sea *	All			

<sup>+</sup> The deleted anomalies are considered to be reliable

\* The deleted anomalies are considered to be unreliable

It was found that if 100 meters is the minimum depth for the acceptance of the point predictions, all the abnormal anomalies, except those in Hudson Bay and Hudson Strait and in areas  $\phi < -63^\circ$ , will not be accepted. However, as indicated in table 7.1, 3 large anomalies in the Alleutian Islands and 6 large anomalies in the Hawaii Islands will be lost due to such criterion. For our global analysis, a variable criterion apparently will not be practical, thus we must sacrifice these large, reliable anomalies. For local analysis, the depth criterion could be flexible and good information may not be lost.

Based on the above analysis, it is recommended that the point value can be accepted if the following criteria are fulfilled:

- (i) Depth  $> 100$  meters
- (ii)  $\phi \geq -63^\circ$
- (iii)  $(\phi, \lambda)$  not in area  $(45^\circ < \phi < 65^\circ, 250^\circ < \lambda < 296^\circ)$

Using the three criteria, we summarize the statistics of the global eighth degree data in Table 7.2.



Table 7.2

The statistics of the global eighth degree data based on the point value acceptance criteria  
(unit: anomaly: mgals, SSH: meters)

	Anomaly	Sea surface height
No. of points predicted	2322080	2322080
No. of points accepted	2048487	2048487
maximum	439.70	82.66
minimum	-368.50	-107.54
mean	-1.74	2.36
RMS	30.06	31.02
mean predicted std. dev.	12.11	0.08
RMS value of predicted std. dev.	12.43	0.09

In addition, Table 7.3 gives the frequency distribution of the anomalies and Table 7.4 provides the standard deviation distribution of the anomalies.

Table 7.3

Frequency distribution of the accepted anomalies

Range (mgals)	No. of points
-400 to -350	8
-350 to -300	97
-300 to -250	482
-250 to -200	1643
-200 to -150	4461
-150 to -100	9675
-100 to -50	56944
-50 to 0	1038565
0 to 50	866448
50 to 100	58109
100 to 150	8735
150 to 200	2485
200 to 250	623
250 to 300	157
300 to 350	43
350 to 400	11
400 to 450	1

Table 7.4

Standard deviation distribution of the accepted anomalies

Range (mgal)	No. of points
5 to 10	410063
10 to 15	1420338
15 to 20	188123
20 to 25	17503
25 to 30	7495
30 to 35	2873
35 to 40	1235
40 to 45	483
45 to 50	186
50 to 55	82
55 to 60	55
60 to 100	51

The maximum anomaly 439.70 mgals was located at  $\phi = 27.625$ ,  $\lambda = 142.250$  which is in the Bonin Trench and the minimum anomaly -368.50 mgals was located at  $\phi=19.250$ ,  $\lambda = 294.25$ , which is in the Puerto Rico Trench; The maximum sea surface height 82.66 meters was found at  $\phi = -5.250$ ,  $\lambda 150.125$ , which is close to the New Britain Trench and the minimum sea surface height -107.54 meters was found at  $\phi = 4.625$ ,  $\lambda=150.125$ , which is close to the Mid Indian Basin. Furthermore, a group of large anomalies was also found in the Japan-Kuril Trench, Bonin Trench, Mariana Trench, Caroline Islands, Hawaiian Islands and the Puerto Rico Trench.

In Chapter 5, we have demonstrated the excellent agreement between the production result and the "ground truth" - ship data. The capability of recovering high frequency anomalies in the Bermuda area was also shown in Figure 5.28. Comparison of 2011 ship measured anomalies and those from production result showed an RMS difference of 15.9 mgals even though the gravity field surrounding the Bermuda Island is rough. Since Rapp (ibid.) used  $3^\circ \times 3^\circ$  as the prediction cell and the number of data points were only 300, the gravity field has been considerably smoothed. Therefore, the result from Rapp (ibid.) showed an RMS difference of 47 mgals when comparison was performed for Rapp's result and the ground truth at the same 2011 points.

Many unrecovered sea mounts in Rapp's (ibid.) work now also clearly showed up in the anomaly maps. Three areas were selected for the comparison.

The areas are:

1. The Mariana Trench and Magellan Seamounts:  $12^\circ \leq \phi \leq 24^\circ$ ,  $145^\circ \leq \lambda \leq 155^\circ$   
(Figures 7.1, 7.2, 7.3, and 7.4)
2. The Hawaiian Islands:  $15^\circ \leq \phi \leq 25^\circ$ ,  $196^\circ \leq \lambda \leq 206^\circ$  (Figures 7.5, 7.6, 7.7, 7.8)
3. The New England Seamount area:  $36^\circ \leq \phi \leq 40^\circ$ ,  $295^\circ \leq \lambda \leq 299^\circ$  (Figures 7.9, 7.10, 7.11, 7.12)

From the comparisons, several conclusions can be drawn:

- (i) Both of the results indicated the same trends of anomalies and sea surface heights, but the current result provides more detailed information.
- (ii) Substantial difference in anomalies are found, but the change in SSHs sometimes is not significant. The current result clearly separates the signature which is indistinguishable in Rapp's result. This implies that better resolution has been achieved by the current result.
- (iii) Shifts (with respect to Rapp's result) of the current estimated SSHs can take place in areas where significant sea surface topography corrections were made. The effect of various corrections made for the SSHs as described in chapter 2 will also cause the shifts.
- (iv) The current result has better capability of reflecting the mass distribution of the seamounts. This aspect can be illustrated by the figures in the third area and the corresponding bathymetry map (from TUG87 5 minute elevations), Figure 7.4. It is also true that the sea surface heights do not reveal the seamount's signature as much detail as the anomalies do and this holds for both the current result and Rapp's result.

A final remark will be related to the accuracy estimates of the current result. The computation of accuracy estimates was performed in a rigorous way by the equation (in case of a point anomaly):

$$M_g^2 = \alpha \left[ C_{gg} - C_{gh} \left( C_{hh} + \frac{1}{\alpha} C_{nn} \right)^{-1} C_{gh} \right] \quad (7-1)$$

where  $\alpha$  is the scaling factor to the global covariances  $C_{gg}$ ,  $C_{gh}$  and  $C_{hh}$ .  $\alpha$  is determined by the residual SSHs and has a minimum value of 0.425 (see eq. (3-21)). In case of erroneous SSHs and a rough gravity field,  $\alpha$  could be relatively large. However, the standard deviations are mainly governed by the data distribution and the number of points used but not the  $\alpha$  value. For example, in the Bermuda area, the standard deviations are on the order of 10 to 20 mgals for a prediction cell that has a large  $\alpha$  value (typically 2 to 3); nevertheless, under the same  $\alpha$  value the standard deviations could be raised to 60 mgals for a prediction cell located in the Kuril Trench area. As mentioned in Chapter 5, in a prediction cell of  $0.5^\circ \times 0.5^\circ$  (with  $0.25^\circ$  border), the number of data points can reach 1430 in the Bermuda area while only 120 points might be found in the Kuril Trench area. (Note: 120 could reduce to half due to minimum spacing criterion). From this example, we therefore can see how the altimeter data distribution and density can effect the standard deviations of the prediction.

Instead of rigorously computing the accuracy estimates, Rapp (1985) used an empirical formula to obtain a set of uniform accuracy within a prediction cell (i.e.,  $3^\circ \times 3^\circ$ ). He then showed that the mean value of the predicted standard deviations was  $\pm 12.2$  mgal and the root mean square standard deviation was  $\pm 12.7$  mgal. For the current result, the corresponding values are  $\pm 12.1$  mgal and  $\pm 12.4$  mgal, respectively. Thus, it can be justified that Rapp's empirical formula is appropriate in a global sense.

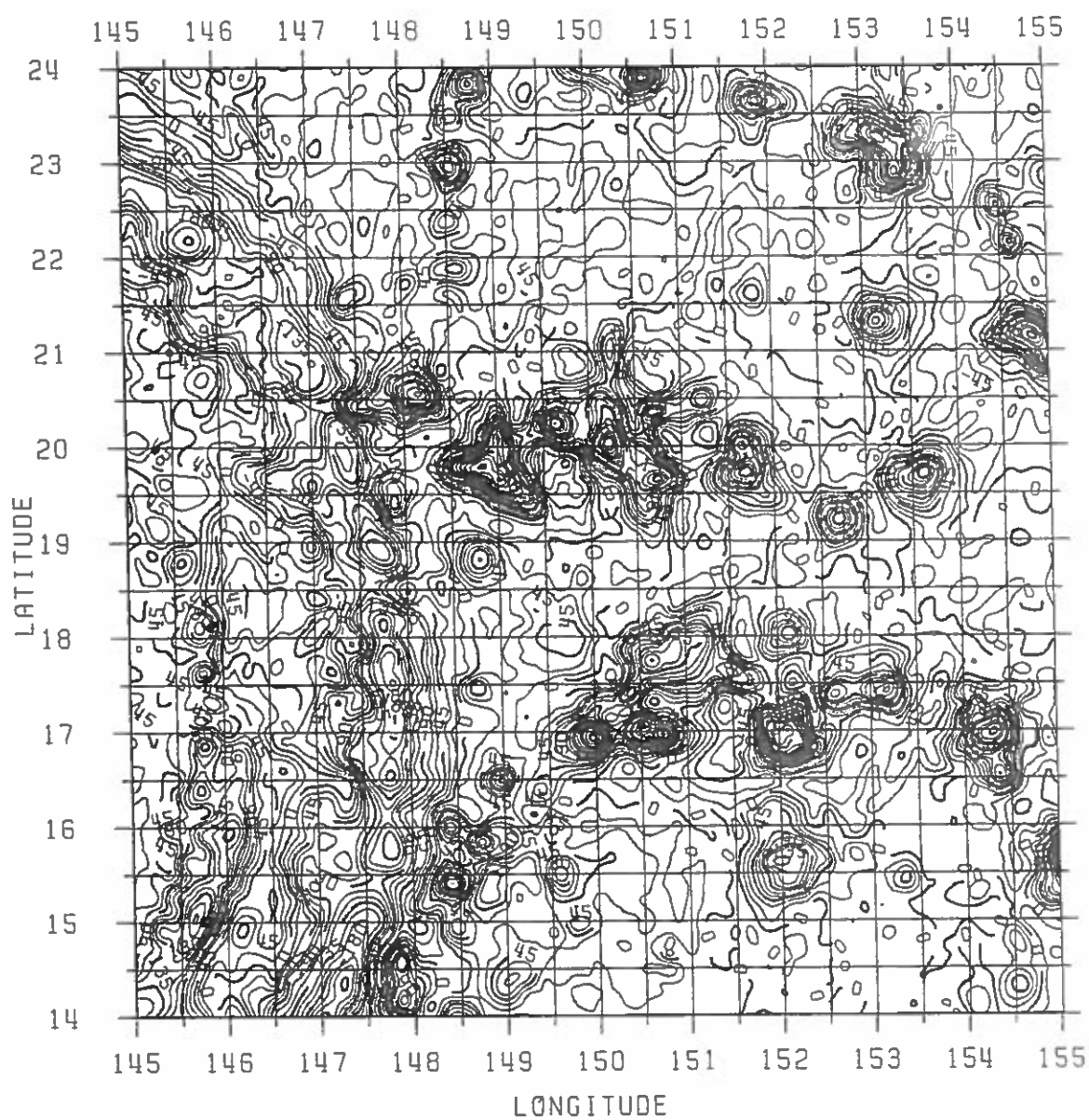


Figure 7.1 Anomaly map from the current production result in the Mariana Trench and Magellan Seamounts. CI = 15 mgals

