



**UCGE Reports
Number 20185**

Department of Geomatics Engineering

**An Analysis on the Optimal Combination of Geoid,
Orthometric and Ellipsoidal Height Data**

(URL: <http://www.geomatics.ucalgary.ca/links/GradTheses.html>)

by

Georgia Fotopoulos

December 2003



UNIVERSITY OF
CALGARY

UNIVERSITY OF CALGARY

An analysis on the optimal combination of geoid, orthometric and ellipsoidal height data

by

Georgia Fotopoulos

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

DECEMBER, 2003

© Georgia Fotopoulos 2003

ABSTRACT

The main objective of this research is to present a detailed analysis of the optimal combination of heterogeneous height data, with particular emphasis on (i) modelling systematic errors and datum inconsistencies, (ii) separation of random errors and estimation of variance components for each height type, and (iii) practical considerations for modernizing vertical control systems. Specifically, vertical control networks consisting of ellipsoidal, orthometric and geoid height data are investigated. Although the theoretical relationship between these height types is simple in nature, its practical implementation has proven to be quite challenging due to numerous factors that cause discrepancies among the combined height data. To address these challenges a general procedure involving empirical and statistical tests for assessing the performance of selected parametric models is developed. In addition, variance component estimation is applied to the common adjustment of the heterogeneous heights. This leads to an in-depth analysis of the effects of correlation among heights of the same type, provisions for computing non-negative variance factors, and the intrinsic connection between the proper modelling of systematic errors and datum inconsistencies with the estimated variance components. Additional numerical studies include the calibration of geoid error models (both regional and global), scaling the GPS-derived ellipsoidal height covariance matrix, and evaluating the accuracy of orthometric heights obtained from national/regional adjustments of levelling data. Ultimately, one of the main motivations for this work is embedded in the eminent need to introduce *modern* tools and techniques, such as GNSS-levelling, in establishing vertical control. Therefore, part of this research is aimed at bringing to the forefront some of the key issues that affect the achievable accuracy level of GNSS-levelling. Overall, the analysis of the optimal combination of the heterogeneous height data conducted herein provides valuable insight to be used for a variety of height-related applications.

PREFACE

This is an unaltered version of the author's Doctor of Philosophy thesis of the same title. This thesis was accepted by the Faculty of Graduate Studies in December, 2003. The faculty supervisors of this work were Dr. M.G. Sideris and Dr. N. El-Sheimy, and the other members of the examining committee were Dr. J.A.R. Blais, Dr. Y. Gao, Dr. L. Lines, and Dr. A.H.W. Kearsley.

ACKNOWLEDGEMENTS

I would like to express my appreciation and thank my supervisors, Dr. M.G. Sideris and Dr. N. El-Sheimy, for their support and guidance throughout my graduate studies. Their continuous encouragement and advice were greatly appreciated. The other member of my supervisory committee, Dr. J.A.R. Blais, is also acknowledged for answering various questions throughout my studies at the University of Calgary.

I would also like to thank Dr. A.H.W. Kearsley and Dr. W.E. Featherstone for hosting my visit as a research fellow in Australia during my studies. It was an excellent opportunity and an inspiring place to work. My friend and colleague, Christopher Kotsakis, is also thanked for his thorough proofreading of the initial draft of this thesis.

Funding for my PhD studies was provided from numerous sources, including NSERC, iCORE, GEOIDE, Alberta Heritage Foundation, Killam Scholarship Fund, the University of Calgary, and the Department of Geomatics Engineering.

Urs Marti from the Swiss Mapping Authority is gratefully acknowledged for making the Swiss data available. Also, Marc Véronneau, Mike Craymer, and André Mainville from the Geodetic Survey Division at NRCAN are thanked for providing the Canadian data and answering all of my questions. Christian Malmquist is also thanked for his work on the covariance matrix for EGM96.

Finally, I would like to thank my brothers Costa, Alex and Chris for keeping me grounded during the writing of this manuscript. Last, but not least, I thank my parents for their endless support and encouragement. I could not have done it without them.

Στους γονείς μου,

Αθανάσιο και Ειρήνη.

TABLE OF CONTENTS

Approval Page	ii
Abstract	iii
Preface	iv
Acknowledgements	v
Table of Contents	vii
List of Tables	x
List of Figures	xii
List of Symbols	xvi
List of Abbreviations	xix
1 Introduction	1
1.1 Background	1
1.2 Thesis Objectives	5
1.3 Thesis Outline	7
2 Heights, Vertical Datums and GNSS-Levelling	9
2.1 Geoid heights	9
2.2 Orthometric heights	19
2.3 Ellipsoidal heights	26
2.4 Why combine geoid, orthometric and ellipsoidal height data?	35
2.4.1 Modernizing regional vertical datums	36
2.4.2 Global vertical datum	41
2.4.3 GNSS-levelling	46
2.4.4 Refining and testing gravimetric geoid models	48
3 Combined Height Adjustment and Modelling of Systematic Effects	51
3.1 General combined adjustment scheme	51
3.2 Role of the parametric model	59
3.3 Modelling options	63
3.4 Selecting a parametric model	69
3.5 Assessing the parametric model performance	72
3.5.1 Classical empirical approach	73
3.5.2 Cross-validation	75
3.5.3 Assessing the goodness of fit	76
3.5.4 Testing parameter significance	80
3.6 Summary	84

4	Results for the Parametric Model Surface Fits	86
4.1	Results for the Switzerland test network.....	87
4.2	Results for the Canadian test network.....	95
4.3	Fitting a gravimetric geoid model to the Australian height datum via GPS	106
4.4	Remarks on results	113
5	Overview of Variance Component Estimation	115
5.1	Introduction	115
5.2	Review of methods for estimating variance components.....	119
5.2.1	Functional models	121
5.2.2	Stochastic models.....	123
5.2.3	Selection of variance component estimation procedures	126
5.3	The MINQUE method.....	131
5.4	Application of MINQUE to the combined height adjustment problem	133
5.5	Remarks on VCE and the combined height adjustment problem	136
5.6	Iterative minimum norm quadratic unbiased estimation.....	138
5.7	Iterative almost unbiased estimation	142
6	Case Studies: Estimation of Variance Components via a Mixed Adjustment of Ellipsoidal, Geoid and Orthometric Heights	144
6.1	Description of the Swiss test network data	145
6.1.1	Initial covariance matrix for the GPS heights	146
6.1.2	Initial covariance matrix for the geoid heights.....	147
6.1.3	Initial covariance matrix for the orthometric heights.....	148
6.2	Description of the southern British Columbia/Alberta test network data	151
6.2.1	Initial covariance matrix for the GPS heights	152
6.2.2	Initial covariance matrix for the geoid heights.....	153
6.2.3	Initial covariance matrix for the orthometric heights.....	154
6.3	Case Study I - Testing a-priori covariance matrices for the height data	156
6.3.1	Scaling the covariance matrix for the ellipsoidal heights	159
6.3.2	Calibration of geoid error models	161
6.3.3	Assessing the accuracy of orthometric heights	168
6.4	Case Study II - Non-negative variance components	169
6.5	Case Study III - Effects of correlations	175
6.6	Case Study IV - Role of the type of parametric model	177
6.7	Summary	180

7	Practical Considerations for Modernizing Vertical Control	182
7.1	Introduction	182
7.2	GNSS-levelling: What are the issues?	184
7.2.1	Computing the accuracy of the levelling data.....	185
7.2.2	Computing the accuracy of the GPS data	188
7.2.3	Computing the accuracy of the geoid height differences.....	190
7.2.4	Achievable accuracy of GNSS-levelling for a new baseline	192
7.2.5	Tests in northern Canada.....	199
7.3	Combining heterogeneous heights at sea	204
7.4	Other practical issues	212
7.5	Summary	214
8	Conclusions and Recommendations	216
8.1	Conclusions	216
8.2	Recommendations for future work.....	221
	References	224

LIST OF TABLES

2.1: Expected commission and omission errors after GOCE.....	18
4.1: Statistics of residuals at points used in the adjustment for various parametric models (Swiss network, units: cm)	91
4.2: Statistics of cross-validation tests for various parametric models (Swiss network, units: cm)	92
4.3: Condition numbers for various parametric models in Switzerland.....	93
4.4: Summary of selection criteria for Swiss test network.....	94
4.5: Statistics of residuals at points used in the adjustment for various parametric models (southern BC/AB network, units: cm).....	100
4.6: Condition numbers for various parametric models in southern BC/AB	101
4.7: Statistics of cross-validation tests for various parametric models (southern BC/AB network, units: cm)	102
4.8: Statistics of residuals at points used in the adjustment for orthonormalized versions of various parametric models (southern BC/AB network, units: cm)	104
4.9: Summary of selection criteria for southern BC/AB test network	105
4.10: Statistics of the residuals before any fit (units: cm).....	108
4.11: Description of Australian regional network geometry	108
4.12: Statistics of residuals at the points used in the adjustment (upper row) and statistics obtained through cross-validation (bottom row) for Australia (units: cm)	110
4.13: Condition numbers for various parametric models for the Australian network.....	111
4.14: Final selected models per region and corresponding RMS obtained through cross-validation	112
5.1: Timeline of key VCE developments in geodetic literature	120
5.2: Number of GPS-levelling benchmarks in different regions.....	137
6.1: Initial CV matrix characteristics for the Swiss network	150

6.2:	Initial CV matrix characteristics for the southern BC/AB network	156
6.3:	Estimated variance components and average standard deviations for local and global geoid models.....	165
6.4:	Estimated variance components for orthometric height data and average standard deviations for C_H	169
6.5:	Estimated variance components using the IMINQUE and IAUE methods	173
6.6:	Effect of correlations on estimated variance factors (Swiss network).....	175
6.7:	Effect of correlations on estimated variance factors (south BC/AB network).....	176
6.8:	Estimated variance components for the Swiss test network using various parametric models	178
6.9:	Estimated variance components for the southern BC/AB test network using various parametric models.....	178
7.1:	Statistics of various test network configurations (d denotes baseline length in km).....	194
7.2:	Standard error of relative height components (in cm) as a function of baseline length.....	197
7.3:	Results of baselines moving northward from the test network area (units: cm)....	202

LIST OF FIGURES

1.1: Relationship between ellipsoidal, geoid and orthometric heights.....	2
2.1: Computation of regional geoid models using heterogeneous data.....	13
2.2: Orthometric height of a point.....	21
2.3: Approximation in orthometric heights by neglecting the difference in lengths between the curved plumb line and the ellipsoidal normal.....	24
2.4: Reference ellipsoid and geodetic coordinates	27
2.5: Zenith dependence of station height, receiver clock and tropospheric delay parameters	30
2.6: Sources of errors for global navigation satellite systems.....	31
2.7: Establishment of a reference benchmark height	38
3.1: Illustrative view of GPS/geoid levelling and the role of the corrector surface	62
3.2: First (a), second (b), third (c) and fourth (d) order bivariate polynomial fits	71
3.3: Example of height misclosures before and after parametric model fit	74
3.4: Classical empirical testing approach.....	74
3.5: Cross-validation procedure on a point-by-point basis.....	76
3.6: Relationship between R^2 and R_α^2	78
3.7: Stepwise procedure for testing parameter significance.....	84
3.8: Flowchart of major steps for assessing the performance of parametric corrector surface models	85
4.1: Swiss test network of GPS-on-benchmarks and original height misclosures	87
4.2a: Low-order corrector surface model fits for the Swiss network.....	89
4.2b: High-order corrector surface model fits for the Swiss network.....	90
4.3: Statistical measures of goodness of fit for various parametric models (Swiss test network)	91

4.4: Classic 4-parameter corrector surface fit for the Swiss test network.....	94
4.5: Canadian GPS-levelling benchmark network	95
4.6: Test network of GPS-on-benchmarks in southern BC/AB and original height misclosures	96
4.7a: Low-order corrector surface model fits for the southern BC/AB network	98
4.7b: High-order corrector surface model fits for the southern BC/AB network	99
4.8: Statistical measures of goodness of fit for various parametric models (southern BC/AB test network).....	100
4.9: Empirical test results for the southern BC/AB region.....	102
4.10: 7-parameter differential similarity fit for the southern BC/AB test network.....	106
4.11: GPS-levelling data on the Australian mainland.....	107
4.12: Examples of the 4-parameter corrector surface fit (left) and the 10-parameter MRE fit (right) for the Australian mainland	109
4.13: Selected parametric models for regional fits in Australia (units: m)	112
5.1: Iterative variance component estimation computational scheme.....	139
5.2: Example of estimated variance components at each iteration step	140
6.1: Swiss test network and error bars for $\ell = h - H - N$	145
6.2: Plot of initial covariance matrix for GPS heights (Swiss network)	146
6.3: Plot of initial covariance matrix for geoid heights (Swiss network).....	148
6.4: Plot of initial covariance matrix for orthometric heights (Swiss network).....	149
6.5: Southern BC/AB test network and error bars for $\ell = h - H - N$	152
6.6: Plot of initial covariance matrix for GPS heights (southern BC/AB network).....	153
6.7: Plot of initial covariance matrix for geoid heights (southern BC/AB network) ...	154
6.8: Plot of initial covariance matrix for orthometric heights (southern BC/AB network)	155

6.9: Effect of different a-priori variance factors (Swiss network, $\sigma_{H_o}^2 = 1$)	157
6.10: Effect of different a-priori variance factors (BC/AB network, $\sigma_{N_o}^2 = 1$).....	158
6.11: Effect of different a-priori variance factors (Swiss network, $\sigma_{h_o}^2 = 1$).....	161
6.12: Plot of covariance matrix for EGM96 over the Swiss network area.....	167
6.13: Behaviour of estimated variance components vs. redundancy for the Swiss network (note: all three variance components considered).....	171
6.14: Behaviour of estimated variance components vs. redundancy for the southern BC/AB network (note: all three variance components considered)	172
6.15: Iterations and estimated variance components using the IMINQUE and IAUE methods.....	174
7.1: Optimal combination of ellipsoidal, orthometric and geoid height data.....	183
7.2: Test network area and distribution of the GPS control benchmarks.....	184
7.3: Baseline configuration for the levelling network.....	186
7.4: Relative measurement accuracy for the levelling baselines.....	187
7.5: Baseline configuration for the GPS network.....	189
7.6: Relative measurement accuracy for the height component of GPS baselines	190
7.7: Relative geoidal height accuracy.....	191
7.8: Example of the locations of newly established baselines within the test network area.....	195
7.9: Test network configurations for the multi-data adjustment	196
7.10: Accuracy contribution of the corrector surface model to GPS-levelling.....	198
7.11: Test network coverage area and locations of newly established baselines in the north	201
7.12: Flowchart outlining main components of software program designed to compute the accuracy of orthometric height differences via GNSS-levelling.....	203

7.13: Satellite altimetry measurements and the relationship between various height reference surfaces.....	205
7.13: Example of a mixed network of terrestrial benchmarks and tide gauge stations...	207

LIST OF SYMBOLS

a	semi-major axis of the geocentric reference ellipsoid
$\mathbf{a}^T \mathbf{x}$	correction term for the systematic errors and datum inconsistencies
b	semi-minor axis of the geocentric reference ellipsoid
\mathbf{C}_h	covariance matrix for ellipsoidal heights
\mathbf{C}_H	covariance matrix for orthometric heights
\mathbf{C}_N	covariance matrix for geoidal heights
$\mathbf{C}_{\hat{\mathbf{x}}}$	covariance matrix for the unknown parameters of the corrector surface
$\mathbf{C}_{\Delta h}$	covariance matrix for the ellipsoidal height differences
$\mathbf{C}_{\Delta H}$	covariance matrix for the orthometric height differences
$\mathbf{C}_{\Delta N}$	covariance matrix for the geoidal height differences
$d\lambda$	difference in longitude of a point “ i ” and the mean longitude ($\lambda_i - \bar{\lambda}$)
$d\varphi$	difference in latitude of a point “ i ” and the mean latitude ($\varphi_i - \bar{\varphi}$)
e^2	eccentricity for geocentric reference ellipsoid
h	ellipsoidal height
H	orthometric height
H^{dyn}	dynamic height
H^*	normal height
\tilde{H}	orthometric height at GPS-levelling stations and SST at tide gauge stations
k	number of unknown variance components
ℓ	observations
\mathbf{l}	vector of observations
N	geoidal undulation
N_{GM}	geoid component due to global geopotential model
$N_{\Delta g}$	geoid component due to local gravity anomaly data
N_H	topographic indirect effect

\mathbf{Q}_h	cofactor matrix for ellipsoidal heights
\mathbf{Q}_H	cofactor matrix for orthometric heights
\mathbf{Q}_N	cofactor matrix for geoidal heights
\mathbf{Q}_{SST}	cofactor matrix for SST values
$\mathbf{Q}_{\Delta h}$	cofactor matrix for ellipsoidal height differences
$\mathbf{Q}_{\Delta H}$	cofactor matrix for orthometric height differences
$\mathbf{Q}_{\Delta N}$	cofactor matrix for geoidal height differences
\mathbf{R}	redundancy matrix
\Re	set of real numbers
R^2	coefficient of determination
R_α^2	adjusted coefficient of determination
SST_i	sea surface topography at a tide gauge station “ i ”
$\hat{\mathbf{v}}$	vector of adjusted residuals (total)
$\hat{\mathbf{v}}_h$	vector of adjusted residuals for ellipsoidal heights
$\hat{\mathbf{v}}_H$	vector of adjusted residuals for orthometric heights
$\hat{\mathbf{v}}_N$	vector of adjusted residuals for geoidal heights
$\hat{\mathbf{v}}_{SST}$	vector of adjusted residuals for SST values
$\hat{\mathbf{x}}$	vector of unknown parameters of the corrector surface model
Δg	gravity anomaly
Δh	ellipsoidal height difference
ΔH	orthometric height difference
ΔN	geoidal height difference
$\boldsymbol{\theta}$	vector of unknown variance components
λ	geodetic longitude
$\bar{\lambda}$	mean geodetic longitude for a network of points
μ	mean
σ	standard deviation
σ_{ave}	average standard deviation

$\hat{\sigma}_h^2$	estimated variance component for ellipsoidal heights
$\hat{\sigma}_H^2$	estimated variance component for the orthometric heights
$\hat{\sigma}_N^2$	estimated variance component for the geoid heights
$\hat{\sigma}_{SST}^2$	estimated variance component for SST values
$\sigma_{i_0}^2$	a-priori variance factor for $i = h, H, N$
φ	geodetic latitude
$\bar{\varphi}$	mean geodetic latitude for a network of points

LIST OF ABBREVIATIONS

AHD	Australian Height Datum
ANOVA	analysis of variance
AUE	almost unbiased estimation
BC/AB	British Columbia/Alberta
BIQUE	best invariant quadratic unbiased estimation
BLIMPE	best linear minimum partial bias estimation
BQMBNE	best quadratic minimum bias non-negative estimation
BQUNE	best quadratic unbiased estimation
CGG2000	Canadian Gravimetric Geoid 2000
CGVD28	Canadian geodetic vertical datum 1928
CHAMP	challenging minisatellite payload
CV	covariance
DEM	digital elevation model
EGM96	Earth Geopotential Model 1996
GPS	global positioning system
GNSS	global navigation satellite system
GRACE	gravity recovery and climate experiment
GRS80	global reference surface 1980
GOCE	gravity field and steady-state ocean circulation Explorer
GSD95	geodetic survey division 1995 (geoid model)
IAUE	iterative almost unbiased estimation
IGS	International GPS Service
IMINQUE	iterative minimum norm quadratic unbiased estimation
LEO	low Earth orbit
MINQUE	minimum norm quadratic unbiased estimation
MRE	multiple regression equation
MSL	mean sea level
MSST	mean sea surface topography
NAVD88	North American vertical datum 1988

RMS	root mean square
SLR	satellite laser ranging
SST	sea surface topography
VCE	variance component estimation
VLBI	very long baseline interferometry
WHS	world height system

Chapter 1

Introduction

1.1 Background

Observed elevation differences between points on the Earth's surface are traditionally obtained through spirit-levelling (and/or its variants such as trigonometric, barometric levelling, etc.). For over a century the vertical control needs of the geodetic, cartographic, surveying, oceanographic and engineering communities have been well served by this technique (Vaniček *et al.*, 1980). Due to the nature and practical limitations of spirit-levelling most vertical control points are located in valleys and along roads/railways, which restricts the spatial resolution of control networks and confines the representation of the actual terrain. On the other hand, horizontal control networks have historically been established using triangulation and trilateration, which required that points be situated on hilltops or high points (Davis *et al.*, 1981). As a result, most countries have completely separate networks for horizontal and vertical control with few overlapping points. However, with the advent of satellite-based global positioning systems (GPS, GLONASS, and the upcoming GALILEO) and space-borne/airborne radar systems (satellite altimetry, LIDAR, SAR) the ability to obtain accurate heights at virtually any point on land or at sea has in fact been revolutionized.

The fundamental relationship, to first approximation, that binds the ellipsoidal heights obtained from global navigation satellite system (GNSS) measurements and heights with

respect to a vertical geodetic datum established from spirit-levelling and gravity data is given by (Heiskanen and Moritz, 1967)

$$h - H - N = 0 \quad (1.1)$$

where h is the ellipsoidal height, H is the orthometric height and N is the geoidal undulation obtained from a regional gravimetric geoid model or a global geopotential model. The geometrical relationship between the triplet of height types is also illustrated in Figure 1.1.

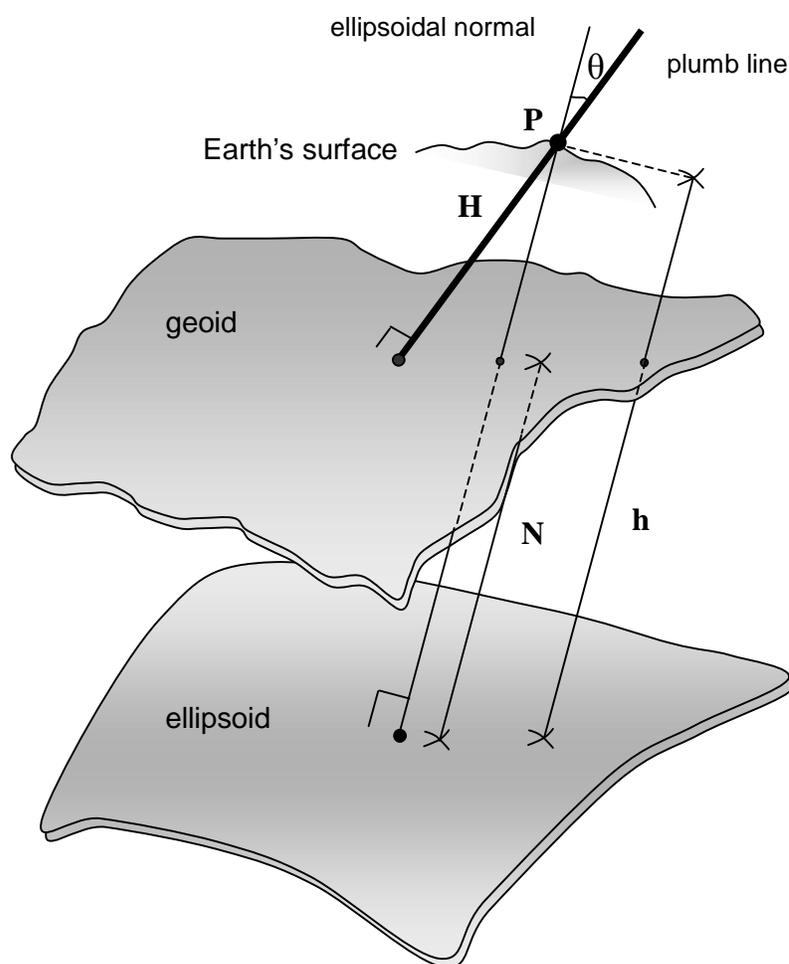


Figure 1.1: Relationship between ellipsoidal, geoid and orthometric heights

For the relative case, where height differences are considered, we simply have

$$\Delta h - \Delta H - \Delta N = 0 \quad (1.2)$$

where Δh , ΔH , and ΔN refer to the ellipsoidal, orthometric and geoid height differences, respectively.

The inherent appeal of this seemingly simple geometrical relationship between the three height types is based on the premise that given any two of the heights, the third can be derived through simple manipulation of Eq. (1.1), or similarly Eq. (1.2) for the relative case. In practice, the implementation of the above equation(s) is more complicated due to numerous factors that cause discrepancies when combining the heterogeneous heights (Rummel and Teunissen, 1989; Kearsley *et al.*, 1993; Schwarz *et al.*, 1987). Some of these factors include, but are not limited to, the following:

- random errors in the derived heights h , H , and N
- datum inconsistencies inherent among the height types, each of which usually refers to a slightly different reference surface
- systematic effects and distortions primarily caused by long-wavelength geoid errors, poorly modelled GPS errors (e.g., tropospheric refraction), and over-constrained levelling network adjustments
- assumptions and theoretical approximations made in processing observed data, such as neglecting sea surface topography effects or river discharge corrections for measured tide gauge values
- approximate or inexact normal/orthometric height corrections
- instability of reference station monuments over time due to geodynamic effects and land subsidence/uplift (de Bruijne *et al.*, 1997; Poutanen *et al.*, 1996)

More details are provided in the sequel.

The major part of the aforementioned discrepancies is usually attributed to the systematic errors and datum inconsistencies. The task of dealing with these effects has been designated to the incorporation of a parametric model (or corrector surface as it is commonly termed) in the combined adjustment of the heights. Numerous studies have been conducted using this approach with several different types of parametric models from a simple bias, a bias and a tilt, higher order polynomials with different base functions (Shretha *et al.*, 1993), finite element models (Jäger, 1999), Fourier series (Haagmans *et al.*, 1998) and collocation-based approaches (Forsberg and Madsen, 1990). It is evident from these studies that the appropriate type of corrector surface model will vary depending on the height network data and therefore a universal model applicable in all cases is not practical.

The unknown parameters for a selected corrector surface model are obtained via a common least-squares adjustment of ellipsoidal, orthometric and geoid height data over a network of co-located GPS-levelling benchmarks. A key issue in this type of common adjustment is the separation of errors among each height type, which in turn allows for the improvement of the stochastic model for the observational noise through the estimation of variance components. There are numerous reasons for conducting such variance component estimation (VCE) investigations. For example, consider the case of optimally refining/testing existing gravimetric geoid models using GPS-levelling height data. Such a comprehensive calibration of geoid error models is essential for numerous applications such as, mean sea level studies (Klees and van Gelderen, 1997), connection of different continental height systems (Rummel, 2000), and establishing vertical control independent of spirit-levelling (Schwarz *et al.*, 1987), to name a few. The latter application is especially important in mountainous terrain and remote areas without existing vertical control (e.g., northern Canada).

An additional important area that will benefit from the implementation of VCE methods is the assessment of the a-posteriori covariance matrix for the height coordinates derived from GPS measurements. Specifically, it will allow for a means to test the accuracy

values for the ellipsoidal heights provided from post-processing software packages, which are often plagued with uncertainty (and usually overly-optimistic). Furthermore, it will allow for the evaluation of the accuracy information provided for orthometric heights obtained from national/regional adjustments of conventional levelling data.

1.2 Thesis Objectives

The main objective of this research is to present a detailed analysis of the optimal combination of heterogeneous height data, with particular emphasis on datum inconsistencies, systematic effects and data accuracy. Specifically, vertical control networks consisting of ellipsoidal, orthometric and geoid height data are investigated. As evidenced from the previous discussion, the combination of these heterogeneous height data is complicated by a number of outstanding issues, including (i) modelling systematic errors and datum inconsistencies, (ii) separation of random errors and estimation of variance components for each height type, and (iii) practical considerations for modernizing vertical control systems.

The selection of the most appropriate parametric model for a particular mixed height data set is complicated and rather arbitrary as it depends on a number of variables such as data distribution, density and quality, which varies for each case. Therefore, it was deemed necessary to focus investigations on the *determination and implementation of valid procedures* for assessing model performance. In this manner, an established general methodology may be implemented that offers the flexibility of being applied with any candidate parametric model and data set.

In addition to a proper parametric model, the suitability of the stochastic model used in the combined network adjustment of the ellipsoidal, orthometric and geoid height data must also be carefully evaluated. This is an important element for the reliable least-squares adjustment of the geodetic data that is often neglected in practical height-related problems. The chosen approach for testing and improving the stochastic model is the

well-known statistical tool of variance component estimation. Many different algorithms for VCE have been investigated with regards to geodetic data analysis, however limited studies have been conducted on its use with the combined height problem as it is posed herein. Therefore, a major emphasis of this research is placed on the determination of the most appropriate VCE algorithm to be applied in this case and its evaluation using real height data and the corresponding data covariance matrices. An in depth analysis of the effects of various factors such as correlation among heights of the same type, provisions for computing non-negative variance factors and the connection (if any) between the proper modelling of systematic errors and datum inconsistencies with the estimated variance components must also be studied.

Ultimately, one of the main motivations for this research is embedded in the need to introduce *modern* tools and techniques in the establishment of vertical control. The manipulation of Eq. (1.1) or (1.2), such that orthometric heights (or height differences, respectively) are obtained using ellipsoidal and geoid height data is called GNSS-levelling and is a procedure that is commonly used in practice and will undoubtedly dominate the future of vertical control. Part of this research is aimed at bringing to the forefront some of the key issues that affect the achievable accuracy level of GNSS-levelling. Considerations for height-related information expected to be more readily available in the near future (such as sea surface topography and land uplift models) must also be made.

An important aspect in all of the investigations mentioned is to test the developed procedures using real data sets that are representative of a variety of practical vertical networks. Overall, the analysis of the optimal combination of heterogeneous height data conducted herein, with particular emphasis on datum inconsistencies, systematic effects and data accuracy, will provide valuable insight and practical results to be used for a variety of height-related applications.

1.3 Thesis Outline

The analysis and results of this research are presented in chapters 2 through 8. An outline of the essential structure of this thesis is discussed below.

In chapter 2, the necessary background information regarding the height data types used in this research is presented. The discussion focuses on the main error sources affecting the computation of geoid, orthometric and ellipsoidal heights. The final section addresses the issue of why to combine the height data. Although the applications for the optimal combination of the heterogeneous height types are innumerable, a short-list of the most prevalent geodetic applications is discussed in this chapter. In particular, the concept of regional vertical datums and modernization issues are considered. Attention is also given to the goal of unifying existing regional datums for a global vertical datum. This is a major topic on its own that will rely heavily on the combined height adjustment solution. Furthermore, the process of GNSS-levelling and the refinement/testing of gravimetric geoid models using GNSS-levelling benchmark data are also described.

In chapter 3, the combined least-squares adjustment scheme implemented throughout this work is described in detail. The formulation is provided for the case where absolute height data values are available, as in Eq. (1.1), and for the relative case, where height differences for baselines are available, as in Eq. (1.2). The second part of the chapter focuses on modelling systematic effects using a parametric corrector surface model. The procedure developed for assessing the parametric model performance is also described in detail.

In chapter 4, the methodology outlined in the previous chapter is implemented using real data sets from three different test networks, namely Switzerland, parts of Canada and Australia, in order to test its effectiveness and to demonstrate the adaptive nature of the parametric model selection process. Each test network poses a different set of challenges for selecting and testing candidate parametric models.

In chapter 5, a review of numerous variance component estimation schemes applied to geodetic applications over the years is provided. In particular, several variance component estimation algorithms are scrutinized for use in the combined height network adjustment problem. The main reasons for selecting the iterative minimum norm quadratic unbiased estimation (IMINQUE) and iterative almost unbiased estimation (IAUE) schemes are provided as well as the detailed algorithms as they pertain to the problem at hand.

In chapter 6, the results of numerous case studies are described, whereby some of the key issues related to the implementation of the IMINQUE (and IAUE) method are analyzed, demonstrating its use for practical height-related applications. Specifically, tests are conducted on changing a-priori covariance matrices for the height data, obtaining non-negative variances, effects of correlations among heights of the same type and the role of the parametric model type on the final estimated variance components. The VCE case studies are conducted using covariance information provided for the Swiss and the southern British Columbia/Alberta test networks.

In chapter 7, some of the practical considerations for modernizing vertical control are presented. This chapter essentially focuses on three main issues that have not been discussed thus far, namely how to evaluate the achievable accuracy of GNSS-levelling, how to incorporate height data at sea (or more appropriately at the coasts), and disseminating parametric model information to users. The complete procedure outlined in this thesis for the optimal combination of geoid, orthometric and ellipsoidal heights is modified in order to incorporate sea surface topography information at tide gauge stations. Finally, a brief discussion at the end of the chapter outlines some of the key concepts for providing users of GNSS with the proper information to transform ellipsoidal heights to heights with respect to a local vertical datum.

Chapter 8 summarizes the main conclusions of this research. Finally, recommendations for future work in this area are also provided.

Chapter 2

Heights, Vertical Datums and GNSS-Levelling

The purpose of this chapter is to provide the necessary background information regarding the type of data and terminology used throughout this thesis. In particular, focus will be placed on describing the major error sources that affect the geoidal undulations, orthometric and ellipsoidal heights. The discussion will also provide insight into the problems and challenges encountered when attempting the optimal combination of these heterogeneous height data. The final section of this chapter addresses the question of, "*Why combine geoid, orthometric and ellipsoidal heights?*" This was briefly introduced in chapter 1 and will be elaborated on herein.

2.1 Geoid heights

The geoid height (or geoidal undulation) can be defined as the separation of the reference ellipsoid with the geoid surface measured along the ellipsoidal normal (see Figure 1.1). The classical Gauss-Listing definition of the geoid is given as an equipotential surface of the Earth's gravity field that coincides with the mean sea level. Today, it is well known that this is not a strictly correct definition as mean sea level departs from the equipotential surface by up to two metres due to various oceanographic phenomena, such as variable temperature, salinity, instantaneous sea surface topography, to name a few

(Vaníček and Krakiwsky, 1986). Over the years, many methods have been developed for determining the geoid, including astronomical levelling, gravimetric geoid determination using Stokes' or Molodensky's approach, and optimal operational schemes for combining heterogeneous data such as least-squares collocation (see Heiskanen and Moritz, 1967 and Moritz, 1980 for more details on these formulations). The focus of this section is to provide a general overview of the main errors affecting the determination of the geoid heights in practice. Details on formulations and techniques for precise geoid determination are numerous and can be found in Vaníček and Christou (1994) as well as the aforementioned sources and will not be dwelled on herein.

One practical procedure for regional geoid determination, which provides insight into the main errors inhibiting the accuracy of the computed geoidal height values (N) or relative geoidal heights (ΔN) is the classic "remove-compute-restore" technique (see, e.g., Rapp and Rummel, 1975; Mainville *et al.*, 1992; Sideris *et al.*, 1992). The underlying procedure can be summarized as follows:

- 1) *Remove* a long-wavelength gravity anomaly field (determined by a global spherical harmonic model) from terrain-reduced gravity anomalies that are computed from local surface gravity measurements and digital elevation data.
- 2) *Compute* "residual co-geoid undulations" $N_{\Delta g}$. This can be done by a spherical Fourier representation of Stokes' convolution integral using the residual gravity anomalies (see, e.g., Haagmans *et al.*, 1993).
- 3) *Restore* a long-wavelength geoid undulation field N_{GM} (determined by a global spherical harmonic model) to the residual co-geoid undulations, and add a topographic indirect effect term N_H (computed from digital elevation data) to form the final geoidal undulations.

The above three steps can be combined in a single formula as follows:

$$N = N_{GM} + N_{\Delta g} + N_H \quad (2.1)$$

The computation of the long-wavelength geoid component N_{GM} is usually made on a grid (e.g., $5' \times 5'$), within the appropriate geographical boundaries for the region of interest. Currently, the most widely used global geopotential model is EGM96 (Lemoine *et al.*, 1998), complete to degree and order 360. The coefficients of the global geopotential models are determined from measurements of satellite orbits, satellite altimetry and gravity anomalies (Pavlis, 1988). In spherical approximation, its contribution is computed according to the following formula (Heiskanen and Moritz, 1967):

$$N_{GM}(\varphi, \lambda) = R \sum_{n=2}^{360} \sum_{m=0}^n (\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\sin \varphi) \quad (2.2)$$

where \bar{P}_{nm} are fully normalized Legendre functions, \bar{C}_{nm} and \bar{S}_{nm} are the fully normalized unitless coefficients of the geopotential model (from which the contribution of a normal gravity field, based on the GRS80 geodetic reference system, has been subtracted), and R is the mean radius of the Earth.

The medium-wavelength contributions to the total geoid heights can be computed from the available local gravity anomaly data according to Stokes' formula (Heiskanen and Moritz, 1967)

$$N_{\Delta g}(\varphi_P, \lambda_P) = \frac{R}{4\pi\gamma} \int_{\lambda_Q} \int_{\varphi_Q} \Delta g(\varphi_Q, \lambda_Q) S(\psi_{PQ}) \cos \varphi_Q d\varphi_Q d\lambda_Q \quad (2.3)$$

where $S(\psi_{PQ})$ is the Stokes' function, ψ_{PQ} is the spherical distance between the computation point P and the running point Q , γ is normal gravity, and Δg are the local

gravity anomaly data. In mountainous regions, such as Canada, Δg are residual Faye anomalies which, when Helmert's second condensation method is used for the terrain effects, are obtained from

$$\Delta g = \Delta g_{FA} + c - \Delta g_{GM} \quad (2.4)$$

In the last equation, Δg_{FA} are the free-air anomalies, c is the classic *terrain correction* term (Heiskanen and Moritz, 1967; Mainville *et al.*, 1994), and Δg_{GM} is the removed long-wavelength contribution of the global geopotential model, which is computed from the expression,

$$\Delta g_{GM} = G \sum_{n=2}^{360} (n-1) \sum_{m=0}^n (\bar{C}_{nm} \cos m\lambda + \bar{S}_{nm} \sin m\lambda) \bar{P}_{nm}(\sin \varphi) \quad (2.5)$$

Detailed discussions of the treatment and computation of all the necessary reductions (atmospheric, free-air gradient, downward continuation, terrain reduction) applied to the original gravity data can be found in Véronneau (1997), Mainville *et al.* (1994), and Vaníček and Christou (1994).

The shorter wavelength information for the regional geoid model is usually obtained from the computation of the indirect effect term N_H , induced by using Helmert's second condensation method for the gravity data reduction on the geoid surface. There are numerous ways of modelling the effects of the topography, however, in this discussion Helmert's second condensation method has been used to illustrate the process. In general, the formulation of the topographic indirect effect on the geoid (according to Helmert's second condensation method) is made in terms of a Taylor series expansion from which only the first three terms are usually considered. Wichiencharoen (1982) should be consulted for all the detailed formulas related to the three terms. The computation of the indirect effect involves height data obtained from a digital elevation model (DEM),

which covers the regional area of interest. In relatively flat areas, the indirect effect is sometimes neglected, which introduces a small error depending on the terrain. The process described by Eq. (2.1) is illustrated in Figure 2.1.

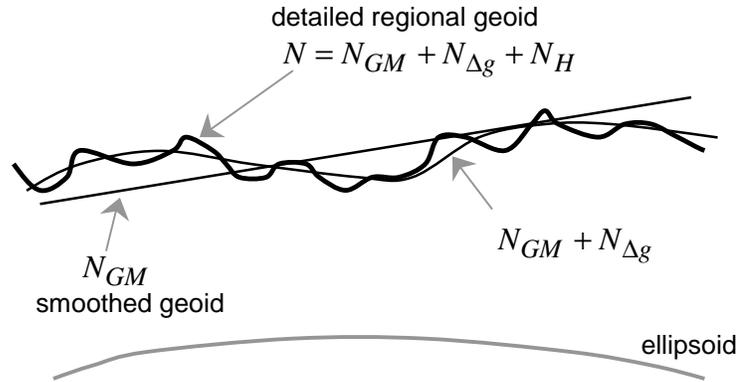


Figure 2.1: Computation of regional geoid models using heterogeneous data (according to Schwarz *et al.*, 1987)

Following this formulation, it is evident that the accuracy of the computed geoid undulations depends on the accuracy of the three components in Eq. (2.1), namely N_{GM} , $N_{\Delta g}$ and N_H (Schwarz *et al.*, 1987 and Sideris, 1994).

Errors due to N_{GM}

The global geopotential model not only contributes to the long wavelength geoid information but also introduces long-wavelength errors that originate from insufficient satellite tracking data, lack of terrestrial gravity data and systematic errors in satellite altimetry. The two main types of errors can be categorized as either omission or commission errors. Omission errors occur from the truncation of the spherical harmonic series expansion (Eq. 2.2), which is available in practice ($n_{\max} < \infty$). The error from these neglected terms can be computed as follows for the absolute geoid heights (Jekeli, 1979 and de Min, 1990):

$$\|\delta N_o\|^2 = \left(\frac{R}{2\gamma}\right)^2 \sum_{n=n_{\max}+1}^{\infty} Q_n^2(\psi) c_n \quad (2.6)$$

and the error for the geoid height differences between two points separated by a distance ψ_{PQ} is given by:

$$\|\delta \Delta N_o\|^2 = 2 \left(\frac{R}{2\gamma}\right)^2 \sum_{n=n_{\max}+1}^{\infty} Q_n^2(\psi) c_n (1 - P_n(\cos \psi_{PQ})) \quad (2.7)$$

where,

$$Q_n(\psi) = \int_{\psi_o}^{\pi} S(\psi) P_n(\cos \psi) \sin \psi d\psi \quad (2.8)$$

where n_{\max} is the maximum degree to which the geopotential model coefficients are used (non-truncated terms) and c_n denotes the gravity anomaly degree variances computed by (Heiskanen and Moritz, 1967):

$$c_n = \gamma^2 (n-1)^2 \sum_{m=0}^n (\bar{C}_{nm}^2 + \bar{S}_{nm}^2) \quad (2.9)$$

The other major contributing error type is due to the noise in the coefficients themselves and is termed commission errors. This can be computed as follows:

$$\|\delta N_c\|^2 = \left(\frac{R}{2\gamma}\right)^2 \sum_{n=2}^{n_{\max}} Q_n^2(\psi) \sigma_n^2 \quad (2.10)$$

where σ_n^2 denotes the error degree variances, see de Min (1990) for detailed formulations. An alternative method for estimating the commission errors for the spherical harmonic coefficients is to use least-squares collocation and is aptly described in Tscherning (2001). As the maximum degree n_{\max} , of the spherical harmonic expansion increases, so does the commission error, while the omission error decreases. Therefore, it is important to strike a balance between the various errors. In general, formal error models should include both omission and commission error types in order to provide a realistic measure of the accuracy of the geoid heights computed from the global geopotential model.

Errors due to $N_{\Delta g}$

The errors contributing to the $N_{\Delta g}$ component are due to the data coverage, density and accuracy of the local gravity data. An efficient and practical algorithm for computing the effect of the errors in the mean gravity anomalies (assumed to be available on an $M \times N$ equi-angular spherical grid) on the computed geoidal heights is given by the following formula (She, 1993):

$$\sigma_{N_{\Delta g}}^2 = \left(\frac{R\Delta\varphi\Delta\lambda}{4\pi\gamma} \right)^2 \sum_{\varphi_Q=\varphi_1}^{\varphi_M} \sum_{\lambda_Q=\lambda_1}^{\lambda_N} (S(\psi_{PQ}) \cos \varphi_Q)^2 \sigma_{\Delta g(\varphi_Q, \lambda_Q)}^2 \quad (2.11)$$

where it is assumed that the errors of the gravity anomaly data are uncorrelated and the associated a-priori variances, denoted by $\sigma_{\Delta g(\varphi_Q, \lambda_Q)}^2$, are known. $\Delta\varphi$ and $\Delta\lambda$ refer to the grid spacing in latitude and longitude, respectively and M , N are the number of parallels and meridians in the grid, respectively. This formulation is provided as it can be efficiently evaluated using 1D FFT (see, She, 1993 for details).

Obviously, higher accuracy is implied by accurate Δg values distributed evenly over the entire area with sufficient spacing, however there are some systematic errors, which influence the quality of the gravity anomalies. Following Heck (1990), the major sources affecting terrestrial gravity anomalies, which cause both systematic and random errors in absolute and relative geoidal heights include:

- inconsistencies in the gravity datum(s)
- inconsistencies in the vertical datum(s)
- inconsistencies in the horizontal datum(s)
- inconsistencies in the types of height systems
- approximation errors due to the use of a simplified free-air reduction formula

Since gravity anomalies depend on the horizontal, vertical and gravity network datums, any inconsistencies of these local datums will introduce biases into the computed (free-air) gravity anomaly values. The wavelengths of the datum inconsistency effects range from hundreds to thousands of kilometres, which means that the medium-to-low frequency spectral components of the gravity field (and thus the geoid) are affected. More details and practical evaluations of most of the effects described can be found in Heck (1990).

Other influences besides the datum inconsistencies and types of height systems include approximation errors due to the use of simplified reduction formulas for computed free-air gravity anomalies (i.e., neglecting non-linear terms) and limited capsize in Stokes' integration (see Forsberg and Featherstone, 1998). The truncation error caused by the limited area of the integration of the terrestrial gravity anomalies to a spherical cap can be reduced by a suitable modification of Stokes' kernel (Véronneau, 2002). In Li and Sideris (1994), the error caused in the geoid heights by the use of different approximations of Stokes' kernel were investigated for parts of Canada and found to be a maximum at the metre-level.

Errors due to N_H

The shorter wavelength errors in the geoid heights are introduced through the spacing and quality of the digital elevation model used in the computation of N_H . Improper modelling of the terrain is especially significant in mountainous regions (e.g., Rocky mountain area in western Canada), where terrain effects contribute significantly to the final geoid model in addition to errors due to the approximate values of the vertical gravity gradient (Forsberg, 1994; Sideris and Forsberg, 1991; Sideris and Li, 1992). Improvements in geoid models according to the computation of N_H , will be seen through the use of higher resolution (and accuracy) DEMs, especially in mountainous regions.

Comments

Overall, the most significant error contribution in the total error budget for the computed geoid heights is due to the global geopotential model used as a reference in the *remove-compute-restore* technique. The situation is expected to drastically improve with unprecedented global satellite-based gravity coverage promised from low earth orbiting (LEO) missions such as CHAMP, GOCE and GRACE. Specifically, the first two missions, CHAMP and GOCE, will considerably improve the long and medium wavelength information, while GRACE will allow for time variations in the gravity field to be accounted for (ESA, 1999). It is expected that the 'new' global geopotential model derived by incorporating these satellite data from the aforementioned missions will provide a total error (commission and omission) of $\pm 15\text{ cm}$ (Pavlis and Kenyon, 2002). This is a factor of 10 improvement compared to the current EGM96 model (Tscherning *et al.*, 2000). Such improvements will start to emerge from the measurements made by the Gravity Field and Steady-State Ocean Circulation Explorer (GOCE), which will fly at a low altitude ($\sim 250\text{ km}$) and high precision gradiometers on-board lead to improved higher degree harmonics. In addition, the precise orbit determination and high-low satellite tracking made possible by on-board GPS/GLONASS receivers will improve the lower degree harmonics (ESA, 1999 and Tscherning *et al.*, 2000). Estimates of the

expected commission and omission errors, in terms of standard deviation after GOCE for various maximum degrees of expansion are provided in the Table 2.1. All values are given for the geoid height difference between two arbitrary points approximately 100 km apart.

Table 2.1: Expected commission and omission errors after GOCE
(according to Rummel, 2000)

Degree of GM model	Commission standard error (cm)	Omission standard error (cm)
150	0.07	38
200	0.2	28
250	0.8	23
300	1.3	18

The level of achievable accuracy and global homogeneous coverage provided by GOCE will not only improve the global geopotential model, but also contribute to the establishment of a global unified vertical datum (Rummel, 2000; see also section 2.4.2). Although these satellite mission data will produce new global geopotential models with a more homogeneous error spectrum, thus reducing systematic effects on geoid heights, terrestrial data is still needed for medium frequency improvements. Measurements on the surface of the Earth or airborne measurements typically provide better short wavelength resolution than satellite-based measurements (see Li, 2000 for more details).

Regional geoid model accuracy also varies depending on the computational methodology (assumptions used) and available data in the region of interest. In areas where regional models exist, they should be used as they are more accurate compared to global models. However, many parts of the globe do not have access to a regional geoid model usually due to lack of data. In these cases, one may resort to applying global geopotential model values. An alternative approach to determining discrete geoid height values is the geometric approach, which utilizes the relationship between GPS-derived ellipsoidal

heights and orthometric heights with respect to a local vertical datum to provide discrete point values for N . In this case, a number of systematic datum issues enter the problem, which emphasizes the requirement that the same ellipsoidal reference surface is used for both the global geopotential model and residual gravity anomaly computations. Discussions in section 2.4.4 will provide more insight into this practical problem.

2.2 Orthometric heights

Height differences between points on the Earth's surface have traditionally been obtained through terrestrial levelling methods, such as spirit-levelling (and/or barometric levelling, trigonometric levelling, etc). For over a century, the vertical control needs of the geodetic, cartographic, oceanographic, surveying and engineering communities have been well served by this system. Although costly and labourious, spirit-levelling is an inherently precise measurement system whose procedural and instrumental requirements have evolved to limit possible systematic errors. Associated random errors in levelling originate from several sources, such as refractive scintillation or 'heat waves', refraction variation between readings, vibrations of instrument due to wind blowing, and movement of rod or non-verticality of rod caused by wind, terrain and unsteadiness of surveyor, to name a few (see Gareau, 1986 for details). These errors are generally dealt with through redundancy and minimized in the least-squares adjustment process (Vaniček *et al.*, 1980). However, it should be realized that national networks of vertical control established in this way involve large samples of measurements collected under inhomogeneous conditions, such as variable terrain, environments, and instruments, with different observers and over different durations. This results in a number of errors/corrections that must be made to the measurements (see Davis *et al.*, 1981, pp.118-187 for details).

The problem with using only the elevation differences obtained from spirit-levelling for height-related applications is that the results are not unique as they depend on the path taken from one point to the other (due to non-parallelism of the equipotential surfaces).

Thus, a number of different height systems can be defined, which use the measurements of vertical increments between equipotential surfaces along a path from spirit-levelling (dn) and measurements of gravity (g), as given by:

$$C_P = \int_{P_o}^P g \cdot dn \quad (2.12)$$

where C_P is the geopotential number and represents the difference in potential between the constant value at the geoid, W_o , and the potential at the point, P , on the surface, W_P , as follows:

$$C_P = W_o - W_P \quad (2.13)$$

All points have a unique geopotential number with respect to the geoid and it can be scaled by gravity in order to obtain a height coordinate with units of length, as we have become accustomed to using for describing heights. Depending on the type of 'gravity' value used to scale the geopotential number, different types of heights can be derived. In this section, the focus will be placed on describing orthometric heights as they will be used throughout this thesis for numerical computations. However, two other common height systems (dynamic and normal heights) will also be briefly described in order to provide a basis for discussion and comparison.

Orthometric heights are defined as the distance along the plumb line between the geoid and the point of interest on the surface of the Earth. Eq. (2.12) along the plumb line becomes

$$C_P = \int_{P_o}^P g \cdot dH \quad (2.14)$$

where dH is the differential element along the plumb line between the geoid and the point on the Earth's surface. Eq. (2.14) can be re-written in a more practical form as follows (Heiskanen and Moritz, 1967):

$$H_P = \frac{C_P}{\bar{g}_P} \quad (2.15)$$

which provides the orthometric height of the point P on the Earth's surface denoted by H_P (see Figure 2.2).

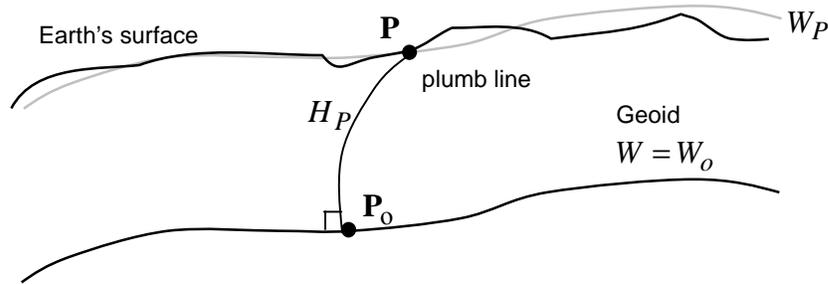


Figure 2.2: Orthometric height of a point

The mean value of gravity along the plumb line is given by

$$\bar{g}_P = \frac{1}{H_P} \int_{P_0}^P g \cdot dH \quad (2.16)$$

The exact computation of \bar{g}_P would require complete knowledge of the mass density of the crust, which is not practically available. Therefore, approximations must be made in order to obtain the corresponding orthometric height values. It is evident from Eqs. (2.15

and 2.16) that one may obtain any number of orthometric height systems depending on the selected value for \bar{g}_P (Bomford, 1971). This is an important point to understand when dealing with orthometric heights in practice as the providers of the data may not explicitly define the fundamental reductions and approximations made in computing \bar{g}_P . Therefore, one must be cautious especially when combining different types of height data or when working with different national databases of orthometric heights.

One of the most common orthometric height systems is Helmert heights, which are based on the Poincaré-Prey reduction model. The approximate value for gravity inside the crust is obtained in three steps as follows:

1. remove the Bouguer plate of uniform density
2. apply free-air downward continuation using the normal gradient of gravity
3. restore the Bouguer plate

The mean value of gravity along the plumb line is computed from the average of the gravity at the endpoints as follows (Heiskanen and Moritz, 1967):

$$\bar{g}_P = g_P - 2\pi k \rho H_P + \frac{1}{2} \frac{\partial \gamma}{\partial h} H_P \quad (2.17)$$

where k is Newton's gravitational constant ($66.7 \times 10^{-9} \text{ cm}^3 \text{ g}^{-1} \text{ sec}^{-2}$). By substituting nominal values for the density, $\rho = 2.67 \text{ g/cm}^3$, and the normal gravity gradient, $\frac{\partial \gamma}{\partial h} = 0.3086 \text{ mGal/m}$, we obtain the simplified expression given below:

$$\bar{g}_P = g_P + 0.0424 H_P \quad (2.18)$$

where the units associated with the 'factor' of 0.0424 are mGal/m and the height H_P is given in metres. When Eq. (2.18) is substituted into Eq. (2.15), one obtains the Helmert

heights often used in practice for numerical computations of heights above the geoid as follows:

$$H_P = \frac{C_P}{g_P + 0.0424 H_P} \quad (2.19)$$

It should be noted that the computation of the mean gravity along the plumb line in this manner requires H_P , therefore Eq. (2.18) is usually solved through iteration or by solving the quadratic represented in Eq. (2.15) with Eq. (2.17) and neglecting terms beyond second-order (see Jekeli, 2000 for full formulations). Additional approximations for the computation of the mean gravity may be made in practice, which also introduces some error in the computation of the orthometric heights (see description of real datasets in chapters 4 and 6).

Throughout this discussion, orthometric heights have been geometrically defined as the distance along the plumb line from the geoid to the corresponding point on the Earth's surface (also referred to as Pizzetti's projection; see Heiskanen and Moritz (1967), p. 180). In practice and for all numerical computations/transformations, a simplification is made whereby the orthometric height is described as the distance along the ellipsoidal normal, referred to as Helmert's projection, as shown in Figure 2.3. The error caused by neglecting the difference in the length of the curved plumb line and the ellipsoidal normal (or the difference between Pizzetti's and Helmert's projection) is negligible for all topographic heights on the Earth's surface. This is demonstrated explicitly in Figure 2.3 where θ is the deflection of the vertical whose practical values range from the arcsecond-level to a maximum of one arcminute. The effect on the height value can be approximated by $\delta h \cong h \sin \theta \tan \theta$, which takes on a maximum value at the sub-millimetre level (i.e. assume $\theta = 1'$, $h = 10,000m$ then $\delta h = 0.8mm$). Therefore, for all theoretical formulations and numerical tests conducted throughout the sequel the classical Helmert projection will be used.

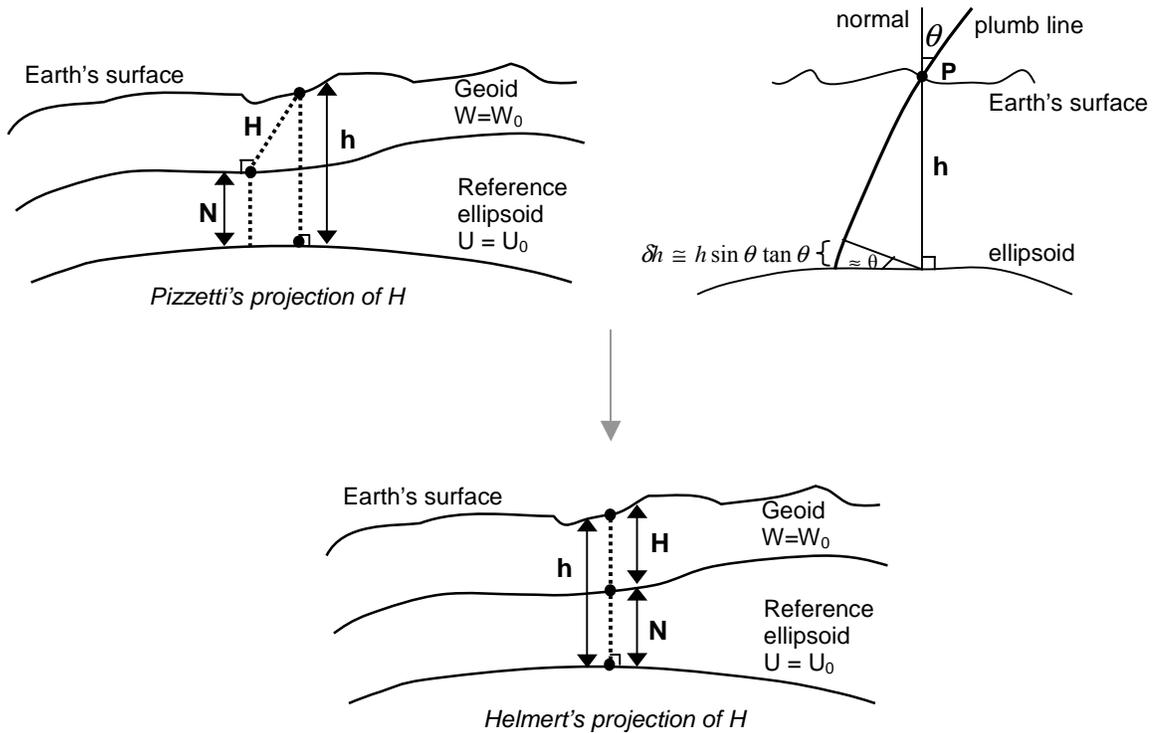


Figure 2.3: Approximation in orthometric heights by neglecting the difference in lengths between the curved plumb line and the ellipsoidal normal

Normal and dynamic heights

Although not directly used in this work, it is important to provide a brief overview of the normal height system as it is the basis of heights in many regions worldwide. If the value for gravity in Eq. (2.15) is replaced by the mean normal gravity along the plumb line, $\bar{\gamma}_P$, then we obtain normal heights denoted by H_P^* and computed via

$$H_P^* = \frac{C_P}{\bar{\gamma}_P} \quad (2.20)$$

Both orthometric and normal heights have a clear geometrical interpretation, with the key difference being that the normal heights refer to the telluroid. Computationally, there is

no need to make approximations for the density of the Earth's crust in order to compute $\bar{\gamma}_P$, and therefore H_P^* can be computed exactly. The reader is referred to Heiskanen and Moritz (1967) for more details. The equivalent form of Eq. (1.1), which gives the geometrical relationship between the ellipsoidal height, h , normal height, H^* , and height anomaly, ζ , is given by

$$h - H^* - \zeta = 0 \quad (2.21)$$

In this case, the geoid surface is replaced by the quasi-geoid, which is closely related to the geoid and in fact coincides with the geoid in the open seas. An important distinction between the geoid and the quasi-geoid is that the latter is not considered to be an equipotential surface of the Earth's gravity field (see Heiskanen and Moritz, 1967 for more details).

The final height system described is the dynamic heights defined as follows:

$$H_P^{dyn} = \frac{C_P}{\gamma_o} \quad (2.22)$$

where the value for gravity in Eq. (2.15) is replaced by γ_o , representing the normal gravity for a fixed latitude, usually taken to be 45° . Unlike the orthometric (Eq. 2.15) and normal heights (Eq. 2.20), dynamic heights have no geometrical interpretation and are merely a conversion of the geopotential number to units of length. Thus, although these heights are not geometrically meaningful like orthometric and normal heights, dynamic heights are the only type from the three, which are *physically meaningful*. That is, they will indicate the direction of water flow. In contrast, two points with identical orthometric heights generally lie on different equipotential surfaces and water will flow from one point to the other. The selection of the type of height to use in the combined height

problem (i.e., orthometric versus normal) depends on the available national databases, which traditionally prefer one system over the other. Essentially, each of the height systems described provides a unique definition of the vertical coordinate of a point on the Earth's surface based on levelling and gravity information. Since they are linked through the geopotential number, it is theoretically possible to convert between any of the three height types.

Comments

Although levelling measurements are very precise (i.e., at the mm-level depending on the order or class of levelling), it is often the regional or national network adjustments of vertical control points that leads to the greatest source of (systematic) error. If the vertical datum (see section 2.4.1) of a height network is based on fixing a single point (e.g., a tide gauge station), then the adjusted orthometric heights will contain a constant bias over the entire network area. The situation is somewhat complicated when an over-constrained network adjustment is performed (i.e., fixing more than one tide gauge station), which introduces distortions throughout the network.

2.3 Ellipsoidal heights

The physical shape of the Earth can be approximated by the mathematical surface of a rotational ellipsoid defined by a semi-major axis, a , and flattening, f . All other ellipsoidal shape and size defining quantities can be subsequently derived from these parameters (semi-minor axis b , eccentricity e^2 , and the curvature in the prime vertical R_N). Because of its smooth well-defined surface, the ellipsoid offers a convenient reference surface for mathematical operations and is widely used for horizontal coordinates (Seeber, 1993). The geodetic latitude φ , and longitude λ , are defined in Figure 2.4, where it is assumed that the centre of the ellipsoid coincides with the Earth's centre of mass, its minor axis is aligned with the Earth's reference pole and the p-axis is the intersection of the meridian plane with the equatorial plane.

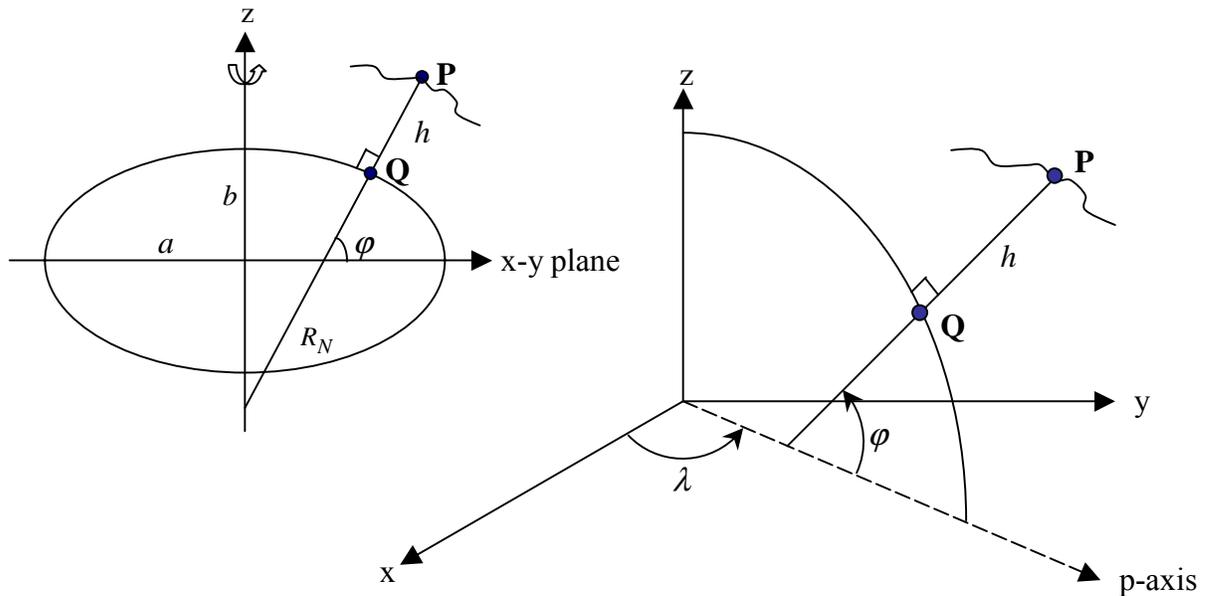


Figure 2.4: Reference ellipsoid and geodetic coordinates

The straight-line distance between a point **P** on the surface of the Earth and its projection along the ellipsoidal normal onto the ellipsoid, denoted by **Q**, is the ellipsoidal height h . The location of the point **P** can also be defined in terms of Cartesian coordinates (x, y, z) , which has greatly benefited from the advent of satellite-based methods, such as GPS. Using GPS (or another global navigation satellite system), three-dimensional coordinates of a satellite-signal receiver can be determined within the same reference frame used to determine the coordinates of the satellites.

Curvilinear geodetic coordinates (φ, λ, h) offer an intuitive appeal that is lacking for Cartesian coordinates and are therefore preferred by users for describing locations on the surface of the Earth. The transformation of geodetic coordinates to a Cartesian coordinate system where the origin coincides with the centre of the ellipsoid and the z -axis aligns

with the semi-minor axis of the ellipsoid, is given by the closed form equations as follows

$(\varphi, \lambda, h) \rightarrow (x, y, z)$:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} (R_N + h) \cos \varphi \cos \lambda \\ (R_N + h) \cos \varphi \sin \lambda \\ [(1 - e^2)R_N + h] \sin \varphi \end{bmatrix} \quad (2.23)$$

where, the curvature in the prime vertical is computed from

$$R_N = \frac{a}{\sqrt{1 - e^2 \sin^2 \varphi}} \quad (2.24)$$

The transformation from the Cartesian coordinates to the geodetic coordinates, $(x, y, z) \rightarrow (\varphi, \lambda, h)$, is not as simple to compute as there is no linear relationship. Therefore, the computation involves the analytical solution of a complicated fourth order equation or an iterative solution using the formulations below.

$$\begin{aligned} \tan \lambda &= \frac{y_p}{x_p} \\ h &= \frac{p}{\cos \varphi} - R_N \\ \varphi &= \arctan \left(\frac{z_p}{p} \left(1 - e^2 \frac{R_N}{R_N + h} \right)^{-1} \right) \end{aligned} \quad (2.25)$$

where $p = \sqrt{x_p^2 + y_p^2}$. This method is preferred and often implemented in GPS processing software packages. As evidenced from the equations, the process requires initial estimates for h and φ and an appropriate convergence criterion, which can be set

for h or R_N at the resolution level of the initial Cartesian coordinates. Details on these formulations and additional transformations between global geodetic systems and local geodetic systems can be found in numerous texts (Seeber, 1993). In particular, Soler (1998) and Soler and Marshall (2002) describe the coordinate transformations encountered when dealing with GPS data. In general, it is important to be aware of the type of ellipsoidal reference surface implied by the data and its relation to the geoid surface and the local vertical datum, which may vary from region to region.

Using today's available technology and techniques, ellipsoidal heights can be obtained from a number of difference systems, such as very long baseline interferometry (VLBI), satellite laser ranging (SLR), and navigation based systems such as DORIS, GPS, and GLONASS. In the near future it is expected that the European contribution to global navigation satellite systems, GALILEO (ESA, 2002), will be used. Furthermore, satellite altimetry measurements are used to obtain ellipsoidal heights over the oceans, which cover more than 70% of the Earth's surface. Thus, although the most popular method in use today is GPS, the alternatives are set to broaden in the near future. This being said, all new global satellite-based navigation systems will benefit greatly from the experience gained by researchers and users working with GPS. In fact, many of the challenges and error sources that affect the quality of the positioning coordinates will still have to be dealt with. Therefore, it is appropriate to discuss some of the main error sources affecting the determination of ellipsoidal heights using GPS, as it is the main tool used to obtain ellipsoidal height data for all of the numerical tests used throughout this dissertation. Comprehensive overviews of the fundamental concepts, measuring, and processing procedures for GPS can be found in many textbooks such as Hofmann-Wellenhof *et al.* (1992), Parkinson and Spilker (1996a/b), and Kaplan (1996) and will not be dwelled on herein.

The computation of ellipsoidal heights using GPS measurements is in general more challenging than estimating horizontal coordinates. Although, the common error sources affecting the quality of the positions affect all three coordinates, there are a few key

differences, which result in poorer height values (by approximately two to three times), namely (Rothacher, 2001):

- satellite geometry/configuration can only be observed in one hemisphere above the horizon (i.e., there will never be satellites 'below' the receiver antenna)
- need to estimate receiver clock corrections at every epoch
- estimation of tropospheric zenith delay parameters (every hour)

The most limiting factor remains the very high correlation of receiver clock corrections and tropospheric zenith delay parameters with the ellipsoidal height. The estimation of these effects significantly hinders the achievable accuracy of the height component, even in the absence of other errors and biases (Santerre, 1991). A suggested means for partially decorrelating the height from the receiver clock and tropospheric delay is to take advantage of the zenith dependence and process GPS data at low elevation cut-off angles (Rothacher, 2001). Figure 2.5 (originally published in Rothacher, 2001) depicts the zenith dependence of the height $\delta\rho_h(z)$, receiver clock $\delta\rho_{clk}(z)$ and tropospheric zenith delay $\delta\rho_{trop}(z)$ parameters, where c is the speed of light and δt_R is the receiver clock correction. Of course, lowering the elevation cut-off introduces other problems with data processing as the noise level increases significantly. Therefore, due to the nature of the satellite configuration and the need to estimate receiver clocks (even differences), the height component will always be less accurate than the horizontal positions (*ibid.*).

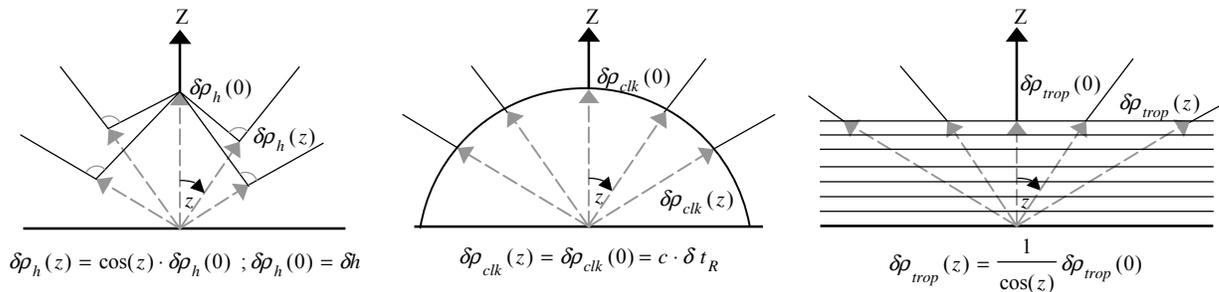


Figure 2.5: Zenith dependence of station height, receiver clock and tropospheric delay parameters (according to Rothacher, 2001)

The errors affecting GPS measurements originate from three sources, namely satellite errors, signal propagation errors and receiver errors (see Figure 2.6). All three types of error sources affect the quality of the estimated ellipsoidal heights and the most significant will be discussed herein.

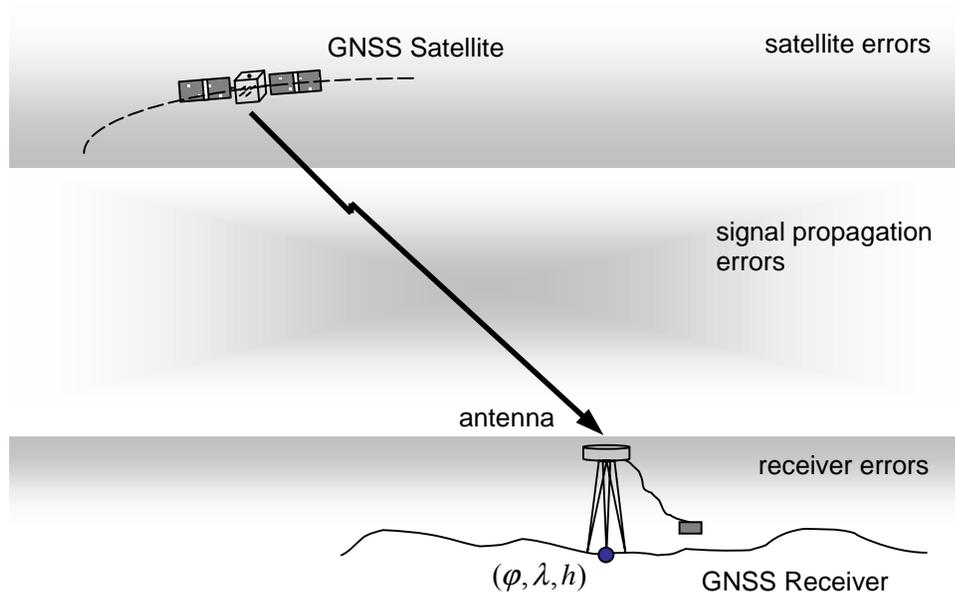


Figure 2.6: Sources of errors for global navigation satellite systems

Orbital Errors

At the satellite level, the most predominant source of error for ellipsoidal height determination is the orbital errors. For short baselines, the orbit error is cancelled when differential processing is performed, however the effect is spatially correlated and therefore the level of cancellation/reduction is dictated by the baseline length. A conservative and perhaps even pessimistic estimate of the decorrelation of satellite orbit errors based on the baseline length is provided by the following linear relationship (Seeber, 1993, p. 297):

$$\frac{\sigma_b}{b} \cong \frac{\sigma_\rho}{\rho} \quad (2.26)$$

where σ_b is the baseline error for a baseline length b . The satellite range is represented by ρ (approximately 22,000 km for GPS satellites) and used to compute the orbit error σ_ρ . In general, the vertical coordinate is affected more than the horizontal coordinates because the largest orbit errors are in the along-track direction, which results in a *tilting* of the network. The best means to deal with this error source is to use precise ephemeris information provided by the International GPS Service (IGS), which has a significantly lower σ_ρ than the broadcast ephemeris information provided in the broadcast navigation message.

Tropospheric Delay Errors

Atmospheric errors account for a large part of the error sources affecting the satellite signals as they propagate towards the receiver(s) located on the surface of the Earth. The signal travels through two parts of the atmosphere, namely the ionosphere and the neutral atmosphere. This neutral part ranges from 0 km to 40 km above the surface of the Earth and is considered a key deteriorating factor for height determination. Specifically, signals traveling through the troposphere suffer the effects of tropospheric attenuation, delay and short-term variations (scintillation). The magnitudes of these effects are a function of satellite elevation and atmospheric conditions such as temperature, pressure and relative humidity during signal propagation. Furthermore, the troposphere is a non-dispersive medium for GPS frequencies, which means that the tropospheric range errors are not frequency dependent and therefore cannot be cancelled through the use of dual-frequency measurements (unlike the ionospheric effects). The most damaging part is the relative tropospheric bias which is caused by errors in tropospheric refraction at one of the stations in a baseline configuration. The general estimate of the bias caused in ellipsoidal height difference measurements, Δh , is given by (Beutler *et al.*, 1987):

$$\Delta h = \frac{\Delta \rho_r^\circ}{\cos(z_{\max})} \quad (2.27)$$

where $\Delta \rho_r^\circ$ is the relative tropospheric zenith correction and z_{\max} is the zenith angle of the observation. Using this estimation, an error or unmodelled differential tropospheric delay of 1 cm in $\Delta \rho_r^\circ$ at a moderate satellite elevation angle of 20° ($z_{\max} = 70^\circ$) yields an error of 3 cm in the estimated ellipsoidal height difference. Two methods that can be applied to estimate tropospheric refraction include, modelling tropospheric parameters simultaneously with all other GPS parameters (clock, latitude, longitude, height, ambiguities), or independent modelling of the troposphere using water vapour radiometers and ground meteorological observations. Ultimately, the best means to deal with tropospheric effects for high precision height determination is by improving the measurements and models for water vapour content (Dodson, 1995).

Multipath

Multipath is a signal propagation error, which occurs when a signal arrives at a receiver via multiple paths (Braasch, 1996). It is caused by the reflection and diffraction of the transmitted signal by objects in the area surrounding the receiver antenna. In Elósegui *et al.* (1995), the magnitude of the multipath error on the vertical coordinate was estimated and found to be strongly dependent on the satellite elevation angle. For instance, a variation from 5° to 10° in elevation cut-off changed the estimates of the ellipsoidal height from tens of millimetres to several centimetres (*ibid.*). From a processing point of view, the problem is juxtaposed as lowering the elevation cut-off (i.e., for VLBI and GPS measurements) helps to decorrelate the tropospheric and height parameters, but at the same time may cause an increase in multipath effects. Over the past decade, there have been numerous improvements to receiver and antenna technology (choke rings, ground planes), which aid in mitigating the effects of multipath. Despite these technological advances, the best method for most GNSS users to mitigate multipath effects is to simply avoid it by carefully selecting receiver station sites that are free of any reflective

obstructions. Selecting a low multipath environment is an important consideration that must be adhered to when establishing permanent vertical control stations. Existing levelling stations may not be optimally located for such measures and therefore biases may exist in the ellipsoidal heights of co-located GPS-levelling benchmarks.

Atmospheric and Ocean Loading

For most combined height applications discussed in this thesis, a large network (e.g., of national scale) consisting of *accurate* ellipsoidal height determinations is used. In such cases, it is important to consider the vertical motion of the Earth's crust caused by differential loading effects of the atmosphere and ocean tides. In general, the deformation of the crust as a reaction to changing atmospheric pressure is at the level of 1 to 2 cm (Van Dam *et al.*, 1994). The larger displacement is due to ocean loading, which is more difficult to model and may cause height changes of more than 10 cm for stations situated near the coasts (Baker *et al.*, 1995). This is important for GPS monitoring of tide gauge stations, which may be incorporated into vertical datum definitions (see section 2.4.1).

The best means to deal with these effects are to apply corrections to the estimated heights based on global models (in conjunction with higher resolution local models, if they are available), which are designed to predict the response to loads. The accuracy of the global ocean load models may vary depending on the location, with more accurate predictions in the open oceans and degrading accuracy approaching the coastal areas. With GPS measurements, making observations over a 24-hour period averages out most of the error. However, shorter occupation times may lead to significant biases in the estimated ellipsoidal heights if appropriate corrections are not applied. It is important for users of vertical control stations to be aware of the type of 'corrections' that have been made to the supplied ellipsoidal heights. Furthermore, when combining ellipsoidal and geoid heights, it is imperative that both geoid and GPS-derived heights are reduced in a consistent manner (Poutanen *et al.*, 1996).

Antenna Phase Centre Offsets

At the receiver level, the antenna phase centre offsets are of great concern for accurate ellipsoidal height estimates. GPS measurements are actually made with respect to the point in the antenna known as the phase centre, not the survey mark. Corrections must be applied to reduce the measurement to the unknown point. It has been shown that the antenna phase centre is not fixed and varies depending on the elevation of the satellite and also the frequency of the propagated signal. For the combined height networks used in this work, complications arise from the mixing of different antenna types, which may produce errors in the ellipsoidal heights of up to 10 cm (Rothacher, 2001). Estimated tropospheric parameters are also highly correlated with antenna phase centre patterns, which may be incorrectly interpreted in processing software, resulting in amplified errors, especially in the height component. Thus, it is important to use the same antenna make and model for network surveys in order to reduce the errors caused by antenna phase centre offsets. Although the mitigation of this error source seems simple compared to the complicated modelling of other error sources, this is a difficult task to manage, particularly for large networks as in Canada and Australia.

2.4 Why combine geoid, orthometric and ellipsoidal height data?

The optimal combination of geoidal undulations, orthometric and ellipsoidal heights is well suited for a number of applications. This is exemplified by the simple geometrical relationship that exists between the triplet of heights, expressed in Eq. (1.1) and depicted in Figure 1.1. Traditional methods for establishing vertical control, although precise, are very labourious, costly and impractical in harsh terrain and environmental conditions. On the other hand, ellipsoidal heights can be efficiently and relatively inexpensively be established with dense coverage over land (i.e., using global navigation satellite systems) or over the oceans (i.e., using satellite altimetry), albeit at a poorer accuracy level. The main problem with these techniques is that the heights refer to a fictitious reference ellipsoid approximating the true shape of the Earth and therefore do not embody any physical meaning. The link between geometrically-defined ellipsoidal heights and heights

with respect to a local vertical datum (i.e., geoid) is provided by geoidal heights. Recognizing the inherent advantages and limitations of each type of height system, it is clear that a proper combination of the heterogeneous heights, with proper error analysis, will benefit innumerable applications, not necessarily restricted to geodesy (e.g., oceanography, cartography, geophysics). For the purposes of the discussion in this dissertation, a number of important geodetic application areas that will benefit from the optimal combination of the heterogeneous height data have been identified, namely:

- modernizing regional vertical datums
- unifying national/regional vertical datums for a global vertical datum
- transforming between different types of height data
- refining and testing existing gravimetric geoid models

The following four sections provide an overview of these important application areas.

2.4.1 Modernizing regional vertical datums

A vertical datum is a reference surface to which the vertical coordinate of points is referred. At a national level, some of the practical uses and benefits of a consistent regional vertical datum include, but are not limited to, the following (Zilkoski *et al.*, 1992):

- improved coastal/harbour navigation
- accurate elevation models for flood mitigation
- accurate elevation models for environmental hazards
- enhanced aircraft safety and aircraft landing
- accurate models for storm surges and coastal erosion
- improved models for chemical spill monitoring
- improved understanding of tectonic movement
- improved management of natural resources

Traditionally, geodesists have used three different types of vertical datums (according to Vaníček, 1991), namely (a) geoid, (b) quasi-geoid, and (c) reference ellipsoid. All of these reference surfaces can be defined either globally or regionally, such that they approximate the entire Earth's surface or some specified region, respectively.

With no official global vertical datum definition, most countries or regions today use regional vertical datums as a local reference height system. This has resulted in over 100 regional vertical datums being used all over the world (Pan and Sjöberg, 1998). The datums vary due to different types of definitions, different methods of realizations and the fact that they are based on local/regional data. A common approach for defining regional vertical datums is to average sea level observations over approximately 19 years, or more precisely, ~ 18.6 years, which corresponds to the longest tidal component period (Melchior, 1978; Smith, 1999), for one or more fundamental tide gauge. This average sea level value is known as mean sea level (MSL) and is used because it was assumed that the geoid and MSL coincided (more or less). This assumption is obviously false, as it is known today that the MSL and the geoid differ by approximately ± 2 metres (Klees and van Gelderen, 1997). Also, the geoid is by definition an equipotential surface, whereas MSL is not, due to numerous meteorological, hydrological, and oceanographic effects (Groten and Müller, 1991). This discrepancy between MSL and the geoid is known as mean dynamic sea surface topography, hereinafter denoted by MSST. With the current demands for a cm-level accurate vertical datum, the discrepancy between the geoid and MSL cannot be ignored.

Figure 2.7 depicts a typical scenario for the establishment of a reference benchmark to define a regional vertical datum. The tide gauge records the instantaneous sea level height H_{ISL} and these values are averaged over a long term in order to obtain the mean value of the local sea level H_{MSL} . The height of the tide gauge is also measured with respect to a reference benchmark that is situated on land a short distance from the tide gauge station. Then the height of the reference benchmark above mean sea level H_{BM} is computed by:

$$H_{BM} = H_{MSL} + \Delta H_{BM-TG} \quad (2.28)$$

Levelling begins from this benchmark and reference heights are accumulated by measuring height differences along levelling lines. The accuracy of the reference benchmark height derived in this manner is dependent on the precision of the height difference ΔH_{BM-TG} and the value for mean sea level H_{MSL} . If one assumes that the value for mean sea level is computed over a sufficiently long period of time which averages out all tidal period components and any higher frequency effects such as currents, then the accuracy depends on ΔH_{BM-TG} .

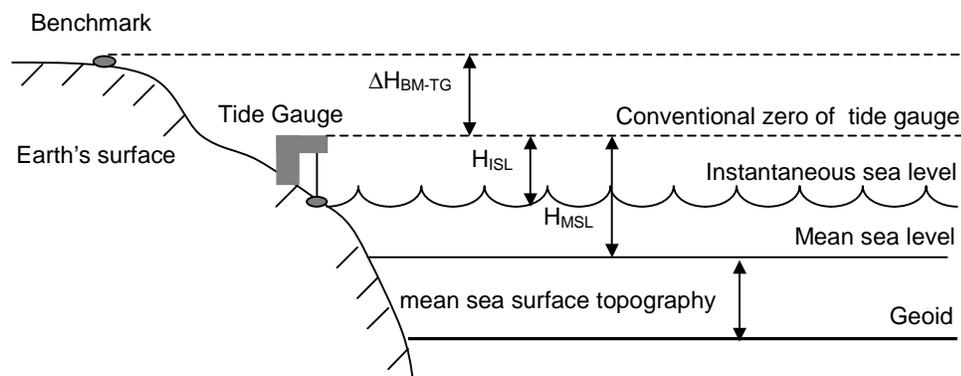


Figure 2.7: Establishment of a reference benchmark height

For highly accurate heights as those needed for a cm-level vertical datum, the tide gauges cannot be assumed to be vertically stable. It is well known that land motion at tide gauges is a source of systematic error, which causes distortion in the height network if it is not corrected for. Land motion at tide gauges and reference benchmarks may be caused abruptly by earthquakes or by erosion or more subtle changes such as post-glacial rebound and land subsidence. The solution to this problem is to include an independent space-based geodetic technique such as GPS (or DORIS, GLONASS and in the future GALILEO) in order to estimate the land motion at these tide gauges. However, at this

point in time, there are still too few measurements available at tide gauges to provide an accurate assessment of the global situation (Mitchum, 2000).