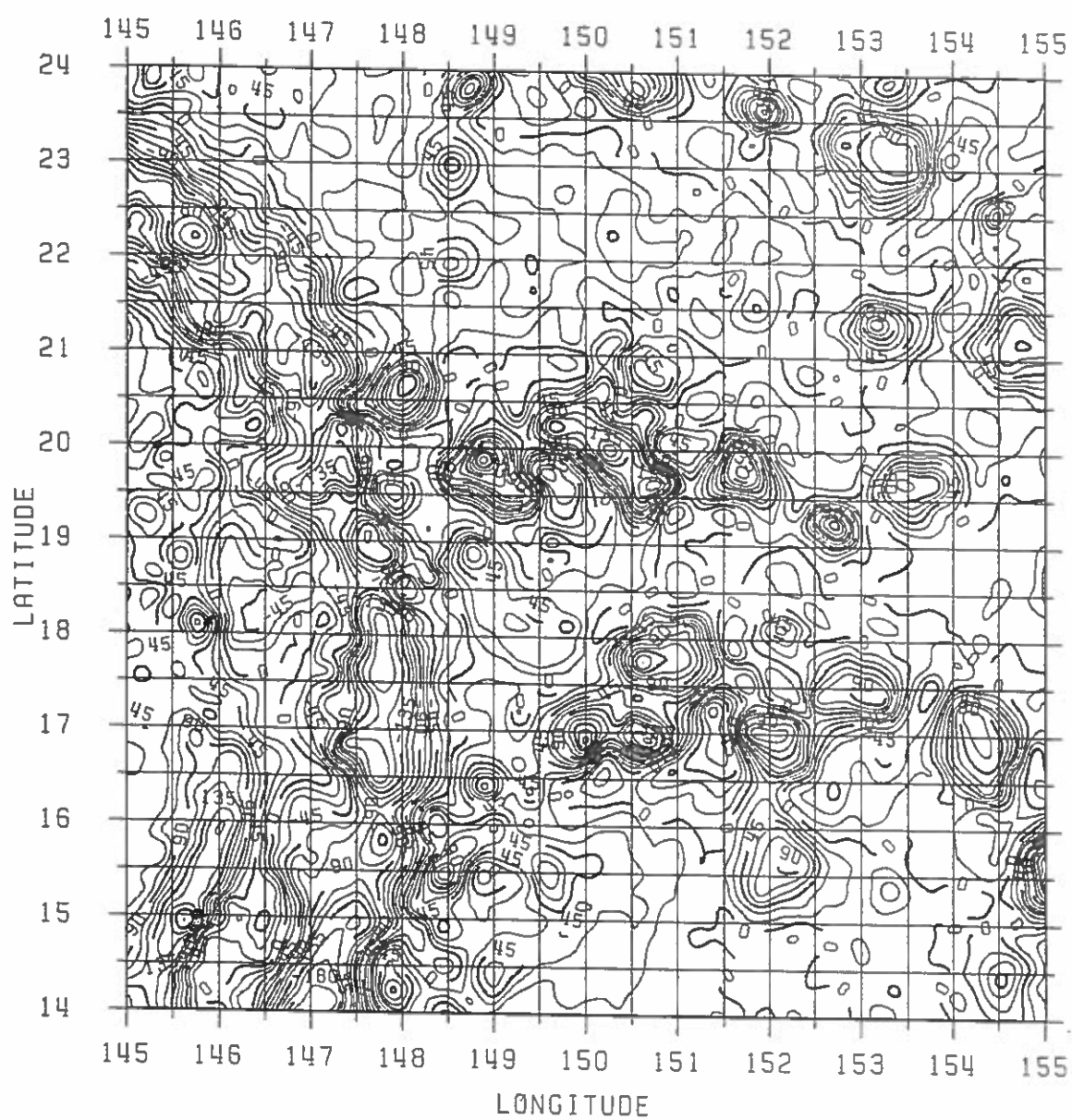


Figure 7.2 Anomaly map from Rapp's (1985) result in the Mariana Trench and Magellan Seamounts. CI = 15 mgals

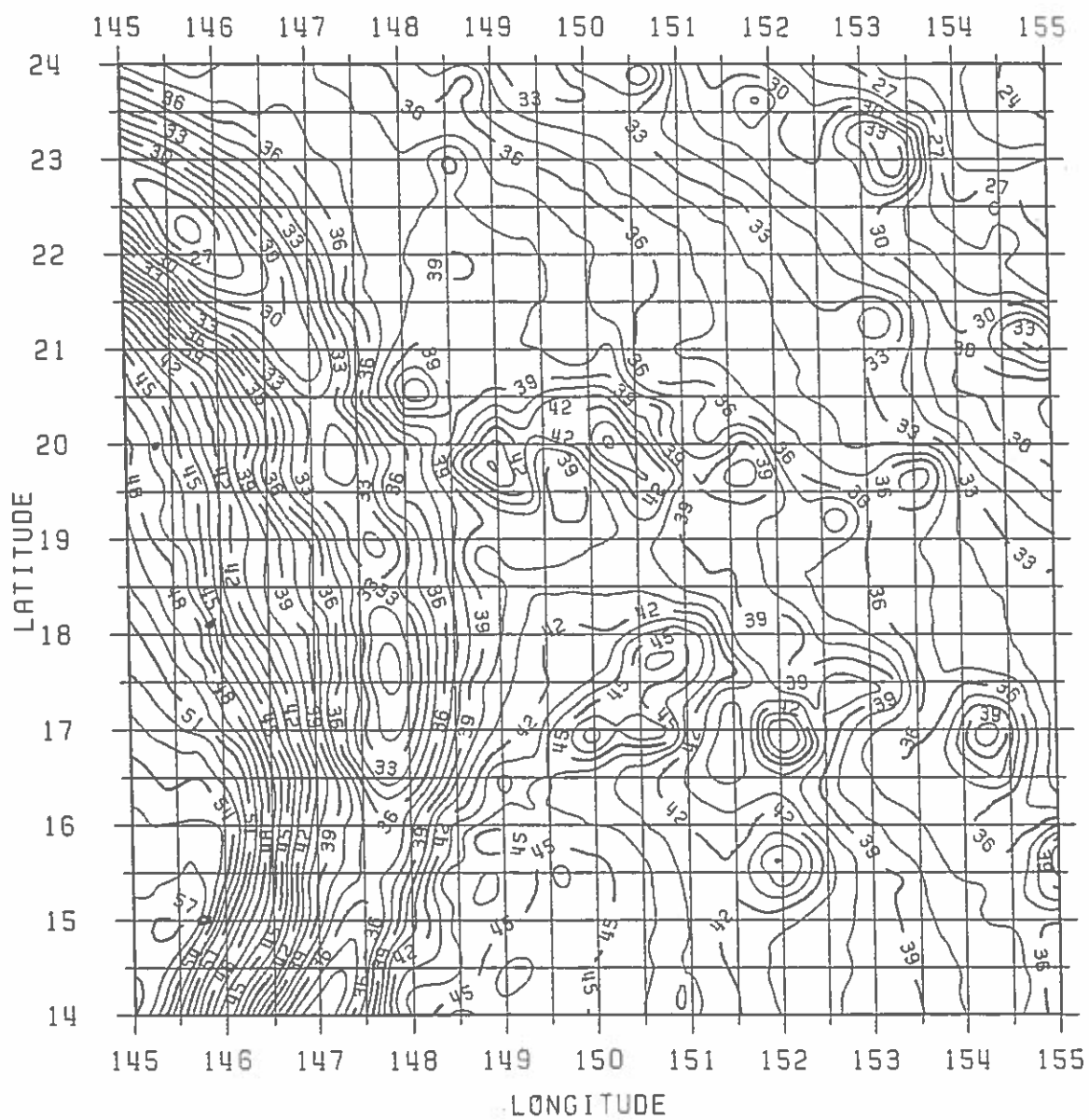
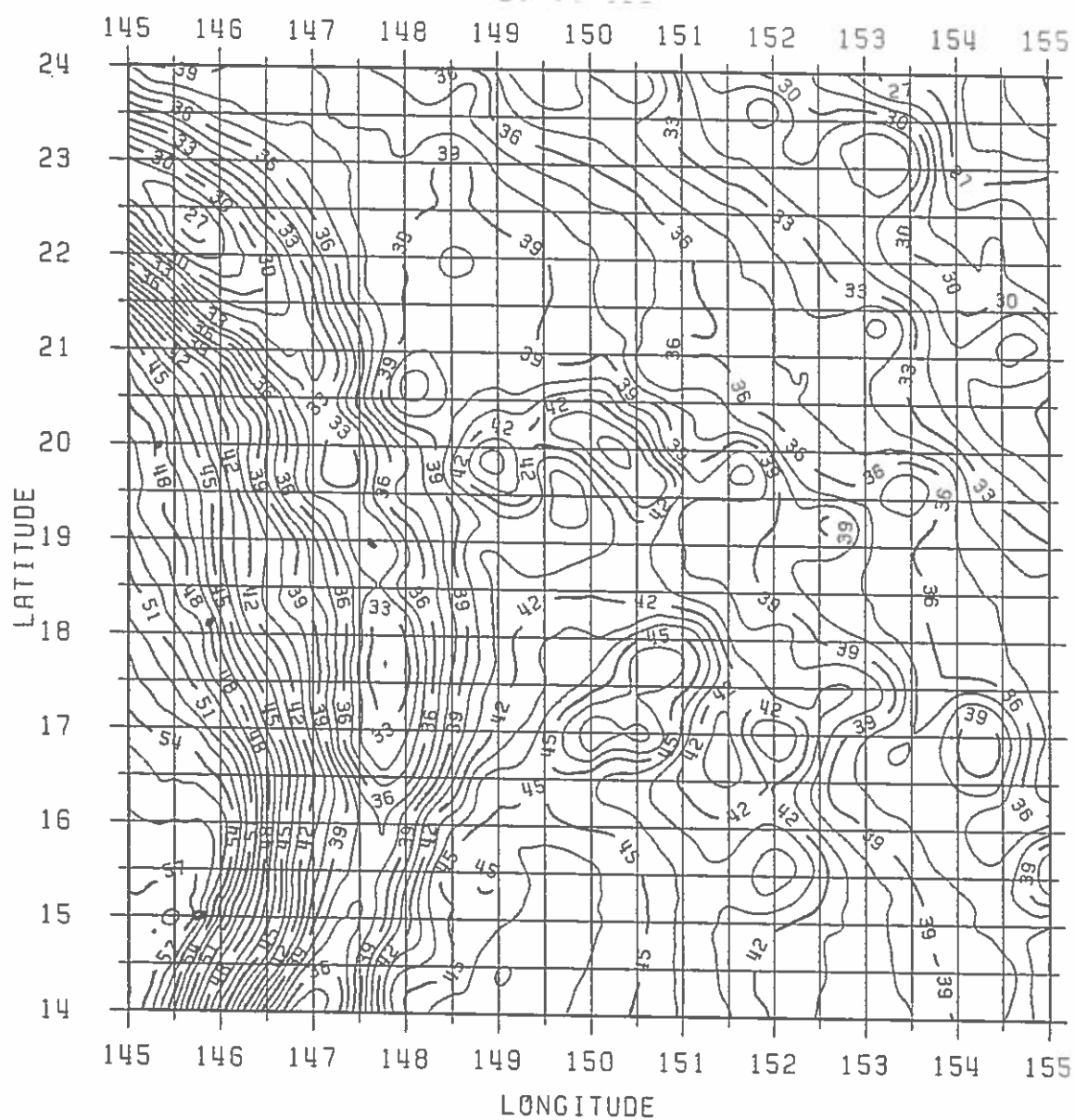


Figure 7.3 Sea surface height map from the current production result in the Mariana Trench and Magellan Seamounts. CI = 1 meter





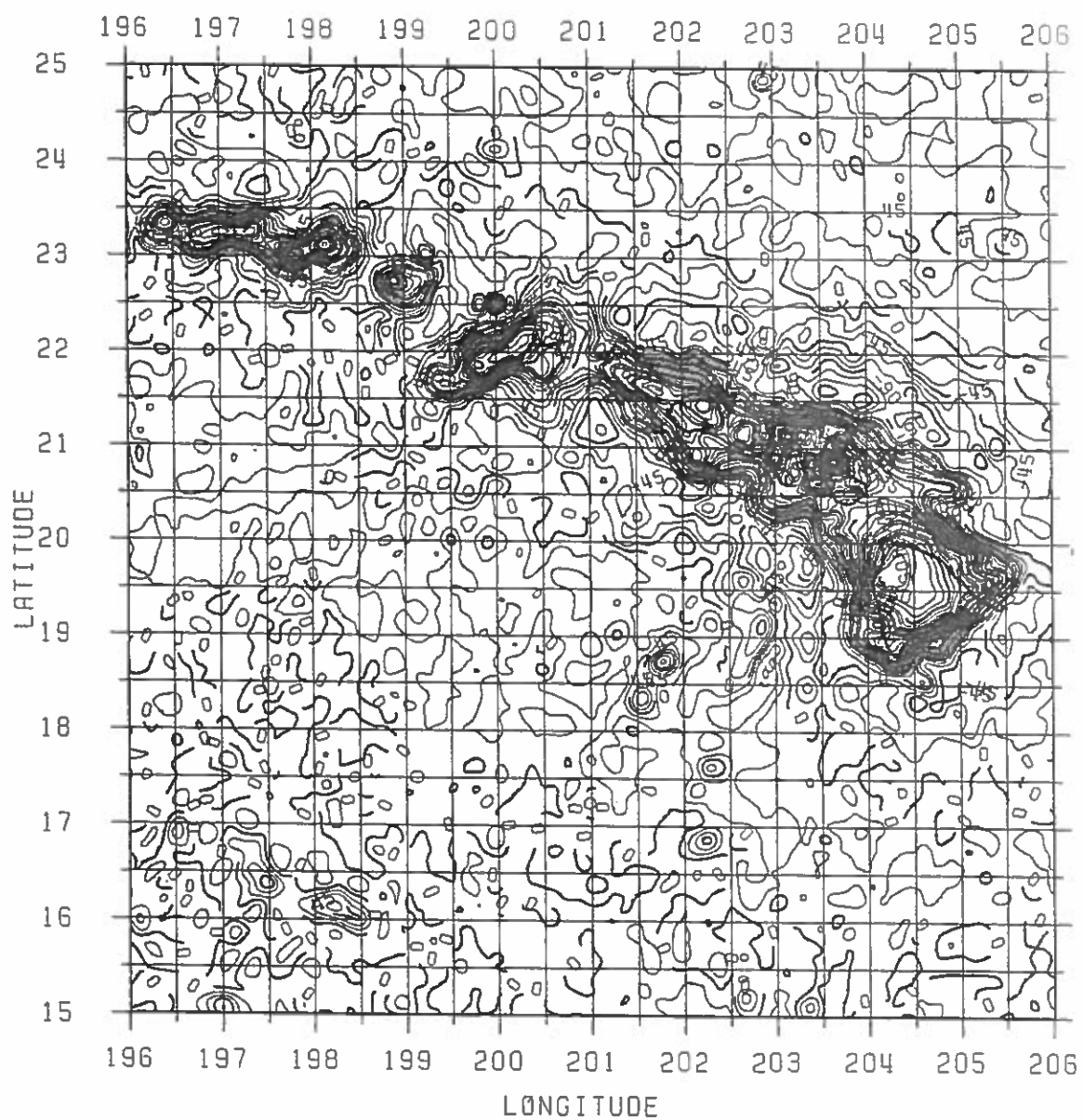


Figure 7.5 Anomaly map from the current result in the Hawaiian Islands. CI = 15 mgals

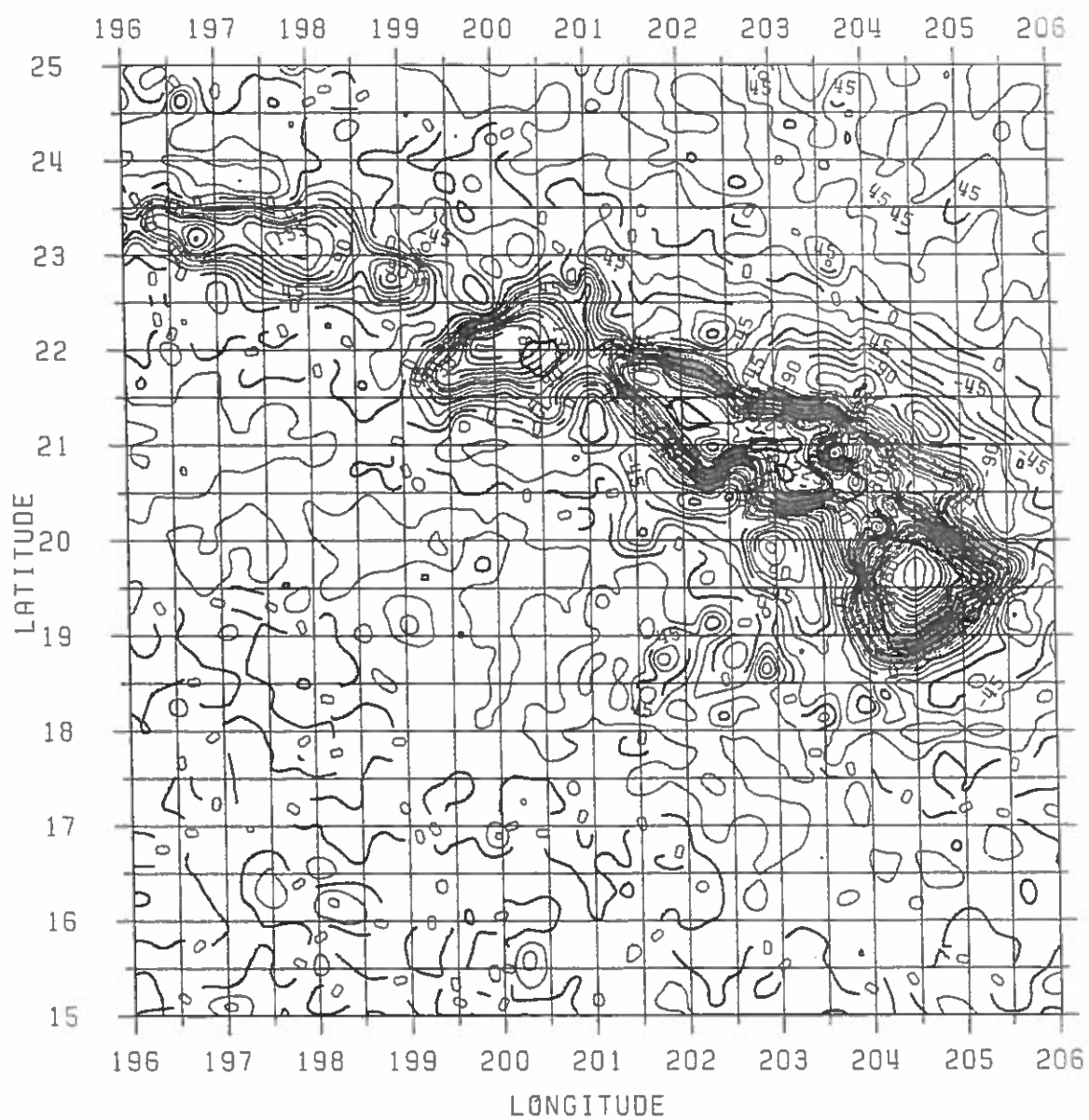


Figure 7.6 Anomaly map from Rapp's (1985) result in the Hawaiian Islands. CI=15 mgals

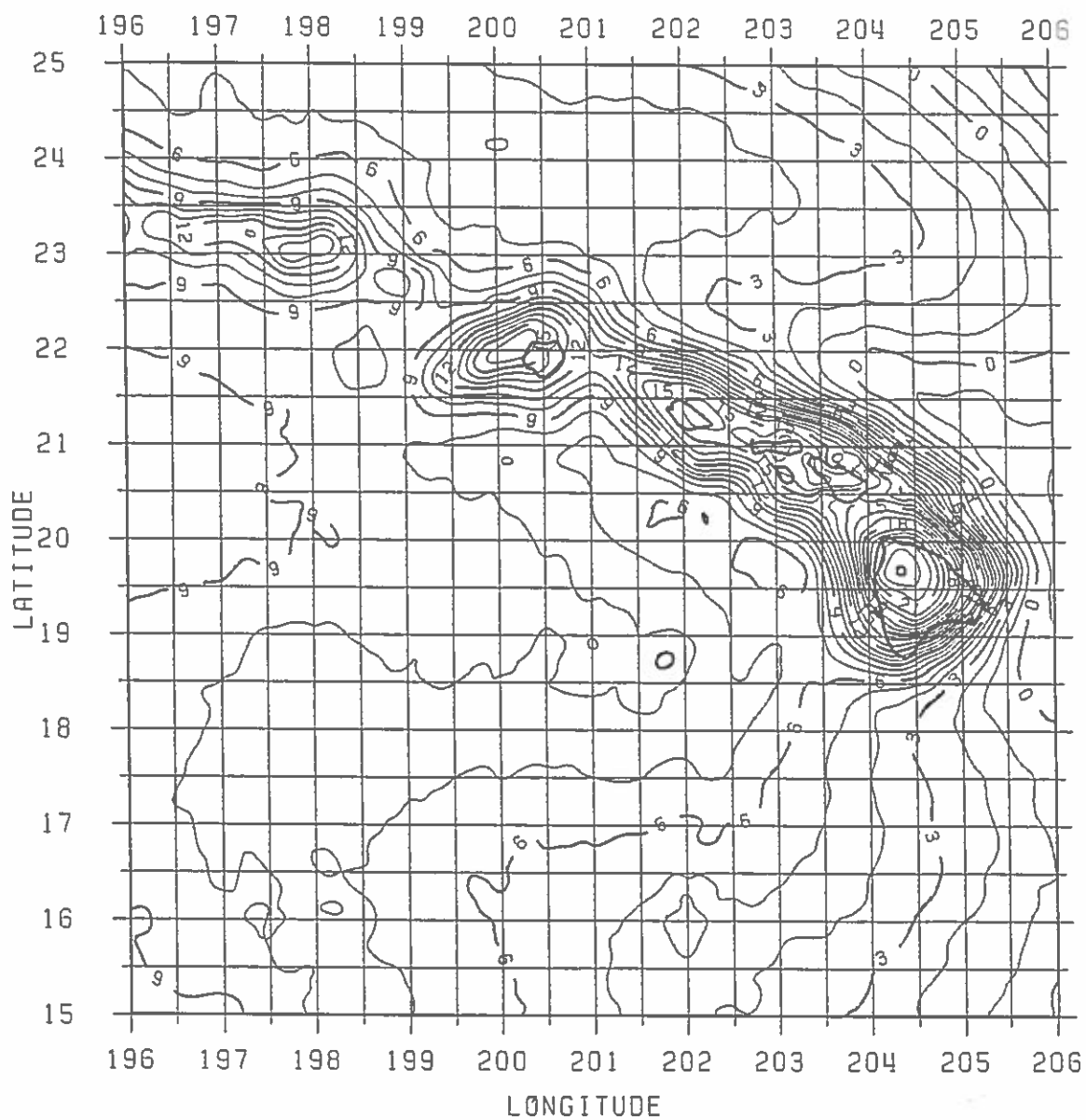


Figure 7.7 Sea surface height map from the current production result in the Hawaiian Islands. CI = 1 meter