As an example of a regional vertical datum, let us consider the most recent general adjustment of the North American Vertical Datum conducted in 1988, known as NAVD88. NAVD88 consisted of a minimally constrained adjustment of observations in Canada, the United States and Mexico. The tide gauge benchmark at a point in Father Point/Rimouski, Quebec was held fixed at zero. This point was chosen due to its precisely documented position, stable location and it provided a link to the International Great Lakes Datum of 1985 (IGLD85). A detailed account of the adjustment can be found in Zilkoski *et al.* (1992). Although the United States has accepted NAVD88 as an official vertical datum (despite some differences between the east and west coasts and some remaining unexplained problems), Canada has not officially adopted NAVD88 as a national vertical datum. In fact, Canada still uses the Canadian Geodetic Vertical Datum of 1928 (CGVD28). This official Canadian height system was adjusted by constraining five tide gauges, three on the Atlantic (Yarmouth, Halifax and Pointe-au-Père) and two on the Pacific (Vancouver and Prince-Rupert) to the MSL of 1928. The heights are determined as Helmert orthometric heights as in the case of NAVD88, however for CGVD28 the actual gravity is replaced by normal gravity (Geodetic Survey Division, 1998).

As new methodologies and techniques evolve to the point where cm-level (and even sub-cm-level) accurate coordinates are needed, the distortions in traditionally-defined regional vertical networks are no longer acceptable. With this in mind, five main approaches have been identified by Vaníček (1991) for the realization of a "modern" regional vertical datum. These options are summarized below, with some additional remarks (see also Kearsley *et al.*, 1993).

- (i) *Define the geoid by mean sea level as measured by a network of reference tide gauges situated along the coastlines of the country and fix the datum to zero at these stations.* As stated previously, this approach will result in distorted heights as

MSL is not an equipotential surface and it varies from the geoid on the order of a few metres. Also, by fixing the datum to zero at these tide gauge stations, one is assuming that the gauge measurements are errorless or any error inherent in the measurements is acceptable. This is also a boldly incorrect assumption. For instance, consider the case in Canada where not only are some tide gauges poorly situated (sites of river discharge), but also the land to which the tide gauges are stationed is moving due to post-glacial rebound. It is known that regions such as Canada and the Scandinavian countries are rebounding or subsiding up to 1-2 mm/year. If these tide gauge motions are neglected, the error propagates into the levelled heights referred to the regional vertical datum and causes distortions and inconsistencies in the final orthometric heights.

- (ii) *Define the vertical datum by performing a free-network adjustment where only one point is held fixed. A correction factor (shift) is applied to the resulting heights from the adjustment so that the mean height of all tide gauges equals zero. This modified version of option (i) above relies heavily on the measurements from a single tide-gauge, while ignoring the observations for MSL made at all other stations.*
- (iii) *Use the best model available to estimate sea surface topography at the tide gauge stations and then adjust the network by holding MSL-MSST to zero for all tide gauges. This approach does eliminate most of the shortcomings identified in options (i) and (ii) above, however there are some practical limitations in terms of accuracy. Tide gauges are situated near coastal areas and even with the use of satellite altimetry, which has revolutionized sea surface observations and greatly improved SST models in open oceans, the performance in coastal areas is still quite poor. Global ocean circulation models derived from satellite altimetry data and hydrostatic models may reach accuracies of 2-3 cm in the open oceans, but the models fall apart in shallow coastal areas giving uncertainties on the order of tens of centimetres (Shum *et al.*, 1997). Therefore, with significant problems still looming in the coastal regions, distortions will be evident in heights referred to a vertical datum that is defined with low accuracy SST models.*

- (iv) *Define the vertical datum in the same manner as option (iii), but allow the reference tide gauges to 'float' in the adjustment by assigning them realistic a-priori weights (estimates of errors).* This approach can incorporate all of the information for MSL and SST at the reference tide gauges. With improvements in models obtained from satellite altimetry and a better understanding of the process of tide gauge observations (e.g., reference benchmark stability, changes in position), estimates of the accuracy of the observations can be made.
- (v) *As in option (iv), but use estimates of orthometric heights from satellite-based ellipsoidal heights and precise gravimetric geoidal heights.* One of the main advantages of this approach is that it relates the regional vertical datum to a global vertical reference surface (since the satellite-derived heights are referred to a global reference ellipsoid). This aids in the realization of an internationally accepted World Height System (WHS) or global vertical datum (Colombo, 1980; Balasubramania, 1994).

A slightly modified version of option (v) which combines all relevant observations from tide gauges, SLR, GPS, satellite altimetry, and global geopotential models along with appropriate estimates of the accuracy is described by Kearsley *et al.* (1993) and tested using Australian data. It is also one of the most promising approaches to be pursued for future work as more accurate data becomes available.

2.4.2 Global vertical datum

A global vertical datum can be defined as a height reference surface for the whole Earth. The concept of a global vertical datum has been a topic of great research and debate over the past century and has yet to be established as an international standard although numerous proposals from the geodetic community have been made (see for example Burša *et al.*, 2001 and Grafarend and Ardalan, 1997). Despite these efforts and the impending need, a widely used practical WHS has not been officially established for all heighting applications. The establishment of an accurate, consistent and well-defined

global vertical datum has many positive implications. Some arguments that support the global definition and practical realization for a global vertical datum include:

- provides a consistent and accurate method for connecting national and/or regional vertical datums
- inconsistencies in gravity anomalies and heights resulting from the use of different datums can be removed by referring measurements to a common geopotential surface
- results of geodetic levelling and oceanographic procedures for determining the sea surface over long distances can be compared (Balasubramania, 1994)

Over the past several decades, another area where a global vertical datum has been deemed necessary is in the study of global change applications, such as, global change monitoring, mean sea level changes, instantaneous sea surface models, polar ice-cap volume monitoring, post-glacial rebound studies and land subsidence studies. All of these applications require a global view of the Earth with measurements not only on land, but over the oceans as well. The most modern tool available to geodesists and oceanographers for measurements over the oceans, other large bodies of water, and even polar ice-sheets (Ekholm, 1998), is satellite altimetry, which can be used to provide ellipsoidal heights.

By studying the various approaches for the realization of regional vertical datums, significant insight can be obtained for the global vertical datum solution. In fact, numerous studies have investigated the connection of regional datums into one as a viable solution (Colombo, 1980; Balasubramania, 1994; and van Onselen, 1997). In Colombo (1980) a combination of geometric and geophysical data was used to define the global vertical datum. Specifically, three-dimensional geocentric coordinates and the geoidal undulation from a high degree geopotential model of at least one fundamental point in each connecting vertical datum are used to achieve an accuracy of ~ 50 cm for the connection of vertical datums between continents. In Balasubramania (1994), the achievable accuracy for datum connection ranges from ± 5 cm to ± 23 cm. Recent results

using more accurate satellite and terrestrial data are given in van Onselen (1997), where depending on the regions to be connected, estimated errors of up to 80 cm occur when using satellite-only geoid models and 20 cm with the incorporation of terrestrial data (see also, Klees and van Gelderen, 1997). An excellent example of the latter is the successful connection of the land-based gravimetric geoid of The Netherlands with the geoid or MSL at sea (North Sea), which has been performed by fitting a correction model through the TOPEX/POSEIDON data and GPS-levelling data (de Bruijne *et al.*, 1997). Other datum connections that have been fueled by the improved accuracy of sea level change estimates derived from satellite altimetry mission data include the precise determination of the sea surface in the region of the Indonesian Archipelago, which enabled the connection and unification of height datums of the Indonesian islands (Naeije *et al.*, 1998).

Although the connection of regional vertical datums into one is a plausible solution to the global vertical datum problem, it is not viable at this time due to the accuracy requirements of a cm-level datum. A datum connection at this level requires very accurate geoid determination over varying wavelengths (depending on the spatial distance between regional height systems) as well as consistency between regions. This has not been achieved by most regions and is certainly not available on a global scale (Klees and van Gelderen, 1997). In the future, this situation is expected to improve with the promising results awaited from CHAMP, GRACE, and GOCE.

In general four strategies for solving the global vertical datum problem have been identified (Heck and Rummel, 1990; Lehmann, 2000), as follows:

- a) *Pure oceanographic approach.* The main problem with connecting regional vertical datums between continents that are separated by the ocean is the sea surface topography. The oceanographic, hydrologic and meteorological processes that cause SST are complex as they deal with a fluid medium that is a dynamic surface with varying salinity, temperature, density, currents, wind stress, and air pressure, that are difficult to accurately model. However, oceanographers develop models for the

- differences of the gravity potential of the sea surface using geostrophic and steric levelling techniques. As with the case of satellite altimetry, these techniques are least reliable/accurate in the low-lying coastal areas. This is an unfortunate predicament as it is in the coastal boundaries where the 'datum connection' is made.
- b) *Satellite altimetry combined with geostrophic levelling.* It is advantageous to combine traditional oceanographic techniques such as geostrophic levelling with modern tools, such as satellite altimetry. In this case, the most useful types of global information that can be derived from satellite altimetry are a global geopotential model and observations for the mean sea surface, which aid in the determination of a marine geoid. The mean sea level for the global vertical datum definition must be available as a two-dimensional surface all over the oceans. Using satellite altimetry this is obtained by determining MSL at discrete points along the ground tracks of the satellite altimeter and then using these values to interpolate wherever information is desired. However, geostrophic levelling is still required for the extrapolation of sea surface topography at the tide gauges since the temporal coverage of satellite altimetry is still quite poor. To date, there is only approximately 20 years of altimetric observations available, whereas over 100 years of tide gauge observations have been made.
- c) *Geodetic boundary-value problem.* The vertical datum issue can be approached from the point of view of a geodetic boundary value problem (Moritz, 1980). A detailed discussion of a possible solution using this approach is given in Lehmann (2000), and involves the combined use of terrestrial data and global geopotential models where the available data types change across coastlines. The treatment of the datum issues in this manner is contained in a family of problems that is known as the altimetry-gravimetry boundary value problem (AGBVP). It should be noted that although this approach may be theoretically defined and the equations can be derived for its solution, it lacks in practicality as the solution often assumes unified data coverage all over the Earth. To deal with this practical limitation, certain assumptions are made

and approximations follow in the formulations. To date, however, a practical solution using this approach has yet to be realized.

- d) *Satellite positioning combined with gravimetry.* In this approach, the connection between the geometric heights obtained from GNSS and levelled heights referred to a certain local vertical datum are utilized (see Figure 1.1). This method is one of the most promising approaches for the vertical datum problem, however no country has adopted it. This is mainly due to the limitation in achievable accuracy, which depends on the accuracy of the ellipsoidal heights and the internal precision of the gravimetric geoid model. Given the capabilities of such an approach, it is evident that further studies in this area are warranted and therefore studied in detail in this thesis.

In Balasubramania (1994), two options for establishing a global vertical datum based on the adjustment of heterogeneous data (related to option (d) above) were explored. The first approach assumed that the required data was uniformly distributed and available all over the Earth. The four kinds of data required for this method, are (i) free-air gravity anomalies, (ii) precise heights of stations above a regional vertical datum, (iii) an accurate global geopotential model, and (iv) accurate ellipsoidal heights of stations. This approach is limited however, by the fact that the data are not uniformly distributed, nor are they available at the required precision at all parts on the globe. In general, Europe and North America are the only continents that have close to the sufficient amount of accurate information. Due to this lack of information on most parts of the Earth, only first results from a single iteration are provided using this method. These results showed that a global vertical datum can be realized to an accuracy of ± 5 cm.

The second approach, is presented in terms of a more practical or operational realization of the global vertical datum by depending on GPS/DORIS tracking networks and accurate geoid models. The orthometric heights computed via Eq. (1.1) and the incorporation of a corrector surface model refer to a conceptual surface (the geoid), which is not associated with any specific MSL. By using this approach, any direct

reference to MSL is eliminated. In order to implement this method globally a geoid that is precise all over the globe is required.

A global vertical datum can also be realized by choosing a geopotential value, W_0 , either by adopting a geopotential for a regional vertical datum, or some arbitrary value. Variations in geopotential values will exist due to the local data used, and some possible values are provided in de Bruijne *et al.* (1997) and Grafarend and Ardalan (1997). One of the most recent studies conducted by Burša *et al.* (2001) proposes the definition of a global vertical datum by adopting the W_0 value that is averaged over the seas. Today, this value is directly obtained from satellite altimetry measurements with an accuracy of approximately 5 cm (using TOPEX/POSEIDON observations). In addition to the altimetric data, other required data for the solution are: GPS-levelling heights referred to a specified regional vertical datum, a global geopotential model, the geocentric gravitational constant, angular velocity of the Earth's rotation, and the second zonal harmonic coefficient (see Burša *et al.*, 2001 for more details).

Further to the aforementioned approaches, it is also important to supply users with the practical ability to connect regional vertical datums and transform from regional to global vertical datums. This can be done using the method proposed by Burša *et al.* (2001) provided GPS-levelling sites with ellipsoidal and orthometric/normal heights or more preferably geopotential numbers are available (also see Pan and Sjöberg, 1998).

2.4.3 GNSS-levelling

The inherent appeal of the seemingly simple linear geometrical relationship between the three height types is based on the premise that given any two of the height types, the third can be derived through simple manipulation of Eq. (1.1.). There are several issues hidden within this statement that will be uncovered in the sequel (e.g., datum inconsistencies, systematic errors, data accuracy). The optimal combination of GPS-derived ellipsoidal

heights with gravimetrically-derived geoid undulations for the determination of orthometric heights above mean sea level, or more precisely with respect to a vertical geodetic datum is referred to as GPS-levelling*. The process can be described as follows for the absolute and relative cases (height difference between two points i and j), respectively:

$$H = h - N \quad (2.29a)$$

$$H_j - H_i = (h_j - h_i) - (N_j - N_i), \quad \Delta H = \Delta h - \Delta N \quad (2.29b)$$

This procedure has been the topic of several studies over the years (see for e.g., Engelis *et al.*, 1985, Forsberg and Madsen, 1990, and Sideris *et al.*, 1992) and demonstrated to provide a viable alternative over conventional levelling methods for lower-order survey requirements. A major limitation of using GPS-levelling as a means for establishing heights or height differences with respect to a local vertical datum is that it is dependant on the achievable accuracy of the ellipsoidal and geoid height data. In practice, the relationship given by Eq. (2.29) is never fulfilled due to numerous errors, systematic distortions and datum inconsistencies inherent among the triplet of height data (Jiang and Duquenne, 1995; van Onselen, 1997; Ollikainen, 1997; Fotopoulos *et al.*, 2001a). Thus, a more rigorous treatment for the integration of these different height types requires the incorporation of a parametric corrector surface model in Eq. (2.29). The role of such a model is to absorb the datum inconsistencies and any systematic distortions that exist in the height data sets (Shrestha *et al.*, 1993). More details are provided in the next two chapters.

In practice, the GPS-levelling technique has become quite common and used often erroneously or with a poor understanding of the transformations between reference surfaces and systematic errors involved. As accuracy requirements increase, the incorrect

* The terms GPS-levelling, GPS/geoid levelling and GNSS-levelling are used interchangeably.

application of Eq. (2.29) has more severe implications. Therefore, it is important to develop proper procedures for combining the heterogeneous height data and a means to convey this information to users. This is a major part of the work presented in the remaining parts of this thesis and will be discussed specifically in chapter 7 for numerical case studies in northern Canada, where vertical control is difficult to establish and GNSS-levelling offers a viable alternative.

2.4.4 Refining and testing gravimetric geoid models

Another common manipulation of Eq. (1.1) is the combined use of co-located ellipsoidal and orthometric heights (or height differences) in order to compute geoidal height values at the GPS-levelling benchmarks. The form of the equation for absolute and relative values (between points i and j), respectively, is given by

$$N = h - H \quad (2.30a)$$

$$N_j - N_i = (h_j - h_i) - (H_j - H_i) \quad \Delta N = \Delta h - \Delta H \quad (2.30b)$$

These GPS-derived geoid heights are invariably different from the values interpolated from a gravimetrically-derived geoid model and are influenced by the datum inconsistencies, biases and errors associated with the independently derived ellipsoidal and orthometric height data. For instance, a gravimetrically computed geoid model, obtained from the *remove-compute-restore* process described in section 2.1, will (theoretically) refer to the geocentric reference system implicit in the used geopotential model. This reference system will in turn correspond to the adopted coordinate set for the satellite tracking stations used in the global geopotential solution. This coordinate set will not necessarily agree with the adopted reference system for the ellipsoidal heights obtained from the GPS measurements. Furthermore, the local levelling datum to which

the orthometric heights refer will not likely correspond to the reference potential value of the geopotential model or the GPS reference system.

From this brief discussion, the complexity involved with combining the different height data becomes evident and it is imperative that users are knowledgeable about these matters in order to apply Eq. (2.30) properly. In practice, the major applications of Eq. (2.30) include:

- external independent evaluation of gravimetric geoid accuracy,
- incorporation of GPS-derived geoid heights into the gravimetric geoid solution as a constraint, and
- densification of networks that have already been positioned by conventional horizontal and vertical methods,

which are all described in more detail below.

Comparisons between different geoid solutions provide insight into the accuracy of the geoid determination techniques (Sideris *et al.*, 1992). To date, comparisons of gravimetrically-derived geoid model values interpolated at GPS-on-benchmarks with geometrically computed geoid values (derived from Eq. 2.30) provide the best external means of evaluating the geoid model accuracy. In order for this method to provide an indication of the 'accuracy' of the gravimetric geoid model, it is important that the GPS-levelling data used for testing is not incorporated in the original geoid solution. This is an obvious statement, but often neglected in practice.

Long-wavelength errors present in gravimetrically-derived geoid models (described in section 2.1) may be reduced by constraining the geoid solution to observed geoid values at GPS-levelling benchmarks (Forsberg and Madsen, 1990). This is a common procedure implemented in many recent national geoid models through the use of least-squares collocation procedures, and shown to give positive results (see Tscherning *et al.*, 2001 and Featherstone, 2000). In Roman and Smith (2000), this approach, referred to as the

hybird methodology, is used to compute the geoid model for the continental United States.

The use of geoid values derived from Eq. (2.30) has also offered the opportunity for government agencies and mapping authorities responsible for national height databases to densify existing vertical control networks, as proposed in Engelis *et al.* (1985). Given the improvements and availability of satellite-based positioning techniques, it is certain that this procedure will continue and become increasingly popular in the near future.

Chapter 3

Combined Height Adjustment and Modelling of Systematic Effects

In this chapter, the algorithms and methodology used in the combined adjustment of the ellipsoidal, orthometric and geoid height data is presented. The first section describes the combined least-squares adjustment scheme for both absolute and relative height input values. The unknown parameters solved for in this adjustment are the 'coefficients' of some parametric model selected for dealing with the systematic errors and datum inconsistencies inherent among the heterogeneous height data. As there are numerous options available for the form of the parametric model, an overview of the main choices is also provided. Finally, the key elements of a procedure developed for assessing the parametric model performance is described in detail.

3.1 General combined adjustment scheme

In this section, a description of the observation equations and mathematical models used for the multi-data adjustment of ellipsoidal, geoidal and orthometric heights, is provided. The formulation herein forms the basis for all of the results presented in the remaining chapters, with particular emphasis placed on the role of the systematic and random errors inherent among the heterogeneous height types. It should be recognized that there are a number of options available for combining these height data (see, e.g., Kearsley *et al.*,

1993, Kotsakis and Sideris, 1999, and Dinter *et al.*, 2001). The algorithm used herein is an amalgamation of the procedures described in Kotsakis and Sideris (1999), Kearsley *et al.* (1993), de Bruijne *et al.* (1997), and others and it has been implemented and tested extensively using real data sets by the author; see, e.g., Fotopoulos *et al.* (2001a,b). It was selected as the most appropriate adjustment scheme as it offers a practical solution to the problem given the data currently available.

Problem formulation using relative height data

Given a network of points with known ellipsoidal and orthometric height values and the availability of a gravimetric geoid model, a combined 1D least-squares adjustment of GPS, levelling and geoid height data can be performed. In practice, the height differences for each data type are formed with respect to some selected initial point/station. Therefore, the “observed” input values are Δh_{ij} , ΔH_{ij} , and ΔN_{ij} for each pair of points (i, j) forming a baseline in the test network. The corresponding observation equation model is given by

$$\Delta h_{ij} = h_j^\alpha - h_i^\alpha + f_{ij}^{\Delta h} + v_{ij}^{\Delta h} \quad (3.1a)$$

$$\Delta H_{ij} = H_j^\alpha - H_i^\alpha + f_{ij}^{\Delta H} + v_{ij}^{\Delta H} \quad (3.1b)$$

$$\Delta N_{ij} = N_j^\alpha - N_i^\alpha + f_{ij}^{\Delta N} + v_{ij}^{\Delta N} \quad (3.1c)$$

where the superscript α denotes the true values of the various heights at each point. The $f_{ij}^{(\cdot)}$ terms describe the systematic errors and datum inconsistencies in the height data sets. The true values h_i^α , H_i^α , N_i^α , h_j^α , H_j^α , and N_j^α refer to a common geodetic datum such that the following conditions, based on Eq. (1.1), are satisfied at each station:

$$h_i^\alpha - H_i^\alpha - N_i^\alpha = 0 \quad (3.2a)$$

$$h_j^\alpha - H_j^\alpha - N_j^\alpha = 0 \quad (3.2b)$$

which can be extended for the relative case to include a condition for each baseline combination as follows:

$$\Delta h_{ij}^\alpha - \Delta H_{ij}^\alpha - \Delta N_{ij}^\alpha = 0 \quad (3.2c)$$

Using Eqs. (3.2a), (3.2b) and (3.2c), a 'new' observation equation for each baseline in the test network can be formed according to the following expression:

$$\ell_{ij} = f_{ij} + v_{ij}^{\Delta h} - v_{ij}^{\Delta H} - v_{ij}^{\Delta N} \quad (3.3)$$

where the 'observed' height misclosure value for each baseline is given by

$$\ell_{ij} = \Delta h_{ij} - \Delta H_{ij} - \Delta N_{ij} \quad (3.4)$$

The f_{ij} term in Eq. (3.3) refers to the total (combined) correction term for the systematic errors and datum inconsistencies in the multi-data test network. It can be modeled according to a deterministic parametric form

$$f_{ij} = (\mathbf{a}_j^T - \mathbf{a}_i^T) \mathbf{x} = \mathbf{a}_{ij}^T \mathbf{x} \quad (3.5)$$

where $\mathbf{a}_j, \mathbf{a}_i$ are $n \times 1$ vectors of known coefficients that usually depend on the horizontal location of the network points i, j , and \mathbf{x} is an $n \times 1$ vector of unknown parameters. The role of the corrector surface model in this combined multi-data height adjustment must also be investigated and is discussed further in section 3.2. Furthermore,

the choice of the parametric form of the corrector surface model will be discussed in more detail in sections 3.3 and 3.4.

The $v_{ij}^{(\cdot)}$ terms in Eqs. (3.1a-c) describe the zero-mean random errors in the ellipsoidal, orthometric and geoid height differences. Their second order statistical properties, for all baselines in the multi-data test network, are provided by the covariance (CV) matrices

$$E \left\{ \mathbf{v}_{\Delta h} \mathbf{v}_{\Delta h}^T \right\} = \mathbf{C}_{\Delta h} \quad (3.6a)$$

$$E \left\{ \mathbf{v}_{\Delta H} \mathbf{v}_{\Delta H}^T \right\} = \mathbf{C}_{\Delta H} \quad (3.6b)$$

$$E \left\{ \mathbf{v}_{\Delta N} \mathbf{v}_{\Delta N}^T \right\} = \mathbf{C}_{\Delta N} \quad (3.6c)$$

where E denotes the mathematical expectation operator and $\mathbf{v}_{\Delta h}$, $\mathbf{v}_{\Delta H}$, and $\mathbf{v}_{\Delta N}$ are the vectors that contain the unknown random errors for the height differences for all network baselines.

Multi-data adjustment using absolute height data

Given the observation equation set-up presented above, the adjustment of observations is performed for the relative height data case. This approach will be revisiting in chapter 7, where it is used with simulated test data to assess the accuracy of orthometric height differences obtained from GPS-levelling. In this section, the observation equation formulations are provided for the absolute case where the 'observed' input data to the adjustment are h_i , H_i and N_i for a common set of co-located GPS-levelling benchmarks over a network area. This approach is used throughout as it was found to be the most accessible form of the real height data sets. Thus, the following discussion will focus on the typical scenario of a triplet of height information at *each* control point in the network where the observation equation for each station in the network is given by

$$\ell_i = f_i + v_i^h - v_i^H - v_i^N \quad (3.7)$$

where,

$$\ell_i = h_i - H_i - N_i \quad (3.8)$$

The f_i term in Eq. (3.7) refers to the total (combined) correction term for the systematic errors and datum inconsistencies in the multi-data test network and can be modeled according to a deterministic parametric form

$$f_i = \mathbf{a}_i^T \mathbf{x} \quad (3.9)$$

In the absolute case, the covariance matrices are denoted by

$$E \{ \mathbf{v}_h \mathbf{v}_h^T \} = \mathbf{C}_h \quad (3.10a)$$

$$E \{ \mathbf{v}_H \mathbf{v}_H^T \} = \mathbf{C}_H \quad (3.10b)$$

$$E \{ \mathbf{v}_N \mathbf{v}_N^T \} = \mathbf{C}_N \quad (3.10c)$$

where \mathbf{v}_h , \mathbf{v}_H , and \mathbf{v}_N are the vectors that contain the unknown random errors for each height type.

The general linear functional model used for the multi-data (combined) adjustment of the heterogeneous height data described above is given as follows:

$$\mathbf{l} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} \quad (3.11a)$$

$$E \{ \mathbf{v} \} = 0 \quad (3.11b)$$

$$E \{ \mathbf{v}\mathbf{v}^T \} = \mathbf{C}_v \quad (3.11c)$$

where the $m \times 1$ vector of observations \mathbf{l} is composed of the height 'misclosure' at each GPS-levelling benchmark as given in Eq. (3.8). It should be noted that the absolute height values, h , H , and N are the results of previous adjustments conducted using baseline information for each of the ellipsoidal and orthometric height networks and similarly in the computations for a gravimetric geoid model. Therefore, the results of these original adjustments are used as input into this secondary combined least-squares adjustment. The $m \times u$ design matrix, \mathbf{A} , depends on the parametric model type (see section 3.3). \mathbf{B} is the block-structured matrix denoted by

$$\mathbf{B} = [\mathbf{I} \quad -\mathbf{I} \quad -\mathbf{I}] \quad (3.12)$$

where each \mathbf{I} is an $m \times m$ unit matrix. \mathbf{x} is a $u \times 1$ vector containing the unknown parameters corresponding to the selected parametric model and \mathbf{v} is a vector of random errors with zero mean (Eq. 3.11b), described by the following formula:

$$\mathbf{v} = \left[\mathbf{v}_h^T \quad \mathbf{v}_H^T \quad \mathbf{v}_N^T \right]^T \quad (3.13)$$

where $\mathbf{v}_{(\cdot)}$ is an $m \times 1$ vector of random errors for each of the h, H, N data types. The corresponding covariance matrix is described in general by Eq. (3.11c).

Applying the least-squares minimization principle of

$$\mathbf{v}^T \mathbf{P} \mathbf{v} = \mathbf{v}_h^T \mathbf{P}_h \mathbf{v}_h + \mathbf{v}_H^T \mathbf{P}_H \mathbf{v}_H + \mathbf{v}_N^T \mathbf{P}_N \mathbf{v}_N = \text{minimum} \quad (3.14)$$

where the block diagonal weight matrix \mathbf{P} is

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_h & 0 & 0 \\ 0 & \mathbf{P}_H & 0 \\ 0 & 0 & \mathbf{P}_N \end{bmatrix} = \begin{bmatrix} \mathbf{C}_h^{-1} & 0 & 0 \\ 0 & \mathbf{C}_H^{-1} & 0 \\ 0 & 0 & \mathbf{C}_N^{-1} \end{bmatrix} \quad (3.15)$$

one can solve for the unknown parameters (i.e., coefficients) of the corrector surface model by

$$\hat{\mathbf{x}} = [\mathbf{A}^T (\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1} \mathbf{A}]^{-1} \mathbf{A}^T (\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1} \mathbf{l} \quad (3.16)$$

The combined adjusted residuals from the adjustment are given by

$$\mathbf{B}\hat{\mathbf{v}} = \hat{\mathbf{v}}_h - \hat{\mathbf{v}}_H - \hat{\mathbf{v}}_N \quad (3.17)$$

where we can explicitly solve for the **separate** adjusted residuals, according to height data type, by applying the well known formulation (Mikhail, 1976)

$$\hat{\mathbf{v}} = \mathbf{P}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{P}^{-1} \mathbf{B}^T)^{-1} (\mathbf{w} + \mathbf{A} \hat{\mathbf{x}}) \quad (3.18)$$

where $\mathbf{w} = \mathbf{l}$ and is also shown in Kotsakis and Sideris (1999) as follows:

$$\hat{\mathbf{v}}_h = \mathbf{C}_h (\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1} \mathbf{M} \mathbf{l} \quad (3.19a)$$

$$\hat{\mathbf{v}}_H = \mathbf{C}_H (\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1} \mathbf{M} \mathbf{l} \quad (3.19b)$$

$$\hat{\mathbf{v}}_N = \mathbf{C}_N (\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1} \mathbf{M} \mathbf{l} \quad (3.19c)$$

where the \mathbf{M} matrix is expressed by

$$\mathbf{M} = \mathbf{I} - \mathbf{A}(\mathbf{A}^T(\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1}\mathbf{A}^T(\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1}) \quad (3.20)$$

The accuracy of the corrector surface parameters can be computed by

$$\mathbf{C}_{\hat{\mathbf{x}}} = [\mathbf{A}^T(\mathbf{C}_h + \mathbf{C}_H + \mathbf{C}_N)^{-1}\mathbf{A}]^{-1} \quad (3.21)$$

This formulation provides us with the interesting opportunity to evaluate the contribution of each of the height types through the evaluation of $\mathbf{C}_{\mathbf{v}}$, which can also be represented by the following expression:

$$\mathbf{C}_{\mathbf{v}} = \begin{bmatrix} \mathbf{C}_h & 0 & 0 \\ 0 & \mathbf{C}_H & 0 \\ 0 & 0 & \mathbf{C}_N \end{bmatrix} = \begin{bmatrix} \sigma_h^2 \mathbf{Q}_h & 0 & 0 \\ 0 & \sigma_H^2 \mathbf{Q}_H & 0 \\ 0 & 0 & \sigma_N^2 \mathbf{Q}_N \end{bmatrix} \quad (3.22)$$

A detailed discussion of the stochastic model and random errors is given in chapters 5 and 6 and will not be dwelled on in this chapter.

Note on an alternate formulation of the problem

An alternative, yet equivalent formulation of the problem, can be stated by replacing the functional model in Eq. (3.11a) with

$$\mathbf{l} = \mathbf{Ax} + \mathbf{v}^*, \quad E\{\mathbf{v}^*\} = 0 \quad (3.23)$$

where $\mathbf{v}^* = \mathbf{Bv}$, or equivalently stated by

$$\mathbf{v}^* = \mathbf{v}_h - \mathbf{v}_H - \mathbf{v}_N \quad (3.24)$$

It can easily be shown that the final equations corresponding to this alternative formulation are identical to those derived thus far.

The unique perspective obtained by implementing this combined adjustment approach as described herein is embedded in two main areas, namely:

- the evaluation of the contribution of the $f_i = \mathbf{a}_i^T \mathbf{x}$ term, which refers to the total (combined) correction term for the systematic errors and datum inconsistencies in the multi-data test network, and
- the separation of residuals according to the height data types, which allows for the refinement of data covariance matrices.

In the past, numerous studies have been conducted which focused on the first issue. Despite these efforts, a consistent approach for implementing and assessing the model performance has not been widely proposed or accepted. The concept of the multi-data one-dimensional adjustment should not be trivialized, as there are many sources of systematic and random errors involved that have to be dealt with properly in order to rigorously combine all of the data and obtain meaningful and optimal (in the least-squares sense) results. This problem will not only be addressed in the sequel from a theoretical point of view, but it will also be approached from the practical point of view where real-world data limitations are taken into account.

3.2 Role of the parametric model

The main factors that cause discrepancies when combining the heterogeneous heights include the following (Rummel and Teunissen, 1989; Kearsley *et al.*, 1993; Schwarz *et al.*, 1987):

– *Random errors in the derived heights h , H , and N*

The covariance matrices for each of the height types (absolute or differences) are usually obtained from separate network adjustments of the individual height types. The main errors affecting the data were described in chapter 2.

– *Datum inconsistencies inherent among the height types*

Each of the triplet of height data refers to a different reference surface. For instance, GPS-derived heights refer to a reference ellipsoid used to determine the satellite orbits. Orthometric heights, computed from levelling and gravity data, refer to a local vertical datum, which is usually defined by fixing one or more tide-gauge stations (see discussion in sections 2.2 and 2.4.1). Finally, the geoidal undulations interpolated from a gravimetrically-derived geoid model refer to the reference surface used in the global geopotential model, which may not be the same as the one for the gravity anomalies Δg .

– *Systematic effects and distortions in the height data*

These systematic effects have been described in chapter 2 and are mainly caused by the long wavelength geoid errors, which are usually attributed to the global geopotential model (i.e., EGM96). Biases are also introduced into the gravimetric geoid model due to differences between data sources whose adopted reference systems are slightly different. In addition, systematic effects are also contained in the ellipsoidal heights, which are a result of poorly modelled GPS errors, such as atmospheric refraction (especially tropospheric errors). Although spirit-levelled height differences are usually quite precise, the derived orthometric heights for a region or nation are usually the result of an over-constrained levelling network adjustment, which introduces distortions.

– *Assumptions and theoretical approximations made in processing observed data*

Common approximations include neglecting sea surface topography (SST) effects or river discharge corrections for measured tide gauge values, which results in a

significant deviation of readings from mean sea level. Other factors include the use of approximations or inexact normal/orthometric height corrections and using normal gravity values instead of actual surface gravity values in computing orthometric heights (Véronneau, 2002). The computation of regional or continental geoid models also suffers from insufficient approximations in the gravity field modelling method used (de Min, 1990).

– *Instability of reference station monuments over time*

Temporal deviations of control station coordinates can be attributed to geodynamic effects such as post-glacial rebound (e.g., see, de Bruijne *et al.*, 1997), crustal motion and land subsidence. Most GPS processing software eliminate all tidal effects when computing the final coordinate differences. To be consistent, the non-tidal geoid should be used (Ekman, 1989). More details on the error caused by mixing ellipsoidal heights referring to a non-tidal crust and orthometric heights whose reference surface is the mean or zero geoid is given in Poutanen *et al.* (1996).

The combined effects of these factors and others result in poorly estimated height values and more importantly inaccurate assessments, if any, of the results achievable by GPS-levelling. Thus far, the burden of dealing with most of these factors (mainly the systematic errors and datum inconsistencies) has been designated to the use of a corrector surface model. Before continuing with a description of modelling options, it is important to pause for a moment and take a closer look at the role of the parametric model in the GPS/geoid levelling problem. Given the theoretical relationship among the three types of height data and the incorporation of an appropriate corrector surface model, the orthometric height for a *new* point (not belonging to the original multi-data network) is obtained as follows:

$$H = h - N - a^T \hat{\mathbf{x}} \quad (3.25)$$

The question that must be addressed is *to which vertical reference system does the computed value H refer?* To answer this question, we refer to Figure 3.1, which provides an illustrative view of the various reference surfaces embedded in the different height data sets.

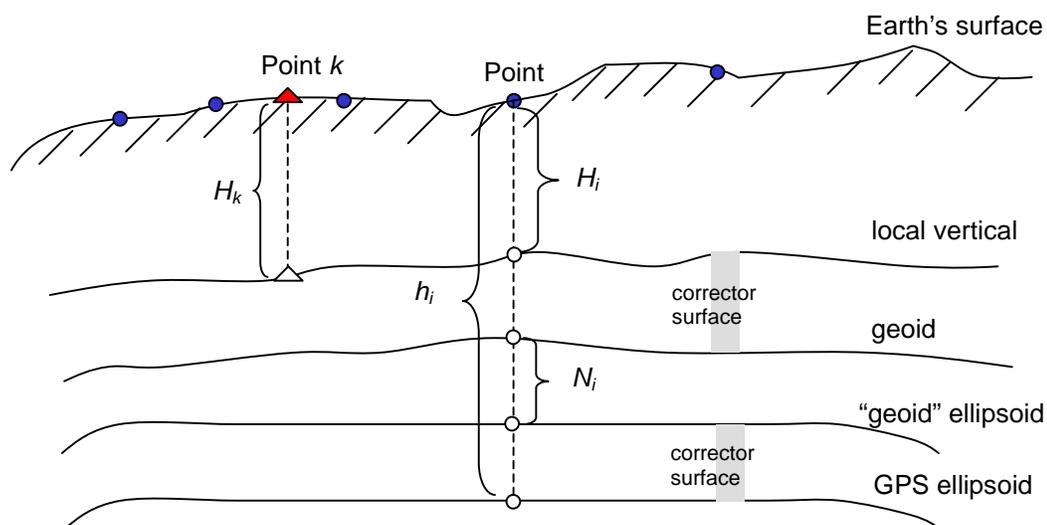


Figure 3.1: Illustrative view of GPS/geoid levelling and the role of the corrector surface

In this figure, the points on the Earth's surface represented by a solid circle belong to the multi-data control network and the point denoted by a triangle is the 'new' point for which the orthometric height is to be computed via GPS/geoid levelling. For the sake of this discussion, if one ignores the systematic effects, and concentrate on the datum inconsistencies, one can see from the figure that the role of the corrector surface is twofold. In general, the datum discrepancies occur between (i) the local vertical datum and the geoid model (both of which are supposed to represent different equipotential surfaces of the Earth's gravity field) and (ii) the two ellipsoids to which the GPS measurements and geoid undulations refer to. These discrepancies are typically not constant biases as depicted in the figure, but they may take on a more complicated form.

In order to obtain the orthometric height through GPS/geoid levelling that refers to the local vertical datum for the new point, H_k , a connection between the different height surfaces must be made. This connection is embedded in the corrector term f (see Eq. 3.9) and can take on many forms, depending on the model selection. It should be cautioned, however, that the corrector surface model will provide a consistent connection between the heights derived from GPS/geoid levelling and the official local vertical datum, only if the orthometric heights used in the multi-data adjustment also refer to the official local vertical datum.

It is evident that the parametric model plays a major role in the combined height adjustment process, which highlights the importance of the following issues:

- (i) selecting the appropriate type of model
- (ii) selecting the extent/form of the model
- (iii) assessing the performance of the chosen model

To date, the first two issues have been given a significant amount of attention in research and will be reviewed in the following section. However, the latter issue, which is equally as important, has often been neglected. Therefore, it will be one of the key aims of this thesis to investigate all of these issues with relevance to practical problems using real data.

3.3 Modelling options

The choice of the parametric form of the corrector surface model is not a trivial task. In fact, the list of potential candidates for the 'corrector' surface is extensive. Arguably, the selection process is arbitrary unless some physical reasoning can be applied to the discrepancies between the GPS-derived geoid heights N^{GPS} , and the geoid heights from the gravimetric geoid model N^{grav} , which fulfills

$$\ell_i = h_i - H_i - N_i = a_i^T x + v_i \quad (3.25)$$

$$a_i^T x = N^{GPS} - N^{grav} \quad (3.26)$$

In the past, researchers have often utilized a simple tilted plane-fit model, which in several cases has satisfied accuracy requirements. However, as the achievable accuracy of GPS and geoid heights improves, the use of such a simple model may not be sufficient. The problem is further complicated because selecting the proper model type depends on the data distribution, density and quality, which varies for each case.

In general, the most common approach to the bilinear term in Eq. (3.26) is to employ a parameterized trend with a finite set of unknown parameters represented in its linear form as follows:

$$p = b_1 f_1 + b_2 f_2 + \dots + b_q f_q \quad (3.27)$$

where b_1, b_2, \dots, b_q are the unknown coefficients to be solved for in the combined least-squares approach discussed in section 3.1 and f_1, f_2, \dots, f_q are known base functions. The type of base functions may vary. One possibility is a polynomial (of various orders), also represented by the multiple regression equation (MRE)

$$\mathbf{a}_i^T \mathbf{x} = \sum_{m=0}^M \sum_{n=0}^N (\varphi_i - \bar{\varphi})^n (\lambda_i - \bar{\lambda})^m x_q \quad (3.28)$$

where $\bar{\varphi}, \bar{\lambda}$ are the mean latitude and longitude of the GPS-levelling points, respectively, and x_q contains the q unknown coefficients. The parameter q varies according to the number of terms up to a maximum of $q = (N + 1) \cdot (M + 1)$. An example of the use of this

bivariate MRE type of model is included in Bin *et al.* (1995) for the Singapore region. Other types of base functions include trigonometric, harmonic, Fourier series, and wavelets. In Featherstone (2000), the use of continuous curvature splines in tension was investigated for parts of Australia.

In some cases, two or more different types of base functions may be appended/merged. This was the case for the recent North Sea region model where the selected models can be represented by the following equations (Haagmans *et al.*, 1998):

$$a + b\lambda + c\varphi + d\varphi\lambda \quad (3.29)$$

$$\sum_{i=1}^I \sum_{j=1}^J a_{ij} \cos(i\lambda) \cos(j\varphi) + b_{ij} \sin(i\lambda) \cos(j\varphi) + c_{ij} \cos(i\lambda) \sin(j\varphi) + d_{ij} \sin(i\lambda) \sin(j\varphi) \quad (3.30)$$

Eq. (3.29) is a bilinear trend function and Eq. (3.30) is a trigonometric function based on Fourier analysis, which was used in different parts to model the long-wavelengths.