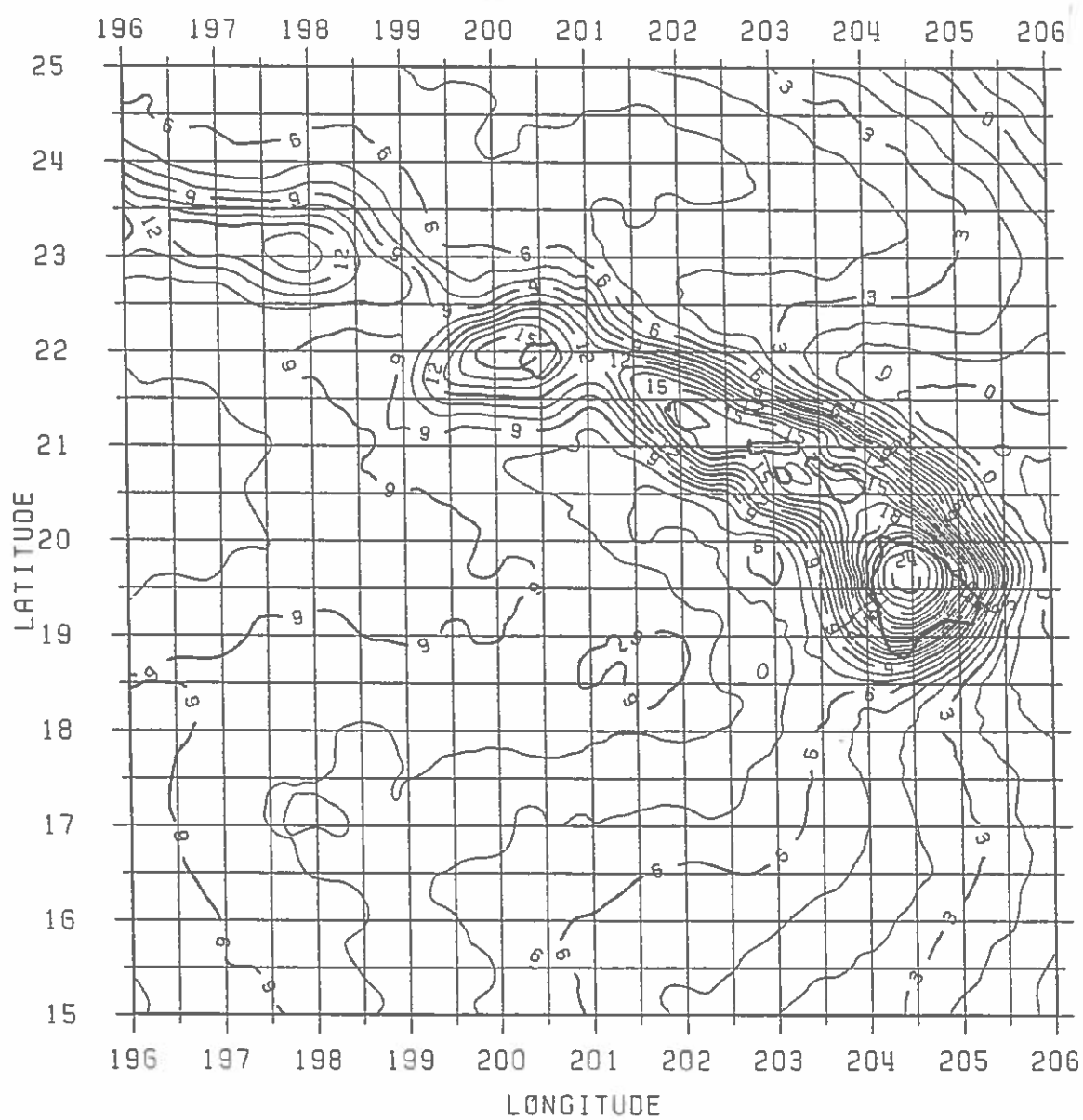


Figure 7.8 Sea surface height map from Rapp's (1985) result in the Hawaiian Islands. CI = 1 meter

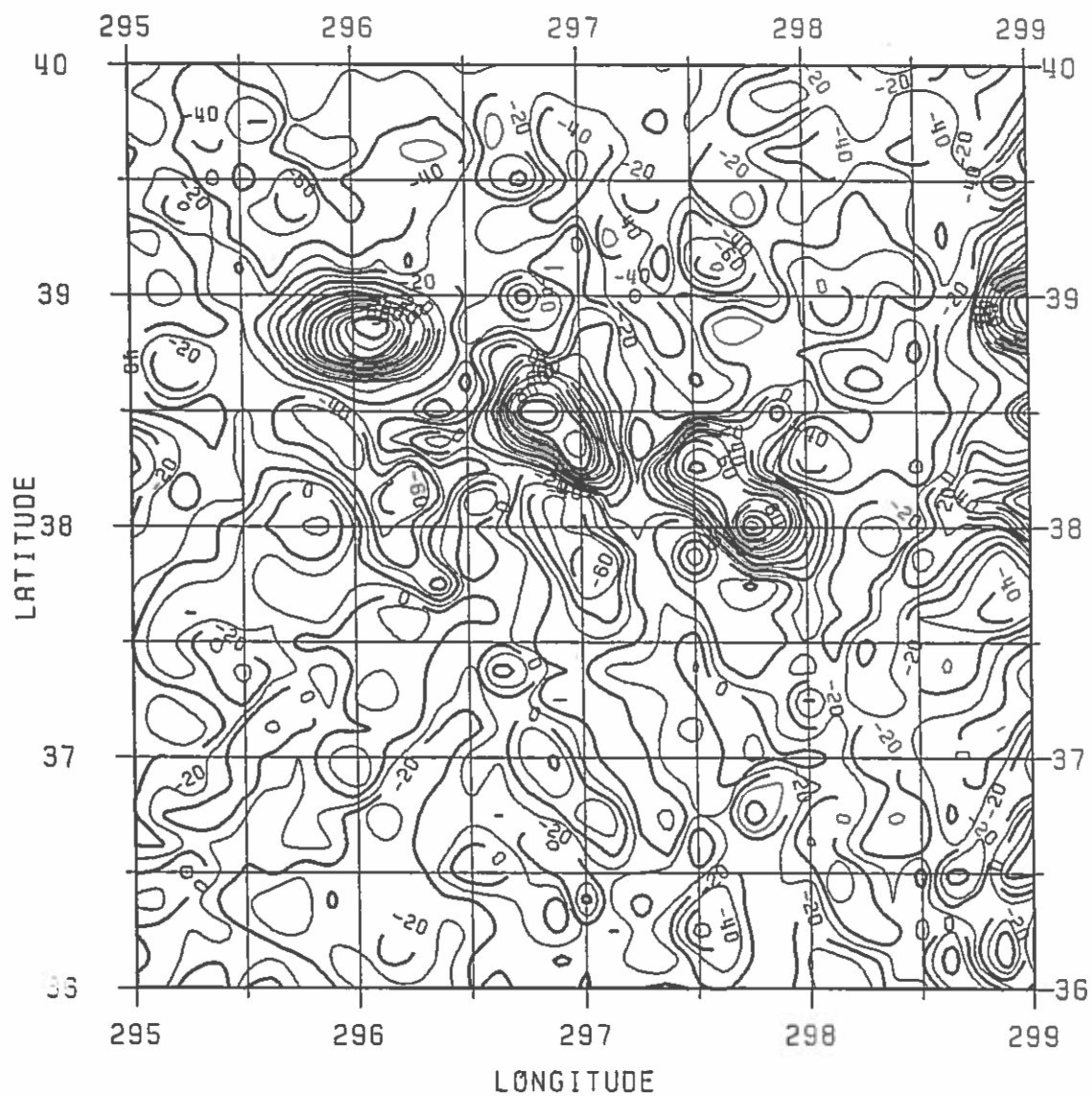


Figure 7.9 Anomaly map from the current production result in the New England Seamount area. CI = 10 mgals

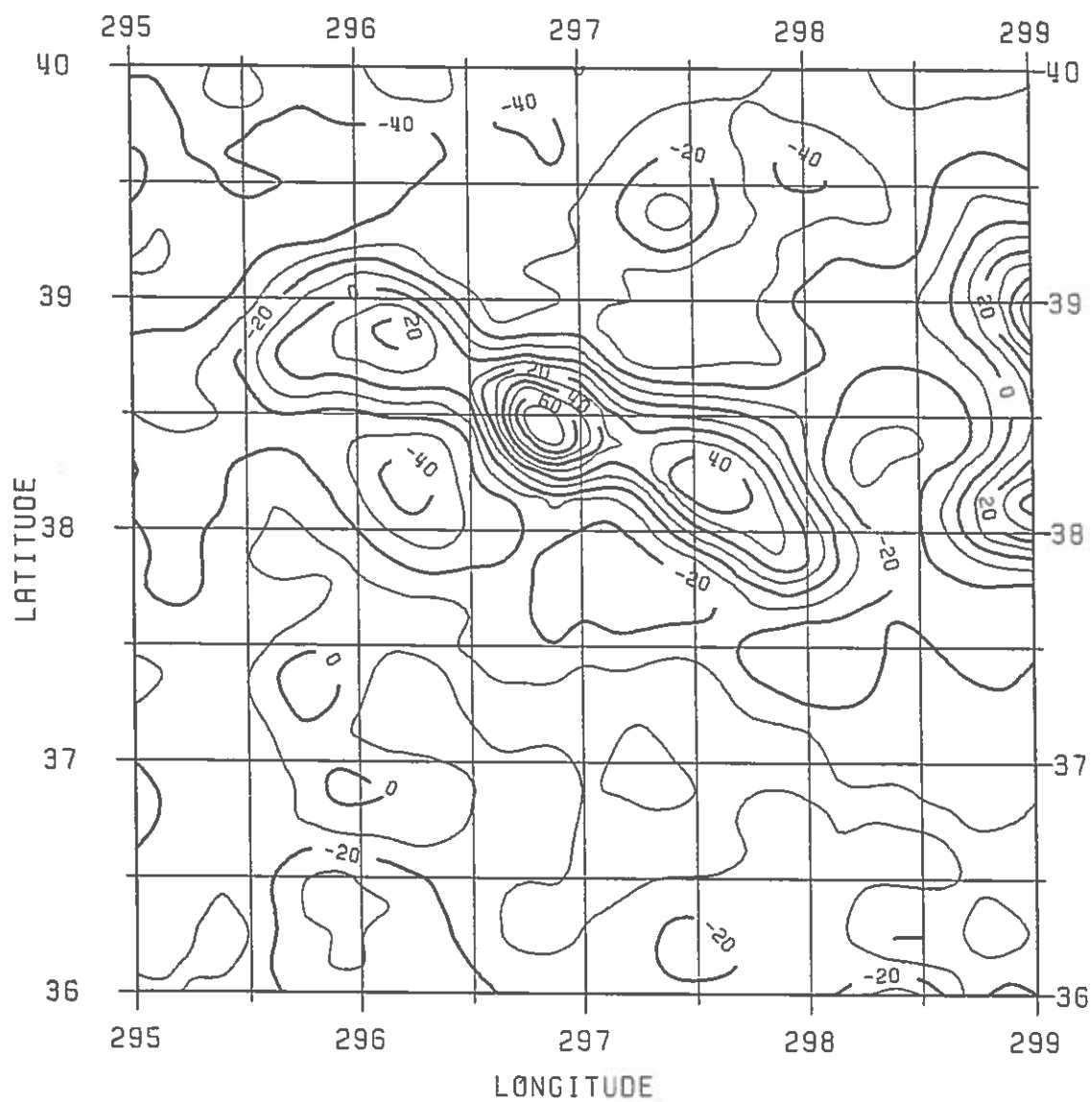


Figure 7.10 Anomaly map from Rapp's (1985) result in the New England Seamount area. CI = 10 mgals

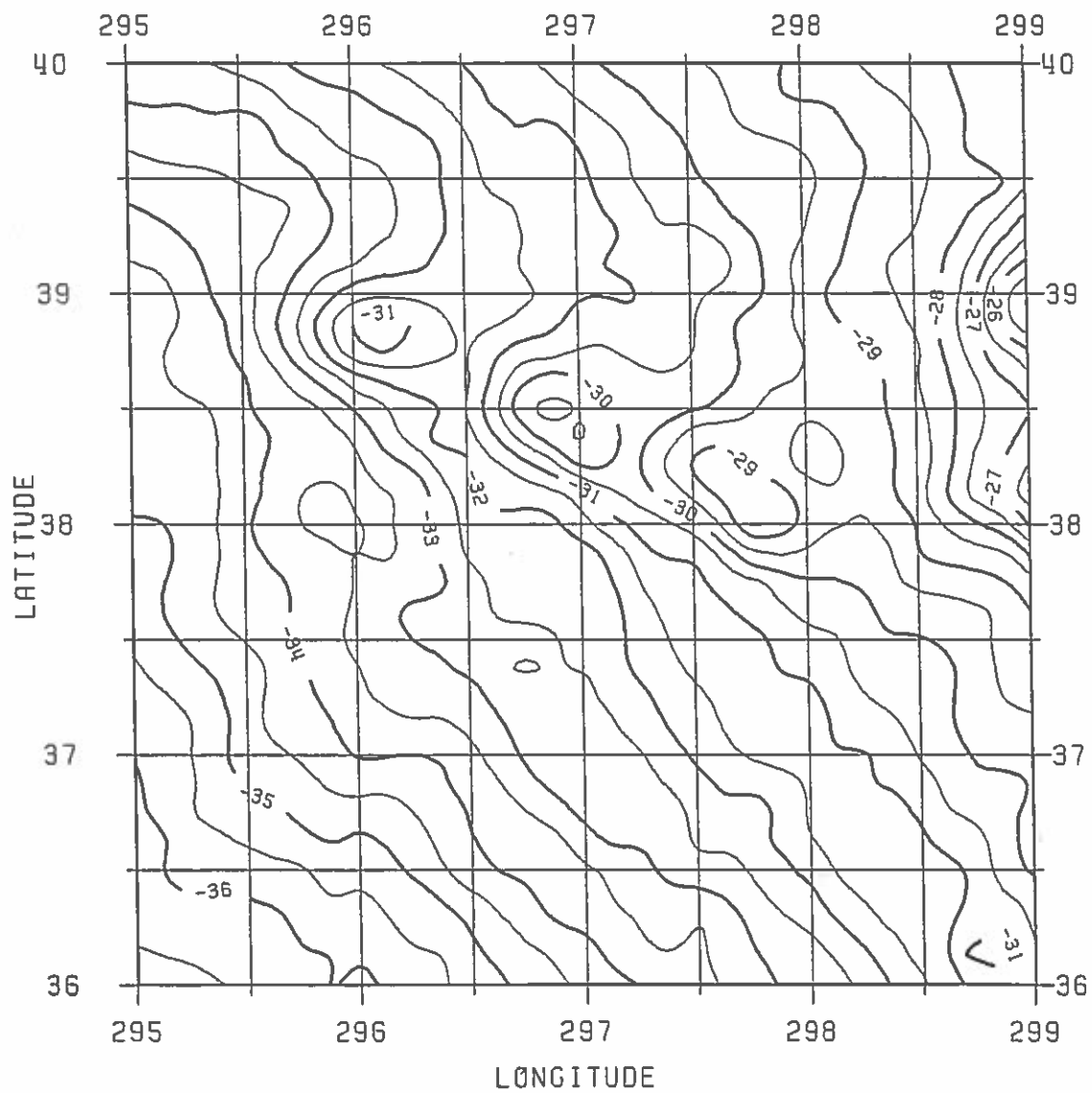


Figure 7.11 Sea surface height map from the current result in the New England Seamount area. CI = 1 meter

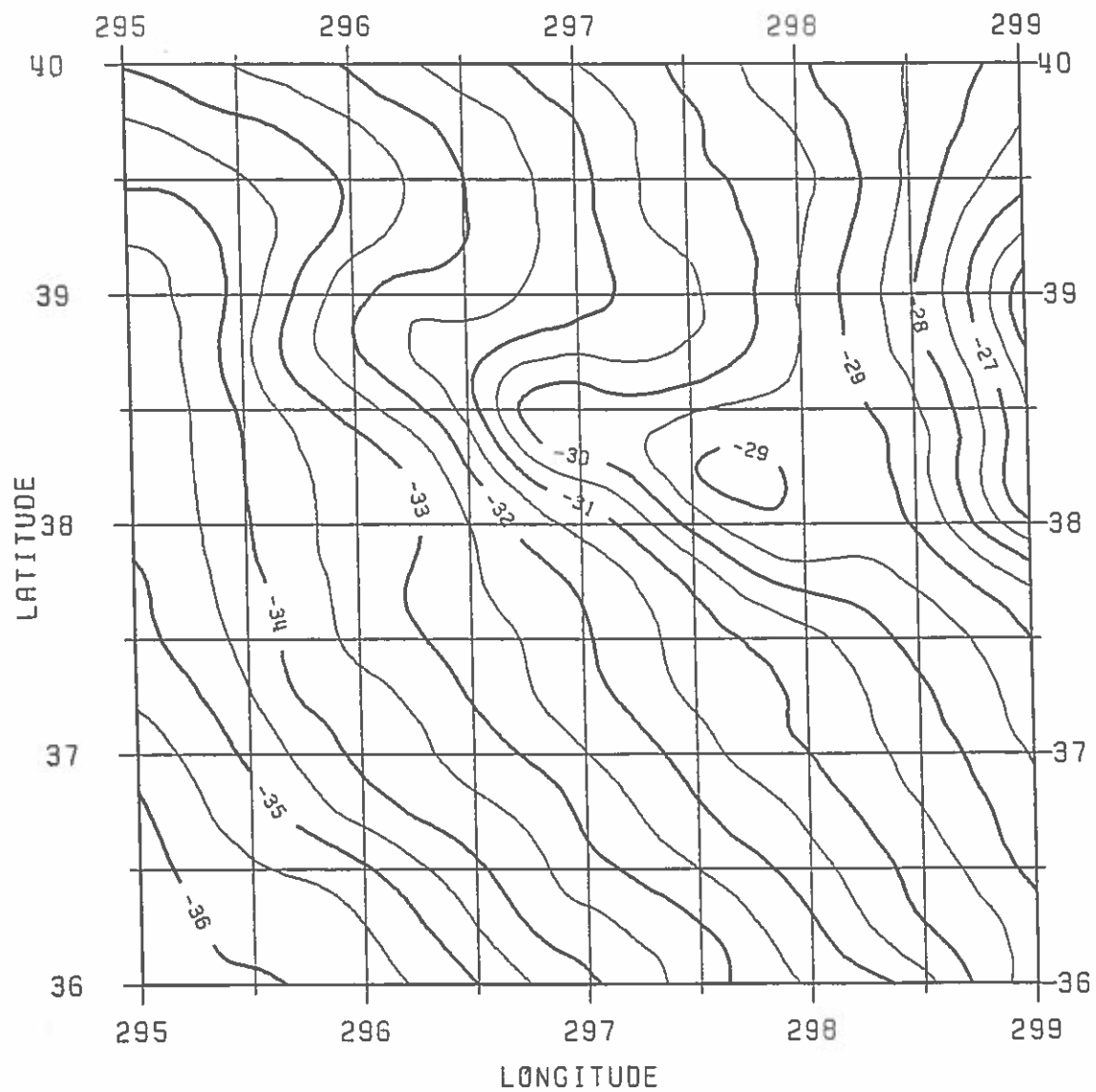


Figure 7.12 Sea surface height map from Rapp's (1985) result in the New England Seamount area. CI = 1 meter