Another family of models is based on the general 7-parameter similarity datum shift transformation, with the simplified classic 4-parameter model (Heiskanen and Moritz, 1967, chapter 5) given by

$$\mathbf{a}_i^T \mathbf{x} = x_1 + x_2 \cos \varphi_i \cos \lambda_i + x_3 \cos \varphi_i \sin \lambda_i + x_4 \sin \varphi_i \quad (3.31)$$

where φ_i, λ_i are the latitude and longitude, respectively, of the GPS-levelling points. The full form of the design matrix would be given as follows:

$$\mathbf{A}_{m \times 4} = \begin{pmatrix} 1 & \cos \varphi_1 \cos \lambda_1 & \cos \varphi_1 \sin \lambda_1 & \sin \varphi_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \cos \varphi_{m-1} \cos \lambda_{m-1} & \cos \varphi_{m-1} \sin \lambda_{m-1} & \sin \varphi_{m-1} \\ 1 & \cos \varphi_m \cos \lambda_m & \cos \varphi_m \sin \lambda_m & \sin \varphi_m \end{pmatrix} \quad (3.32)$$

An extended version of the above model is given with the inclusion of a fifth parameter as follows:

$$\mathbf{a}_i^T \mathbf{x} = x_1 + x_2 \cos \varphi_i \cos \lambda_i + x_3 \cos \varphi_i \sin \lambda_i + x_4 \sin \varphi_i + x_5 \sin^2 \varphi_i \quad (3.33)$$

It should be noted that the parameters from such a 'datum shift transformation' do not represent the *true* datum shift parameters (translations, rotations and scale) because other long-wavelength errors inherent in the data (such as those in the geoid heights) will be interpreted as tilts and be absorbed by the parameters to some degree. Recently, a more complicated form of the differential similarity transformation model was developed and tested in the Canadian region and is given by (Kotsakis *et al.*, 2001):

$$\begin{aligned} \mathbf{a}_i^T \mathbf{x} = & x_1 \cos \varphi_i \cos \lambda_i + x_2 \cos \varphi_i \sin \lambda_i + x_3 \sin \varphi_i + x_4 \left(\frac{\sin \varphi_i \cos \varphi_i \sin \lambda_i}{W} \right) + \\ & x_5 \left(\frac{\sin \varphi_i \cos \varphi_i \cos \lambda_i}{W} \right) + x_6 \left(\frac{1 - f^2 \sin^2 \varphi_i}{W} \right) + x_7 \left(\frac{\sin^2 \varphi_i}{W} \right) \end{aligned} \quad (3.34)$$

where $W = \sqrt{1 - e^2 \sin^2 \varphi_i}$, e^2 is the eccentricity and f is the flattening of the reference ellipsoid.

Many researchers have opted for applying anyone of the aforementioned trend surfaces and then modelling the remaining residuals using least-squares collocation (Forsberg, 1998) as follows:

$$\tilde{\mathbf{v}} = \mathbf{C}_{sr} \left(\mathbf{C}_{rr} + \sigma^{-1} \mathbf{I} \right)^{-1} \mathbf{r} \quad (3.35)$$

where \mathbf{r} is a vector of known residuals with variance σ , to be predicted at another location(s), denoted by s . The above equation is usually implemented using a second-order Markov covariance model of the form

$$\mathbf{C}(d) = \mathbf{C}_o \left(1 + \frac{d}{\alpha} \right) e^{-d/\alpha} \quad (3.36)$$

where d is the correlation distance and \mathbf{C}_o and α are empirically determined from the actual data for each case (Forsberg, 1998). The use of other types of covariance functions has also been investigated; see, for example, Milbert (1995) and Denker *et al.* (2000).

Using a mosaic of parametric models

Thus far, the discussion on the type of model has been based on the use of a single model to represent an entire region. This approach is sometimes limiting as it assumes that a homogeneous set of discrepancies exist over an entire region, regardless of its extent and data distribution. Consider for instance, the task of selecting a single model to adequately model all of the discrepancies across large regions such as Canada and Australia, where comparatively sparsely distributed sets of GPS-levelling control points are available (Véronneau, 2002; Johnston and Luton, 2001). An additional limitation of this approach is that it relies on a single model to deal with both long and short wavelength discrepancies.

One way to deal with this is to divide the region into a number of smaller sub-regions and fit the appropriate model to that region using, for example, any of the aforementioned models. The type of model or extent of the model (i.e., order of polynomial) may vary for each sub-region. The new problems that arise when implementing this approach are (i) how to divide the region and (ii) how to connect across adjacent sub-regions. Extensive

studies that test this method are described by Jiang and Duquenne (1996). In this research, the long and short wavelength errors are dealt with separately by using different models. A global transformation model is applied (based on a trigonometric function) to deal with the general transformation of reference systems. Several polynomial models (usually in planar representation) are applied to the divided sub-regions in order to deal with local deformations. The combined adjustment employs a set of constraint equations for common points in the neighbouring sub-regions. For example, the following constraint conditions can be used for the absolute (pointwise) and relative (baseline) cases:

$$\Delta N = \Delta N_o + a(\varphi - \varphi_o) + b(\lambda - \lambda_o) \quad (3.37)$$

$$\Delta N_{ij} = a(\varphi_j - \varphi_i) + b(\lambda_j - \lambda_i) \quad (3.38)$$

where ΔN is the observed geoidal undulation as given with respect to the geoid, $\varphi_o, \lambda_o, \Delta N_o$ are the mean values of latitude, longitude and geoidal undulation, respectively, in the sub-region, ΔN_{ij} is the difference of geoidal undulation between points i and j , and the coefficients to be determined from the adjustment of the common points are denoted by a and b . It should be noted that in Jiang and Duquenne (1996), equivalent forms of Eqs. (3.37) and (3.38) were derived for observed height anomalies, $\Delta \zeta$, given with respect to the quasi-geoid. The common points belonging to each neighbouring sub-region should have the same adjusted values. For more details and results using test network data in France, see Duquenne *et al.* (1995) and Jiang and Duquenne (1996).

A similar test study was conducted for Australia where it was determined that it was appropriate to use a mosaic of models in four sub-regions in order to model the discrepancies (see Fotopoulos *et al.*, 2002). Test results for this region are described in more detail in chapter 4. A final note on this method is the specifications used to

determine the sub-region boundaries. This will vary from region to region, but may be based on guidelines such as geographical location, topography, data distribution/sources. The process can be likened to the Helmert block approach often used in the adjustment of large geodetic networks to distribute the computational burden (see for example Wolf, 1978).

An alternative approach to modelling the discrepancies between the ellipsoidal, orthometric and geoid heights that is worth mentioning is the use of a finite element model (FEM). In the approach described by Jäger (1999), a FEM is parameterized by sets of bivariate polynomials with continuity conditions employed to deal with discontinuities between surface meshes. All of the combined height information ($h, H, \Delta H, \Delta h$, and N) contributes to the overall adjustment, which is used to compute the coefficients of the finite element model. A key difference between the designation of the meshes in this approach as compared to the sub-regions in the previous discussion is that the nodes and edges of the meshes are independent of the location of the data points. This affords the user a greater degree of flexibility in representing the corrector surface.

3.4 Selecting a parametric model

As evidenced by the brief overview in the previous section, there is a plethora of different models available for the corrector surface model. The type of parametric model suited for a particular set of control points may be completely incompatible for a different region. Therefore, the importance of empirical tests with real data cannot be stressed enough. However, experience has shown that some general guidelines can be followed in order to filter through the different models and test only a few for each case. The most important aspect is to avoid over-parameterization. A high degree surface may give unrealistic extrema in data voids where control points are missing. This is an important factor for the combined height problem in particular, as one of the most favourable locations to utilize

GNSS-levelling are in areas where it is difficult to establish vertical control and therefore data gaps are prevalent (see chapter 7).

In order to demonstrate the consequences of over-parameterization, consider, as an example, a bivariate polynomial fit of various orders (see the multiple regression equation in Eq. 3.28). The coefficients for four different orders of the model (first, second, third and fourth) were computed for the same set of vertical control points in the western Canadian provinces of Alberta and Saskatchewan. The results of the computed fits are depicted in Figure 3.2. Two sets of figures for each model are shown, with the figures on the left showing the behaviour of the computed corrector surface and the figures on the right depicting the control network with an overlay of the corrector surface. As evidenced by the dispersion of the data, in the northern parts of the provinces there are virtually no GPS-levelling benchmarks. This is particularly evident in the northeastern quadrant (north of 56°N and east of 100°W). Focusing on this area, it is clear that as the order of the polynomial fit increases from first to fourth, the more exaggerated are the artifacts produced by the model. At a much lower level, the same behaviour is evidenced in the southwestern corner where there is also a gap in the data distribution.

Although it is obvious, from visual inspection, that over-parameterization leads to poor results in some areas, it often goes unnoticed in practice as models are usually tested by comparing the computed values to existing control. In these areas, all models seem to perform well and there is no erratic behaviour. In fact, as a general rule of thumb for prediction of MRE surfaces, one should aim to derive the simplest model that adequately fits the data (Lancaster and Šalkauskas, 1986).

Details on the different tests that can be conducted to assess model performance will be provided in the following section. However, it is appropriate to describe some model characteristics that may help or hinder the assessment process. In theory, the decision on the degree of the polynomial/MRE surface should be reached by hypothesis testing (Dermanis and Rossikopoulos, 1991).

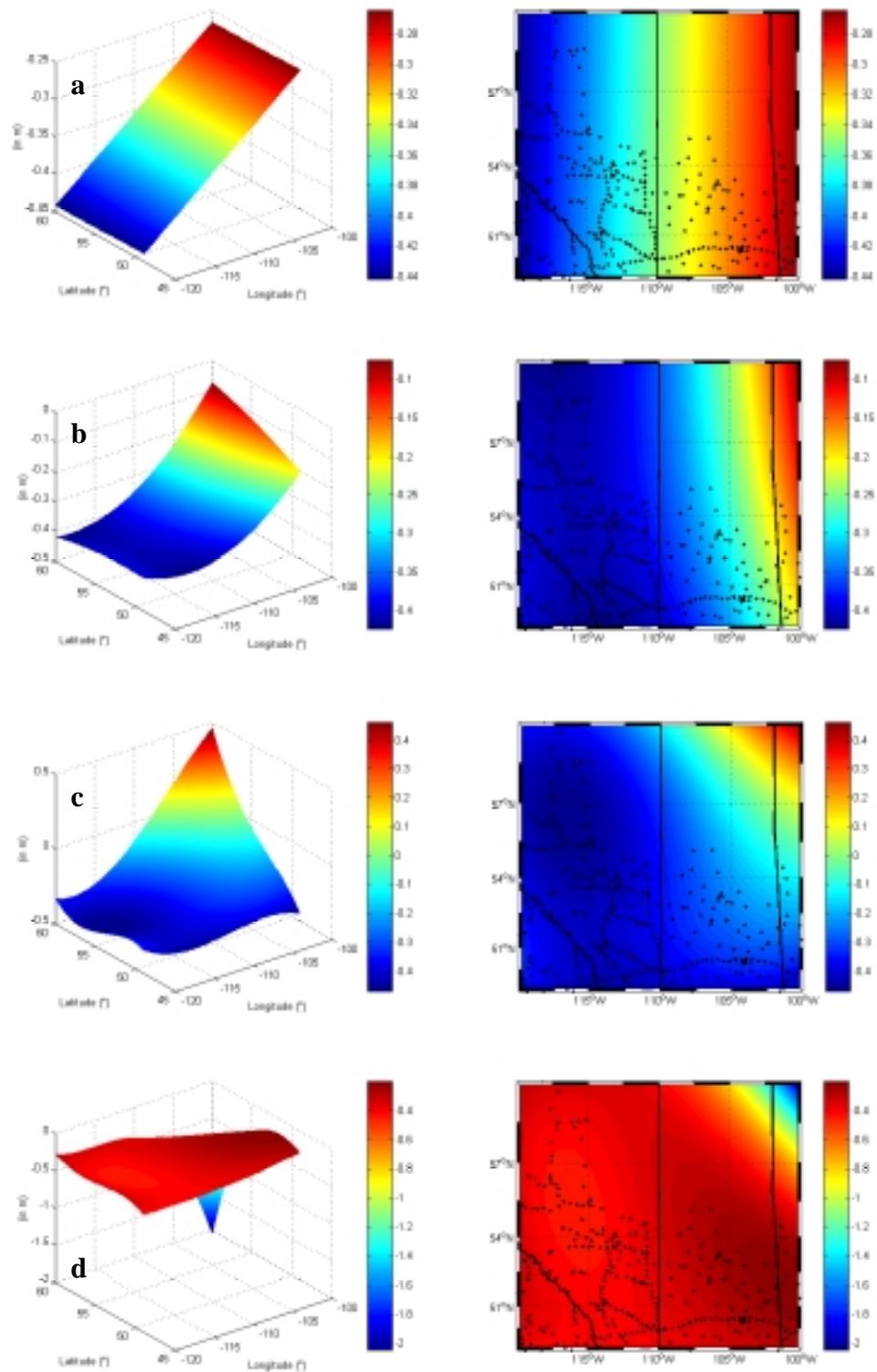


Figure 3.2: First (a), second (b), third (c) and fourth (d) order bivariate polynomial fits

However, the results of such statistical tests are often hindered by the fact that independent coefficients generated by a polynomial series are usually correlated with one another. Therefore, it is worth considering models with orthogonal base functions, which ensures no correlation between coefficients. Some classic orthogonal polynomials include Legendre, Tschebyscheff of first and second kind, Jacobi, Laguerre and Hermite (Davis, 1963, ch. 10). If the application of these models is not suitable or too complex for practical use, then one can also apply orthogonalization/orthonormalization procedures to decorrelate existing base functions. A common orthonormalization procedure that is relatively simple to implement in practice is the Gram-Schmidt orthonormalization method (see Carroll and Green, 1976; ch. 3 for more details), which is employed in this work and the results will be discussed in chapter 4.

Finally, a very useful guideline to follow, if possible, is to select a set of nested models as opposed to non-nested models. Two models are nested if one can be derived from the deletion of some of the terms in the other model. The imposition of such a criterion for a set of models to be tested greatly facilitates the assessment process as demonstrated in the following section.

3.5 Assessing the parametric model performance

In general, the process applied for selecting the best parametric model in a particular region suffers from a high degree of arbitrariness in both choosing the model type and in assessing its performance. In order to address this daunting issue, several tests are presented, which can be applied to the results of the combined least-squares adjustment of the ellipsoidal/orthometric/geoid heights. More specifically, they include the following:

- classic empirical approach
- assessing the goodness of fit
- cross-validation
- testing parameter significance

Each of these steps will be described in detail below. It is assumed throughout the process that reliable information for the statistical behaviour of the ellipsoidal, geoid and orthometric height data is available and any gross errors/blunders have been detected and removed from the observational data in order for the results to be meaningful.

3.5.1 Classical empirical approach

The most common method used in practice to assess the performance of the selected parametric model(s) is to compute the statistics for the adjusted residuals after the least-squares fit. The adjusted residuals for each station in the network, \hat{v}_i , are computed as follows:

$$\hat{v}_i = h_i - H_i - N_i - \mathbf{a}_i^T \hat{\mathbf{x}} \quad (3.39)$$

The model that results in the smallest set of residuals is deemed to be the most appropriate ('best' fit). Figure 3.3 depicts a typical series containing the original height misclosures (as computed from Eq. 3.8) and the adjusted residuals after the fit (Eq. 3.39). Of note is the reduction in the average value to zero imposed by the least-squares adjustment. In effect, these values give an assessment of the *precision* of the model as they indicate how well the data sets fit each other.

In Figure 3.4, this classic empirical approach is illustrated. One of the main problems encountered when using this empirical method as the sole means for selecting between different models is that the lowest RMS usually corresponds to the highest order model. In fact, as the number of parameters in the corrector surface model increases, the associated root mean square (RMS) decreases. This is expected as the parameters absorb more of the differences (Fotopoulos *et al.*, 2001b). Therefore, this method is valid for testing the precision of the model, but it should not be interpreted as the *accuracy* or the prediction capability of the model.

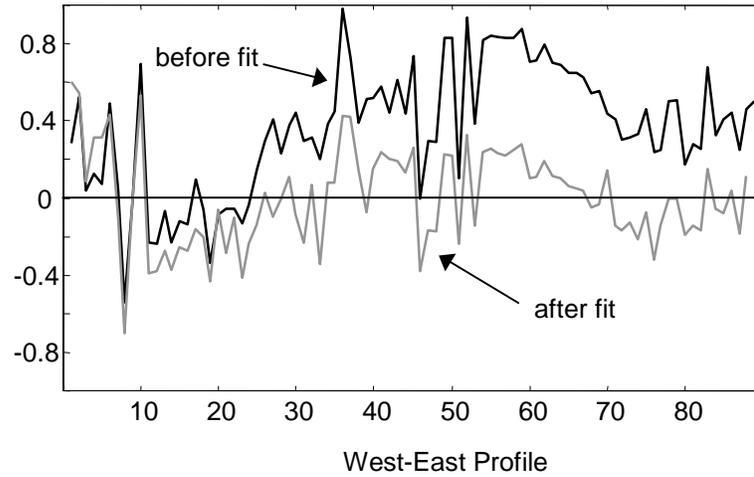


Figure 3.3: Example of height misclosures before and after parametric model fit

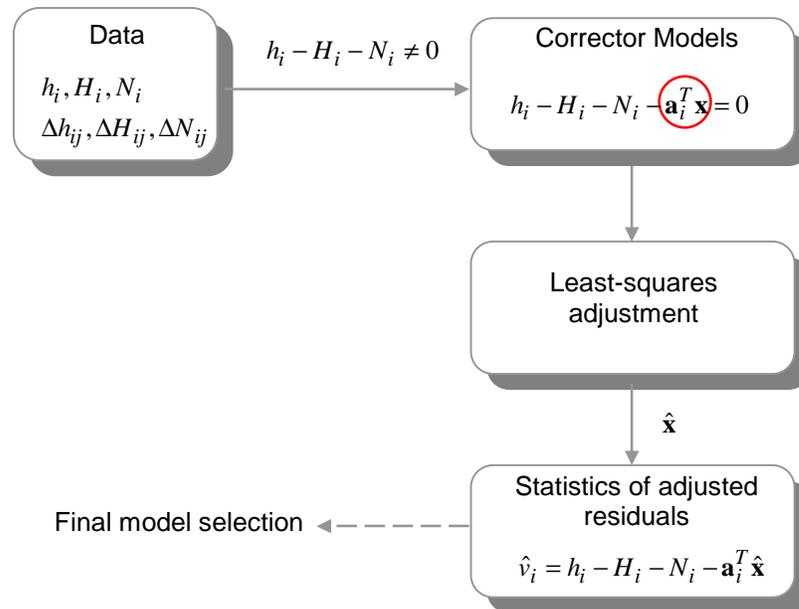


Figure 3.4: Classical empirical testing approach

3.5.2 Cross-validation

An additional empirical approach that can be used to complement the previous method and obtain a more realistic measure of the accuracy of the models is known as cross-validation. The general process can be summarized in four steps:

- (i) select a subset of vertical control points in the area of interest
- (ii) use the selected points in the combined least-squares adjustment to compute the model parameters, $\hat{\mathbf{x}}$
- (iii) use the computed model to predict the residual values at new points, not included in the original subset
- (iv) compare the predicted values from step (iii) with the 'known' height misclosures

One important practical problem with using this approach is that the results are dependent on the accuracy of the subset of points used for the comparison. Often the accuracy of the points are not known. Also, it is preferable to use as much data as possible in order to compute the unknown parameters. To alleviate these issues, the approach adopted herein involves a slight modification of the above steps as follows:

- (i) select all but one point, i.e., point P in Figure 3.5
- (ii) use the subset of points (encircled in Figure 3.5) in the combined adjustment to compute the model parameters, $\hat{\mathbf{x}}$
- (iii) use the computed model to predict the residual value at point P
- (iv) compare the predicted values with the known height misclosure at P ,

$$\Delta \hat{v}_p = h_p - H_p - N_p - \mathbf{a}_p^T \hat{\mathbf{x}}$$

- (v) repeat (i)-(iv) for each GPS-leveilling benchmark in the network and compute the

average RMS by $\frac{1}{m} \sum_{i=1}^m \sqrt{\mu_i^2 + \sigma_i^2}$

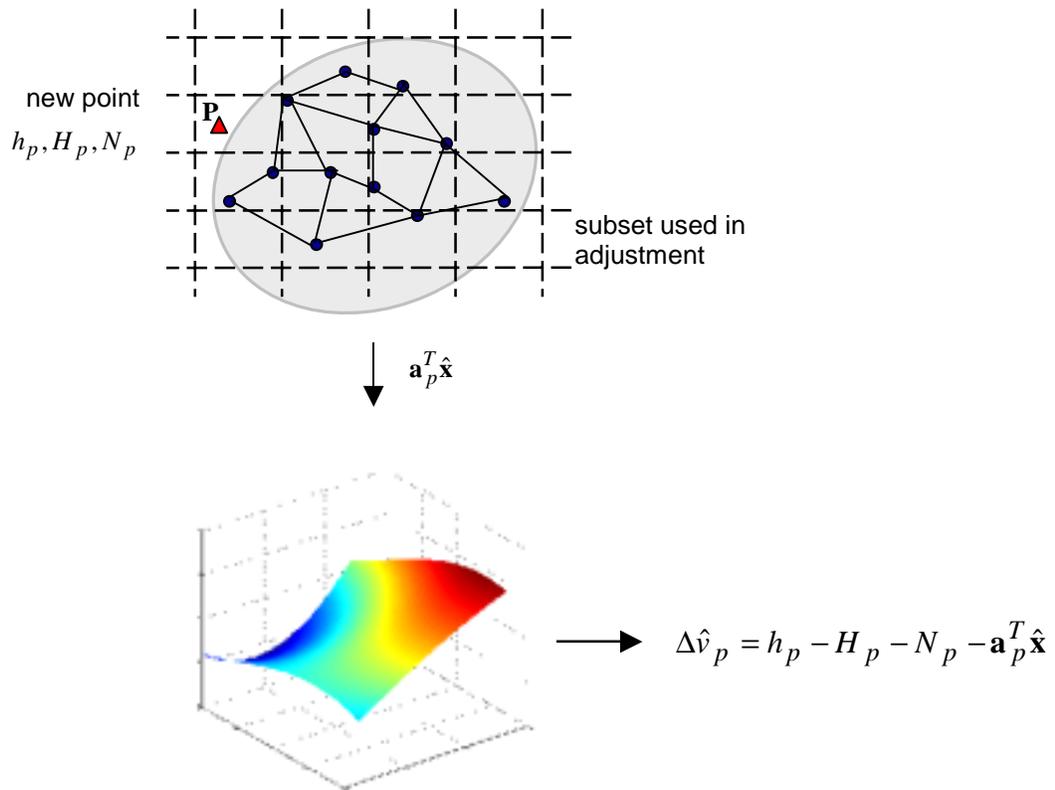


Figure 3.5: Cross-validation procedure on a point-by-point basis

The average RMS computed in this manner provides a more realistic indication of the *accuracy* of the selected parametric model and its performance as a prediction surface for a new point. It is the preferred empirical testing scheme, as it does not rely exclusively on the accuracy of a single point or a small subset of points. It also maintains high data redundancy to compute the parameters in the combined least-squares adjustment.

3.5.3 Assessing the goodness of fit

A statistical measure of the goodness of the parametric model fit for a discrete set of points is given by the *coefficient of determination*, denoted by R^2 . It can be described as

the ratio of the sum of the squares due to the fit, to the sum of the squares about the mean of the observations, as follows (Sen and Srivastava, 1990):

$$R^2 = 1 - \frac{\sum_{i=1}^m (\ell_i - \hat{v}_i)^2}{\sum_{i=1}^m (\ell_i - \bar{\ell}_i)^2} \quad (3.40)$$

where m observations are given by Eq. (3.8), the adjusted residuals for each station in the network are computed using Eq. (3.39), and $\bar{\ell}_i$ is the mean value of the observations.

In the extreme case where the parametric model fit is perfect, $\sum_{i=1}^m (\ell_i - \hat{v}_i)^2 = 0$ and $R^2 = 1$. The other extreme occurs if one considers the variation from the residuals to be nearly as large as the variation about the mean of the observations resulting in the fractional part in Eq. (3.40) to approach unity and $R^2 \rightarrow 0$. Thus, the coefficient of determination varies between 0 and 1 ($0 \leq R^2 \leq 1$) and the closer the value is to one, the smaller the residuals and hence the better the fit.

At first glance, R^2 may be regarded as a single-value indicator of the goodness of fit. However, it is important to realize that R^2 is a statistic and as with all statistics its values are somewhat governed by chance and peculiarities in the data (Wesolowsky, 1976). A relevant example to consider is the case where the data redundancy (or degrees of freedom) is small. In such cases, it is possible to obtain an erroneously large R^2 value, regardless of the quality of the fit. In fact, as the number of explanatory variables in the model (i.e., coefficients) increases, so does R^2 (*ibid.*). To deal with this limitation, a new statistic can be computed which is corrected for the degrees of freedom. This value, denoted by R_{α}^2 , is called the *adjusted coefficient of determination* and is computed as follows (Sen and Srivastava, 1990):

$$R_{\alpha}^2 = 1 - \frac{\left[\sum_{i=1}^m (\ell_i - \hat{v}_i)^2 \right] / (m-u)}{\left[\sum_{i=1}^m (\ell_i - \bar{\ell}_i)^2 \right] / (m-1)} \quad (3.41)$$

where u is the number of parameters in the model. The relationship between the two statistics is easily derived from relating Eqs. (3.40) and (3.41) and determined to be

$$R_{\alpha}^2 = 1 - \frac{(m-1)}{(m-u)} (1 - R^2) \quad (3.42)$$

This relationship between the two statistics is also depicted in Figure 3.6 for a sample of $m = 100$, where the coefficient of determination takes on constant values of 0.2, 0.5 and 0.9.

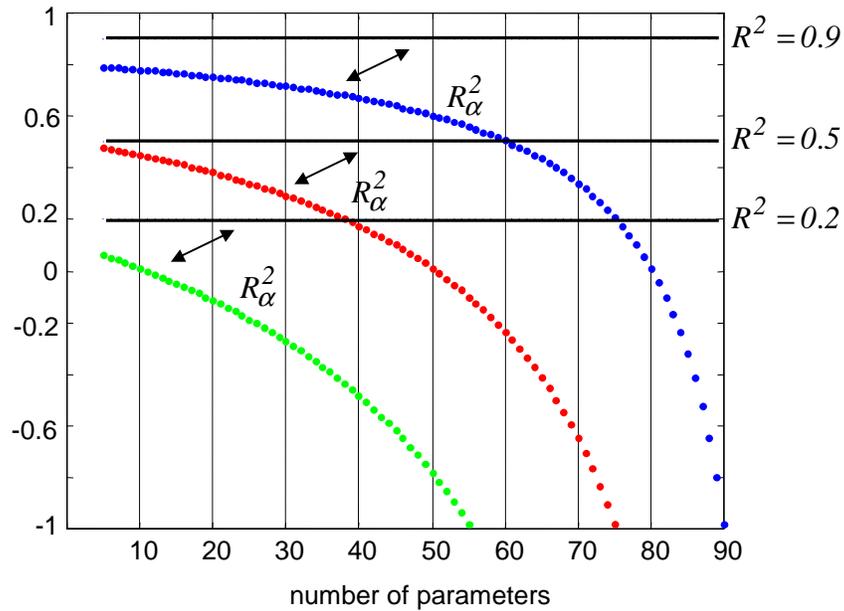


Figure 3.6: Relationship between R^2 and R_{α}^2

It is evident that as the number of parameters increases, the gap between R^2 and R_α^2 also increases. In general, the adjusted coefficient of determination also varies as $0 \leq R_\alpha^2 \leq 1$, with a good fit implied as R_α^2 approaches 1. However, there are cases where R_α^2 may give a negative value (also shown in Figure 3.6). This is a confusing result that may lead to incorrect interpretations.

Given the limitations of both measures of fit, it is important to not rely exclusively on these values. Instead, the values should be computed and accompanied by a reasonable interpretation and additional tests, such as the empirical procedures described in the previous two sections. Realizing that the coefficient of determination may give a high or low value as a result of chance or peculiarities in the data, the user can be more critical of its result. For instance, another common effect encountered in practice is the result of a low R^2 due to the fact that there was not enough variation in the observations to justify a 'good' or 'bad' fit. Therefore, these statistical measures can be a powerful tool in pointing out inappropriate models rather than establishing the validity of the model, which can be further tested by empirical cross-validation.

It is important to recognize that all statistical measures/tests depend on the geometry of the data (i.e., \mathbf{A} matrix). Therefore, the results of the tests will vary as the network configuration changes. The emphasis is therefore placed on establishing a *procedure* that allows for the selection of the most suitable model. Examples of relevant research in this area include de Bruijne *et al.* (1997), where statistical tests are applied in tandem with empirical fits to the data to find the best representation for the discrepancies between the gravimetric and GPS-derived geoid models in the North Sea region. In another case, using Australian data, the deciding factors/procedure for assessing the goodness of fit involved statistics of the differences (sections 3.5.1 and 3.5.2), number of outliers, mean discrepancy over all possible baselines and the number of height residuals beyond the acceptable errors defined for third order spirit-levelling (Featherstone, 2000). In any case,

some user intervention is required as it is important to evaluate all sets of criteria with regards to the purpose/objectives of the combined height adjustment. In the following section, the final module of the selection procedure is presented.

3.5.4 Testing parameter significance

As discussed previously, one of the key issues in the selection of a parametric model is the avoidance of over-parameterization. In line with this frame of thought is the principle of parsimony commonly referred to in statistical literature, where one should not use any more entities, beyond what is necessary, to explain anything. In this case, for the sake of simplicity, computational efficiency, and to avoid the effects of over-parameterization, the significance of each parameter in the selected model should be tested. Unnecessary terms may bias other parameters in the model, which will hinder the capability to assess the model performance (Wesolowsky, 1976). Therefore, the main focus of this section is to develop a procedure for testing, identifying and eliminating unnecessary terms to form a more simplified trend model. Due to the number of possibilities and the computational effort involved in such a task, it was determined that an automated procedure for testing parameter significance would be sought. In general, there are three schemes that can be implemented, namely (i) backward elimination, (ii) forward selection and (iii) stepwise procedure. All three procedures require that the models tested are nested as described in section 3.4.

Backward elimination

In the backward elimination procedure, one begins by fitting to the data the most extended (highest order) form of the model. The next step is to test if a parameter or set of parameters in the model are significant. The vector of parameters can be separated and denoted by

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_I \\ \mathbf{x}_{(I)} \end{bmatrix} \quad (3.43)$$

where \mathbf{x}_I is the set of parameters to be tested and $\mathbf{x}_{(I)}$ are the remaining parameters (complement) in the model. The test is specified by the null hypothesis (H_o) that states which parameter(s) are insignificant versus the alternative hypothesis (H_a) that declares these parameters to be significant, denoted as follows:

$$H_o : \mathbf{x}_I = \mathbf{0} \quad \text{vs.} \quad H_a : \mathbf{x}_I \neq \mathbf{0} \quad (3.44)$$

The statistic used to test this null hypothesis is the *F-statistic* computed as a function of the observations (Dermanis and Rossikopoulos, 1991)

$$\tilde{F} = \frac{\hat{\mathbf{x}}_I^T \mathbf{Q}_{\hat{\mathbf{x}}_I}^{-1} \hat{\mathbf{x}}_I}{k \hat{\sigma}^2} \quad (3.45)$$

where, $\mathbf{Q}_{\hat{\mathbf{x}}_I}^{-1}$ is the corresponding sub-matrix of the inverse of the normal equations,

$\mathbf{Q}_{\hat{\mathbf{x}}} = \mathbf{N}^{-1}$, k is the number of parameters tested, and $\hat{\sigma}^2$ is the a-posteriori variance factor.

The null hypothesis is accepted when

$$\tilde{F} \leq F_{k,f}^\alpha \quad (3.46)$$

$F_{k,f}^\alpha$ is computed from standard statistical tables for a confidence level α and degrees of freedom f (see Papoulis, 1990 and Koch, 1999 for details). If Eq. (3.46) is fulfilled then the corresponding parameters are deleted from the model. If the contrary is true, i.e.,

$$\tilde{F} > F_{k,f}^\alpha \quad (3.47)$$

the 'tested' parameters remain in the model. The procedure is repeated until all of the remaining parameters in the model pass the F -test or the user is satisfied with the final model.

An alternative equation for computing the F -statistic is given by (Wesolowsky, 1976):

$$\tilde{F} = \frac{\left[\sum (\ell - \hat{v})_{partial}^2 - \sum (\ell - \hat{v})_{full}^2 \right] / k}{\left[\sum (\ell - \hat{v})_{full}^2 \right] / m - u} \quad (3.48)$$

where the subscripts *full* and *partial* denote the values computed using all of the parameters in the model (highest order) and the *partial* denotes the values computed if the 'tested' parameters are eliminated. This statistic, termed the *partial F-test*, is commonly implemented for testing regression parameters. However, in this case, Eq. (3.45) was preferred as it allows for the significance of parameter(s) to be scrutinized and eliminated without the need to repeat the combined least-squares height adjustment.

Forward selection

The forward selection procedure begins with the most simple model (lowest order) and follows the same testing procedure as above, but for parameter addition. Thus, in this case the pair of hypotheses are given by

$$H_o : \mathbf{x}_* = \mathbf{0} \quad \text{vs.} \quad H_a : \mathbf{x}_* \neq \mathbf{0} \quad (3.49)$$

where, \mathbf{x}_* are parameters not included in the model. If the *F-test* passes, then the additional parameters are not added to the model. If, however, it fails, then the alternative hypothesis is accepted and the parameters are added to the model. The procedure is repeated until no more additional parameters can be added, or the user is satisfied that all useful parameters have been incorporated. An alternative test for determining if the parameter(s) should be added is to compute the coefficient of determination and add the parameter corresponding to the highest R^2 value.

Stepwise procedure

The stepwise procedure is a combination of both the backward elimination and forward selection procedures. Essentially one begins with the most simple model, as in the forward selection procedure, and then selects parameters one-by-one or several at a time. After inclusion, each parameter is examined for significance using the *F-test* described above. Embedded in this process is the backward elimination procedure for deleting parameters. An important issue is to select the same α -level for both the forward and backward schemes, otherwise the overall process will be counter-productive. The procedure is summarized in Figure 3.7. The challenging aspect in implementing any of these parameter significance tests is the fact that there is no unique answer. The final selection will be based on where one starts. Also, depending on the selected level of significance, different conclusions can be drawn.

In general, a higher α is set to err on the side of inclusion, whereas a lower value is selected if one wants to be careful about not including inappropriate parameters. The most limiting obstacle is parameter correlation, which may skew results and causes a number of problems, namely the standard errors of regression coefficients increase, computational/numerical difficulties arise, and biased estimates for the remaining regression parameters result if the missing parameters were correlated with the others.

One way to deal with this is to detect and delete highly correlated parameters. Another viable alternative, opted for herein, is to orthogonalize the model. If this procedure is applied properly, statistical tests can be performed without the consequences of multicollinearity.

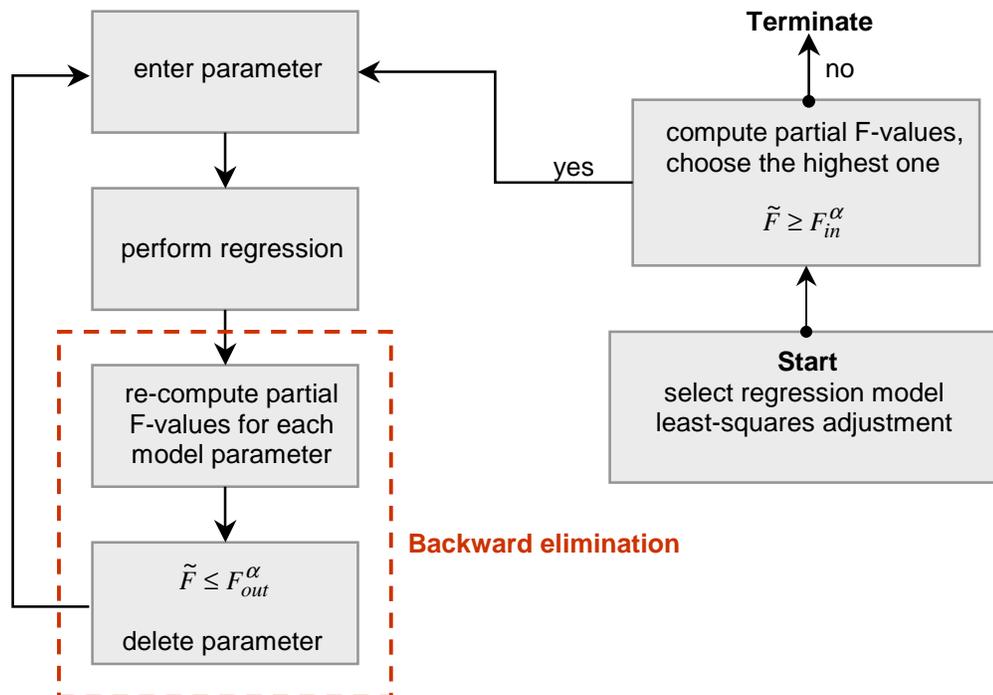


Figure 3.7: Stepwise procedure for testing parameter significance

3.6 Summary

The combined least-squares adjustment scheme described in section 3.1 has been implemented in a software program. The candidate parametric models employed include Eq. (3.28), Eq. (3.31), Eq. (3.33) and Eq. (3.34). The program is designed such that modifications to include additional models can easily be made. An amalgamation of all the steps described above to assess the selected parametric model performance (sections 3.5.1, 3.5.2, 3.5.3 and 3.5.4) has also been implemented in a (semi)-automated program.

The term *(semi)-automated* is used to describe the procedure as some user intervention is required. The major steps of the parametric model testing procedure (also used for the numerical tests in the next chapter) are summarized in Figure 3.8.

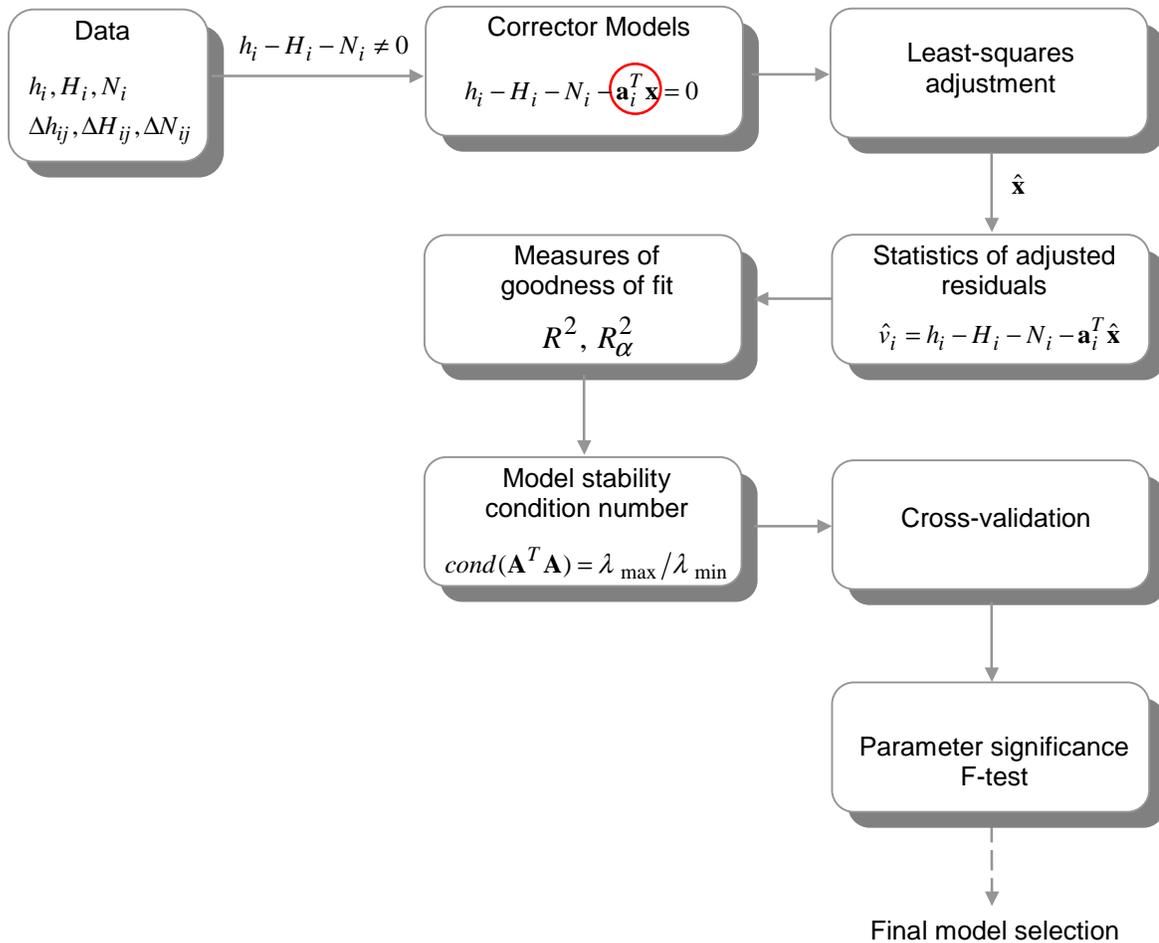


Figure 3.8: Flowchart of major steps for assessing the performance of parametric corrector surface models

Chapter 4

Results for the Parametric Model Surface Fits

The focus of this chapter is to apply the combined adjustment scheme and parametric model testing procedure described in the previous chapter with real data sets from various networks. The results for three different networks, namely Switzerland, Canada and Australia, are analyzed. Each test network poses a different set of challenges and the results reveal practical aspects of the problem of deriving optimal transformation models. Consequently, the discussion of the results for each test network will focus on the unique aspects/issues involved in each case.

For instance, the results for the Swiss network, described in section 4.1, will focus on the application of the main procedural steps summarized in Figure 3.8, with the exception of testing parameter significance as it was determined that it was unnecessary. In section 4.2, the results for the Canadian regional test network demonstrate the utility of testing parameter significance in addition to all of the other empirical and statistical tests. Finally, the results for the Australian national network (section 4.3) illustrate the need for regional analysis and the solution of creating a mosaic of corrector surfaces corresponding to different regions, motivated by the variable data distribution. Ultimately, a new vertical reference system, which utilizes heterogeneous height data from both satellite- and land-based methods, should be established in all networks. However, for the immediate and intermediate needs, and as a preliminary step, an

efficient method for transforming heights between different reference surfaces (even distorted ones) is required.

4.1 Results for the Switzerland test network

A test network consisting of 111 co-located GPS-levelling benchmarks distributed throughout Switzerland (a $330 \text{ km} \times 210 \text{ km}$ region) was used for the numerical tests. The data distribution is depicted in Figure 4.1 along with the height misclosures at each point as computed from Eq. (3.8). The average value is 1.1 cm ranging from a minimum of -4.9 cm and a maximum of 19 cm with an overall RMS of 4 cm.