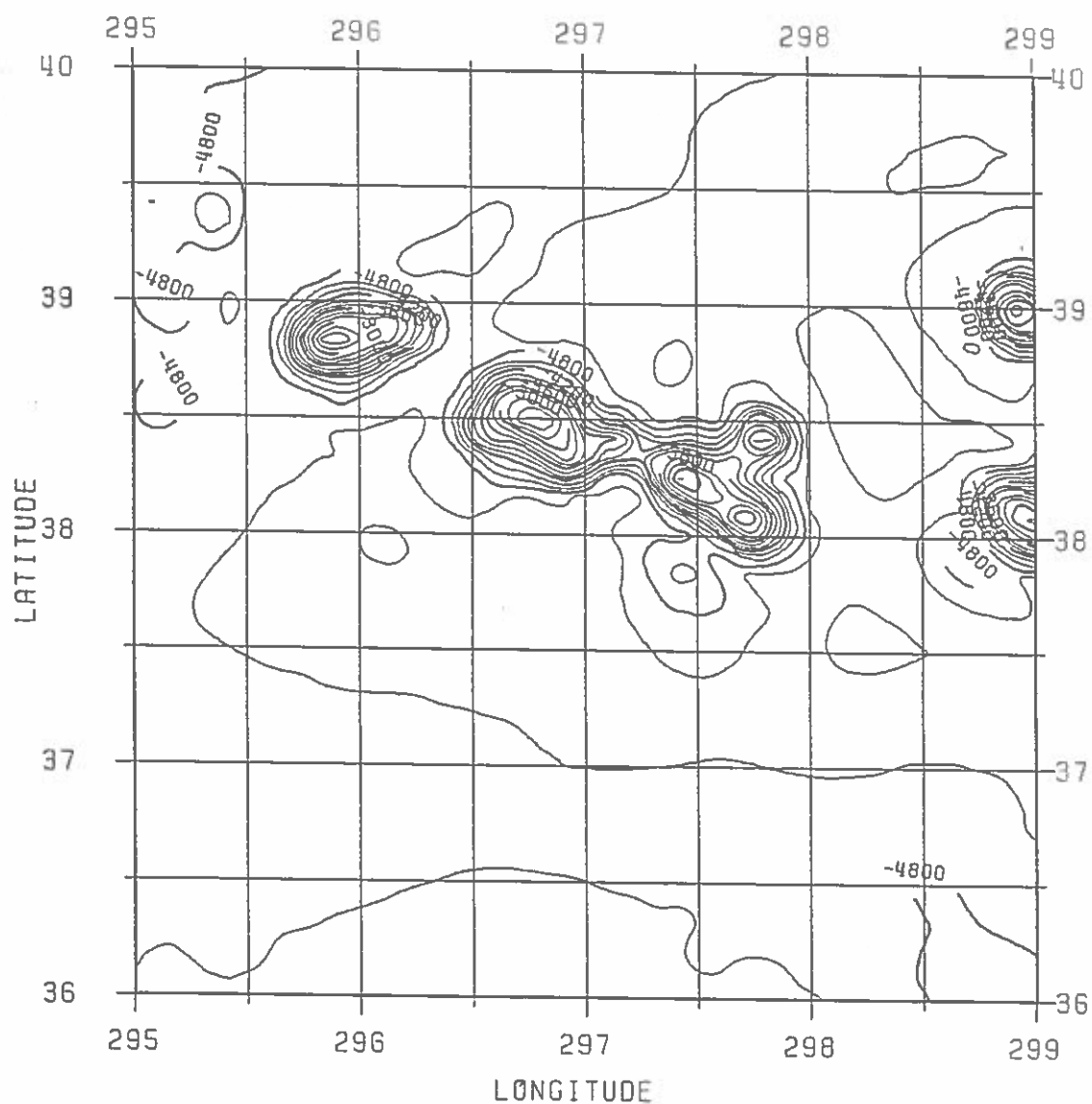


Figure 7.13 Bathymetry map in the New England Seamount area. CI = 200 meters

Another area of discussion is the spectral behavior of the eight degree data. Based on the analysis by Brammer and Sailor, and the FFT analysis in eight  $60^\circ$  - length profiles, Rapp (*ibid.*) concluded that the approximate point interval should be  $0^\circ.19$  instead of  $0^\circ.125$ . However, with the use of the new procedure in this study, the resolution of  $0^\circ.19$  becomes questionable. In the previous section, we have shown the intensified signatures in some seamount areas obtained from the current result. Also, in

Section 5.7 we have shown the excellent agreement between the predicted anomalies and ship determined anomalies in the rough gravity field area (see Figure 5.28). Therefore, with the data selection and the prediction procedure used in this study, we expect that the resolution will be higher than  $0.19^\circ$ , or 21.5 km.

An indirect proof of higher resolution can be carried out by the analysis of potential degree variances. Following Rapp (ibid.), the unitless anomaly degree variances,  $\sigma_l$ , can be computed from the power spectrum density of the anomaly covariance function by

$$\sigma_l = 5.02 \times 10^{-17} \frac{1 + \frac{3}{2}}{(l-1)^2} \phi_{gg} \left[ \omega = \frac{1 + \frac{3}{2}}{R} \right] \quad (7-2)$$

where  $\phi_{gg}(\omega)$  is the psd defined by the Hankel transform

$$\phi_{gg}(\omega) = 2\pi \int_0^\infty s C(s) J_0(\omega s) ds \quad (7-3)$$

where in planar case  $s$  is the distance,  $C(s)$  is the anomaly covariance function,  $\omega$  is the frequency function,  $J_0$  is the Bessel function of zero order. In equation (7-2), we have assumed that  $\phi_{gg}$  is given in  $(\text{mgal deg})^2$  and  $1^\circ$  spherical distance corresponds to 111 km on a mean sphere with  $R = 6371$  km.

Using Rapp's spectrum classification of the gravity field (Rapp, appendix D, 1985), we choose 3 areas for computing  $\sigma_l$ . Information of interest is indicated in Table 7.5.

Table 7.5

Three areas where potential  
degree variances are computed

area	spectrum classification	location
1	Rough	$15^\circ < \phi < 30^\circ, 285^\circ < \lambda < 300^\circ$
2	Mild	$15^\circ < \phi < 30^\circ, 180^\circ < \lambda < 195^\circ$
3	Smooth	$-30^\circ < \phi < -15^\circ, 250^\circ < \lambda < 265^\circ$

the reference anomalies based on GEMT1 potential coefficients to degree 20 are subtracted from 121 x 121 point anomalies. Program F491V1 in Rapp's program library is then used to compute  $\sigma_l$  for both the current result and Rapp's (1985) result. The corresponding spherical harmonic degree in such a computation is

$$l = 24i, \quad 1 \leq i \leq 60 \quad (7-4)$$

which is also based on Rapp (ibid.).

The resulting potential degree variances ( $\sigma_l$ ) are presented in Figure 7.5. In Figure 7.5, we also include the  $\sigma_l$  based on Kaula's rule and Tscherning/Rapp's model. Kaula's rule is

$$\sigma_l = (2l + 1) \left[ \frac{10^{-5}}{l^2} \right]^2 \quad (7-5)$$

and Tscherning/Rapp's model is

$$\sigma_l = \frac{A}{(l-1)^2 \gamma} \frac{(l-1)}{2(l-2)(l+24)} s^{l+2} \quad (7-6)$$

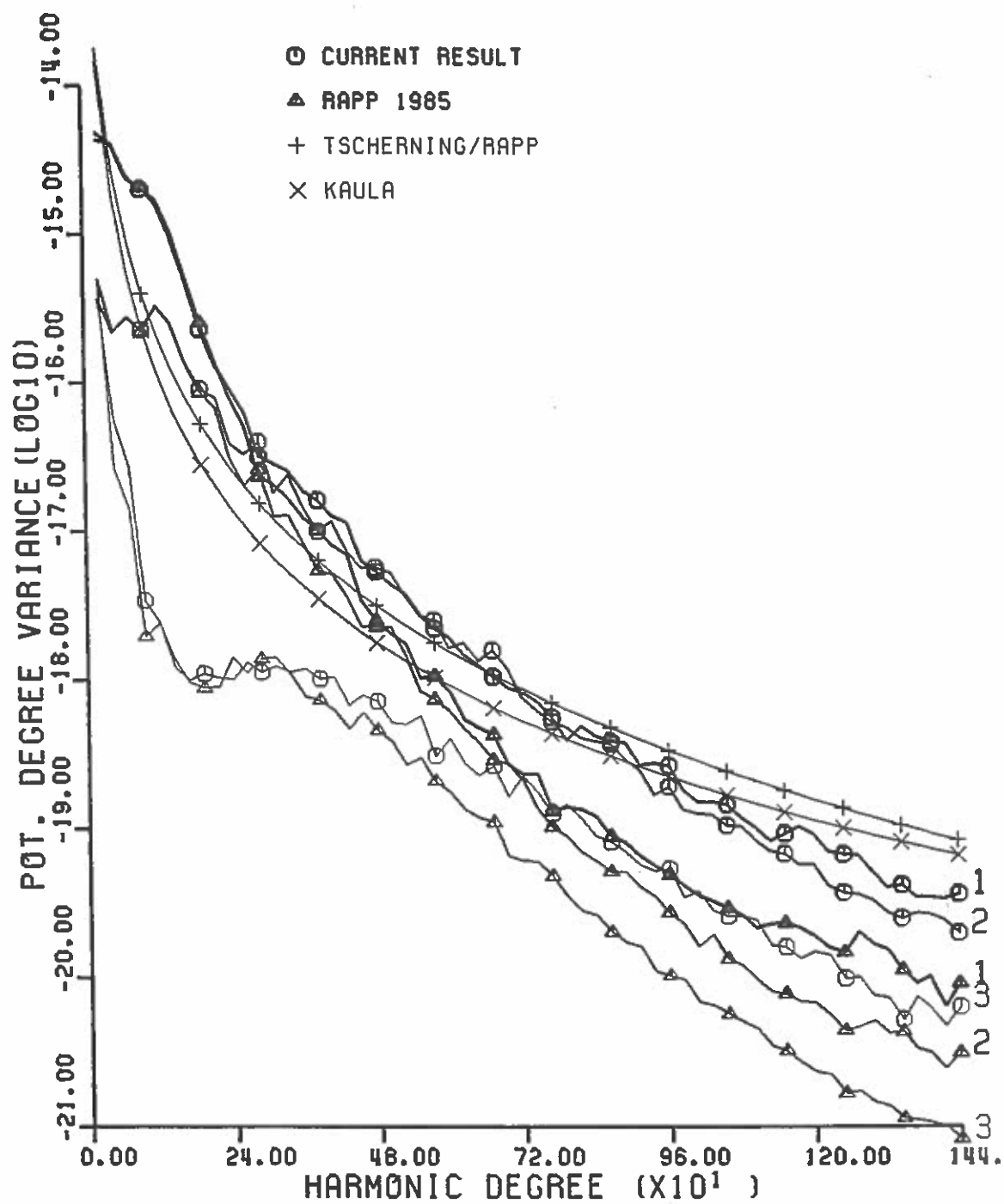
where  $A = 425.28 \text{ mgal}^2$  and  $s = 0.999617$ .

Theoretically, the highest harmonic degree is 1440 though it will be restricted to 900, if the resolution is  $0.19^\circ$ . To show the possibility of higher resolution, the potential

degree variances corresponding to degrees higher than 900 are also included in Figure 7.14. Two conclusions can be drawn from Figure 7.14.

(i) For the rough and mild areas (area 1 and area 2, respectively), the  $\sigma_1$  values from the current result agree very well with those from Rapp's before approximately degree 180. After degree 180, the deviation between the two sets of values begins to increase, but the latter one decays faster. This shows that more power has been obtained in the high degree part from the current result.

(ii) For the smooth area (area 3), small discrepancies between the  $\sigma_1$  values from the current result and Rapp's result can be found before approximately degree 300. Substantial deviation between the two sets of values starts from degree 300 and the latter one decays faster.



area 1:  $15^\circ < \phi < 30^\circ$ ,  $285^\circ < \lambda < 300^\circ$   
 area 2:  $15^\circ < \phi < 30^\circ$ ,  $180^\circ < \lambda < 195^\circ$   
 area 3:  $-30^\circ < \phi < -15^\circ$ ,  $250^\circ < \lambda < 265^\circ$

Figure 7.14 Potential degree variances from various sources in selected areas

## 7.2 Correlation of the eighth degree data and determination of 30'x30' mean anomalies and sea surface heights

For the purpose of high degree spherical harmonic expansion, it is necessary to compute the mean values from the current result. For the 30' x 30' mean value case, the simplest way is to take the straight mean of the values falling in the 30' x 30' block by

$$\bar{\Delta g} = \frac{\sum_{i=1}^n \Delta g_i}{n}, \quad 13 \leq n \leq 25 \quad (7-7)$$

$$\bar{N} = \frac{\sum_{i=1}^n N_i}{n}, \quad 13 \leq n \leq 25 \quad (7-8)$$

where  $\Delta g_i$  and  $N_i$  are the eighth degree point anomaly and sea surface height, respectively. The minimum number of points that can form a mean value is 13 and obviously the maximum number of points in a 30' x 30' block is 25 points.

The accuracy estimates of 30' x 30' mean anomalies for such a procedure is theoretically given by

$$\sigma_{\Delta g}^2 = \frac{1}{n^2} \left( \sum_{i=1}^n \sigma_{\Delta g_i}^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \sigma_{\Delta g_{ij}} \right) \quad (7-9)$$

where  $\sigma_{\Delta g_{ij}}$  is the error covariance between  $i$  and  $j$  points. Since our data does not contain this value, we use the following approximate method:

Let

$$\bar{\sigma}_{\Delta g_{xy}} = \beta \left( \frac{\sum_{i=1}^n \sigma_{\Delta g_i}^2}{n} \right) \quad (7-10)$$

be the mean error covariance for the points in a 30' x 30' block.  $\beta$  is the average correlation coefficient. This follows the usual definition of correlation between points  $x$  and  $y$  in the fundamental statistics:

$$\sigma_{xy} = \rho_{xy} \sigma_x \sigma_y \quad (7-11)$$

In equation (7-10) the product  $\sigma_x \sigma_y$  is replaced by the mean variance.

The estimated value for  $\sigma_{\Delta g}^2$  is then

$$\sigma_{\Delta g}^2 = \frac{1}{n^2} \left[ \sum \sigma_{\Delta g_i}^2 + (n^2 - n) \left( \beta \frac{\sum \sigma_{\Delta g_i}^2}{n} \right) \right] \quad (7-12)$$

which takes into account the error covariances through the use of an average correlation coefficient  $\beta$ . Eq. (7-12) reduces to

$$\sigma_{\Delta g}^2 = (1 + \beta (n - 1)) \frac{\sum \sigma_{\Delta g_i}^2}{n^2} \quad (7-13)$$

Thus

$$\beta = \frac{1}{n-1} \left( \frac{n^2 \sigma_{\Delta g}^2 - \sum \sigma_{\Delta g_i}^2}{\sum \sigma_{\Delta g_i}^2} \right) \quad (7-14)$$

To determine  $\beta$ , we need the rigorously computed  $\sigma_{\Delta g}^2$  in the collocation process. In other words, if we preserve error covariance matrix in the collocation, then  $\sigma_{\Delta g}^2$  can be obtained through

$$\sigma_{\Delta g}^2 = \frac{1}{n} A E_{\Delta g} A^T \quad (7-15)$$

where  $A$  is  $1 \times n$  matrix containing elements "1" alone.  $E_{\Delta g}$  is  $n \times n$  error covariance matrix whose diagonal elements are error variances of the predicted point anomalies. Equation (7-15) is equivalent to equation (7-9).

A set of altimeter data is tested to determine a  $\beta$  value. The area is located at  $30^\circ < \phi < 34^\circ$ ,  $293^\circ < \lambda < 297^\circ$ , which covers the Bermuda Islands. This area is excellent for testing since rough and smooth gravity fields were found here. In addition, we have a variety of choices of numbers of points used for the predictions, hence the change of  $\beta$  value due to the use of different numbers of points can be easily detected. In the tests, it is found that the error correlation between predicted anomalies is lower than that between predicted sea surface heights in a  $30' \times 30'$  prediction cell. The reduction of the points used for prediction also raises the correlation between predicted quantities. Table 7.6 shows the various  $\beta$  values for the case where an average of  $N$  points were used for prediction.

Table 7.6  
Average correlations of predicted anomalies and sea surface heights  
based on various numbers of points

N	$\beta_{\Delta g}$	$\beta_N$
263	0.02	0.06
247	0.02	0.06
153	0.02	0.10
45	0.05	0.18

For the oceanwide computation of  $30' \times 30'$  mean values, a varying  $\beta$  for each block is not possible. Therefore, we adopt an average  $\beta$  value for the computational purpose:

$$\beta_{\Delta g} = 0.04 \quad (7-16)$$

$$\beta_N = 0.12 \quad (7-17)$$



Thus, the accuracy estimates for 30' x 30' mean values are:

$$\sigma_{\Delta g}^2 = \left(1 + (n-1) \beta_{\Delta g}\right) \frac{\sum_{i=1}^n \sigma_{\Delta g_i}^2}{n^2} \quad (7-18)$$

$$\sigma_N^2 = \left(1 + (n-1) \beta_N\right) \frac{\sum_{i=1}^n \sigma_{N_i}^2}{n^2} \quad (7-19)$$

where  $\sigma_{\Delta g_i}$  and  $\sigma_{N_i}$  are available in the eighth degree data set.

Equations (7-18) and (7-19) are then investigated by comparing the estimated  $\sigma_{\Delta g}$  and  $\sigma_N$  from these two empirical formulas and those from equation (7-15) in the same test area (i.e.,  $30^\circ < \phi < 34^\circ$ ,  $293^\circ < \lambda < 297^\circ$ ) and area where bias removal tests were performed. (see Chapter 5). The agreement of  $\sigma_{\Delta g}$  from the two results is on the order of 1 to 2 mgals and the agreement of  $\sigma_N$  is on the order of 1 to 2 cm. This shows that equations (7-18) and (7-19) are applicable for our oceanwide computation of 30' x 30' mean values.

Using equations (7-7), (7-8), and (7-9), the 30' x 30' mean anomalies and sea surface heights and their accuracy estimates are computed from the current eighth degree data base. In the computations, we still apply the point acceptance criteria except that now we adopt the point values if the depths are greater than 0 meters (in Section 7.1, 100 meters is the minimum acceptable depth). Furthermore, in order to compare the current 30' x 30' mean values with Rapp's (1985), the point values below  $\phi = -63.0$  are also taken into account. Briefly speaking, the following criteria are accepted for computing the 30' x 30' mean values:

- (i) Depth > 0 meters
- (ii)  $-72^\circ \leq \phi \leq 72^\circ$
- (iii)  $(\phi, \lambda)$  not in area  $(45^\circ < \phi < 65^\circ, 250^\circ < \lambda < 296^\circ)$

It should be pointed out that the 0 meter depth criterion is based on the consideration of the smoothing effect of averaging point values in a 30' x 30' block, but the mean values in shallow water are still not reliable. Based on this effect, the statistics for the 30' x 30' mean anomalies has excluded the mean anomalies whose mean depths are smaller than 100 meters, Table 7.7 shows the statistics.

Table 7.7

The statistics for the 30' x 30' mean anomalies computed from the current result.  
(mgals)

No. of mean anomalies	134486
mean value	-1.89
mean predicted standard deviation	3.49
RMS predicted standard deviation	3.57
global mean anomaly variance	25.67
maximum anomaly	309.77
minimum anomaly	-287.89
maximum predicted standard deviation	18.22
minimum predicted standard deviation	2.51

The comparison of the current mean values with Rapp's (ibid.) mean values yields the results in Table 7.8.

Table 7.8

Comparison of 30' x 30' altimeter derived mean anomalies between the current result and Rapp's (1985) result

	case 1 <sup>+</sup>	case 2 <sup>*</sup>
No. of blocks compared	132881	127489
mean difference in anomaly	-0.57	-0.65
mean difference in std.	-2.15	-2.04
RMS difference in anomaly	6.47	5.43
RMS difference in std.	2.73	2.46
maximum difference in anomaly	184.95	98.15
minimum difference in anomaly	-199.63	-80.35
maximum difference in std.	8.17	5.06
minimum difference in std.	-28.11	-28.11

<sup>+</sup> Comparison made between  $-71^{\circ}5 \leq \phi \leq 72^{\circ}0$

<sup>\*</sup> Comparison made between  $-63^{\circ}0 \leq \phi \leq 72^{\circ}0$

In Table 7.8, we have considered two cases: case 1 includes the comparison made between  $-71.5 \leq \phi \leq 72.0$ , case 2 includes the comparison made between  $-63.0 \leq \phi \leq 72.0$ . The substantial differences of the maximum and minimum differences in the two cases were caused by the instability of prediction in the area below  $\phi = -63.0$ . Thus we believe that the predictions in the area below  $\phi = -63^\circ$  are not reliable.

Large differences of mean anomalies between the two results are also found in high latitude areas and areas with strong signature of anomalies. In case 2, the block of maximum difference is located in the Bermuda area, at (north-west corner)  $\phi = 32.5$ ,  $\lambda = 295.0$ ; The block of minimum difference is located in a trench area, at (north-west corner)  $\phi = -5.0$ ,  $\lambda = 145.5$ . In order to present the differences completely, the  $30' \times 30'$  blocks where the absolute differences are greater than 15 mgal are shown in Figure 7.15.

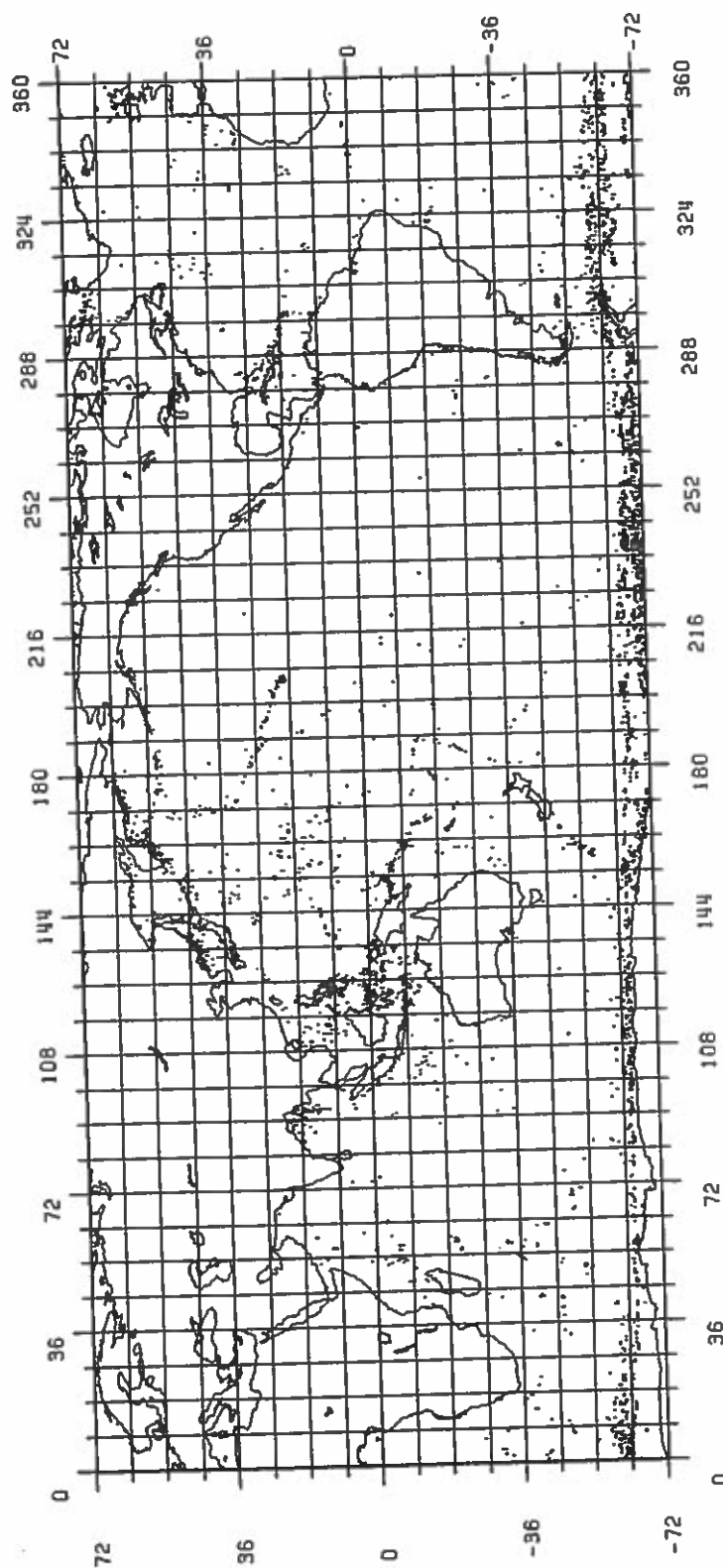


Figure 7.15 Distribution of 30' x 30' cells where the absolute differences between the 30' x 30' mean anomalies from the current result and Rapp's (1985) result are greater than 15 mgals.