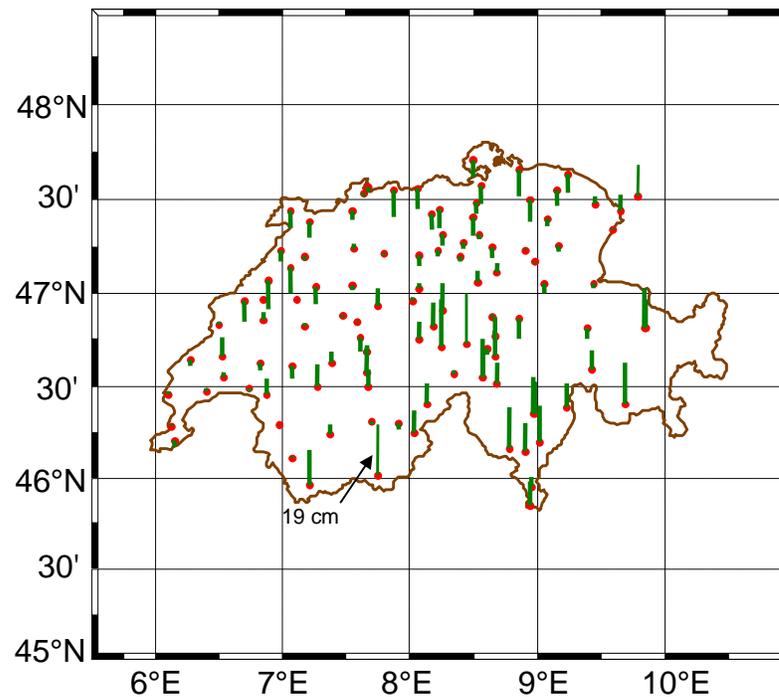


Figure 4.1: Swiss test network of GPS-on-benchmarks and original height misclosures

Details about the differences between the height systems and studies on the development of a consistent national height system in Switzerland are given in Marti *et al.* (2000) and Marti (2002), respectively. The network provides an excellent testbed for different parametric models as the data distribution is relatively consistent throughout the country with an average spacing of approximately 20 km between control points. In general, the height misclosures for the Switzerland network are not as variable as in other networks (see, e.g., Canadian region, section 4.2 and Australian network, section 4.3), with a range of almost 24 cm.

The parametric model assessment procedure summarized in chapter 3 is tested using two pre-specified families of corrector surfaces, namely:

- (i) nested bivariate polynomial series models up to fourth degree, e.g.

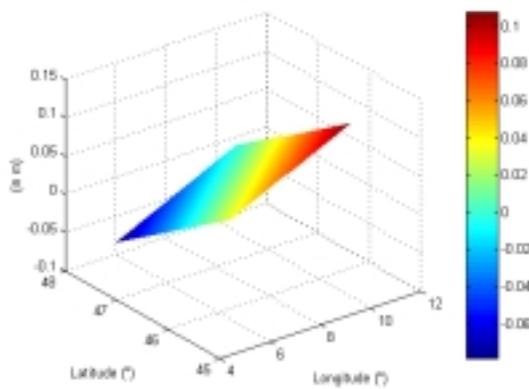
$$p(\varphi, \lambda) = x_1 + x_2\Delta\varphi + x_3\Delta\lambda + x_4\Delta\varphi\Delta\lambda + x_5\Delta\varphi^2 + x_6\Delta\lambda^2 + x_7\Delta\varphi^2\Delta\lambda + x_8\Delta\varphi\Delta\lambda^2 + x_9\Delta\varphi^3 + x_{10}\Delta\lambda^3 + x_{11}\Delta\varphi^2\Delta\lambda^2 + x_{12}\Delta\varphi^3\Delta\lambda + x_{13}\Delta\varphi\Delta\lambda^3 + x_{14}\Delta\varphi^4 + x_{15}\Delta\lambda^4$$

- (ii) similarity-based transformation model and its more simplified forms (Eqs. 3.31, 3.33 and 3.34).

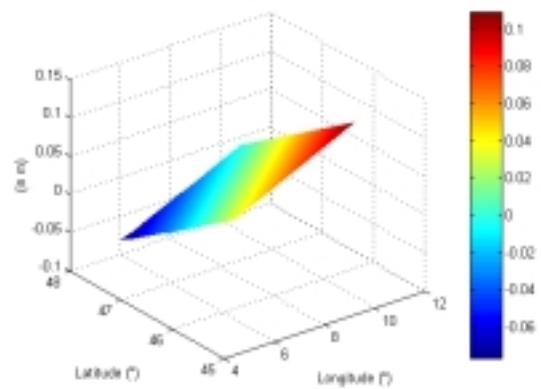
The process works best if the most extended form of the corrector surface is first estimated and then the nested variations of the model can be assessed by properly applying the statistical and empirical tests. The computed corrector surfaces corresponding to all seven models for the Switzerland network are given in Figures 4.2a and 4.2b (note the different scales). The lower-order models (first and second) are provided in Figure 4.2a and the higher-order models (third and fourth) are plotted in Figure 4.2b, with the corresponding base functions for each model also noted next to each plot.

Table 4.1 summarizes the statistics referring to the internal precision of each model, as described in section 3.5.1. This also indicates the precision of the gravimetric geoid model according to the GPS-levelling benchmarks. Predictably, as the number of terms in

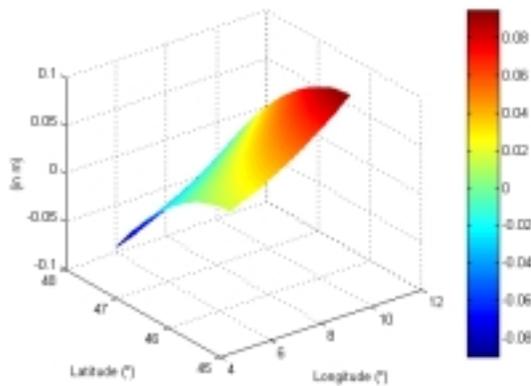
the model increases, the RMS of fit decreases/improves. From these initial results, the full fourth-order polynomial model seems to provide the lowest RMS value of 2.0 cm improving the original RMS with no fit by approximately 50%. In general, the results do not show a significant difference between models. Therefore, based on these first results all models seem to perform satisfactorily providing an average RMS of about 2.3 cm.



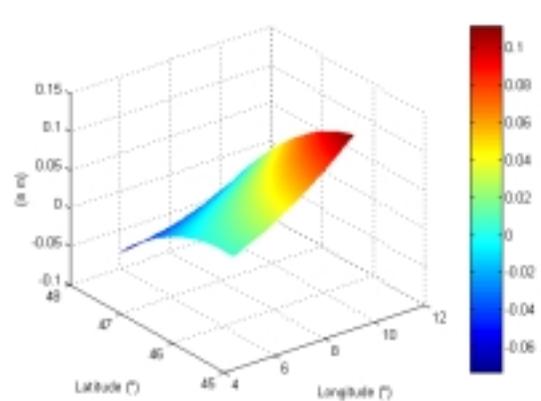
MODEL A. $1 \Delta\varphi \Delta\lambda$



MODEL B. $1 \cos\varphi \cos\lambda \cos\varphi \sin\lambda \sin\varphi$

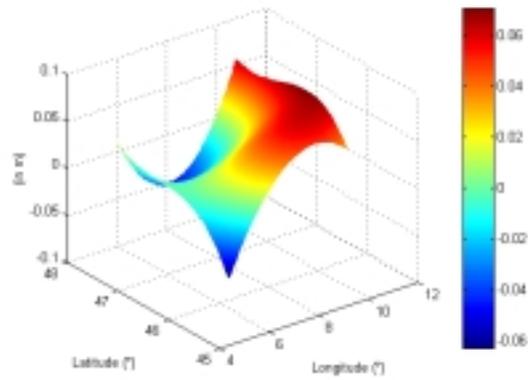


MODEL C. $1 \cos\varphi \cos\lambda \cos\varphi \sin\lambda \sin\varphi \sin^2\varphi$



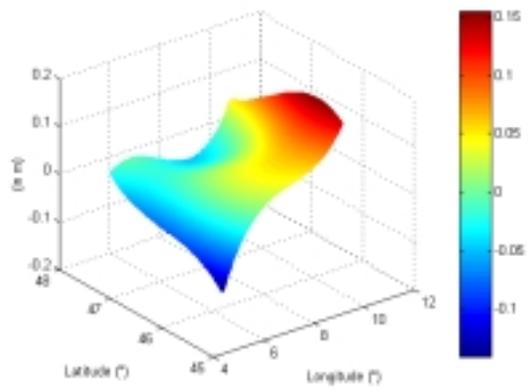
MODEL D. $1 \Delta\varphi \Delta\lambda \Delta\varphi\Delta\lambda \Delta\varphi^2 \Delta\lambda^2$

Figure 4.2a: Low-order corrector surface model fits for the Swiss network



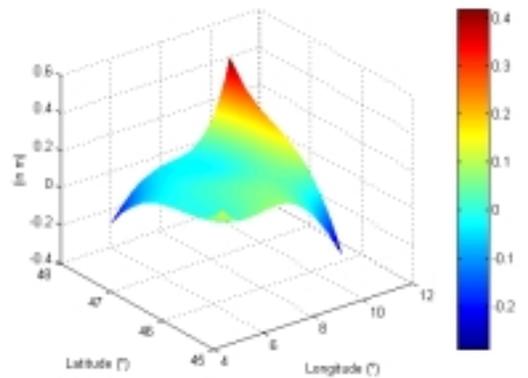
MODEL E.

$$\frac{\cos \varphi \cos \lambda}{\sin \varphi \cos \varphi \sin \lambda} \quad \frac{\cos \varphi \sin \lambda}{\sin \varphi \cos \varphi \cos \lambda} \quad \frac{\sin \varphi}{1 - f^2 \sin^2 \varphi} \quad \frac{\sin^2 \varphi}{W}$$



MODEL F.

$$1 \quad \Delta \varphi \quad \Delta \lambda \quad \Delta \varphi \Delta \lambda \quad \Delta \varphi^2 \\ \Delta \lambda^2 \quad \Delta \varphi^2 \Delta \lambda \quad \Delta \varphi \Delta \lambda^2 \quad \Delta \varphi^3 \quad \Delta \lambda^3$$



MODEL G.

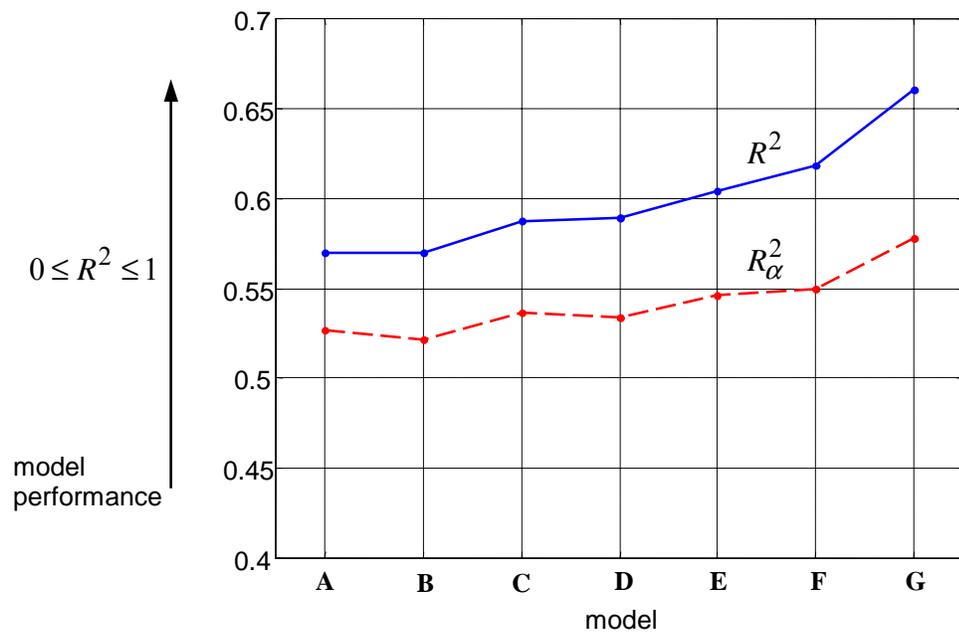
$$1 \quad \Delta \varphi \quad \Delta \lambda \quad \Delta \varphi \Delta \lambda \quad \Delta \varphi^2 \\ \Delta \lambda^2 \quad \Delta \varphi^2 \Delta \lambda \quad \Delta \varphi \Delta \lambda^2 \quad \Delta \varphi^3 \quad \Delta \lambda^3 \\ \Delta \varphi^2 \Delta \lambda^2 \quad \Delta \varphi^3 \Delta \lambda \quad \Delta \varphi \Delta \lambda^3 \quad \Delta \varphi^4 \quad \Delta \lambda^4$$

Figure 4.2b: High-order corrector surface model fits for the Swiss network

Table 4.1: Statistics of residuals at points used in the adjustment for various parametric models (Swiss network, units: cm)

Model	Min	Max	μ	σ	RMS
A. 1 st order polynomial	-5.8	6.7	0	2.3	2.3
B. 4-parameter	-5.8	6.7	0	2.3	2.3
C. 5-parameter	-6.2	6.4	0	2.3	2.3
D. 2 nd order polynomial	-6.1	6.4	0	2.3	2.3
E. differential similarity	-5.9	6.4	0	2.2	2.2
F. 3 rd order polynomial	-5.5	6.4	0	2.2	2.2
G. 4 th order polynomial	-5.6	6.0	0	2.0	2.0

Figure 4.3 shows the computed coefficient of determination and the adjusted coefficient of determination for the seven models tested.



A 1st order polynomial **B** classic 4-parameter **C** classic 5-parameter
D 2nd order polynomial **E** differential similarity **F** 3rd order polynomial
G 4th order polynomial

Figure 4.3: Statistical measures of goodness of fit for various parametric models (Swiss test network)

According to these measures, there is only a slight difference between the performance of the different models with a marginal variability in R^2 between 0.56 to 0.66 and the more indicative R_{α}^2 ranging from 0.53 to 0.57. Again, the one outstanding model is the fourth-order bivariate polynomial (model G), which corresponds to the highest measures of goodness of fit. Therefore, based on the combined results of these first tests, the fourth-order polynomial would be identified as being suitable for this region. Most studies would stop at this point and provide the RMS of fit (i.e., 2.0 cm) as an indication of the accuracy of the best fit model. However, further tests will show that a very different conclusion can be drawn.

Perhaps the most revealing test results are given in Table 4.2, which summarizes the results of the empirical cross-validation (or prediction) at independent control points (as described in section 3.5.2). The models are arranged according to the number of parameters increasing from top to bottom. From these values it is evident that the fourth-order polynomial fit is not the best choice if the model is to be used for computing heights at independent stations (which is usually the case). Furthermore, the results reveal that the optimal choice would be the classic 4-parameter fit (Eq. 3.31) with an overall RMS of 2.4 cm.

Table 4.2: Statistics of cross-validation tests for various parametric models (Swiss network, units: cm)

Model	Min	Max	μ	σ	RMS
1 st order poly.	-6.2	14.4	-0.02	2.8	2.8
4-parameter	-6.3	6.9	0	2.4	2.4
5-parameter	-6.6	13.2	-0.01	2.7	2.7
2 nd order poly.	-6.6	13.1	-0.02	2.8	2.8
differential similarity	-6.5	12.9	-0.01	2.7	2.7
3 rd order poly.	-6.3	11.0	-0.01	2.7	2.7
4 th order poly.	-10.0	7.1	-0.03	2.7	2.7

Evidently, two very different conclusions on the optimal model can be drawn from the same data depending on the criteria for testing. It should be mentioned however, that in this case there is only a minor difference in performance between each model. This is most likely due to the fact that the data is consistently distributed and exhibits rather small variations from point-to-point. Nonetheless, the use of a parametric model does reduce the original RMS of 4 cm by ~ 1 cm to just over 2 cm. Additional values of interest are the minimum and maximum in Table 4.2, which indicate the overall range of values for each parametric model. It is clear that the 4-parameter model is the best choice for reducing the overall range to between -6 cm to 7 cm.

The performance of the corrector surfaces can also be gauged on the numerical stability over the region of interest. Since, there is not a prominent variation in the achievable accuracy for each model, the condition number may provide some insight into the overall performance of the model. Table 4.3 summarizes the condition numbers corresponding to each of the parametric models for the Swiss test network. It is evident from Table 4.3 that, in general, the most stable models are those of the lowest order with fewer unknown parameters. The higher order models tend to be less stable and less accurate when applied at independent control points (as in GPS/geoid levelling).

Table 4.3: Condition numbers for various parametric models in Switzerland

model	number of terms	condition number
1 st order polynomial	3	1.66×10^4
4-parameter	4	6.20×10^8
5-parameter	5	8.53×10^9
2 nd order polynomial	6	2.56×10^8
differential similarity	7	3.71×10^{12}
3 rd order polynomial	10	9.05×10^{13}
4 th order polynomial	15	5.62×10^{16}

The next step in the assessment process would be to determine if any of the model parameters are insignificant using the procedure described in section 3.5.4. The procedure was carried out, however, it soon became obvious that it was not necessary. The selected model is of low order (2^{nd}) and consists of only four terms. Based on the collective results presented above, it was deemed appropriate to make the final decision of the classic 4-parameter fit (Eq. 3.31). The selection criteria are summarized in Table 4.4 and the final fit for the Switzerland network is given in Figure 4.4.

Table 4.4: Summary of selection criteria for Swiss test network

R^2	0.57
R_{α}^2	0.52
$\sqrt{\hat{\mathbf{v}}^T \hat{\mathbf{v}}}$	24.4 cm
condition number	6.20×10^8
RMS after fit	2.3 cm
RMS (prediction/cross-validation)	2.4 cm

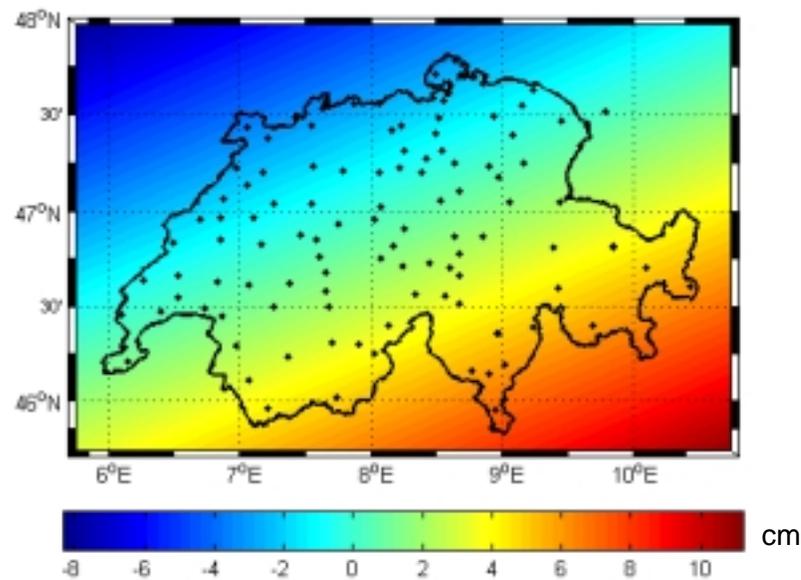


Figure 4.4: Classic 4-parameter corrector surface fit for the Swiss test network

4.2 Results for the Canadian test network

The Canadian national network of GPS-on-benchmarks consists of 1292 points scattered across the country as depicted in Figure 4.5. The harsh terrain and environmental conditions in northern areas hinders the capacity to establish precise vertical control using spirit-levelling techniques (Gareau, 1986). As a result, the distribution of the control points is relatively inconsistent with a high concentration in the southern parts of the country (along the Trans-Canada highway) and the surrounding populated areas.

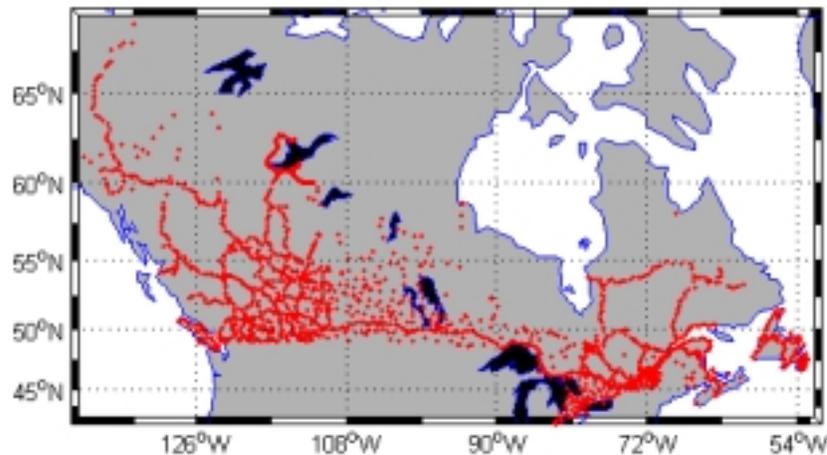


Figure 4.5: Canadian GPS-levelling benchmark network

For the results presented in this section, a subset of the GPS-on-benchmarks in Canada were used. Figure 4.6 shows this 495 km \times 334 km region situated in the southern British Columbia (BC) and Alberta (AB) provinces. In this region, 63 points are unevenly distributed with several gaps containing few or no control points (i.e., mid-western region between 50°N - 51°N and 122°W - 124°W). The original height misclosures for the geoid, orthometric and ellipsoidal heights are also indicated in the figure for each benchmark.

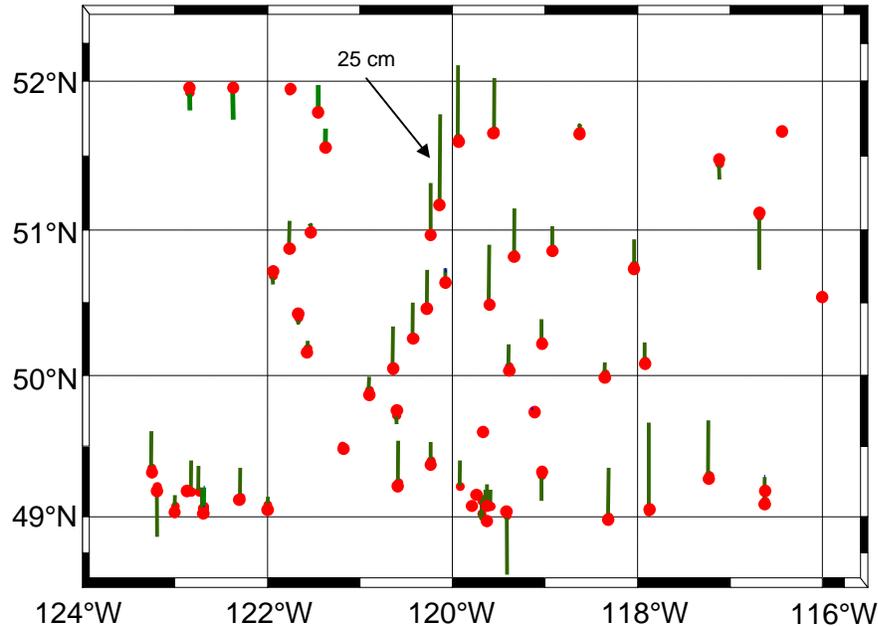


Figure 4.6: Test network of GPS-on-benchmarks in southern BC/AB and original height misclosures

The required geoid heights for this study were obtained through bi-quadratic interpolation of the CGG2000 geoid model at the GPS-levelling benchmark locations. CGG2000 is the latest gravimetric geoid model in Canada available on a $2' \times 2'$ grid. The official validation of the CGG2000 model was carried out using 1902 carefully selected GPS-levelling benchmarks across the country and gives an average value for the differences of -26 cm and a standard deviation of 17.9 cm. Furthermore, the computed discrepancies between the two coasts is 34 cm using this new model, compared to 81 cm with the previous geoid model, GSD95 (Véronneau, 1997). Although the discrepancies may be due to a number of sources (as described in section 3.2), for the Canadian data in particular, it is suspected that a large part of the misfit is due to the systematic errors in the primary levelling network. A detailed report describing the complete computational procedure is provided in Véronneau (2002). Of particular interest is the use of a modified Stokes' kernel, which filters the longer wavelength contributions from the surface gravity measurements that are plagued by systematic errors. These systematic errors in the

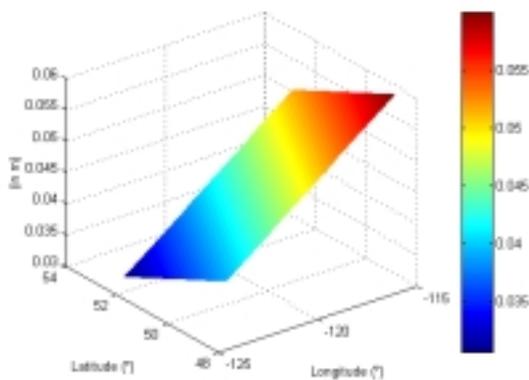
gravity measurements and reductions are attributed as the main cause of the prevalent systematic east-west trend in the GSD95 geoid model (Véronneau, 1997).

Currently, the published orthometric heights in Canada refer to the official local vertical datum, CGVD28. It was established using a network of more than 80,000 spirit-levelled benchmarks. The value for mean sea level was determined in 1928 by averaging sea level measurements at six tide gauge stations (two on the Pacific Ocean coast, two on the Atlantic Ocean coast, one on the St. Lawrence River and one in southern Quebec). Due to the sheer size of the country, levelling survey campaigns often span decades apart. This introduces a number of issues with monument stability, which is considerably affected by post-glacial rebound and sea level changes, to name a few. These combined effects can add up to several centimetres and even decimetres in some parts of the country. The orthometric heights provided for the 63 points in the southern BC/AB test network are Helmert orthometric heights as described in chapter 2.

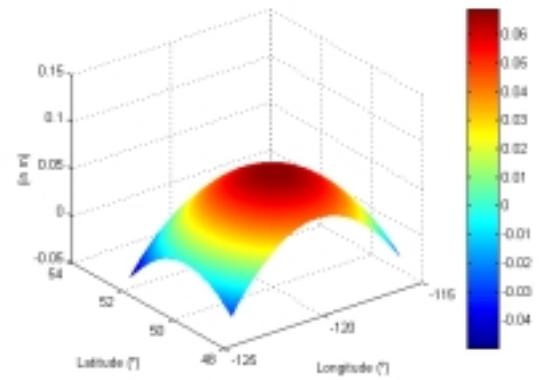
A major issue with the computation of the ellipsoidal heights from GPS campaigns prevalent in most large networks, such as those in Canada, is the inevitable mixture of different types of receivers and antennae, not to mention the measurements at different epochs. In some parts of the country, GPS coordinates have been determined using single frequency receivers, which ultimately provide less accurate ellipsoidal heights than more precise dual frequency techniques. Other factors that deteriorate ellipsoidal height determination, such as antenna phase centre offsets, tropospheric corrections and session length are also common problems with some parts of the national network (see discussion in section 2.3).

An interesting aspect of this test network is that there is a significant variation in the height misclosures, ranging from a minimum of -17.1 cm to a maximum of 25.2 cm with an average of 4.5 cm, standard deviation of 8.1 cm and an overall RMS of 9.3 cm. It is expected that the results of the corrector surface fits will be more variable than those for the Swiss network where the original misclosures were comparatively smooth.

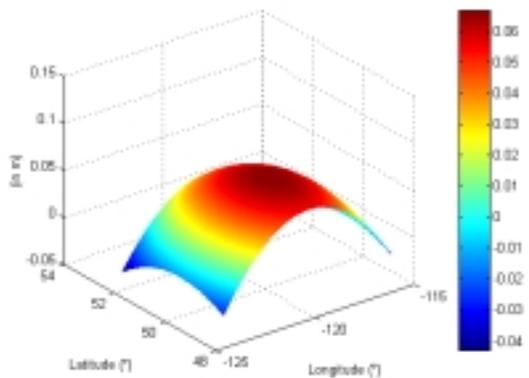
The same two families of parametric models used with the Swiss network were tested in the southern BC/AB region and the resulting corrector surfaces are depicted in Figures 4.7a and 4.7b for low-order and high-order fits, respectively. The statistics of the fits are given in Table 4.5, which shows the 'best-fit' model to be the 7-parameter differential similarity model, the third-order polynomial, and the fourth-order polynomial, with an improved RMS of ~ 6.7 cm compared to the original height misclosure of 9.3 cm.



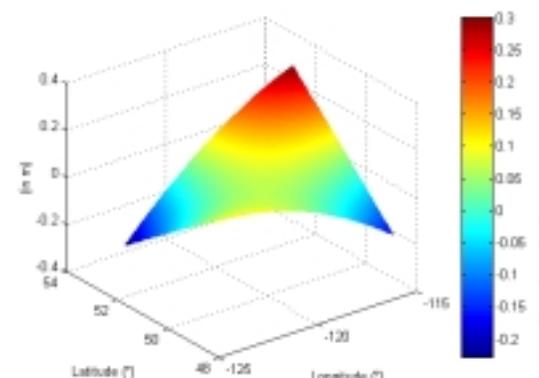
MODEL A. $1 \Delta\varphi \Delta\lambda$



MODEL B. $1 \cos\varphi \cos\lambda \cos\varphi \sin\lambda \sin\varphi$

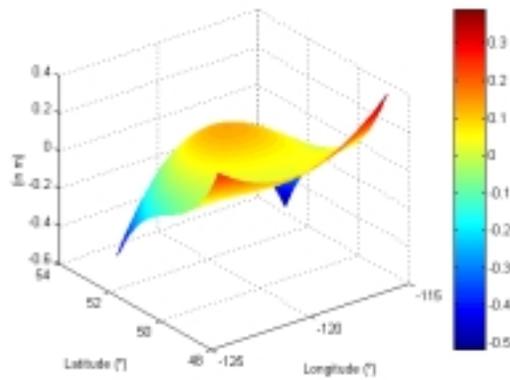


MODEL C. $1 \cos\varphi \cos\lambda \cos\varphi \sin\lambda \sin\varphi \sin^2\varphi$



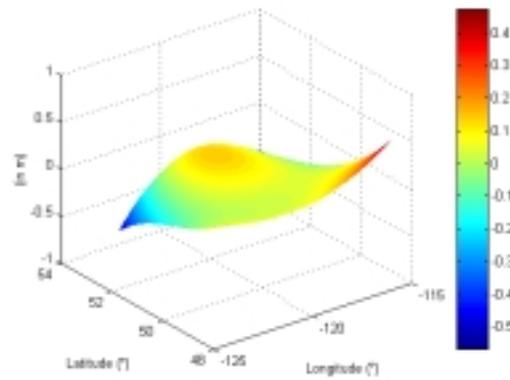
MODEL D. $1 \Delta\varphi \Delta\lambda \Delta\varphi\Delta\lambda \Delta\varphi^2 \Delta\lambda^2$

Figure 4.7a: Low-order corrector surface model fits for the southern BC/AB network



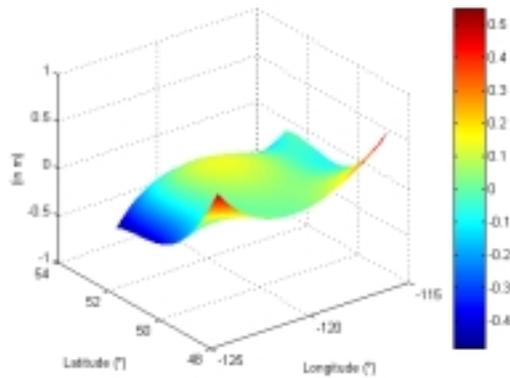
MODEL E.

$$\frac{\cos \varphi \cos \lambda}{W} \quad \frac{\cos \varphi \sin \lambda}{W} \quad \frac{\sin \varphi}{1 - f^2 \sin^2 \varphi} \quad \frac{\sin^2 \varphi}{W}$$



MODEL F.

$$1 \quad \Delta \varphi \quad \Delta \lambda \quad \Delta \varphi \Delta \lambda \quad \Delta \varphi^2 \\ \Delta \lambda^2 \quad \Delta \varphi^2 \Delta \lambda \quad \Delta \varphi \Delta \lambda^2 \quad \Delta \varphi^3 \quad \Delta \lambda^3$$



MODEL G.

$$1 \quad \Delta \varphi \quad \Delta \lambda \quad \Delta \varphi \Delta \lambda \quad \Delta \varphi^2 \\ \Delta \lambda^2 \quad \Delta \varphi^2 \Delta \lambda \quad \Delta \varphi \Delta \lambda^2 \quad \Delta \varphi^3 \quad \Delta \lambda^3 \\ \Delta \varphi^2 \Delta \lambda^2 \quad \Delta \varphi^3 \Delta \lambda \quad \Delta \varphi \Delta \lambda^3 \quad \Delta \varphi^4 \quad \Delta \lambda^4$$

Figure 4.7b: High-order corrector surface model fits for the southern BC/AB network

Table 4.5: Statistics of residuals at points used in the adjustment for various parametric models (southern BC/AB network, units: cm)

Model	Min	Max	μ	σ	RMS
A. 1 st order polynomial	-22.1	20.5	0	8.1	8.1
B. 4-parameter	-22.8	20.8	0	7.9	7.9
C. 5-parameter	-23.2	20.5	0	7.9	7.9
D. 2 nd order polynomial	-21.1	22.2	0	7.6	7.6
E. differential similarity	-21.6	17.3	0	6.7	6.7
F. 3 rd order polynomial	-21.1	17.2	0	6.7	6.7
G. 4 th order polynomial	-20.1	17.1	0	6.6	6.6

The corresponding measures of goodness of fit, R^2 and R_α^2 , depicted in Figure 4.8 clearly indicates the differential similarity model as a better choice than the third or fourth-order polynomial models according to R_α^2 .

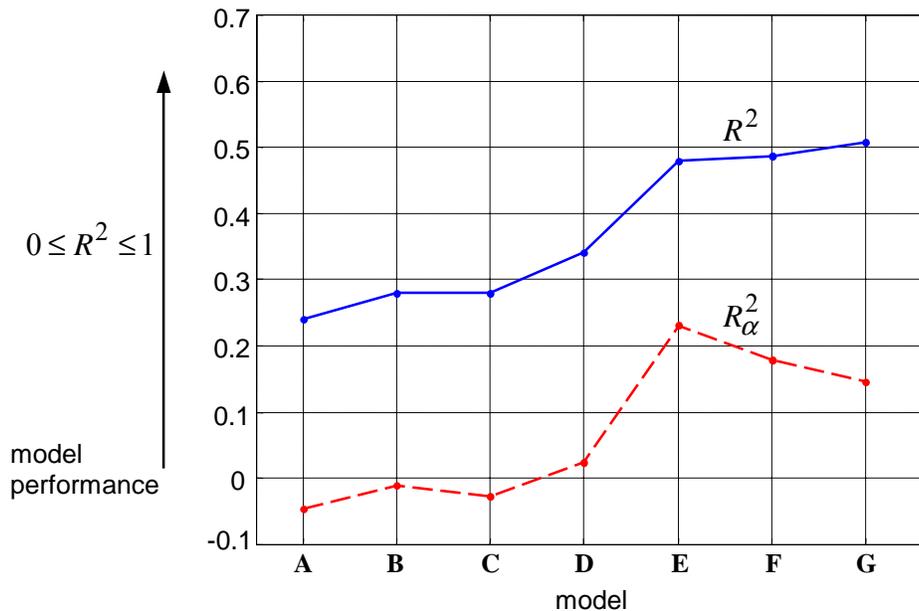


Figure 4.8: Statistical measures of goodness of fit for various parametric models (southern BC/AB test network)

The inflation caused in R^2 by an increase in the number of parameters from 7, 10 and 15 for the differential similarity, third, and fourth-order polynomial models, respectively, is insignificant compared to the jump experienced from the second-order polynomial (6 terms) to the 7-term differential similarity model. The inconclusive negative values obtained for the adjusted coefficient of determination for the lower-order models should also be noted.

The computed condition numbers, indicating the numerical stability of each of the parametric corrector surfaces in this region are provided in Table 4.6. The numerical stability is comparable to the results obtained for the Switzerland data, with the first-order polynomial model consisting of three terms corresponding to the lowest condition number (as expected).

Table 4.6: Condition numbers for various parametric models in southern BC/AB

model	number of terms	condition number
A. 1 st order polynomial	3	3.50×10^3
B. 4-parameter	4	4.53×10^7
C. 5-parameter	5	1.41×10^9
D. 2 nd order polynomial	6	2.37×10^7
E. differential similarity	7	1.52×10^{12}
F. 3 rd order polynomial	10	1.94×10^{11}
G. 4 th order polynomial	15	1.49×10^{15}

The empirical cross-validation results are summarized in Table 4.7. Of particular interest are the visible effects of over-parameterization exhibited by the behaviour of the fourth-order polynomial trend surface, which provides a high RMS of 13 cm. This is even an inferior result to not applying any parametric model. The third-order polynomial model performs close to the original misclosure RMS at the 9 cm-level. The model that gives the best prediction results is the differential similarity with an RMS of 6.7 cm. This 7-parameter model is the extended form of the classic 4-parameter model and takes into

account the change in the ellipsoidal parameters (GRS80 used in this case). Results after the model fit and prediction (cross-validation) empirical tests are illustrated in Figure 4.9, which clearly identifies the differences in performance between the models and the various conclusions that can be drawn regarding the best model depending on the type of test used.

Table 4.7: Statistics of cross-validation tests for various parametric models (southern BC/AB network, units: cm)

Model	Min	Max	μ	σ	RMS
A. 1 st order polynomial	-23.0	21.3	-0.1	8.7	8.7
B. 4-parameter	-23.7	22.3	0	8.5	8.5
C. 5-parameter	-25.3	22.4	-0.1	8.7	8.7
D. 2 nd order polynomial	-25.3	22.8	-0.1	8.8	8.8
E. differential similarity	-23.6	20.0	0	7.9	7.9
F. 3 rd order polynomial	-23.9	26.9	0.2	8.8	8.8
G. 4 th order polynomial	-26.2	51.8	1.3	13.0	13.0

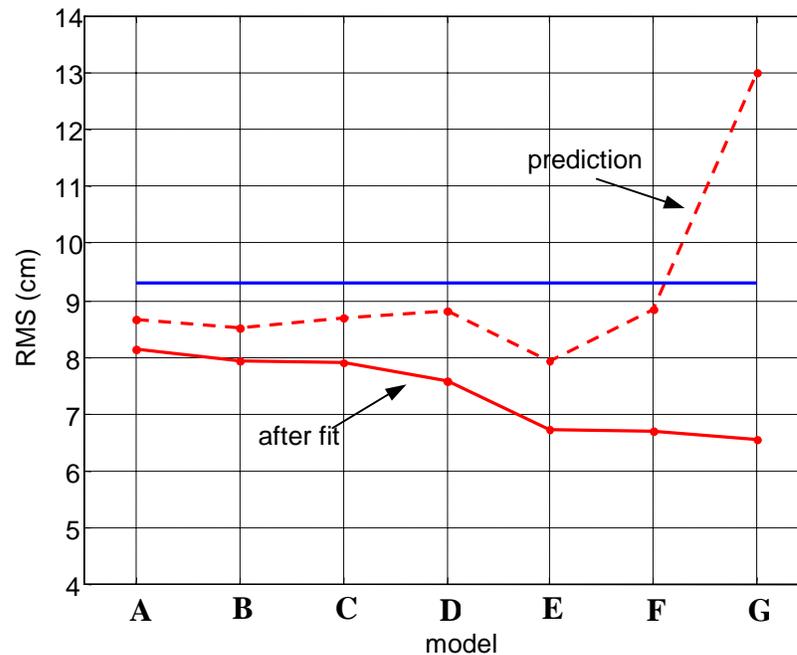


Figure 4.9: Empirical test results for the southern BC/AB region

In most practical cases, the truncated/approximated 4- or 5-parameter versions of the differential similarity model are implemented. To test the significance of the additional parameters in the full 7-parameter version of the model (Eq. 3.34), the backward elimination procedure described in section 3.5.4 was used. The first statistical test is implemented to determine if the fourth, fifth, sixth and seventh parameters are significant. The hypothesis is set up as follows:

$$H_o : [x_4 \ x_5 \ x_6 \ x_7]^T = \mathbf{0} \quad \text{vs.} \quad H_a : [x_4 \ x_5 \ x_6 \ x_7]^T \neq \mathbf{0}$$

where,

$$x_4 = \frac{\sin \varphi \cos \varphi \sin \lambda}{W}$$

$$x_5 = \frac{\sin \varphi \cos \varphi \cos \lambda}{W}$$

$$x_6 = \frac{1 - f^2 \sin^2 \varphi}{W}$$

$$x_7 = \frac{\sin^2 \varphi}{W}.$$

The computed \tilde{F} -value was 6.44 compared to the critical value obtained from the statistical tables of $F_{4,56}^{0.05} = 2.54$ and $F_{4,56}^{0.01} = 3.68$ for different levels of significance. In both cases, $\tilde{F} > F_{k,f}^\alpha$, and therefore the null hypothesis is rejected suggesting that all of the tested terms are significant. Additional *F-tests* were conducted, testing each of the seven parameters individually, i.e. $H_o : x_i = 0, i = 1, \dots, 7$, and the results indicated that all seven parameters are statistically significant.

Previously, it was mentioned that one must be cautious with the interpretation of these results as correlation among the model parameters may distort results. Consequently, the

model was re-formulated with a *new* set of orthonormal base functions using the Gram-Schmidt process (Carroll and Green, 1976; ch. 3), which gives a *new* set of uncorrelated parameters. Each new parameter was tested for significance by applying the same procedure as above. Surprisingly, it was found that for the orthonormal form of the model, only **two** of the total seven parameters were significant at the 99% confidence level ($\alpha = 0.01$). Using a 95% confidence level ($\alpha = 0.05$), **four** of the seven terms were deemed significant. Table 4.8 summarizes the statistics after the fit for the three versions of the orthonormalized parametric models (i.e., 7, 4, 2 terms). The RMS of fit is on the same level as those achieved using the models given in Table 4.5.

Table 4.8: Statistics of residuals at points used in the adjustment for orthonormalized versions of various parametric models (southern BC/AB network, units: cm)

Model	Min	Max	μ	σ	RMS
7-terms ortho.	-21.6	17.3	0	6.8	6.8
4-terms ortho.	-22.3	20.1	0	8.1	8.1
2-terms ortho.	-21.4	20.8	0	8.2	8.2

The main problem encountered for testing these orthonormalized versions of the model any further (i.e., via cross-validation) is the lack of analytical form. That is, the model parameters and system of base functions are derived for the particular data and cannot be interpreted in a specific format for testing at independent control points. In order to avoid this limitation, it is important to start with an orthogonal model/series of base functions (as listed in section 3.4). Then the same process of identifying and eliminating insignificant terms using the *F-test* can be applied as shown. The final 'modified' model, derived in this manner, can be tested at new points since the analytical form is preserved throughout the process. This is an important issue, which shows the power of statistical testing if applied to models consisting of orthogonal systems of bases.

The application of this elimination procedure for insignificant parameters allows the user to describe the behaviour of the 'corrections' over the coverage area with fewer terms. The results provide valuable insight for several practical implementations of heterogeneous height data. For instance, in the near future, GNSS-levelling will become a viable alternative for realizing vertical datums and for (near) real-time height-related applications. In this case, corrector surface parameters may be disseminated to users anywhere in the country similar to wide area differential GPS (WADGPS) corrections. The various dissemination options will be described in more detail in chapter 7.

The final test criteria for the selected model in this region are given in Table 4.9.

Table 4.9: Summary of selection criteria for southern BC/AB test network

R^2	0.48
R_α^2	0.23
$\sqrt{\hat{\mathbf{v}}^T \hat{\mathbf{v}}}$	53 cm
condition number	1.52×10^{12}
RMS after fit	6.7 cm
RMS (prediction/cross-validation)	7.9 cm

Figure 4.10 depicts the selected differential similarity corrector surface over the network. It should be noted that the 4-parameter model also performs well in this area providing an RMS of 8.5 cm when tested at independent points. The condition number for this model was considerably high (surpassed only by the full fourth-order polynomial), which may pose numerical problems, although none were noted throughout the tests conducted thus far.

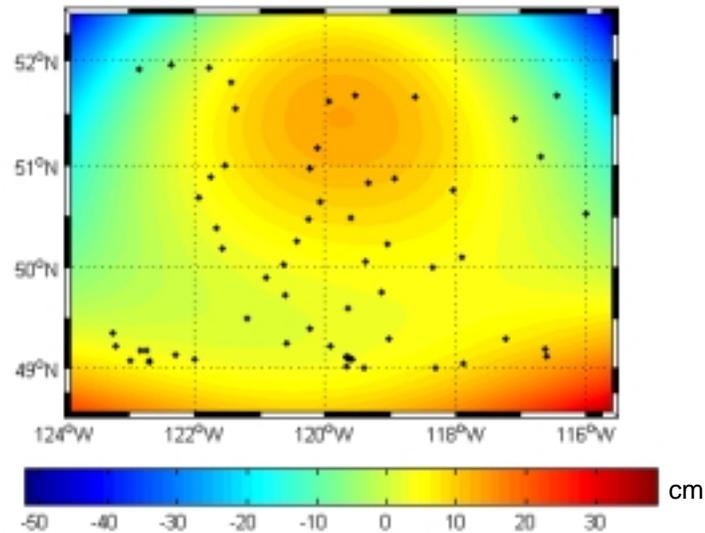


Figure 4.10: 7-parameter differential similarity fit for the southern BC/AB test network

4.3 Fitting a gravimetric geoid model to the Australian height datum via GPS

The most common use of the Australian (quasi)geoid model is to transform GPS-derived ellipsoidal heights to normal-orthometric heights referring to the Australian Height Datum (AHD; Roelse *et al.*, 1971). A properly selected 'corrector' surface model will allow for the direct transformation of GPS-derived ellipsoidal heights to the AHD and *vice versa* (Featherstone, 1998). The numerical data used for the tests comprise the latest Australian gravimetric geoid model, AUSGeoid98, available on a $2' \times 2'$ grid (Featherstone *et al.* 2001) and 963 GPS-levelling points scattered over the Australian mainland (see Figure 4.11; Featherstone and Guo 2001). The gravimetric geoid values N^{grav} were obtained by bi-cubic interpolation of the AUSGeoid98 grid to the GPS-levelling points. The residuals/observations were formed via Eq. (3.8) for all points, and a routine 3-RMS test (i.e., Z-score > 3) was applied to remove outliers. The remaining 953 residuals were used in a least-squares adjustment to solve for the unknown parameters of each pre-selected 'corrector' surface.



Figure 4.11: GPS-levelling data on the Australian mainland (according to Featherstone and Guo, 2001)

The absolute differences between the gravimetric geoid model and the nation-wide set of GPS-levelling control points varies from -2.6 m along the southern coast to a maximum of more than 3.5 m on the eastern coast. Most notably there is a dominant north-south trend of approximately 0.26 mm/km (*ibid.*). Much speculation exists that this may be due to the AHD heights, which were computed by constraining to mean sea level at 30 tide gauges around Australia. At the time, sea surface topography was not taken into account and could therefore be a possible reason for this systematic effect (see discussions in sections 2.2, 2.4.1 and 3.4). In addition, the AHD is not a true (Helmert) orthometric height system, as observed gravity data were not used to apply corrections to the levelling data. It is a normal-orthometric system (Roelse *et al.*, 1971). The differences can be significant with high-frequency content. Allister and Featherstone (2001) calculate a difference of 5 mm over 15 km between Helmert orthometric heights and the normal-orthometric heights used for the AHD. It is also widely acknowledged that the AHD is plagued with distortions, and the currently available GPS data are of a variety of vintages,

with many being collected under sub-optimal conditions (Johnston and Luton, 2001). The statistics for the absolute differences between the gravimetric and GPS-levelling height values is given in Table 4.10.

Table 4.10: Statistics of the residuals before any fit (units: cm)

Data	Min	Max	μ	σ	RMS
Original (963 pts)	-257.2	355.8	1.1	31.7	31.7
After 3RMS (953 pts)	-75.9	93.3	1.3	25.8	25.8

The following results use the 953 data points remaining after the removal of 10 outliers. The results of the numerical tests are presented at both national and regional scales. This is particularly important for a large area, where it is often difficult to have a uniform data distribution and quality. From Figure 4.11, it is evident that the GPS-levelling data are scattered with good coverage in the more populated southeastern parts and relatively poor coverage in the inner and northern parts of the continent. In order to assess the performance of the different models on a more local scale, Australia was divided into four regions. Each region reflects a different density and distribution of the GPS-levelling points. A summary of the four regions is given in Table 4.11.

Table 4.11: Description of Australian regional network geometry

Region	Number of points	Data Distribution
NW	32	consistent
NE	88	mixed
SE	653	good
SW	173	poor

A closer examination of the southwestern quadrant reveals the importance of both data density and distribution. In this area, there are 173 control points, however ~ 100 of them are concentrated in a small region surrounding Perth (32°S , 116°E). On the other hand, the most populated part of the country, the southeastern region, contains 653 points, which are more uniformly distributed.

In this section, the results of the following models will be investigated in detail:

- a) 4-parameter model as given in Eq. (3.31)
- b) 5-parameter model as given in Eq. (3.33)
- c) MRE where $M=N=1$, $M=N=2$, ..., $M=N=6$, as given in Eq. (3.28)
- d) 10-parameter MRE of fourth-order
- e) 16-parameter MRE of sixth-order
- f) 20-parameter MRE of sixth-order

Models (d), (e) and (f) are essentially nested models of the fourth and sixth-order MREs, respectively. The criteria for parameter deletion involved the magnitude of the test statistic, as described in section 3.5.4 and implemented previously with the Canadian data. It should be noted that for testing purposes, the number of parameters in the model (i.e. 10, 16 or 20) was predetermined and therefore parameter deletion/addition was not as flexible as in the general case. Figure 4.12 gives two graphical examples of the corrector surfaces generated for the Australian mainland (excluding Tasmania).

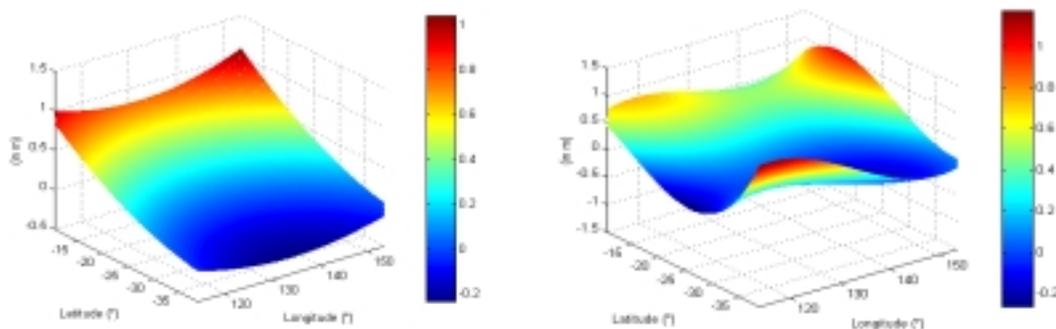


Figure 4.12: Examples of the 4-parameter corrector surface fit (left) and the 10-parameter MRE fit (right) for the Australian mainland

Of note is the order of magnitude of the corrections (± 1.5 m), as compared to those for the Switzerland and regional Canadian networks (see Figures 4.2a/b and Figures 4.7a/b). Table 4.12 gives the statistics of the results for the full forms of the parametric models tested. There are two sets of statistics for each model. The first set (upper row) shows the statistics for the adjusted residuals after the least-squares fit at the points used in the adjustment (as described in section 3.5.1). By looking at these results, the lowest RMS is achieved by the MRE where $M=N=6$ (i.e., a 12th-order polynomial fit). The effect of a decrease in the corresponding RMS as the number of parameters in the models increases is again evident from these results. The results for the cross-validation process (described in section 3.5.2), repeated 953 times, are given in the second set of statistics. The 4-parameter model achieves the lowest RMS. Furthermore, the worst accuracy is given by the MRE where $M=N=6$. This result is significant as it shows the dramatic difference in performance if the model parameters are to be used for estimating values at other points independent of the original adjustment (RMS of 13.5 cm after the fit compared to 86.1 cm obtained from cross-validation). In this case, the sixth-order MRE should not be used. The importance of testing the models with an independent set of control points is clearly emphasized by these results.

Table 4.12: Statistics of residuals at the points used in the adjustment (upper row) and statistics obtained through cross-validation (bottom row) for Australia (units: cm)

Model	Min	Max	μ	σ	RMS
4-parameter	-117.2	61.8	0	19.4	19.4
	-119.0	61.9	0	19.5	19.5
5-parameter	-119.9	61.5	0	19.4	19.4
	-125.0	61.7	0	19.6	19.6
MRE $N=M=1$	-103.8	62.3	0	19.9	19.9
	-76.7	109.6	1.7	25.2	25.3
MRE $N=M=3$	-73.4	53.7	0	16.2	16.2
	-75.9	100.1	1.6	26.2	26.2
MRE $N=M=6$	-82.5	60.6	0	13.5	13.5
	-822.8	1540.6	7.2	85.8	86.1

Table 4.13 summarizes the condition numbers computed for various corrector surfaces. It is evident from the results in the table that, in general, the most stable models are those of lowest order with fewer unknown parameters. The higher order models tend to be less stable and less accurate when compared to independent control points not used in the original adjustment. Therefore, one should be extremely cautious when dealing with higher order MRE corrector surfaces as this may lead to uncontrolled oscillations in the computed values, generating artifacts that are not indicative of the ‘correction’ behavior over the area (consider the >15m maximum value obtained using MRE $N = M = 6$ in Table 4.12). This will cause significant problems when the ‘corrected’ gravimetric geoid model is used to transform GPS-derived ellipsoidal heights to the local vertical datum.

Table 4.13: Condition numbers for various parametric models for the Australian network

Model	Condition Number
4-parameter	2.233×10^4
5-parameter	4.771×10^4
MRE $N=M=1$	6.073×10^3
MRE $N=M=3$	7.640×10^{10}
MRE $N=M=6$	8.529×10^{21}

On a regional level, the *best* model fit varied depending on the control network data (geometry and quality). Figure 4.13 graphically depicts the best local fits determined according to a combination of the results obtained from the estimated accuracy (cross-validation) and numerical stability. Note that the scales on the four plots are different. Based on these criteria, the selected models by region and the associated RMS of the adjusted residuals are provided in Table 4.14. All of the local fits resulted in an improvement over the national RMS of 19.4 cm achieved with the 4-parameter model.

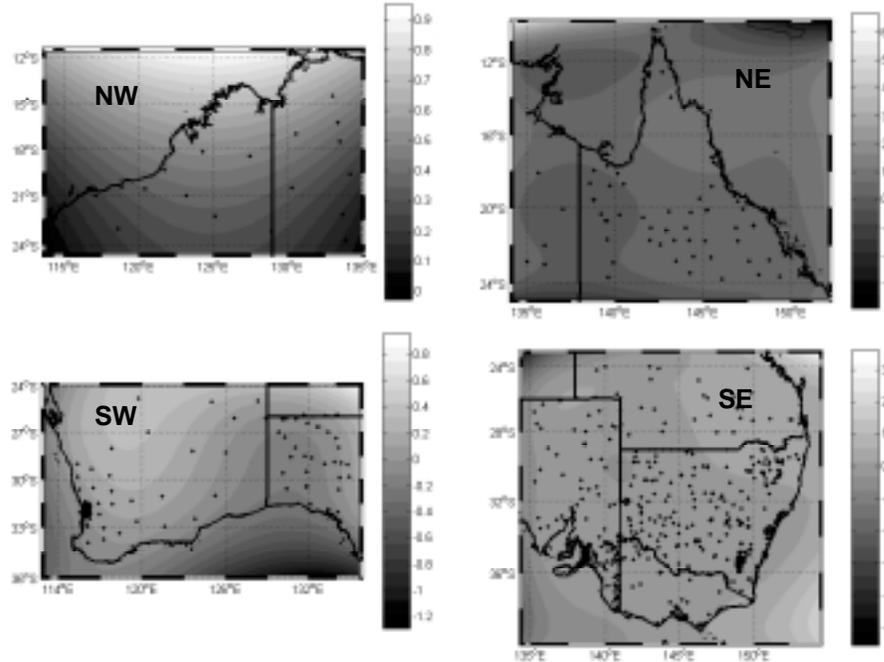


Figure 4.13: Selected parametric models for regional fits in Australia (units: m)

Table 4.14: Final selected models per region and corresponding RMS obtained through cross-validation

Region	Model	RMS (cm)
NW	(b) 5-parameter	15.8
NE	(e) 16-parameter MRE	15.7
SW	(d) 10-parameter MRE	8.3
SE	(f) 20-parameter MRE	13.1

Summary

Evidently, the incorporation of a corrector surface for modeling local discrepancies in the Australian height data is necessary. In practice, local models may be provided to users who wish to fit the gravimetric geoid model to the AHD, however further work must be conducted on how to patch together the different models in a uniform manner across the borders of the neighboring regions. On a national level, the best model was the classic 4-

parameter fit with an RMS of approximately 19 cm. The achievable RMS for local regions, based on empirical cross-validation tests, was found to vary from a minimum of ~ 8 cm in the SW region to ~ 16 cm in the northern regions (see Table 4.14). In addition to determining the most appropriate corrector surfaces, systematic errors in the AHD, such as the apparent north-south trend and the systematic differences between Helmert and normal-orthometric heights, should be considered (Fotopoulos *et al.*, 2002 and Featherstone *et al.*, 2002).

4.4 Remarks on results

The three numerical data sets from Switzerland, Canada and Australia have been used to demonstrate the selection and assessment process for parametric corrector surface models. From these results, it is evident that each region represents a different set of challenges and therefore a completely automated procedure for selecting and testing models should not be used. Instead, it is instructive for users to familiarize themselves with the possible types of systematic errors and datum inconsistencies involved in the combined height adjustment problem. Armed with this knowledge a semi-automated procedure as summarized in Figure 3.8 can be followed. This proposed procedure provides a series of tests that can lead to the selection of the best model from a pre-selected group (defined by the user) or family of corrector surfaces.

In the Swiss case, the GPS-on-benchmark data was well distributed with a small average height misclosure of 1.1 cm, compared to more than 9 cm and 25 cm for the southern BC/AB and Australian data sets, respectively. A significant difference between the various corrector surface fits to the Switzerland data was not evident.

For the southern BC/AB test network, the utility of testing model parameter significance was explored. It was found that the high correlation among model parameters inhibits the statistical testing process and therefore it is recommended that polynomials comprised of orthogonal base functions (such as Legendre, Jacobi, or Hermite) are used.

Orthogonalization procedures can be used to decorrelate parameters of any parametric model, however the results cannot be applied for prediction of height values at new points (GNSS-levelling) due to the lack of an analytical form for the 'orthogonalized' model.

Finally, the results of the Australian test network revealed the option of using a mosaic of different parametric models for different parts of the continent. The practical implementation of such results would also require constraints be imposed for neighbouring 'patches' in order to ensure a smooth transition (see Eqs. 3.37 and 3.38).

The procedure described in this chapter focused on the systematic errors and datum inconsistencies inherent among the height data types. However, the optimal combined least-squares adjustment solution presented in section 3.1 also involves the incorporation of a proper stochastic model for the observational noise (Eq. 3.22), which is the focus of the following two chapters.

Chapter 5

Overview of Variance Component Estimation

5.1 Introduction

One of the key elements required for the least-squares adjustment of geodetic data, in addition to proper modelling of systematic effects, is a realistic stochastic model for the observational noise. Although the effect of the data covariance matrix may not be readily evident in the solution of the unknown parameters (see Marana and Sanso, 1996 for a detailed discussion), a poor covariance matrix may adversely affect decisions based upon statistical testing of hypotheses involving least-squares residuals and the estimated parameters. Additional reasons for using a correct covariance matrix for the observational noise include the examination of the relative magnitudes of the errors in observations due to different factors, the preservation of quality control and the facilitation of efficiently designed surveys (Rao and Kleffe, 1988).

Often, in practice, the data covariance matrix is oversimplified or neglected entirely. In order to provide a better understanding of the inherent limitation of a simplified stochastic model, it is instructive to go through an example with one of the most common functional models used within the context of least-squares adjustment. The general Gauss-Markov model, that was described in chapter 3, is repeated here for the sake of discussion

$$\mathbf{l} = \mathbf{Ax} + \mathbf{v} \quad (5.1)$$

where, \mathbf{l} is an $m \times 1$ vector of observations, \mathbf{x} is a $u \times 1$ vector of unknown parameters, \mathbf{A} is an $m \times u$ known design matrix of full column-rank that relates the unknown parameters with the observations and \mathbf{v} is a vector of random errors with zero mean

$$E\{\mathbf{v}\} = 0 \quad (5.2)$$

and a variance-covariance matrix given by

$$\mathbf{C}_v = E\{(\mathbf{l} - \mathbf{Ax})(\mathbf{l} - \mathbf{Ax})^T\} = E\{\mathbf{vv}^T\} = \sigma_o^2 \mathbf{Q}_v \quad (5.3)$$

where \mathbf{Q}_v is a known symmetric positive definite cofactor matrix and σ_o^2 is an unknown variance factor. The a-priori variance factor can be estimated from the adjusted results according to the formula

$$\hat{\sigma}_o^2 = \frac{\hat{\mathbf{v}}^T \mathbf{P} \hat{\mathbf{v}}}{m - u} \quad (5.4)$$

The adjusted least-squares residuals are denoted by the $m \times 1$ vector $\hat{\mathbf{v}}$, $\mathbf{P} = \mathbf{C}_v^{-1}$ is the data weight matrix. Given the estimated a-posteriori variance factor in Eq. (5.4), the data covariance matrix can finally be 'improved' by a simple scaling

$$\mathbf{C}_v^{improved} = \hat{\sigma}_o^2 \mathbf{Q}_v \quad (5.5)$$

This classic, and somewhat simplified approach, for the stochastic model is limiting as it allows for only one common variance factor of the CV matrix. This may be adequate for adjustments of observations of the same type and similar quality, however it is not

realistic for a variety of geodetic applications, where heterogeneous data types are involved. A more general treatment of the stochastic model includes the ability to estimate more than one variance and/or covariance components to improve the CV information.

In addition to combining different types of data, often observations of the same type but different quality are merged. This is the case with the rigorous national or regional network adjustments of various orders (in terms of precision) of levelling data (Kelly, 1991). Another excellent practical example is the combined adjustment of the multi-national European levelling networks, where an estimated variance component was introduced for each national network in the combined adjustment (Adam *et al.*, 2000). In these and many other cases the simplicity of Eqs. 5.3, 5.4, and 5.5 does not allow for an adequate description of the behaviour of the data.

The chosen approach presented herein for testing and improving the stochastic model is the well known statistical tool of variance component estimation (VCE). Examples of recent geodetic applications where various VCE methods have been successfully implemented include:

- assessment of triangulation network for monitoring tectonic activity observed with different EDM instruments, testing parameter significance for error models of geodetic levelling, and estimating error components and weighting of GPS observations (Chen *et al.*, 1990)
- estimation of the variance components for satellite laser ranging data (Sahin *et al.*, 1992)
- estimation of elements of the measurement CV matrix for precise GPS measurements, which improves reliability and efficiency of positioning results (Wang, 1999)
- estimation of a stochastic model for GPS phase measurements that incorporates temporal correlations (Satirapod *et al.*, 2002)

- construction of an adequate CV matrix for processing of GPS code and phase data (Tiberius and Kenselaar, 2003)
- implications of estimated variance components on the results of spatial deformation trend analysis (Grodecki, 1997)
- the statistical analysis of very long baseline interferometry data (Lucas and Dillinger, 1998)
- estimation of variance components for weighting data of different types and the regularization parameter used in simulated gravity field models from new satellite missions such as GOCE (Koch and Kusche, 2002)

The implementation of VCE techniques in the adjustment of combined height data types, more specifically ellipsoidal, orthometric and geoid heights, has not been suitably addressed in geodetic literature. The motivations for applying VCE techniques to the adjustment of heterogeneous height data are numerous. In particular, it will facilitate studies on the calibration of the geoid error model. Another important area that will benefit from this type of VCE testing is the assessment of the noise in the heights derived from GPS measurements. Furthermore, it will allow for the evaluation of the levelling precision and provide an independent test of the error values associated with various orders of conventional spirit-levelling. Thus, reliable calibration of the height data covariance matrices, via VCE, will enhance our knowledge of the error budget of **all** of the data in the combined adjustment. Finally, but definitely not least, the achievable accuracy of GNSS-levelling can be tested. This will provide users with realistic accuracy measures of an exceedingly popular survey practice.

The selection and application of a suitable VCE technique to the combined height data analysis problem is the topic of the remaining sections in this chapter. Specifically, an overview of some of the existing methods for estimating variance components is provided in section 5.2. Details on the final selected VCE procedures are given in sections 5.3 through 5.7.

5.2 Review of methods for estimating variance components

Over the years, a plethora of VCE methods have been developed. A proper and complete review of all of these methods would be beyond the scope of this text, instead the focus has been placed on identifying and discussing the key developments related to geodetic research (see also Grafarend, 1985). In particular, the discussion will focus on procedures that have been tested with real data and are useful for geodetic applications of heterogeneous data. It should also be noted that although the estimation procedures discussed can be applied for both variance *and* covariance components, provided the stochastic model has been set up as such, the emphasis herein is placed on estimating only *variance* components.

An effective method for characterizing the different VCE procedures is to list them according to certain distinguishable features. In accordance to Crocetto *et al.* (2000), most approaches for estimating variance-covariance components within a least-squares estimation framework can be categorized according to the following:

- functional model
- stochastic model
- estimation procedure
- simplifications, assumptions

Recognizing the utility of such a categorical scheme, a timeline/summary outlining the key developments in VCE theory applicable to geodetic problems is provided in Table 5.1, which uses these distinguishable features as a guideline.

Table 5.1: Timeline of key VCE developments in geodetic literature

Reference	Functional Model	Stochastic Model	Estimation Procedure
Helmert (1924)	Gauss-Markov	$C = \sigma_i^2 Q_i$	Helmert's
Kubik (1970)	Gauss-Helmert	$C = \text{diag}[\sigma_i^2 I_i]$	maximum likelihood
Rao (1971)	Gauss-Markov	$C = \sum_{i=1}^k \sigma_i Q_i$	MINQUE
Sjöberg (1983)	condition-only Gauss-Helmert	$C = \sum_{i=1}^k \sigma_i^2 Q_i$	MINQUE
Sjöberg (1984)	Gauss-Helmert	$C = \sum_{i=1}^k \sigma_i^2 Q_i$	iterative BIQUE, BQMBNE, BQUNE
Grafarend (1984)	condition-only Gauss-Helmert	$C = \sum_{i=1}^k \sigma_i Q_i$	Helmert's
Koch (1986)	Gauss-Markov	no restriction	iterative maximum likelihood
Koch (1987)	Gauss-Markov	no restriction	approximate Bayesian inference
Koch (1988)	Gauss-Markov	$C = \sum_{i=1}^k \sigma_i Q_i$	Bayes estimators
Caspary (1987)	Gauss-Markov	$C = \sum_{i=1}^k \sigma_i^2 Q_i$	iterative BIQUE
Ou Ziqiang (1989)	Gauss-Markov condition-only	$C = \begin{bmatrix} \sigma_{11}^2 Q_{11} & \sigma_{12} Q_{12} & \cdots & \sigma_{1m} Q_{1m} \\ \sigma_{21} Q_{21} & \sigma_{22}^2 Q_{22} & \cdots & \sigma_{2m} Q_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{m1} Q_{m1} & \sigma_{m2} Q_{m2} & \cdots & \sigma_{mm}^2 Q_{mm} \end{bmatrix}$	iterative maximum likelihood
Ou Ziqiang (1991)	Gauss-Markov	$C = \text{diag}[\sigma_i^2 P_i^{-1}]$	approximate Bayes & strict Bayes

Yu (1992)	Gauss-Helmert	$\mathbf{C} = \begin{bmatrix} \sigma_{11}^2 \mathbf{Q}_{11} & \sigma_{12} \mathbf{Q}_{12} & \dots & \sigma_{1m} \mathbf{Q}_{1m} \\ \sigma_{21} \mathbf{Q}_{21} & \sigma_{22}^2 \mathbf{Q}_{22} & \dots & \sigma_{2m} \mathbf{Q}_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{m1} \mathbf{Q}_{m1} & \sigma_{m2} \mathbf{Q}_{m2} & \dots & \sigma_{mm}^2 \mathbf{Q}_{mm} \end{bmatrix}$	BIQUE
Sjöberg (1995)	Gauss-Markov	$\mathbf{C} = \sigma_1^2 \mathbf{I} + \sigma_2^2 \mathbf{F}$	BQMBNE
Yu (1996)	Gauss-Helmert	$\mathbf{C} = \begin{bmatrix} \sigma_{11}^2 \mathbf{Q}_{11} & \sigma_{12} \mathbf{Q}_{12} & \dots & \sigma_{1m} \mathbf{Q}_{1m} \\ \sigma_{21} \mathbf{Q}_{21} & \sigma_{22}^2 \mathbf{Q}_{22} & \dots & \sigma_{2m} \mathbf{Q}_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{m1} \mathbf{Q}_{m1} & \sigma_{m2} \mathbf{Q}_{m2} & \dots & \sigma_{mm}^2 \mathbf{Q}_{mm} \end{bmatrix}$	maximum likelihood
Crocetto <i>et al.</i> (2000)	no restriction	$\mathbf{C} = \sum_{i=1}^k \sigma_i \mathbf{Q}_i$ $\mathbf{C} = \text{diag} \left[\mathbf{C}_i = \sum_{j=1}^k \sigma_{ij}^2 \text{diag}(q_{ij}) \right]$	BIQUE
Schaffrin & Iz (2002)	rank deficient Gauss-Markov	$\mathbf{C} = \sigma_o^2 \mathbf{P}^{-1}$	BLIMPBE
Tiberius and Kenselaar (2003)	Gauss-Markov	$\mathbf{C} = \sum_{i=1}^k \sigma_i \mathbf{Q}_i$ $\mathbf{C} = \mathbf{Q}_o + \sigma_i^2 c_i c_i^T$ $\mathbf{C} = \mathbf{Q}_o + \sigma_{ij} (c_i c_j^T + c_j c_i^T)$	BQUE & AUE
Kusche (2003)	Gauss-Markov	$\mathbf{C} = \sum_{i=1}^k \sigma_i^2 \mathbf{V}_i, \quad \mathbf{V} = \sum_{i=1}^k \mathbf{V}_i = \text{diag}(\mathbf{P}_i^{-1})$	Monte-Carlo

5.2.1 Functional models

For geodetic problems, several different functional models have been adopted for setting up the observational equations in order to relate the observables to the unknown parameters. One common choice is the classic Gauss-Markov model given in Eq. (5.1). As can be seen from the second column of Table 5.1, most of the VCE approaches developed use this functional model.

An alternative method of representing the observations is to form condition equations that the observations have to fulfill. This is known as the condition model and is given by

$$\mathbf{B}\mathbf{v} = \mathbf{w} \quad (5.6)$$

where \mathbf{B} is a design matrix relating the observations among themselves, \mathbf{w} is an $m \times 1$ misclosure vector defined by the specific problem and \mathbf{v} is an $m \times 1$ vector of unknown random errors. Examples of VCE schemes that have been implemented based on this functional model are included in Sjöberg (1983), Grafarend (1984), and Ou Ziqiang (1989), to name a few.

Another equally common functional model which can be viewed as an extension of both models described thus far is the Gauss-Helmert model, also known as the mixed or implicit model and is given by

$$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v} + \mathbf{w} = 0 \quad (5.7)$$

where all of the terms have been described previously. Several VCE procedures have been developed using this functional model as a base; see Kubik (1970) and Yu (1992; 1996). In general, the selection and appropriateness of the functional models described have been given more attention by researchers and are typically well defined according to the application. The suitable form of the stochastic model, however, is often in question and remains a challenging area of research, which must be tested to ensure its appropriateness (Wang, 1999).

5.2.2 Stochastic models

A number of different general stochastic models have emerged as practical and useful in geodetic problems. One of the most common models for the covariance matrix of the observations, \mathbf{C} , is expressed by

$$\mathbf{C} = \sum_{i=1}^k \sigma_i \mathbf{Q}_i \quad (5.8)$$

where the unknown variance and covariance components are denoted by σ_i and \mathbf{Q}_i is the respective known positive-definite cofactor matrix. This stochastic model has been used in many studies, including Rao (1971), Grafarend (1984), Koch (1987, 1988), and Tiberius and Kenselaar (2003). It is important to note that in this case both variance **and** covariance components are sought after.

A more simplified, in terms of computation, stochastic model is given by

$$\mathbf{C} = \sum_{i=1}^k \sigma_i^2 \mathbf{Q}_i \quad (5.9)$$

where only variance components σ_i^2 are to be estimated. Such a model is used extensively in many applications, including Sjöberg (1984), Caspary (1987) and Fotopoulos and Sideris (2003).

Additional simplifications can be made for uncorrelated sets of observations of one variance component each and is described by

$$\mathbf{C} = \text{diag}[\sigma_i^2 \mathbf{Q}_i] = \begin{bmatrix} \sigma_1^2 \mathbf{Q}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_k^2 \mathbf{Q}_k \end{bmatrix} \quad (5.10)$$

This block diagonal form of stochastic model corresponds to vectors of observations that are commonly referred to as *disjunctive*, meaning that they are completely uncorrelated between groups, but can be correlated within the same group. Consider a vector of observations ℓ subdivided into k subsets of observations, ℓ_i , as follows:

$$\ell = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \vdots \\ \ell_k \end{bmatrix} \quad (5.11)$$

The corresponding covariance matrix for each subset of observations is given by

$$\mathbf{C}_i = E\{\mathbf{v}_i \mathbf{v}_i^T\} \quad i = 1, 2, \dots, k \quad (5.12)$$

\mathbf{C}_i is a fully populated positive-definite symmetric matrix. Since no correlation exists between subsets of observations, the cross-covariance matrices are formed as follows:

$$\mathbf{C}_{ij} = E\{\mathbf{v}_i \mathbf{v}_j^T\} = \mathbf{0} \quad \text{where } i \neq j \quad \text{and } i, j = 1, 2, 3, \dots, k \quad (5.13)$$

Depending on the application and type of data involved such a restriction may be suitable. See for instance Ou Ziqiang (1991) and Crocetto *et al.* (2000), where this type of model has been successfully implemented in both simulated numerical examples and real-world applications.

If the disjunctive subsets of observations have the same accuracy, the following stochastic model can be used:

$$\mathbf{C} = \text{diag}[\sigma_i^2 \mathbf{I}] \quad \text{where } i = 1, \dots, k \quad (5.14)$$

In this case the weight matrix is replaced by the identity matrix, \mathbf{I} (see Kubik, 1970 for more details).

A specific stochastic model that has been effective for dealing with large data sets (assumed to be normally distributed) is a block-structured variance matrix described by

$$\mathbf{C} = \begin{bmatrix} \sigma_{11}^2 \mathbf{Q}_{11} & \sigma_{12} \mathbf{Q}_{12} & \dots & \sigma_{1m} \mathbf{Q}_{1m} \\ \sigma_{21} \mathbf{Q}_{21} & \sigma_{22}^2 \mathbf{Q}_{22} & \dots & \sigma_{2m} \mathbf{Q}_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{m1} \mathbf{Q}_{m1} & \sigma_{m2} \mathbf{Q}_{m2} & \dots & \sigma_{mm}^2 \mathbf{Q}_{mm} \end{bmatrix} \quad (5.15)$$

where σ_{ij} corresponds to a variance component if $i = j$ and a covariance component if $i \neq j$. \mathbf{Q}_{ij} is some known positive-definite cofactor matrix. A similar block-structured variance-covariance matrix was effectively implemented to deal with massive amounts of data originating from VLBI measurements (Lucas and Dillinger, 1998).

A final stochastic model worth mentioning, especially for geodetic/survey-related applications, is the additive two-variance component model described by Sjöberg (1995)

$$\mathbf{C} = \sigma_1^2 \mathbf{I} + \sigma_2^2 \mathbf{F} \quad (5.16)$$

where \mathbf{I} is the identity matrix and \mathbf{F} is a positive-definite diagonal matrix.

Examples where Eq (5.16) is applicable include the adjustment of levelling networks where the relative accuracy is typically composed of two parts as follows:

$$\sigma = c + b \cdot \sqrt{d} \quad (5.17)$$

where σ is the standard deviation of the observed height difference, c is an empirically determined constant part of the contributing error, b is the height difference accuracy based on pre-determined (or published) standards for levelling and d is the baseline length corresponding to the levelling segment. Some simple error models affecting GPS signals can also be described using a form similar to Eq. (5.17), where the spatially correlated errors are represented by the $b \cdot \sqrt{d}$ part (i.e., ionospheric, tropospheric, and orbital errors) and the spatially uncorrelated or site/receiver specific error sources are simply assigned some constant value c (i.e., multipath and receiver noise).

5.2.3 Selection of variance component estimation procedures

As far as variance component estimation procedures are concerned, a first solution to the problem was provided by Helmert (1924), who proposed a method for unbiased variance estimates. Much later, an independent solution was derived by Rao (1970), who was unaware of Helmert's method, and was called the minimum norm quadratic unbiased estimation (MINQUE) method. Under the assumption of normally distributed observations, both Helmert's and Rao's MINQUE approach are equivalent. Since then, a plethora of methods have been developed and tested for geodetic applications, with the most common outlined in the final column of Table 5.1.

Maximum likelihood estimation methods

The maximum likelihood estimation procedures involve writing the likelihood function in terms of the variance-covariance components, mean values and observations. The unknown parameters of the likelihood function are then solved by setting partial

derivatives of the unknown parameters equal to zero (Kubik, 1970 and Yu, 1996). Various stochastic models and functional models have been tested with this approach and proven to give valuable results. For instance, in Grodecki (1997) a restricted maximum likelihood method was successfully implemented for deformation trend analysis.

Bayesian methods

A closely related approach is the Bayesian methods, thoroughly investigated by Koch (1987), which also require that the distribution of the vector of observations is specified as with the maximum likelihood methods. The key difference with this approach is that it requires some prior knowledge about the vector of variance-covariance components in the form of a prior probability density function. Strictly speaking, Bayes estimates for variance components cannot be solved analytically. That is, numerical integration must be performed and, if too many unknown parameters are involved, problems with computational difficulty arise and numerical results are difficult (if not impossible) to obtain. This has led to the development of approximate Bayes procedures as described in Ou Ziqiang (1991). In this work, the approximate and strict versions are compared and found to give almost numerically equivalent results. The utility of this approximate method is contained in the splitting of the likelihood function into a product of individual likelihood functions. This results in a much simpler method that requires less computational effort than the strict rigorous Bayes method.

Other methods

The most popular estimation procedures are based on a quadratic estimator of the observational residuals \mathbf{v} as follows

$$\hat{\boldsymbol{\gamma}} = \mathbf{v}^T \mathbf{M} \mathbf{v} \quad (5.18)$$

where \mathbf{M} is a symmetric matrix to be determined via the following minimum trace problem:

$$tr\{\mathbf{MQMQ}\} = \min \quad (5.19)$$

for a linear combination of variance-covariance components described by

$$\boldsymbol{\gamma} = \mathbf{p}^T \boldsymbol{\theta} \quad (5.20)$$

where \mathbf{p} is a known vector and $\boldsymbol{\theta}$ is a vector of the unknown variance and covariance components. The estimator minimizes a certain optimality criterion, such as minimum norm, minimum variance, or mean square error, subject to some constraints such as translation invariance or unbiasedness. In general, a quadratic estimator satisfies translation invariance if the trace minimum problem given in Eq. (5.19) is solved subject to the following constraint (Rao and Kleffe, 1988):

$$\mathbf{MA} = 0 \quad (5.21)$$

Furthermore, the estimator is unbiased if

$$E\{\mathbf{v}^T \mathbf{M} \mathbf{v}\} = \mathbf{p}^T \boldsymbol{\theta} \quad (5.22)$$

which is equivalent to $\mathbf{A}^T \mathbf{MA} = 0$. This latter property is important, since a biased estimate of a variance-covariance component may lead to overly optimistic or pessimistic relative weights depending on the magnitude of the bias (*ibid.*).

The following is a list of some common procedures that use this quadratic-based estimation scheme:

- minimum norm quadratic unbiased estimation, MINQUE
- best invariant quadratic unbiased estimation, BIQUE
- best quadratic minimum bias non-negative estimation, BQMBNE

- best quadratic unbiased non-negative estimation, BQUNE
- best quadratic unbiased estimation, BQUE
- almost unbiased estimation, AUE

Simplifications and iterative procedures have been developed for most of the aforementioned methods. A significant characteristic with all of these various estimation procedures is that numerically (and under some assumptions such as normally distributed data) most of the methods give computationally equivalent results. However, distinguishable characteristics in the formulation of the problem for each algorithm exist.

Selection criteria

Ultimately, the selection of the appropriate technique should rely on the desired estimator properties, such as translation invariance, unbiasedness, minimum variance, non-negativeness, computational efficiency, etc. In some cases all of these properties cannot be retained for a particular estimator and it is necessary to determine which properties should be sacrificed.

A prevalent example can be found in Hartung (1981) where the property of unbiasedness was sacrificed for guaranteed estimation of non-negative variances. In Pukelsheim (1981), the existence of simultaneously unbiased and non-negative estimates of variance components was investigated and it was demonstrated that non-negative minimum norm quadratic unbiased estimators exist only in "*very special cases*". Furthermore, in Sjöberg (1995) it was shown that the properties of unbiasedness and non-negativity are incompatible for the additive two-variance component model (Eq. 5.16). These realizations led to the development of the best quadratic minimum bias non-negative estimation technique. In general, the decisions for which estimator properties to retain/enforce must be made on a case-by-case basis depending on the data and specific application.

Nowadays, the over-riding property that is usually sought after is computational efficiency, which arises due to the massive quantities of data that are used for the estimation of many variance-covariance components. In fact, the main criticism of traditional VCE methods is that they involve repeated inversions of large matrices, intensive computational efforts and large storage requirements for lots of unknowns. For these reasons, one may opt for entirely different estimation procedures, such as the Monte Carlo technique described in Kusche (2003) wherein the analysis of the low-to-medium degree gravity field recovery from simulated orbit perturbations of the GOCE mission was conducted. In other cases, mathematical manipulations or simplifications are made to the rigorous algorithm in order to reduce the computational burden involved with inverting large dimensional matrices (see for instance, Wang, 1999 and Crocetto *et al.*, 2000).

Another aspect to consider when selecting an appropriate VCE algorithm is whether the problem deals with balanced or unbalanced data. All methods require that the data be grouped or classified according to some attribute(s) presumed to characterize the variation in them. This classification may be made according to the type of the data and/or according to the quality of the data (Rao and Kleffe, 1988). If the classification of the data into groups or sets (as described by Eq. 5.11) involves the *same number of observations* in each sub-set, then the observations are said to be balanced. For such cases, the estimation of variance components is relatively easier and can be performed using simple algorithms, such as the well known analysis of variance (ANOVA) methods (Searle *et al.*, 1992). However, in general, balanced data is only a result of a designed experiment and rarely encountered in practice, therefore most methods listed above also deal with unbalanced data. This implies that the length of at least one observational sub-vector, ℓ_i , in Eq. (5.11) is different from those of the other sub-vectors.

A critical review of the methods available for the estimation of variance components and their benefits was conducted and it was found that the MINQUE and the AUE methods were the most appropriate. The selection was based on evaluating the utility of the

following criteria to the common adjustment of ellipsoidal, orthometric and geoid height data problem:

- computational load
- balanced versus unbalanced data
- ease of implementation
- flexibility for modifications
- unbiasedness
- non-negative variance factors

Since identical numerical results can be achieved through the correct implementation of any method discussed and computational efficiency was not an issue (see section 5.5), one is free to select any scheme. The selected MINQUE and AUE algorithms suit the application at hand in addition to providing an allowance for improvements and any modifications that may be desired. The following sections describe the selected MINQUE and AUE algorithms in detail.

5.3 The MINQUE method

In this section, the general theory and algorithms of the minimum norm quadratic unbiased estimation procedure are described (Rao, 1971; Rao and Kleffe, 1988). This statistical estimation method has been implemented and proven useful in various applications for not only evaluating the CV matrix of the observations, but also for modelling the error structure of the observations.

MINQUE is classified as a quadratic-based approach where a quadratic estimator is sought that satisfies the minimum norm optimality criterion. Given the Gauss-Markov functional model $\mathbf{l} = \mathbf{Ax} + \mathbf{v}$ where $E\{\mathbf{v}\} = 0$ and the selected stochastic model for the data CV matrix expressed as follows:

$$E\{\mathbf{v}\mathbf{v}^T\} = \sum_{i=1}^k \sigma_i^2 \mathbf{Q}_i = \mathbf{C}_\ell = \mathbf{C}_\mathbf{v} \quad (5.23)$$

the MINQUE problem is reduced to the solution of the following system

$$\mathbf{S}\hat{\boldsymbol{\theta}} = \mathbf{q} \quad (5.24)$$

where $\hat{\boldsymbol{\theta}}$ is a $k \times 1$ vector containing the unknown variance components¹. \mathbf{S} is a $k \times k$ symmetric matrix that may not be of full rank and therefore its pseudo-inverse can be used for solving Eq. (5.24). Each element $\{s_{ij}\}$ in the matrix \mathbf{S} is computed from the expression

$$s_{ij} = tr(\mathbf{R}\mathbf{Q}_i\mathbf{R}\mathbf{Q}_j), \quad i, j = 1, \dots, k \quad (5.25)$$

where $tr(\cdot)$ is the trace operator, $\mathbf{Q}_{(i)}$ is a positive definite cofactor matrix for each group of observations indicated by i and j . \mathbf{R} is a symmetric matrix defined by

$$\mathbf{R} = \mathbf{C}_\ell^{-1} [\mathbf{I} - \mathbf{A}(\mathbf{A}^T \mathbf{C}_\ell^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_\ell^{-1}] \quad (5.26)$$

where \mathbf{A} is an appropriate design matrix of full column-rank and \mathbf{C}_ℓ is the covariance matrix of the observations. The vector \mathbf{q} contains the quadratic forms

$$\mathbf{q} = \{q_i\}, \quad q_i = \hat{\mathbf{v}}_i^T \mathbf{Q}_i^{-1} \hat{\mathbf{v}}_i \quad (5.27)$$

¹ The vector, $\hat{\boldsymbol{\theta}}$, can contain both unknown variance and covariance components. However, throughout this thesis only variance components will be estimated.

where \hat{v}_i are the estimated observational residuals for each group of observations ℓ_i .