## CHAPTER VIII

### Conclusions and recommendations

In this study, we have attempted to predict the high precision point anomalies and altimeter implied geoid undulations from the combined Geos-3/Seasat data through the use of the optimal prediction procedure. The original 5.9 million altimetric point data base has been used in the predictions, but a small portion of data points have been excluded due to the use of data selection criterion (The percentage of the deleted points depends on the data density and the distribution of the data points). The accuracy estimates were rigorously computed; Furthermore, we have incorporated the local arc adjustment program and we believe that the biases of the arcs having the wavelength of 400 ~ 500 km have been removed.

The method used for predicting the gravity anomalies and altimeter implied geoid undulations is the least squares collocation (lsc) method. The covariance function is based on Tscherning/Rapp's anomaly degree variance model and OSU86C noise model. Using different parameters in T/R's anomaly degree variance model, we found that Jekeli's parameters sometimes caused the instability of the covariance matrix or even singularity of the matrix, especially when data density is high or the observations are located at high latitude area. On the other hand, T/R's parameters gave more stable condition of the covariance matrix except for few cases that were presented in Chapter 5. Therefore, T/R's have been adopted in the covariance model for the global production work.

Before the global production work was put into action, we performed the analysis of special routines such as matrix inversion routines, matrix product routine and matrix-vector product routine. The optimization for the production program, F459PRD, was successfully achieved. The contribution of the supercomputer to the global production work is remarkable. We estimate that approximately 12 CPU hours (including the 8.34 hours tabulated in Table 6.1 and the CPU time for repeated production runs) have been consumed by the entire production work. If the optimization were not made for the production program, we then need about 216 cpu hours on CRAY X-MP/28; If the

production program were executed in IBM machine, we probably would need 1632 CPU hours (see the saving factors in Table 4.3 (b)). Therefore, we conclude that for the type of production work in this study not only the use of supercomputer is necessary but also the optimization (mainly vectorization) of the program is necessary. With the advent of more powerful supercomputer (such as CRAY Y-MP/864 that will be installed at the Ohio Supercomputer Center in July, 1989), we have more confidence on attacking more difficult problems.

Before we started the global predictions, many tests were performed in order to choose the optimal parameters for the prediction procedure. With these optimal parameters, we expected that we can recover gravity field related signature as much detail as possible. From these tests, we concluded that a  $4^\circ \times 4^\circ$  block (with  $0.5^\circ$  border) and a  $0.5^\circ \times 0.5^\circ$  cell (with a  $0.25^\circ$  border) were the sizes of the area for one adjustment and one prediction process, respectively. The maximum number of data points in such a prediction cell was 400 beyond which no substantially improved result can be obtained. The use of 4 km as the minimum spacing between the data points has maintained the stability of the covariance matrix and reduced the computer time significantly. Other important parameters can also be found in Chapter 5.

With the use of optimal parameters, excellent agreement between predicted anomalies and the ship measured anomalies has been achieved in the Bermuda are. The RMS difference based on the comparison at the 2011 points was 15.9 mgals. Considering the large variation of the anomalies (from -46.40 to 358.10 mgals), such a small RMS difference proves that the reliable point anomalies and altimeter implied geoid undulations can be obtained from satellite altimeter data.

One unresolved problem is related to the track errors that were not removed through our adjustment model. In this study, we have used the cross-over model in which the bias is assumed to be a constant within the adjustment block. Although we have shown the capability of bias removal of this model, the revised prediction as described in Chapter 6 has revealed the fact that some track errors still cannot be removed through this model. In Chapter 3, we presented two other adjustment models (models B and C), it will be necessary to investigate whether these two models will have better performances in bias removal for the production work.

A total of 2,322,080 point values were predicted in the production work. However, 273,593 of them are located either in shallow water areas (depth less than 100 meters) or in the areas where we believe that the prediction is not reliable (at  $\phi < -63^\circ$  and Hudson Bay and Hudson Strait). Excluding these points, 2,048,487 were considered to be acceptable. Based on the statistics of the accepted point values, the maximum anomaly is 439.70 mgals and minimum anomaly is -368.50 mgals; The maximum undulation is 82.66 meters and minimum undulation is -107.54 meters. The mean standard deviation of the anomalies is 12.11 mgals, which justifies that Rapp's (1985) empirical formula for estimating the standard deviations is appropriate in a global sense. The spectral analysis in the three selected areas (i.e., rough, mild and smooth areas) for the point values indicated that the potential degree variances obtained from the result in this study are almost coincident with those obtained from Rapp's (ibid.) result before harmonic degree of 180. As the harmonic degree goes higher than 180, the degree variances from the current result have gained more power than Rapp's result.

A total of 134,486  $30' \times 30'$  mean anomalies were computed from the point values. The comparison between this set of  $30' \times 30'$  mean anomalies and the set of  $30' \times 30'$  mean anomalies obtained from Rapp's (1985) result at the 127,489 common  $30' \times 30'$  blocks was made. The comparison showed that the RMS difference is 5.43 mgals, the maximum difference is 98.15 mgals, and the minimum difference is -80.35 mgals. The large differences between the two sets of mean anomalies primarily arose from the use of different data densities in the current prediction work and Rapp's prediction work. A maximum of 300 points in a  $3^\circ \times 3^\circ$  prediction cell (with  $0.5^\circ$  border) was used in Rapp's (ibid.) prediction work, while a maximum of 400 points in a  $0.5^\circ \times 0.5^\circ$  prediction cell (with  $0.25^\circ$  border) was used in the current prediction work, thus the data density in the current work is 21 times of that in Rapp's (ibid.) work. Since we have the tool to handle this type of prediction (i.e., the current one), it is recommended that the data density of the altimeter data used for future work should be as high as possible (of course, the use of minimum spacing for altimeter data points is necessary or the covariance matrix will be singular). With the use of high density of altimeter data, as shown in the Bermuda area, we have greater possibility to derive the "true" gravity anomalies and geoid undulations from the altimeter data.

Based on this study, we have the following recommendations for future altimeter work:

1. Improvement of reference values in some rough area must be made.
2. More altimeter data should be combined with the existing data, especially for the areas with the data gaps in coverage. The data gap problem is more serious in the Southern Hemisphere, e.g., at  $-65^{\circ} \leq \phi \leq -35^{\circ}$  and  $-20^{\circ} \leq \lambda \leq 20^{\circ}$ , and at  $-50^{\circ} \leq \phi \leq -30^{\circ}$  and  $60^{\circ} \leq \lambda \leq 90^{\circ}$ .
3. Better methods for removing short-wavelength track errors (on the order of 10 - 50 km) must be developed, or the data such as those along the tracks in Table 6.2 should be excluded from our existing data base.
4. In order to derive true or nearly true point values from altimeter data, a recommended data density is 400 points or higher in an area of  $110 \times 110 \text{ km}^2$ . The data distribution must be even.



## LIST OF REFERENCES

- Boeing Computer Services, (BCS), Supercomputer Vectorization and Optimization Guide, Engineering Technology Application Division, (1984).
- Caspary, W.F., Concepts of Network and Deformation Analysis, Monograph 12, School of Surveying, The University of New South Wales, Kensington, NSW Australia, (1987).
- Dongarra, J.J., E.B. Moler, J.R. Bunch, and G.W. Stewart, Linpack User's Guide, The Society for Industrial and Applied Math (SIAM), Philadelphia, (1979).
- Engelis, T., Spherical Harmonic Expansion of the Levitus Sea Surface Topography, Report No. 385, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1987).
- Gerald, C.F. and P.O. Wheatley, Applied Numerical Analysis, Addison-Wesley Publishing Inc., (1984).
- Jackson, D.D., The Use of A priori Data to Resolve Non-uniqueness in Linear Inversion, Geophys. J.R. astr. Soc., 57, 137-157, (1979).
- Jekeli, C., An Investigation of Two Models for the Degree Variances of Global Function, Report No. 275, Dept. of Geodetic Science and Surveying, The Ohio State University, (1978).
- Kadir, M., The Recovery of High Frequency Gravity Field Information from Satellite Altimeter Data, Internal Report, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1987).
- Knudsen, P., Adjustment of Satellite Altimeter Data from Cross-over Differences Using Covariance Relatives for the Time Varying Components Represented by Gaussian Functions, Proc. IAG Symposia, TOME II, 617-628, (1987).
- Kolenkiewicz, R., and C.F. Martin, Seasat Altimeter Height Calibration, Seasat Special Issue I, American Geophys. Union, Washington, D.C., 3189-3197, (1982).
- Levitus, S., Climatological Atlas of the World Ocean, NOAA Geophysical Fluid Dynamics Laboratory, Professional Paper 13, Rockville, MD, (1982).
- Liang, C., The Adjustment and Combination of Geos-3 and Seasat Altimeter Data, Report No. 346, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1983).

- Lisitzin, E., Sea Level Changes, Elsevier Oceanography Series, Amsterdam, (1974).
- Moritz, H., Advanced Physical Geodesy, Herbert Wechmann Verlag Karlsruhe, FRG, (1980).
- Rapp, R.H., Geos-3 Data Processing for the Recovery of Geoid Undulations and Gravity Anomalies, J. Geophys. Res., 84, B4, 3784-3792, (1979).
- Rapp, R.H., The Determination of Geoid Undulations and Gravity Anomalies from Seasat Altimeter Data, J. Geophys. Res., 88, 1552-1562, (1983).
- Rapp, R.H., Detailed Gravity Anomalies and Sea Surface Heights derived from Geos-3/Seasat Altimeter Data, Report No. 365, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1985).
- Rapp, R.H. and J.Y. Cruz, The Representation of the Earth's Gravitational Potential in a Spherical Harmonic Expansion to Degree 250, Report No. 372, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1986).
- Rapp, R.H., An Estimate of Equatorial Gravity from Terrestrial and Satellite Data, Geophys. Res. Lett., Vol. 14, No. 7, 730-732, (1987).
- Rapp, R.H., Advanced Gravimetric Geodesy, Classnotes, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1988).
- Rowlands, D., The Adjustment of Seasat Altimeter Data on a Global basis for Geoid and Sea Surface Height Determination, Report No. 325, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1981).
- Sünkel, H., Program Listing for the General Surface Representation Module Designed for Geodesy (GSPP), Internal Report, Dept. of Geodetic Science, The Ohio State University, Columbus, (1980).
- Tscherning, C. and R.H. Rapp, Closed Covariance Expressions for Gravity Anomalies, Geoid Undulation, and Deflections of the Vertical Implied by Anomaly Degree Variance Models, Report No. 208, Dept. of Geodetic Science, The Ohio State University, Columbus, (1974).
- Tscherning, C. and P. Knudsen, Determination of Bias Parameters for Satellite Altimeter by Least-squares Collocation, I Hotine-Marussi Symposium on Mathematical Geodesy, Roma, 3-6 June, 833-851, (1985).
- Uotila, U.A., Notes on Adjustment Computation, Part I, Dept. of Geodetic Science and Surveying, The Ohio State University, Columbus, (1986).
- Wieser, M., The Global Digital Terrain Model TUG87, Internal report on set-up, origin and characteristics, Dept. of Mathematical Geodesy, The Technical University in Graz, (1987).