It is evident from the expression for the \mathbf{R} matrix (Eq. 5.26) that initial estimates for the unknown variance components must be provided as they are embedded in \mathbf{C}_ℓ that is used to compute \mathbf{R} . This introduces one of the main drawbacks or criticisms of the MINQUE approach, which is the fact that it is only a locally best estimator. This implies that ' n ' users with ' n ' different a-priori values for the variance factors have the possibility of obtaining ' n ' different estimates, all satisfying the criteria and properties imposed by the MINQUE procedure (i.e., minimum norm, translation invariance, unbiasedness). This is considered a major issue because if good initial estimates were easily obtainable then there would be limited use in performing variance component estimation to begin with! Remedies for overcoming this shortcoming include the use of an iterative approach in conjunction with ensuring that high redundancy is retained.

5.4 Application of MINQUE to the combined height adjustment problem

In this section, the specific formulations for the functional and stochastic models, as well as for the MINQUE algorithm implemented for the combined height data problem are presented. Most of the formulations have been previously described and therefore only a brief review of the required equations is provided.

Given a network of points with ellipsoidal, orthometric and geoid height data, a combined adjustment can be performed. The general functional model used is given by

$$\ell = \mathbf{Ax} + \mathbf{Bv}, \quad E\{\mathbf{v}\} = \mathbf{0} \quad (5.28)$$

where the vector of observations ℓ is composed of the height 'misclosure' at each GPS-leveilling benchmark " i " as follows:

$$\ell_i = h_i - H_i - N_i \quad (5.29)$$

The selected stochastic model for this particular problem can be described in general as

$$E\{\mathbf{v}\mathbf{v}^T\} = \mathbf{C}_v \quad (5.30)$$

where \mathbf{A} is the design matrix which depends on the parametric model type (see chapter 3), \mathbf{B} is the block-structured matrix $\mathbf{B} = [\mathbf{I} \quad -\mathbf{I} \quad -\mathbf{I}]$, where each \mathbf{I} is an $m \times m$ unit matrix (m is the number of observation equations), \mathbf{x} is a vector containing the unknown parameters corresponding to the selected parametric model and \mathbf{v} is a vector of zero-mean random errors, described by the following formula:

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_h^T & \mathbf{v}_H^T & \mathbf{v}_N^T \end{bmatrix}^T \quad (5.31)$$

where $\mathbf{v}_{(\cdot)}$ is an $m \times 1$ vector of random errors for each of the h, H, N data types.

The estimated adjusted residuals, $\hat{\mathbf{v}}$, used to compute the variance components are disjunctive as the three sub-vectors are correlated among the same height type, but uncorrelated with each other. The data is grouped according to height type, as shown in Eq. (5.31). The separated residuals for each height type are computed as follows:

$$\hat{\mathbf{v}}_{(\cdot)} = \mathbf{Q}_{(\cdot)} \mathbf{R} \ell \quad (5.32)$$

where $\hat{\mathbf{v}}_{(\cdot)}$ is computed for each of h, H, N . The above equation is easily derived from the combined adjustment formulation as described in section 3.1 (see also Kotsakis and Sideris, 1999 for additional details).

The covariance matrix \mathbf{C}_v is described by Eq. (5.30), where the positive definite cofactor matrix \mathbf{Q}_v is scaled by the variance factor σ^2 . For the case of heterogeneous disjunctive observations, a block-diagonal covariance model is used

$$\mathbf{C}_v = \begin{bmatrix} \mathbf{C}_h & 0 & 0 \\ 0 & \mathbf{C}_H & 0 \\ 0 & 0 & \mathbf{C}_N \end{bmatrix} = \begin{bmatrix} \sigma_h^2 \mathbf{Q}_h & 0 & 0 \\ 0 & \sigma_H^2 \mathbf{Q}_H & 0 \\ 0 & 0 & \sigma_N^2 \mathbf{Q}_N \end{bmatrix} \quad (5.33)$$

where $\mathbf{C}_{(\cdot)}$ is the covariance matrix for each of the height types. A linearly additive CV matrix model for the observations can be obtained through error propagation and is shown below

$$\mathbf{C}_\ell = \mathbf{B} \mathbf{C}_v \mathbf{B}^T \quad (5.34)$$

The final expression is given by

$$\mathbf{C}_\ell = \sigma_h^2 \mathbf{Q}_h + \sigma_H^2 \mathbf{Q}_H + \sigma_N^2 \mathbf{Q}_N \quad (5.35)$$

where \mathbf{Q}_h , \mathbf{Q}_H , and \mathbf{Q}_N are known positive definite cofactor matrices for the ellipsoidal, orthometric and geoid height data, respectively. The unknown variance components are σ_h^2 , σ_H^2 , σ_N^2 . The problem, therefore, is to solve for:

- the unknown parameters of the parametric model, \mathbf{x} , and
- the individual variance components for each of the height data types, σ_h^2 , σ_H^2 , σ_N^2 .

Following this formulation, the specific MINQUE algorithm for the solution of the three unknown variance components, one for each height data type, is explicitly stated by the solution of the following:

$$\begin{bmatrix} tr(\mathbf{RQ}_h\mathbf{RQ}_h) & tr(\mathbf{RQ}_h\mathbf{RQ}_H) & tr(\mathbf{RQ}_h\mathbf{RQ}_N) \\ tr(\mathbf{RQ}_H\mathbf{RQ}_h) & tr(\mathbf{RQ}_H\mathbf{RQ}_H) & tr(\mathbf{RQ}_H\mathbf{RQ}_N) \\ tr(\mathbf{RQ}_N\mathbf{RQ}_h) & tr(\mathbf{RQ}_N\mathbf{RQ}_H) & tr(\mathbf{RQ}_N\mathbf{RQ}_N) \end{bmatrix} \begin{bmatrix} \hat{\sigma}_h^2 \\ \hat{\sigma}_H^2 \\ \hat{\sigma}_N^2 \end{bmatrix} = \begin{bmatrix} \ell^T \mathbf{R}^T \mathbf{Q}_h \mathbf{R} \ell \\ \ell^T \mathbf{R}^T \mathbf{Q}_H \mathbf{R} \ell \\ \ell^T \mathbf{R}^T \mathbf{Q}_N \mathbf{R} \ell \end{bmatrix} \quad (5.36)$$

5.5 Remarks on VCE and the combined height adjustment problem

An interesting and advantageous property of the combined height adjustment problem as described herein is the inherent relatively high redundancy. Although the distribution and number of GPS-levelling benchmarks available in a particular country or region varies, the number of unknown variance components (three in this case) remains the same. Table 5.2 provides a brief list of the current approximate number of GPS-levelling benchmarks in various countries. The regions are arranged in terms of geographical coverage area, with the largest area corresponding to Canada and the smallest in the list being that of Switzerland.

The table includes a mixture of 1st, 2nd and in some cases 3rd -order levelling benchmarks, which makes a significant difference in the total number. For instance, in Denmark a total of 22 first order GPS-levelling benchmarks currently exist, however if second order levelling points are included then the total number increases to 416 GPS-levelling benchmarks (*R. Forsberg, personal communication*). Given the eminent economic benefits, high efficiency and improved achievable accuracy of instituting satellite-based vertical control points, it is only expected that the number of GNSS-levelling benchmarks will increase over time. In the meantime, they also provide an independent means for validating gravimetric geoid models. This high redundancy is valuable in any statistical method as it increases the reliability of the estimated parameters.

Table 5.2: Number of GPS-levelling benchmarks in different regions

Region	Area	GPS-levelling benchmarks
Switzerland	41,293 km ²	111
Denmark	43,094 km ²	416
Italy	301,309 km ²	500
Germany	356,955 km ²	675
Australia	7,682,300 km ²	1013
Brazil	8,547,404 km ²	344
United States	9,809,431 km ²	11,020
Canada	9,970,610 km ²	1292

Finally, since the same information is available at each point, namely, h_i, H_i, N_i , the rare occurrence of balanced data is at our disposal. In this case, the computations using the MINQUE algorithm are greatly simplified as the same number of 'observations' is contained in each group or sub-class. Furthermore, under the assumption of balanced normally distributed observations, which is a common submission, the MINQUE algorithm also satisfies the property of minimum variance. Thus, it may be questioned why a relatively complex algorithm such as MINQUE is selected over simpler algorithms, which are available for dealing with balanced data only, such as ANOVA (Searle *et al.*, 1992). The main reason for not selecting a simpler approach is due to the fact that it was desired to retain the capability of re-formulating the problem for the case where unbalanced data are used. For instance, in Kearsley *et al.* (1993), a parametric adjustment with constraints was investigated for the combined adjustment of ellipsoidal, geoid and orthometric height data. In this problem formulation, the case of unbalanced data arises, where the number of elements in each of the three observational sub-sets, $\Delta h, \Delta H, \Delta N$ are different from each other.

An additional advantage offered by the design of this particular problem is a relatively low computational load. Only three variance components are sought. The largest matrix inversion will be of the order of the number of observations, m , which in the absolute case is equivalent to the number of network benchmarks. Thus, unlike many other VCE-related applications, where the main obstacle encountered is the high computational load, the problem here lies in the absence of independently derived and reliable variance estimates for each height type. In such cases where there is a deficiency of good initial estimates, it is preferable to rely on a globally best estimator that provides results independent of the a-priori values rather than a locally best estimator. To achieve this, an iterative procedure must be used. The process is described in detail in sections 5.6 and 5.7 and will be used throughout the remaining parts of this thesis for all of the numerical investigations.

5.6 Iterative minimum norm quadratic unbiased estimation

The computed values from a first run through the MINQUE algorithm, $\hat{\theta}_i^I$ are obtained by specifying a-priori values, $\hat{\theta}_i^o$. The resulting estimates can be used as ‘new’ prior values and the MINQUE procedure is repeated. Performing this process several times is referred to as iterative MINQUE (IMINQUE). It is an important application to consider as it yields less dependence of the estimator on the a-priori values, leading to a globally best estimator, rather than a local best estimator which is only ‘best’ given the initial estimates. Furthermore, this is a useful feature when the a-priori values are unreliable or uncertain at best. In Figure 5.1, the steps involved for the implementation of this approach are depicted. The flowchart provides a general overview of the computational scheme for all iterative procedures, where each $\hat{\theta}_i^\alpha$ value represents a ratio computed at each iteration α .

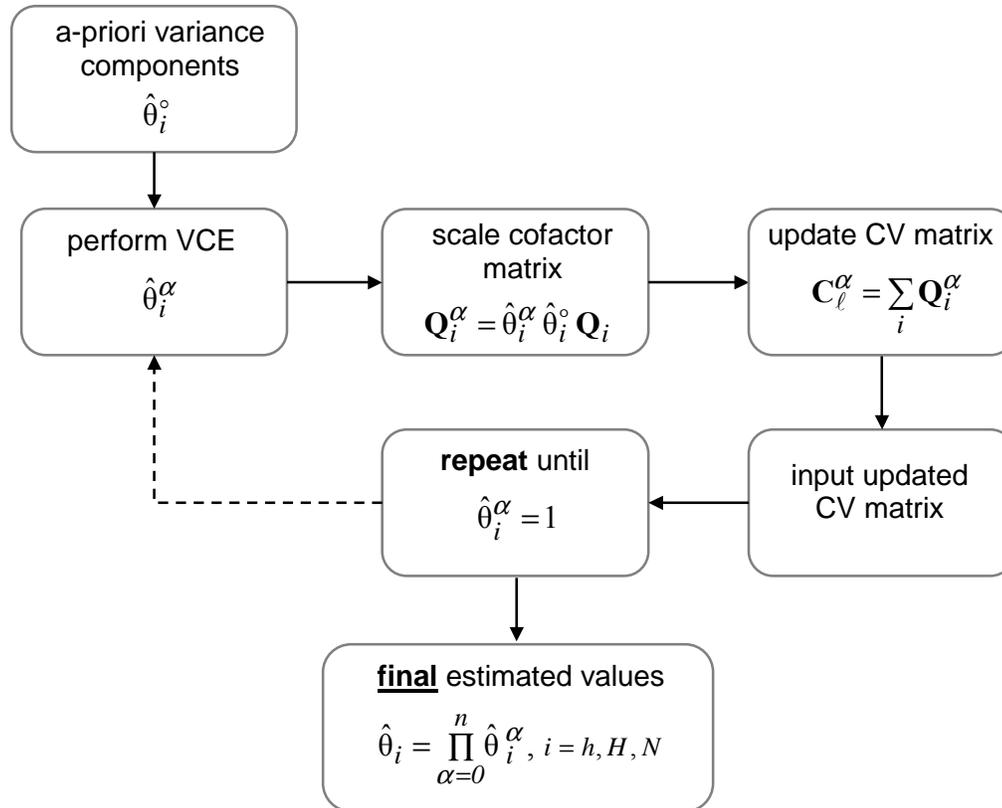


Figure 5.1: Iterative variance component estimation computational scheme

It is evident from this flowchart that all iterative procedures require (i) a-priori values for each of the unknown variance components and (ii) the specification of a convergence criterion to determine when to stop the estimation process. In the work presented herein, the iteration was repeated until all variance factor estimates approached unity, as follows:

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\sigma}_h^2 \\ \hat{\sigma}_H^2 \\ \hat{\sigma}_N^2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (5.37)$$

The implementation of this convergence criterion is performed by checking the difference between the current estimates, $\hat{\boldsymbol{\theta}}^\alpha$, and previous estimates, $\hat{\boldsymbol{\theta}}^{\alpha-1}$, to test if

they are less than a pre-specified small value, i.e., $|\hat{\boldsymbol{\theta}}^\alpha - \hat{\boldsymbol{\theta}}^{\alpha-1}| < 0.0001$. In general, each factor converges at different rates, however the process should not be stopped until **all** of the values satisfy Eq. (5.37). Therefore, the initial values for one group of observations affect the convergence rate, which is essentially a measure of the computational efficiency, of the entire solution. Figure 5.2 shows the convergence behaviour of the estimated variance components for each of the three height data types. The figure was created using the Swiss data set (see 3.6.1) and fully populated cofactor matrices for all three data groups (see chapter 6).

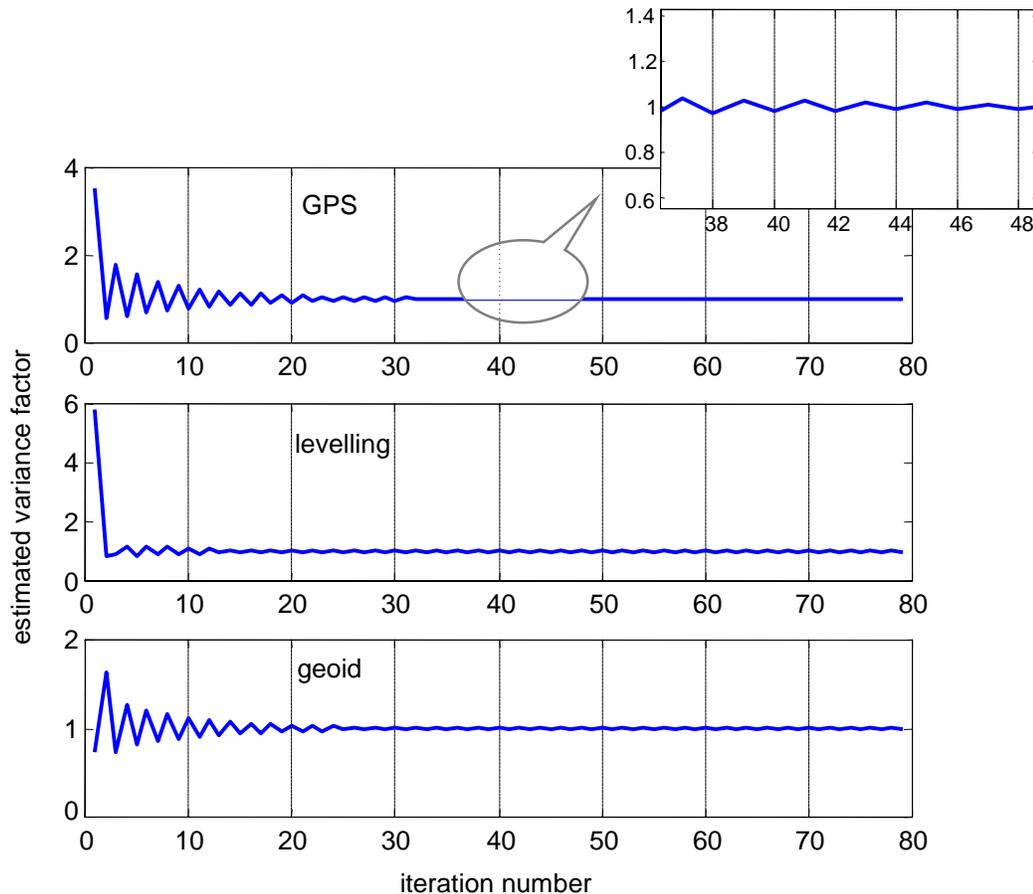


Figure 5.2: Example of estimated variance components at each iteration step

As evidenced in the figure, the initial oscillatory behaviour varies for each component. In this case, the variance component from the levelling data converges the fastest taking approximately 58 iterations for the strict convergence criteria stated above, even though after the first few iterations it is at a higher level than the other two groups of data. The next component to converge is the geoid height variance factor, which levels out after ten more iterations. Finally, the ellipsoidal height variance factor converged, bringing the entire process to take approximately 78 iterations. Due to the resolution of the graph, this is not readily visible, however by zooming in the true number of iterations can be viewed. Of course, if the application allowed for the convergence criteria to be set to less stringent values, the process would be stopped sooner.

The final estimated variance component values are given by

$$\hat{\theta}_i = \prod_{\alpha=0}^n \hat{\theta}_i^{\alpha} \quad i = 1, \dots, k \quad (5.38)$$

where n denotes the total number of iterations. An indication of their accuracy can be computed from the following formula, where the Moore-Penrose pseudo-inverse \mathbf{S}^+ may be used if \mathbf{S} is not of full rank:

$$\mathbf{C}_{\hat{\theta}} = 2\mathbf{S}^{-1} \quad (5.39)$$

In practice, the proper implementation of this iterative process with sufficient observations has worked well. There are, however, some additional theoretical drawbacks associated with the iterative scheme. For instance, it is possible, under certain conditions, that the IMINQUE algorithm does not yield unbiased values (Rao and Kleffe, 1988). Also, there is a possibility that an admissible solution (i.e., positive variances, $\hat{\theta} \in \mathfrak{R}_+^3$) will not be achieved. In fact, the process may not converge, which results in no solution. The following section will describe a procedure designed to avoid such problems.

5.7 Iterative almost unbiased estimation

As mentioned previously, one of the major pitfalls of the described MINQUE algorithm is that no provision has been made to ensure non-negative variance values (i.e., $\hat{\boldsymbol{\theta}} \in \mathfrak{R}_+^3$). In general, negative outcomes of variance components can be attributed to two factors (Sjöberg, 1984):

- an insufficient number of observations compared to unknown parameters (low redundancy), and/or
- an incorrect stochastic model.

Although the actual estimated negative variance values have no meaning, a negative variance outcome does yield important information regarding the problem set-up, information that is lost if the estimator is constrained to give only positive outcomes. Empirically, it has been noticed that negative variance estimates occur less frequently as the number of observations increases. Nonetheless, simplified algorithms exist which ensure that positive estimates are given at each iteration. One such algorithm, known as iterative almost unbiased estimation (IAUE) can be implemented through the following formula (Horn and Horn, 1975):

$$\hat{\theta}_i^{new} = \frac{\hat{\theta}_i^{old} \ell^T \mathbf{R} \mathbf{Q}_i \mathbf{R}^T \ell}{tr(\mathbf{R} \mathbf{Q}_i)} \quad (5.40)$$

where $i = h, H, N$, $\hat{\theta}_i^{new}$ and $\hat{\theta}_i^{old}$ represent the current and previous iteratively-derived variance estimates, respectively. The estimators computed using this algorithm are also invariant with respect to translation of the unknown parameters as with the MINQUE estimates, however the results are ‘almost’ unbiased. In most practical cases, this sacrifice is accepted given the strictly lower mean square error of the AUE compared to either the MINQUE or the sample variance (Rao and Kleffe, 1988).

Several numerical tests were conducted using this algorithm with real data and it proved to give almost identical results to the rigorous IMINQUE (insignificant difference at the sub mm-level). An added benefit of this method is that it is computationally simpler and converges approximately 50% faster than the rigorous approach. Thus, in cases where computational efficiency is an issue, IAUE offers a viable alternative to the rigorous IMINQUE approach. More details from a practical implementation point of view will be presented from the results of the case studies in chapter 6.

Chapter 6

Case Studies: Estimation of Variance Components via a Mixed Adjustment of Ellipsoidal, Geoid and Orthometric Heights

In chapter 5, the theory and methodology involved with the application of VCE in the combined adjustment of heterogeneous height data was discussed. The focus of this chapter is to apply the proposed VCE techniques to real vertical networks and investigate the effects of the following issues on the final estimated variance components:

- modifying the a-priori covariance matrices for the height data,
- estimating non-negative variances,
- ignoring correlations among observations of the same type, and
- assessing the role of the corrector surface model.

The chapter is organized into five main parts, starting with a description of the data used in the estimation of the variance components, namely the a-priori covariance information ($\mathbf{Q}_h, \mathbf{Q}_H$ and \mathbf{Q}_N) and height misclosure data ($\ell = h - H - N$) for each of the test networks. The first test network is located in Switzerland and it was described in detail in section 4.1. The second test network consists of a subset of the Canadian GPS-levelling stations located in southern British Columbia and Alberta as described in section 4.2. Both networks were used for the numerical investigations in this chapter in order to continue the analysis of the vertical networks that started with the modelling of systematic effects in chapter 3. The next four sections of this chapter are dedicated to the

detailed investigation of the four aforementioned items. The main results of these case studies are particularly important for understanding the capabilities and limitations of the potentially powerful statistical tool of VCE in the mixed height network adjustment problem.

6.1 Description of the Swiss test network data

The vertical test network described in section 4.1 consisting of 111 GPS-levelling benchmarks distributed throughout Switzerland was used for the VCE studies in this chapter. As seen in Figure 6.1, there is a very good distribution of vertical control throughout the coverage area.

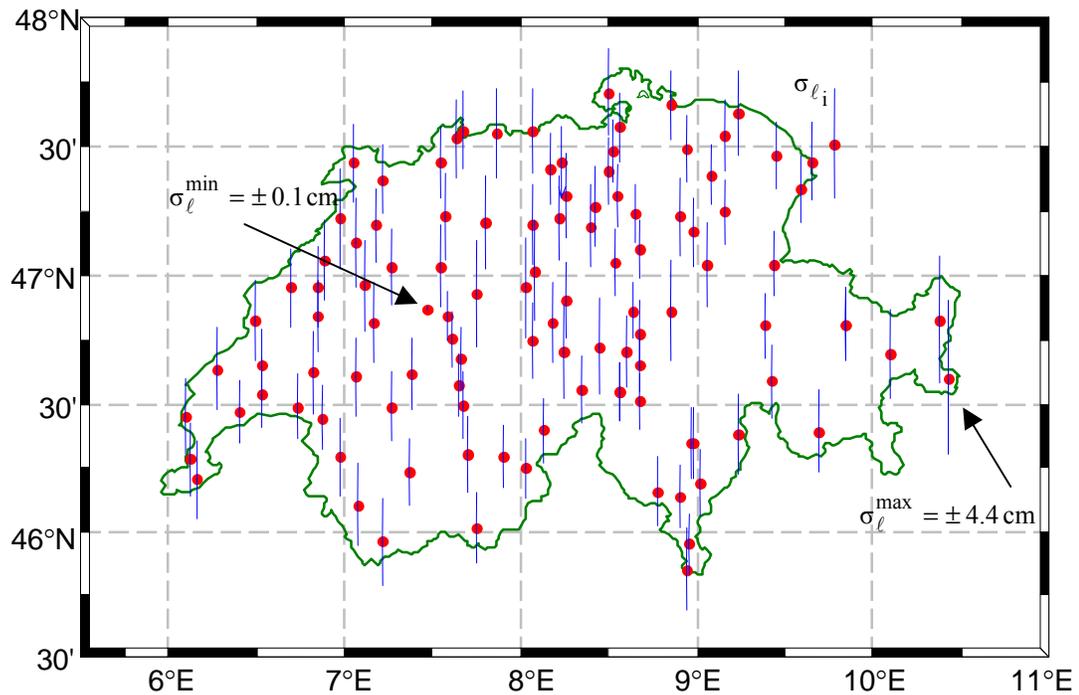


Figure 6.1: Swiss test network and error bars for $\ell = h - H - N$

Figure 6.1 also shows the associated standard deviation of the combined height misclosure, σ_{ℓ_i} , at each point without any fitting of a corrector surface model. It is evident from this figure that there is a moderate accuracy variation, ranging from the mm-level to more than 4 cm. Fully-populated initial CV matrices were also obtained for h , H and N , as described below.

6.1.1 Initial covariance matrix for the GPS heights

The original CV matrix corresponding to the three-dimensional Cartesian geocentric coordinates (x, y, z) was computed from the results of a GPS data post-processing software package. The required information, \mathbf{Q}_h , was extracted after the matrix was transformed into the corresponding CV matrix for the ellipsoidal (φ, λ, h) coordinates (see Marti *et al.*, 2001). Figure 6.2 shows the initial CV matrix for the ellipsoidal heights, which has an average standard deviation of 0.78 cm.

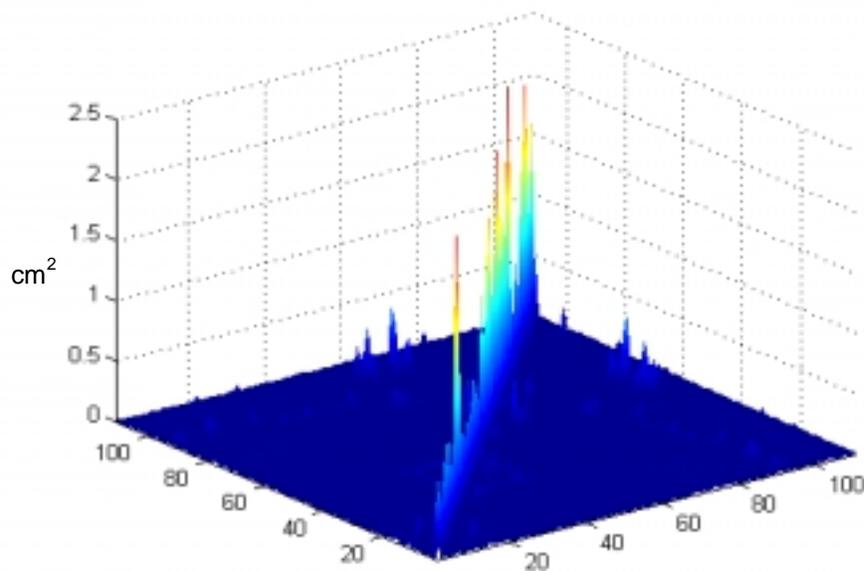


Figure 6.2: Plot of initial covariance matrix for GPS heights (Swiss network)

The average standard deviation is computed as follows:

$$\sigma_{ave} = \sqrt{\text{tr}(\mathbf{Q}_i)/m} \quad (6.1)$$

where, $\mathbf{Q}_i = \mathbf{Q}_h$ for this case. Eq. (6.1) intrinsically provides an optimistic measure of the accuracy as only the diagonal elements in the cofactor matrix are taken into consideration. Typical for GPS-derived coordinates, the output CV matrix was overly-optimistic. In practice, this situation is rectified by *arbitrarily* scaling the CV matrix by some factor. The validity of this all too common practice is addressed in section 6.3.1.

6.1.2 Initial covariance matrix for the geoid heights

The fully-populated CV matrix corresponding to the geoid heights at the GPS-levelling benchmarks was obtained by straightforward application of error propagation to the least-squares collocation equations that were used for the Swiss geoid determination (see Marti *et al.*, 2001 for more details)

$$\mathbf{Q}_N = \mathbf{C}_{NN} - \mathbf{C}_{N\Delta g} \mathbf{C}_{ZZ}^{-1} \mathbf{C}_{N\Delta g}^T \quad (6.2)$$

where \mathbf{C}_{NN} is the covariance matrix of the true unknown geoid heights, $\mathbf{C}_{N\Delta g}$ denotes the cross-covariance matrix between the computed geoid heights, N , and the measured gravity anomalies, Δg , and $\mathbf{C}_{ZZ} = \mathbf{C}_{\Delta g\Delta g} + \mathbf{C}_{nn}$ where n is noise. The computed CV matrix in this manner excludes the uncertainty contribution of the global geopotential model (commission and omission errors) as well as other effects such as terrain reductions and assumptions about the density models. Figure 6.3 depicts the initial fully-populated CV matrix for the geoid heights as computed from Eq. (6.2). As expected, the geoid contributes the most (compared to the other height error components) to the overall accuracy of the height misclosures at the GPS-levelling benchmarks, with an average

standard deviation of 1.92 cm. The comparatively high correlation for neighboring stations is also expected due to the smooth variation of the geoid signal over the network coverage area.

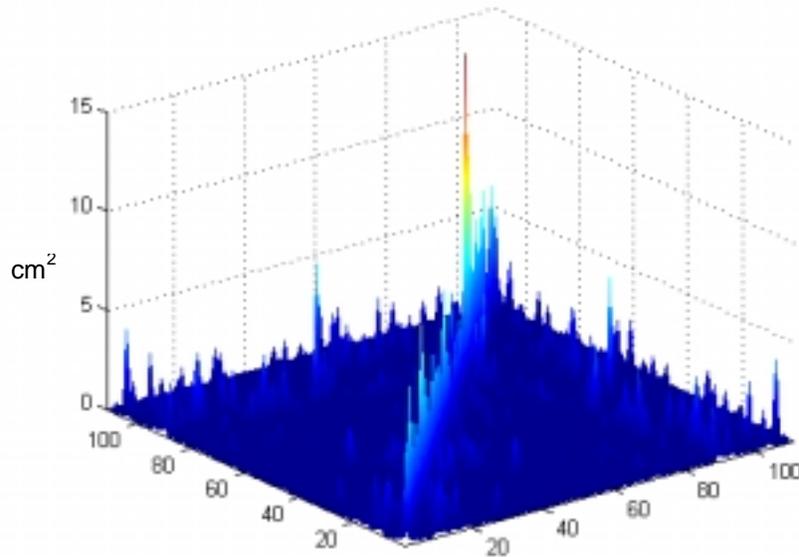


Figure 6.3: Plot of initial covariance matrix for geoid heights (Swiss network)

In practice, regional geoid models are often refined through the incorporation of GPS-levelling data into the gravimetric solution (as described in section 2.4.4). In such cases, the assumption of disjunctive observations is not strictly valid, which may adversely affect calculations of error models. To rectify this situation, the GPS-levelling observations can be excluded from the computation of the gravimetric geoid, which will ensure independence among the CV matrices, satisfying Eq. (5.13).

6.1.3 Initial covariance matrix for the orthometric heights

Of the three CV matrices required in the combined height network adjustment, the CV matrix for the orthometric heights is most readily available as it can usually be extracted

from the network adjustment of the geopotential numbers for a given set of points. This is often conducted through the use of in-house software by the responsible government agencies and provided to users upon request.

In this case, the initial fully-populated covariance matrix for the orthometric heights, \mathbf{Q}_H , comes directly from the rigorous national adjustment of the first and second order levelling measurements since 1902. The orthometric heights were obtained directly from the division of the adjusted geopotential numbers by the mean gravity (computed from surface gravity measurements and a simple 3D density model of the Earth's crust) along the plumb line. Marti (2002) should be consulted for a description of the computation of the orthometric heights and a comprehensive comparison of the differences between Swiss height systems and the approximations made in the computation of the mean gravity. Further information on the national levelling system in Switzerland is provided in Marti *et al.* (2001). Figure 6.3 depicts the initial covariance matrix, which has an associated average standard deviation of approximately 0.74 cm. As expected, the correlation between nearby neighbouring stations is very high.

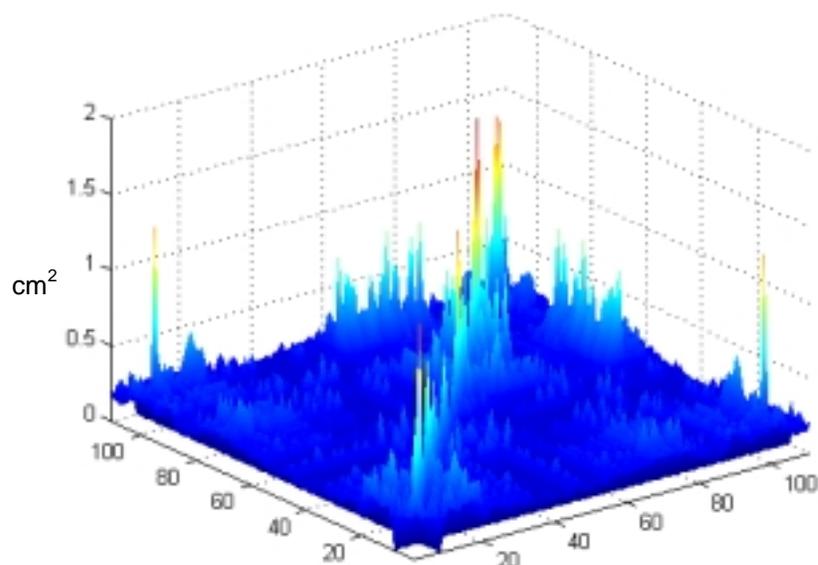


Figure 6.4: Plot of Initial covariance matrix for orthometric heights (Swiss network)

Table 6.1 summarizes some characteristics of the a-priori CV matrices for comparison purposes. The average error is computed from Eq. (6.1) and the condition number is computed from the following equation:

$$\text{cond}(\mathbf{Q}_i) = \frac{\lambda_{\max}}{\lambda_{\min}}, \quad i = h, H, N \quad (6.3)$$

where \mathbf{Q}_i is a positive definite cofactor matrix, and λ_{\max} and λ_{\min} are its maximum and minimum eigenvalues, respectively.

Table 6.1: Initial CV matrix characteristics for the Swiss network

	GPS	Levelling	Geoid
Condition number	146.2	2.23×10^7	4.50×10^5
Average standard deviation (cm)	0.79	0.75	1.93

The comparatively high condition number corresponding to the levelling data matrix is of particular interest. This high value may lead to numerical instability problems when the matrix needs to be inverted. Therefore, a simple ridge regression was applied to alleviate any numerical problems (see Bouman, 1998).

The average standard deviation gives an indication of the **overly-optimistic a-priori accuracy** for the three height types, which is especially evident for the ellipsoidal heights with an average error of less than 1 cm. From the discussion on the various error sources affecting ellipsoidal heights derived from GNSS measurements in section 2.3, it is apparent that this measure is optimistic.

The utility in the covariance matrix characteristics provided in Table 6.1 will become more evident in the numerical case studies, as the effect of ignoring possible (or known) correlations between observations may be dangerous if the condition number of the matrix is large (Strang and Borre, 1997).

6.2 Description of the southern British Columbia/Alberta test network data

Unlike the Swiss test network where a relatively homogeneous coverage of GPS-levelling benchmarks has been established over the entire country, in Canada the coverage is very mixed and too sparse in most parts. Furthermore, it is difficult to cover all parts of the country with existing control over a short period of time using conventional spirit-levelling techniques. This has resulted in a series of ongoing re-adjustment and evaluation of the data. For the numerical studies conducted herein, a 495 km \times 334 km region in southern British Columbia (BC) and Alberta (AB) is used. The Geodetic Survey Division (GSD) of National Resources Canada provided the height data at 63 stations and CV information for the ellipsoidal heights, orthometric heights and geoidal undulations.

Figure 6.5 shows the test network area and the associated standard deviations for the height misclosures at each benchmark. Overall, the average standard deviations, σ_{ℓ_i} , are about three times higher than in the Swiss network case, with a minimum of 3 cm and a maximum of approximately 12 cm. The 63 stations are spaced roughly 30 km apart, compared to about 20 km for the Swiss test network. The relatively poor representation of GPS-levelling benchmarks for the size of the area is a realistic problem encountered in many networks and therefore warrants testing for comparison purposes.

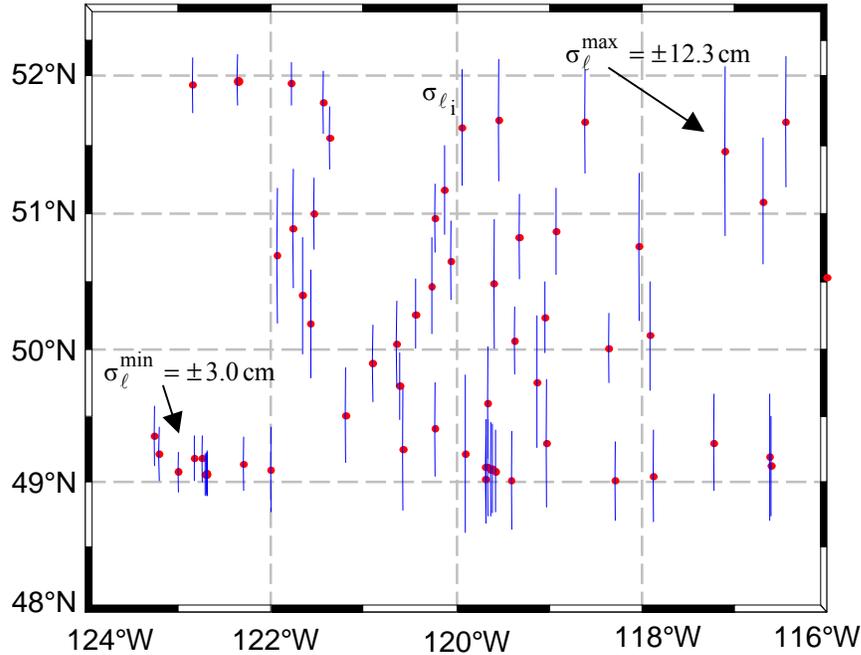


Figure 6.5: Southern BC/AB test network and error bars for $\ell = h - H - N$

6.2.1 Initial covariance matrix for the GPS heights

The weight matrix for the three-dimensional GPS coordinates was obtained from the output of the Geodetic Survey Division GHOST adjustment software (Geodetic adjustment using Helmert blocking of space and terrestrial data; *M. Craymer, personal communication*). The required CV matrix for the adjusted ellipsoidal heights, \mathbf{Q}_h , was extracted from this weight matrix. It is worth mentioning that a slight computational approximation has been made as the weight matrix corresponding to the ellipsoidal heights was extracted first and then inverted in order to get the CV matrix. Thus, any cross-correlation between the horizontal coordinates and the vertical component was lost as it was not involved in the inversion. The following results from the case studies in the next sections may explain the actual effect experienced from this obvious omission.

The a-priori CV matrix for the ellipsoidal heights in this region, \mathbf{Q}_h , is shown in Figure 6.5. The average standard deviation is approximately 3.17 cm, which is notably higher than the corresponding value for the Swiss network. The precision of the vertical coordinates for the Canadian Base Network (a national network of high precision GPS control stations) is approximately 3 cm at the 95% confidence level (Craymer *et al.*, 1997). Therefore, the computed value for this smaller region is on par with the national standards.

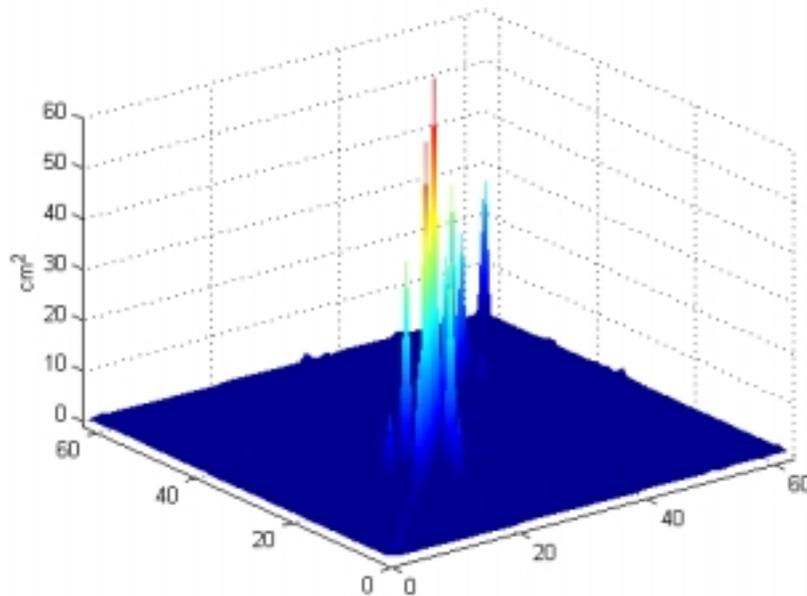


Figure 6.6: Plot of initial covariance matrix for GPS heights (southern BC/AB network)

6.2.2 Initial covariance matrix for the geoid heights

The a-priori error CV matrix, \mathbf{Q}_N , for the geoidal undulations was obtained through error propagation of an error grid of the Helmert gravity anomalies. The variance values corresponding to the 63 stations of interest were obtained through bilinear interpolation of the error grid. The final covariance matrix was provided by GSD in a **diagonal** form with all cross-correlations set to zero. Similar to the Swiss error CV matrix case, the

geoid error model does not include the uncertainty from the global geopotential model, which defines the long wavelength component and the gravity field outside the Stokes' integration cap-size of 6° (Véronneau, 2002). The diagonal a-priori CV matrix is shown in Figure 6.6. The overall average standard deviation is approximately 5.6 cm, which seems to be a more realistic value than that obtained for the Swiss network.

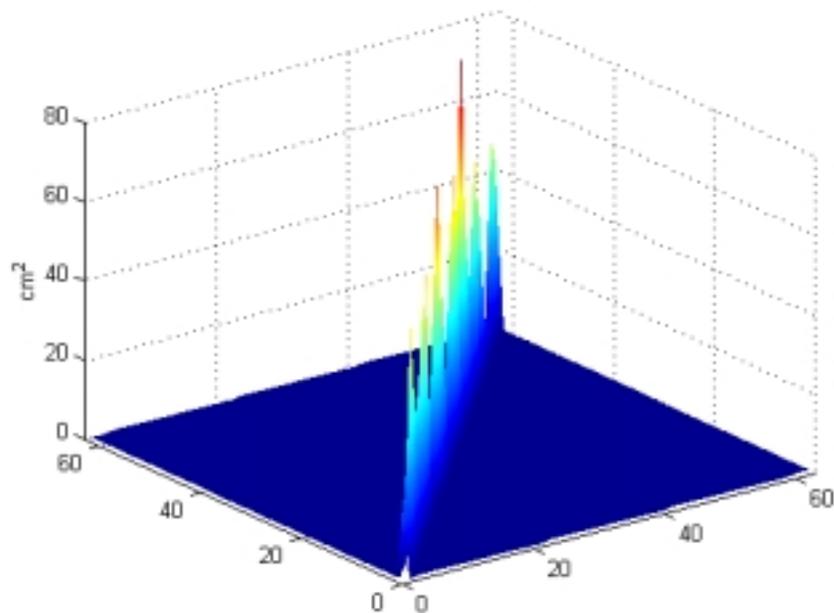


Figure 6.7: Plot of initial error CV matrix for geoid heights (southern BC/AB network)

6.2.3 Initial covariance matrix for the orthometric heights

A fully-populated error covariance matrix, \mathbf{Q}_H , for the orthometric heights in the southern BC/AB region was provided by GSD. The original matrix was in the form of a weight matrix for the corresponding geopotential numbers. In fact, most national network adjustments of levelling networks are performed using geopotential numbers as it provides the flexibility of converting to either orthometric or normal heights depending on the type of gravity used (see discussion in section 2.2). For the Canadian vertical data,

Helmert orthometric heights were used, computed using surface gravity values. All stations in the given region are of first order levelling and were part of a minimally constrained adjustment. The average standard deviation of this a-priori covariance matrix is 0.75 cm, which is coincidentally almost identical to the Swiss levelling covariance matrix.

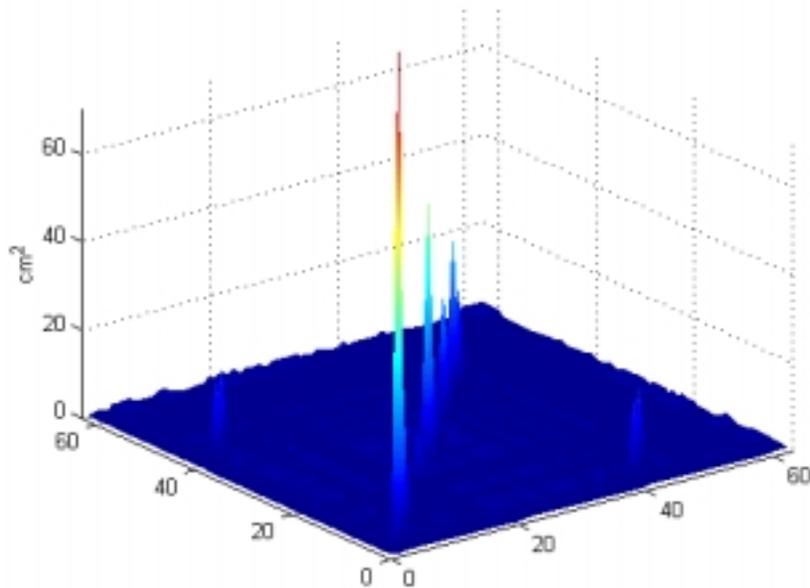


Figure 6.8: Plot of initial CV matrix for orthometric heights (southern BC/AB network)

Table 6.2 summarizes the characteristics of the a-priori covariance matrices for the southern British Columbia and Alberta test network. In general, these covariance matrices indicate more moderate a-priori levels than in the previous network. Once again, the geoid is the major contributor to the overall error budget in the combined height network adjustment. Surprisingly, the numerical stability of the matrices corresponding to the GPS and geoid height data are very good, exhibiting fairly low condition numbers. The effect that this will have on the variance component estimation results will be seen in the numerical case studies in the next sections.

Table 6.2: Initial CV matrix characteristics for the southern BC/AB network

	GPS	Levelling	Geoid
Condition Number	997.5	1.1×10^5	13.7
Average Error (cm)	3.17	2.99	5.57

6.3 Case Study I - Testing a-priori covariance matrices for the height data

The purpose of this case study is to test the effect of different a-priori CV matrices on the final estimated variance components. As stated previously, the IMINQUE scheme, in theory, is a globally best estimator and should provide estimates independent of the a-priori values. However, the procedure does not provide any guarantee of the *correctness* of the final estimated values. Therefore, mere convergence cannot be taken as a positive re-enforcer. To overcome this uncertainty, there are a number of options that can be investigated. For instance, one can compute the covariance matrix for the estimated variance components, $\mathbf{C}_{\hat{\theta}}$, using Eq. (5.39). From this equation, the standard deviations corresponding to each of the estimated variance components can be extracted and, depending on the relative magnitudes, inferences can be made regarding the ‘goodness’ of the estimated values. Another method, though beyond the scope of this study, is the computation of the associated confidence intervals for each of the variance components (Koch, 1987).

In this study, a more empirical approach is followed whereby a number of different a-priori variance values $\sigma_{i_0}^2$ for each height type are tested to see if they all yield the same solution $\hat{\sigma}_i^2$. Interestingly, the results showed that the initial a-priori values for a single group of observations (e.g., orthometric heights) do not have a noticeable effect on the remaining groups, even though theoretically co-dependence is present. Figure 6.9 depicts an example of this procedure using the Swiss test network data, where three different a-priori variance values for the orthometric heights are used as input to the IMINQUE

algorithm and the results for the final estimated variance values for the other two groups of observations (ellipsoidal and geoid heights) remain unchanged. This process was repeated for each of the height types and gave similar results.

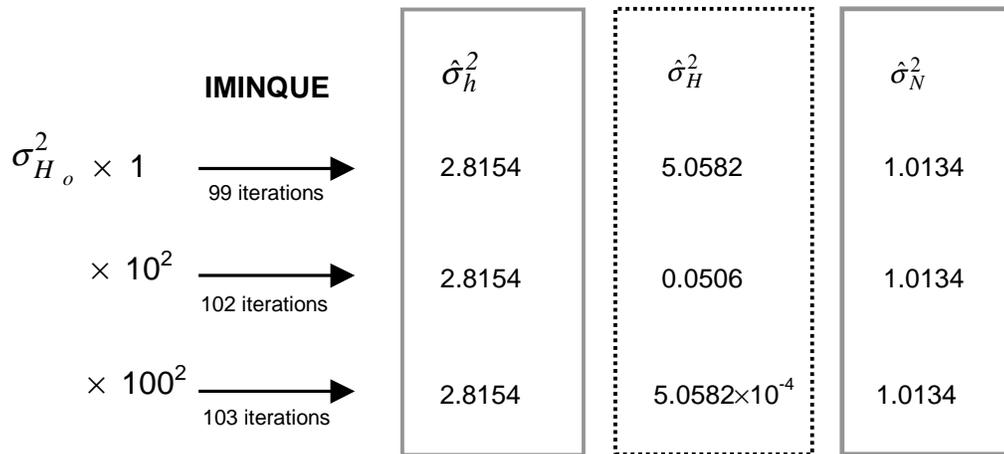


Figure 6.9: Effect of different a-priori variance factors (Swiss network, $\sigma_{H_o}^2 = 1$)

Tests were also conducted using the Canadian data set and the same conclusions could be drawn, as illustrated in Figure 6.10, for different a-priori variance levels corresponding to the geoidal height data. An interesting observation, for both test networks, is that the number of iterations required as the initial values deviate from the 'true' unknown variance factors do not change much. The results also reveal the considerable difference in the number of iterations required for the different test networks with approximately 100 for the Swiss data and 12 for the Canadian regional dataset! This may be related to the numerical stability of the initial covariance matrices, which was noticeable better for the Canadian data, or it may also be an indication of the more realistic measures of the CV information for the latter network as compared to the optimistic height accuracy values for the Swiss network. The Swiss network also contains about twice as many

stations, a fact which increases the computational load, and thus the number of necessary iterations.

$\sigma_{N_o}^2 \times 1$	IMINQUE \longrightarrow 12 iterations	$\hat{\sigma}_h^2$ 7.1965	$\hat{\sigma}_H^2$ 3.9745	$\hat{\sigma}_N^2$ 0.2788
$\times 10^2$	\longrightarrow 14 iterations	$\hat{\sigma}_h^2$ 7.1965	$\hat{\sigma}_H^2$ 3.9745	$\hat{\sigma}_N^2$ 0.0028
$\times 100^2$	\longrightarrow 14 iterations	$\hat{\sigma}_h^2$ 7.1965	$\hat{\sigma}_H^2$ 3.9745	$\hat{\sigma}_N^2$ 2.788×10^{-5}

Figure 6.10: Effect of different a-priori variance factors (BC/AB network, $\sigma_{N_o}^2 = 1$)

In the process of conducting these tests, it was also found that this procedure could be used for more practical purposes, such as:

- testing and scaling the a-posteriori covariance matrix for the height coordinate provided by the GPS post-processing software
- calibrating the geoid error CV matrix for describing the internal accuracy in local or global geoid models
- assessing and evaluating the accuracy information for orthometric heights obtained from national and/or regional network adjustments of levelling data

In this case study, these issues will be investigated in detail and the results provided in the following three sections.

6.3.1 Scaling the covariance matrix for the ellipsoidal heights

It is well known that the covariance matrix for the estimated three-dimensional position parameters obtained from the processing of GPS phase and code data is usually overly-optimistic (Satirapod *et al.*, 2002). This is a direct result of neglecting, even partially, correlations (physical, spatial and temporal) between GPS observables in the creation of the covariance matrix for the input observations. As a result, not only is the usefulness of the computed CV matrix for the GPS coordinates poor (thus giving a false indication of the quality of the results), but it is also unlikely to be appropriate as input into another adjustment (multi-baseline or heterogeneous adjustment with terrestrial data). The solution offered in practice to users is that (Geodetic Survey Division, 1992; page 15):

" ... it is acceptable when justified to scale the formal covariance matrix provided by such software prior to the final adjustment."

Over the years, experience in processing of GPS data has resulted in a variety of different scale factors commonly used, i.e., 30 or 100. In Ollikainen (1997), the formal accuracy for the GPS-derived ellipsoidal heights resulting from the processing with the Bernese software were found to be too optimistic and scaled by empirically derived factors of 8^2 and 10^2 for two different datasets. Obviously, every network is different and the blanket application of a somewhat arbitrary scale factor is not a viable solution, predominantly in high precision surveys. In light of this, much research has been conducted on the difficult task of improving the stochastic model for the observations by incorporating the correlations often neglected. In doing so, the output covariance matrix for the position parameters is less likely to be optimistic.

Among the most significant sources of physical (both spatial and temporal) correlations in GPS phase measurements are the ionospheric and tropospheric delays. In El-Rabbany (1994), an exponential covariance function was derived whereby the temporal physical correlation was taken into account. Numerous other studies have been conducted which include modifying the covariance matrix by modelling individual sources of errors such

as the atmospheric, multipath, receiver noise, and orbital errors, over time and space in order to provide fully-populated covariance matrices for the observables. Needless to say, this is a difficult task and research is ongoing. Recently, a study was conducted whereby VCE (specifically the BQUE scheme) was used to construct an improved CV matrix for the processing of the GPS phase and code observables (Tiberius and Kenselaar, 2003). The stochastic model was refined in steps by estimating variance components for each satellite/channel, covariance components between satellites/channels, covariance components between observation types (L1, L2, C1, P2), and finally the covariances between epochs (temporal correlations).

Despite these efforts, thus far, no commercial post-processing software package openly implements such advanced methods for dealing with correlations and therefore overly-optimistic CV matrices are computed. In fact, some packages do scale the output covariance matrix for the positions, however, rarely is the procedure and/or the scale factor known to the user. For the purposes of the combined adjustment of satellite-derived ellipsoidal height data with other terrestrial height data, it is imperative that the covariance matrix for the GPS height component, used as input into the combined adjustment, be appropriate.

During the course of this study, it was discovered that through proper VCE, the arbitrariness of certain pre-selected factors (such as scaling \mathbf{Q}_h) can be alleviated. For instance, in the Swiss network, a factor of 10^2 was empirically determined as being an appropriate scale for \mathbf{Q}_h (*Urs Marti, personal communication*, see also Figure 6.2). By re-scaling the original covariance matrix with different a-priori factors, as illustrated in Figure 6.11, the independently-derived value was verified as an appropriate scale for \mathbf{Q}_h , which results in an estimated variance component of $\hat{\sigma}_h^2 = 2.8154$. This is an important realization as it ascribes some statistical reasoning towards an otherwise ‘arbitrary’ scaling practice and leads to a better or improved understanding of the true measures of error.

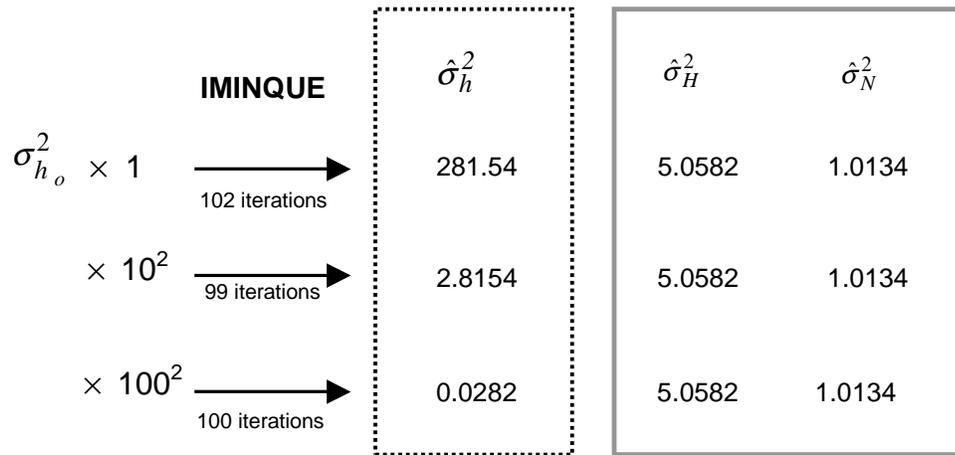


Figure 6.11: Effect of different a-priori variance factors (Swiss network, $\sigma_{h_o}^2 = 1$)

For the Canadian region, the supplied covariance matrix for the ellipsoidal heights did not come with a known pre-selected scaling factor, although it is possible that some scaling may have been applied at the outset. To check if this was the case and to verify the provided matrix, the same procedure was applied to the initial covariance matrix by re-scaling using various a-priori values. It was found that the original matrix was sufficient and did not have to be pre-multiplied by some arbitrarily large factor, such as 100, as was the case for the Swiss data. The final estimated variance component computed for the ellipsoidal heights in this data set was, $\hat{\sigma}_h^2 \cong 7.2$. Thus, the incorporation of the iterative VCE procedure in the combined height adjustment proved to be useful in determining and validating the appropriate scale for the covariance matrix of the ellipsoidal heights.

6.3.2 Calibration of geoid error models

In general, there are two ways for assessing the accuracy of computed gravimetric geoid models. A widely used approach is the comparison of different geoid models over the same region, which have been derived using various computational schemes and data

sets. This is an external way of evaluating the accuracy of existing geoid models and has been implemented in a number of cases on regional and global geoid modelling studies. Closely related to this approach, is the computation of the differences between the gravimetrically-derived geoid values, N_{grav} , with those derived from GPS-levelling, $N_{GPS/levelling}$, at pre-specified discrete locations such that

$$\Delta N = \left| N_{grav} - N_{GPS/levelling} \right| = \left| N_{grav} - h + H \right| \quad (6.4)$$

The statistics of the discrepancies ΔN offer an external empirical evaluation of the gravimetric geoid model accuracy (as described in section 2.4.4). However, there are a number of issues that must be considered before a proper assessment of the geoid model is given using this approach. Firstly, when computing the discrepancies as given by Eq. (6.4) it is often assumed that the GPS-levelling derived geoid values are errorless and therefore used as a basis for comparison. This is obviously a major assumption and should be considered with caution. In addition, it is common practice to use the GPS-levelling data in order to refine the existing gravimetric geoid model. Thus, the computed 'gravimetric' geoid heights have been influenced already by the GPS-levelling values, which are then used as a basis for comparison in assessing the accuracy of the geoid. This approach only provides a means for checking how close the existing model is to the GPS-levelling data, rather than an independent external accuracy assessment. Another major issue in this comparison is the described systematic effects and datum inconsistencies that exist between the three types of height data, h , H and N (see chapter 3 for a detailed discussion). It is important that any systematic effects, if any, are eliminated from the height data so that a reliable description of the behaviour of the random errors can be made. Therefore, the importance of an efficient corrector surface model is stressed for the evaluation of gravimetrically-derived geoidal undulations using GPS-levelling data.

The second method for assessing the geoid error model is through the proper propagation of errors in the used data (i.e. local gravity anomalies, terrain models, etc.) and the

geopotential harmonic coefficients into the final geoid undulations (for detailed formulas, see Sideris and Schwarz, 1987; Strang van Hees, 1986). Other important studies are described in Li and Sideris (1994) where the propagated errors of the terrestrial gravity anomalies and the geopotential coefficients from the global gravity model were combined to form an error model in parts of Canada and the northern United States. In these regions, it was shown that internal propagated variances reasonably reflected the external estimated accuracy conducted via comparisons to GPS-levelling data. Such internal accuracy assessments were used in order to obtain the initial covariance matrices for the described Swiss and Canadian regional test networks; see sections 6.1.2 and 6.2.2, respectively. In the Swiss network, for instance, error propagation was performed on the results of the least-squares collocation formulation (see Eq. 6.2). In the Canadian region, a different approach was taken whereby the errors in the gravity anomalies were propagated in the geoid heights (Véronneau, 2002). It should be noted that for both local covariance matrices the uncertainty of the global geopotential model was not taken into account.

In order to obtain an understanding of the uncertainty of the global geopotential model to the overall error budget, the fully-populated CV matrix corresponding to the EGM96 global geopotential model in the network areas was computed by applying error propagation to the spherical harmonic expansion of the Stokes' formula as given by Eq. (2.2). As input, the fully-populated covariance matrix for the commission errors of the spherical harmonic coefficients ($\bar{C}_{nm}, \bar{S}_{nm}$) up to a maximum degree and order of 70 (Lemoine *et al.*, 1998) was used and denoted by C_{coef} . Although a higher degree of expansion (i.e., 180 or 360) may theoretically recover higher frequency information (thus reducing the aliasing error), the noise is also increased as the number of coefficients increases. Furthermore, only variance values (no covariances) are available for the coefficients for higher degree and order 70.

The final covariance matrix corresponding to the commission error of EGM96 is computed via the following formula:

$$\mathbf{C}_{EGM96} = \mathbf{D}\mathbf{C}_{coef}\mathbf{D}^T \quad (6.5)$$

where the \mathbf{D} matrix can be written as a series of sub-matrices such that

$$\mathbf{D} = [\mathbf{d}_1 \quad \mathbf{d}_2 \quad \dots \quad \dots \quad \mathbf{d}_i]^T \quad (6.6)$$

where \mathbf{d}_i relates the spherical harmonic coefficients to the geoidal undulations for each of the network points, i , and is given by the following expression:

$$\mathbf{d}_i = \begin{bmatrix} P_{20} \cos m\lambda_i & P_{21} \cos m\lambda_i & \dots & P_{nm-1} \cos m\lambda_i & P_{nm} \cos m\lambda_i \\ P_{20} \sin m\lambda_i & P_{21} \sin m\lambda_i & \dots & P_{nm-1} \sin m\lambda_i & P_{nm} \sin m\lambda_i \end{bmatrix}^T \quad (6.7)$$

It should be noted that only the variance values are provided for the zero to second order coefficients, namely $\bar{C}_{20}, \bar{C}_{21}, \bar{C}_{22}, \bar{S}_{21}, \bar{S}_{22}$ (*ibid.*).

Assuming the systematic effects among the different height types have been sufficiently eliminated (e.g., through the use of a proper parametric model), the VCE procedure leads to the refinement or calibration of existing geoid error models through the combined height adjustment process. To illustrate its usefulness, the IMINQUE procedure was used to estimate the a-posteriori variance factors for three different geoid error models in each of the Swiss and Canadian data sets, namely

- (i) local geoid error model derived from propagated errors of terrestrial gravity and height data

- (ii) global geopotential geoid error model derived from propagated variance-covariance information of the spherical harmonic coefficients for the EGM96 model up to degree and order 70 (see Eq. 6.5)
- (iii) diagonal-only version of model (ii).

For the case of the error model of type (i), the covariance matrix for the Swiss data is fully-populated (see section 6.1.2), while for the Canadian regional data only a diagonal covariance matrix is used (see section 6.2.2). The results are summarized in Table 6.3.

Table 6.3: Estimated variance components and average standard deviations for local and global geoid models

Test Network	local geoid model	EGM96 (diagonal)	EGM96 (full)
Switzerland			
original average σ	1.9 cm	25.4 cm	25.4 cm
estimated variance factor	1.02	6.2×10^{-5}	0.27
scaled average σ	1.9 cm	0.2 cm	13.1 cm
Southern BC/AB			
original average σ	5.6 cm	24.7 cm	24.7 cm
estimated variance factor	0.28	0.05 *	0.21
scaled average σ	2.9 cm	5.8 cm	11.4 cm

* computed using the IAUE method

The table shows the computed average standard deviation (Eq. 6.1) both before and after VCE is performed, which indicates the change in the corresponding accuracy level. The final estimated variance components are also provided for each test case. The first column refers to the results obtained when using the original covariance matrices (of type (i)) described in sections 6.1.2 and 6.2.2 for the Swiss and southern BC/AB test networks, respectively. The second column gives the results using a diagonal covariance matrix corresponding to the commission errors of the global geopotential model EGM96 (of type

(iii)) and finally the last column refers to the fully-populated covariance matrix for EGM96 up to degree and order 70 (type (ii)). All of the results were obtained by fitting the classic 4-parameter corrector surface model (Eq. 3.31) to the height residuals in order to deal with the intrinsic systematic effects and then applying the IMINQUE algorithm described in section 5.6. The only exception is the estimated variance component corresponding to the diagonal version of the EGM96 geoid error model in the Canadian region. In this case, a positive-valued solution was not numerically obtainable using IMINQUE and therefore the computed value was computed using the IAUE algorithm (see section 5.7).

By comparing the original average σ for the locally derived geoid error models and that of EGM96 it is evident that the contribution of EGM96 to the total error budget is significant. For the Swiss test network case, the scaled average σ is approximately equal to the original average σ as the estimated variance factor, $\hat{\sigma}_N^2$, is close to one, which validates that the provided CV matrix was reasonable. This does not imply that the *total* uncertainty in the geoid heights in this area is described by the CV matrix as there are many sources of error that have been neglected (as mentioned in section 6.1.2).

The covariance matrix for EGM96 in this region is shown to be pessimistic with an original average σ of 25.4 cm compared to a scaled average σ of 13.1 cm. The gravity coverage in this region is relatively homogeneous in coverage and quality, which possibly contributes to the down-scaling of the CV matrix giving a final estimated variance factor of 0.27. The results computed when the diagonal CV matrix is used reflect the overly-optimistic accuracy values obtained when significant correlations are neglected; see Figure 6.12 for the original fully-populated covariance matrix corresponding to EGM96 in this region.

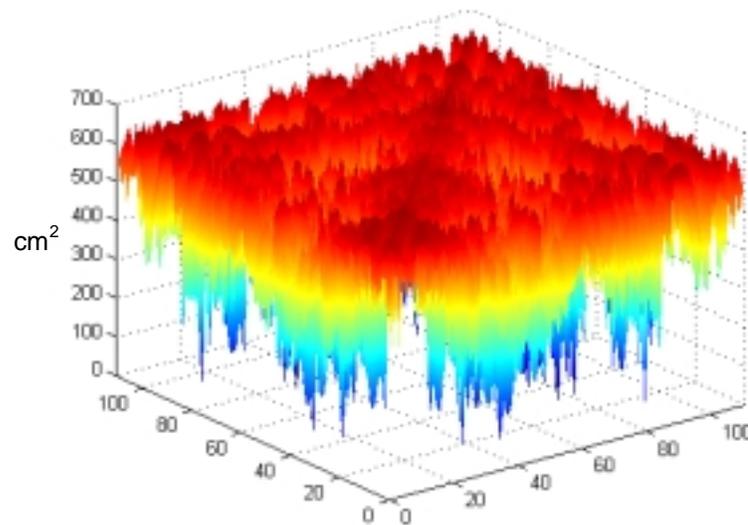


Figure 6.12: Plot of covariance matrix for EGM96 over the Swiss network area

For the Canadian test network, similar conclusions can be drawn regarding the EGM96 error model. However, the locally derived geoid error model, although diagonal-only, was slightly pessimistic at 5.6 cm compared to an average scaled error of approximately 3 cm. The contribution of the global geopotential model to the overall error budget is estimated at approximately 11 cm in this region.

Overall, the test results show the capability of the VCE procedure for calibrating both regional and global geoid error models. In all cases, the estimated variance components were less than one, indicating a down-scaling of the initial CV matrix was necessary. Further to these error models, the global gravity field models expected from the current CHAMP and GRACE and upcoming GOCE missions in the near future will also need to be validated. The proposed VCE technique can be used with terrestrial data (assuming that it is also free of systematic errors) in order to reliably calibrate the newly-derived geoid error models. Thus, the co-dependency between the proper modelling of systematic errors and the effectiveness of the VCE method is evident.

6.3.3 Assessing the accuracy of orthometric heights

In general, the most precise height component in the mixed adjustment are the orthometric heights derived from precise spirit-levelling techniques and gravity data (see section 2.2). Over the years, numerous methods have been tested for evaluating the precision of levelling measurements, including correlation analysis, which involved postulating appropriate covariance functions for the nature of correlations, analysis of variance methods (Kelly, 1991), variance component estimation schemes (Chen *et al.*, 1990), and analysis of section line and loop discrepancies. The studies that have been conducted thus far on assessing the accuracy of levelling have predominantly focused on determining the appropriateness of the variance of a levelling line based on a function of segment length d and height difference ΔH

$$\sigma_{\Delta H}^2 = f(d, \Delta H) \quad (6.8)$$

and can be approximated by the general formulation in Eq. (5.17).

The aforementioned methods usually involve height-difference data of the same type (i.e. levelling) with heterogeneous quality distinguished according to classes or orders of levelling. In this study, the covariance matrix for the absolute orthometric heights is evaluated by testing a number of different a-priori variance factors. The scaled covariance matrix for the orthometric heights can then be used as a more reliable estimate of the accuracy for large or multinational adjustment of levelling networks. Table 6.4 summarizes the estimated variance components for both the Swiss and Canadian test networks, as well as the average standard deviation before and after VCE.

For the case of the Swiss network, the initial CV matrix was extremely optimistic with an average standard deviation of less than 1 cm. The estimated variance factor of 5.06 results in an increased average standard deviation of almost 2 cm.

Table 6.4: Estimated variance components for orthometric height data and average standard deviations for C_H

	Switzerland	Southern BC/AB
original average σ	0.75 cm	2.99 cm
estimated variance factor	5.06	3.97
scaled average σ	1.68 cm	5.96 cm

In the southern BC/AB test network, a more realistic initial measure of accuracy was given with an average standard deviation of ~ 3 cm. The resulting variance factor was 3.97, which provided a final scaled average standard deviation of almost 6 cm for the orthometric heights in this area. Compared to the Swiss test network case, these values are significantly higher, however, the Canadian height network is plagued with numerous systematic and random errors that result in poorer accuracy, as described in chapters 2 and 3.

Overall, the results provide a more realistic measure for the accuracy of the orthometric heights, compared to the optimistic initial values. Such CV testing can also be considered beneficial in assessing national and regional height networks.

6.4 Case Study II - Non-negative variance components

One of the major pitfalls of the MINQUE variance component estimation algorithm is that no provision is in place to ensure that only non-negative variance values ($\hat{\theta} \in \mathfrak{R}_+^3$) are estimated. However, upon closer examination of the problem and after many years of research, it has been found that there are plausible methods for dealing with this unrealistic result, which may also provide valuable insight into the problem at hand. As mentioned previously negative outcomes of variance components can generally be attributed to:

- insufficient number of observations compared to the unknown parameters (low redundancy), and/or
- incorrect stochastic model.

Thus, although the estimated negative variance factor itself is not useful, a negative variance outcome may yield important information regarding the problem set up. This information is lost if the estimator is constrained to give only positive outcomes (Sjöberg, 1984).

The numerical tests conducted using the provided CV matrices for the height data in the Swiss and Canadian test networks have not exhibited negative outcomes. It is suspected that this can partially be attributed to the high redundancy in both cases. To determine the effect of data redundancy on the estimated variance components, using the IMINQUE approach, a test was performed whereby observations were eliminated (one-by-one) and the behaviour (positive or negative) of the estimated components was noted. For the absolute height data formulation, this is equivalent to removing a station from the network.

Figure 6.13 provides a summary of the results for the Swiss network, where it is seen that at least 49 observations are required for convergence and positive-valued final variance components. This corresponds to 44% of the available GPS-on-benchmarks in this region. Furthermore, if fewer than 49 observations were available convergence was not possible. This undesirable outcome is frustrating in practice and has deterred the widespread application of rigorous VCE procedures.

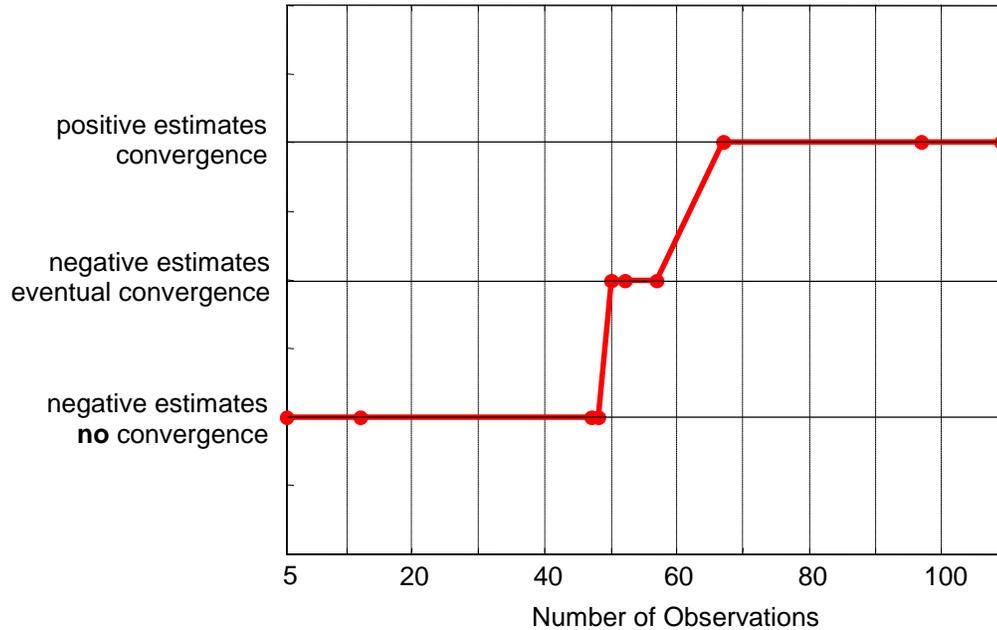


Figure 6.13: Behaviour of estimated variance components vs. redundancy for the Swiss network (note: all three variance components considered)

In the southern BC/AB regional test network positive estimates and definite convergence were achieved when more than 9 observations were used, that is 14% of the total available data in the region. With less than 9 observations, a positive outcome was not achievable using the iterated MINQUE algorithm. By comparing Figures 6.13 and 6.14, it is evident that for the Canadian test network there are more cases where negative outcomes are initially estimated and then gradual convergence to non-negative values is achieved. This may be partially attributed to the overall poorer initial CV information available for the particular region (i.e., diagonal-only covariance matrix for the geoid heights). It is also likely that the inhomogeneous and relatively sparse data distribution plays a significant role, as the removal of each station has a more prominent impact on the overall geometry in this case, as opposed to the more homogeneous data distribution in the Swiss network.

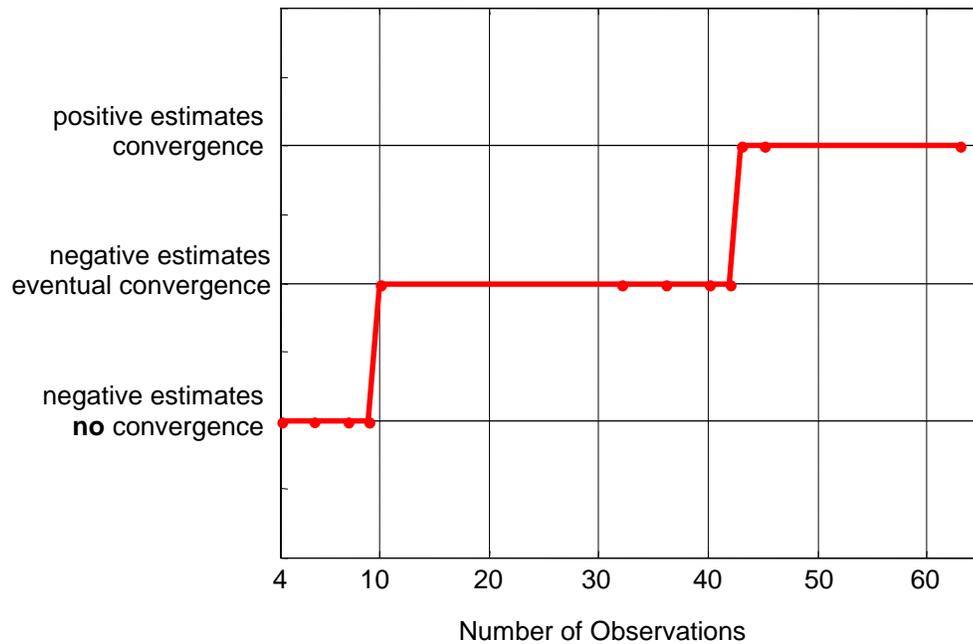


Figure 6.14: Behaviour of estimated variance components vs. redundancy for the southern BC/AB network (note: all three variance components considered)

In effect, these tests raise the interesting question of what is implied by an algorithm that gives negative estimates initially and then converges to positive values. *Should these final values be trusted?* In practice, the only values of interest are the final estimated components. Thus, for the Swiss and Canadian test networks at least 49 and 9 observations (as shown in the figures), respectively, are required. Obviously these values are network dependent and subject to change if the geometry of the network or CV information changes. Theoretically, mapping the behaviour of the estimator as described above provides some insight into the limitations and the achievable results using the unconstrained iterative MINQUE algorithm.

Tests using the IAUE method

As mentioned in section 5.7, simplified algorithms exist which ensure that positive variance component estimates are obtained at each iteration. One such algorithm, known as iterative almost unbiased estimation (IAUE) can be implemented through Eq. (5.40). This algorithm was tested with the real data sets and the results for the estimated variance

components are provided in Table 6.5 along with the rigorous IMINQUE outcomes. All results are computed after the classic 4-parameter model is fit to the height misclosures. Almost unbiased estimators consistently resulted in 'lower' variance factors for all tests. However, the estimated components for both test networks using either the rigorous or the constrained algorithm are very similar, with insignificant differences in all three components.

Table 6.5: Estimated variance components using the IMINQUE and IAUE methods

VCE Method	Switzerland			Southern BC/AB		
	$\hat{\sigma}_h^2$	$\hat{\sigma}_H^2$	$\hat{\sigma}_N^2$	$\hat{\sigma}_h^2$	$\hat{\sigma}_H^2$	$\hat{\sigma}_N^2$
IMINQUE	2.82	5.06	1.02	7.20	3.97	0.28
IAUE	2.69	4.74	0.97	7.00	3.76	0.26

An added benefit of the IAUE method is that it is computationally simpler and converges approximately 50% faster than the rigorous approach. Figure 6.15 shows the estimated variance components at each iteration as they collectively converge to unity using both the IMINQUE and the IAUE algorithms.

The figure shows that the final estimated variance components are computed after 76 iterations using the IMINQUE method, as opposed to only 45 for the IAUE method. These results refer to the Swiss data, but similar results were obtained using the Canadian data, with the IAUE method converging to the values shown in Table 6.5 in only 12 iterations as compared to 34 using the IMINQUE method. Thus, in cases where computational efficiency is an issue, IAUE offers a viable alternative to the rigorous approach; see Rao and Kleffe (1988) for details regarding this simplified positively-constrained algorithm.

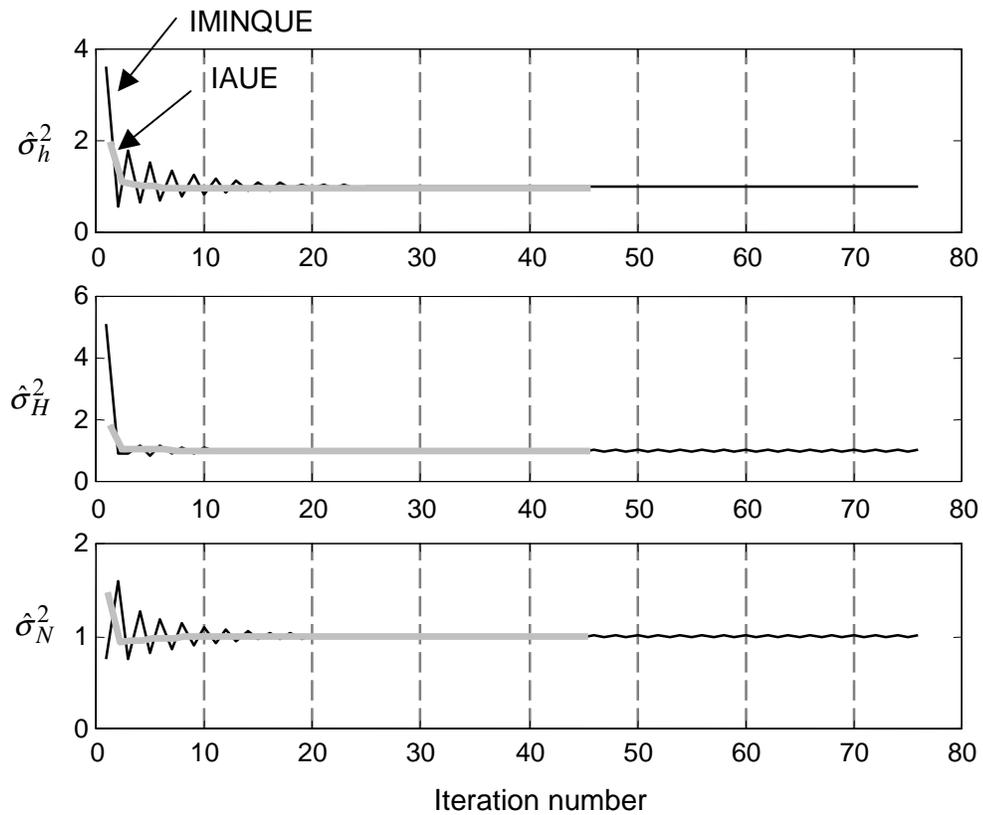


Figure 6.15: Iterations and estimated variance components using the IMINQUE and IAUE methods

When computational efficiency is not an issue, as in the case at hand, then a viable means for avoiding the computation of negative variance components with the IMINQUE method is to increase the degrees of freedom by:

- adding more observations
- reducing the number of unknown variance components

In the problem studied herein, the number of variance components was limited to three, however increasing the number by a few more components will not adversely affect the VCE scheme as a high degree of freedom will be retained.

6.5 Case Study III - Effects of correlations

In practice, fully-populated covariance matrices for each group of height data ($\mathbf{Q}_h, \mathbf{Q}_H, \mathbf{Q}_N$) are not made available to the user or they are difficult to obtain. To overcome this, some users resort to approximating CV information from manufacturer's specifications or measurement accuracy, analysis of physical sources of observational errors, and experience from working with the data. For very large data sets it may be preferable to deal with only the variances, therefore often a diagonal covariance matrix is implemented in secondary adjustments. To test the effect of correlations between observations of the same type on the estimated variance components, numerical experiments were conducted with fully-populated and diagonal covariance matrices. The results for the Swiss test network are given in Table 6.6.

Table 6.6: Effect of correlations on estimated variance factors (Swiss network)

Covariance Matrices	$\hat{\sigma}_h^2$	$\hat{\sigma}_H^2$	$\hat{\sigma}_N^2$	Number of iterations
Full	2.82	5.06	1.02	99
Diagonal	0.71	3.63	1.07	152

It is evident from the results in the table that due to the correlation among the heights, a diagonal covariance matrix is further from the 'true' CV matrix and therefore requires more iterations to obtain the final estimated values, i.e., 99 compared to 152 for the diagonal-only covariance matrices. Also, by neglecting the off-diagonal elements, overly-optimistic CV matrices are obtained compared to the fully-populated case. This is particularly demonstrated in the computed $\hat{\sigma}_h^2$ value where relatively high correlations existed between neighbouring stations for the ellipsoidal heights, thus underestimating the variance component by a factor of four. An exception is shown in the computed $\hat{\sigma}_N^2$ factor where the estimated values corresponding to the fully-populated and diagonal-only

CV matrices are essentially the same, with a slight increase for the diagonal-only CV matrix. Results will vary depending on the degree of correlation, however it is clear that unrealistically ‘good’ results are obtained when correlations are ignored, as expected.

The test was repeated using the southern BC/AB test network data and the results are provided in Table 6.7. In this case, only \mathbf{Q}_h and \mathbf{Q}_H were fully-populated with an original \mathbf{Q}_N in diagonal form. Similar results are exhibited in this case, as with the Swiss data, with the interesting exception of the estimated variance factor for the geoid heights. Although the original CV information provided for the geoidal undulations does not change from one scenario to the other (remains diagonal-only form), the fact that the prior CV matrices for the remaining types of heights is modified to remove the effect of correlations results in a change in $\hat{\sigma}_N^2$ from 0.28 to 0.48. Both values are overly-optimistic for this mountainous region, however the important realization is the effect that changing the form of the CV matrices has on the final estimated variance factors. This 'connection' was not noticed in the results of case study I, where different a-priori scaling factors for the CV matrices were tested.

Table 6.7: Effect of correlations on estimated variance factors (south BC/AB network)

covariance matrices	$\hat{\sigma}_h^2$	$\hat{\sigma}_H^2$	$\hat{\sigma}_N^2$	number of iterations
Full $\mathbf{C}_h, \mathbf{C}_H$	7.20	3.97	0.28	12
Diagonal \mathbf{C}_N				
Diagonal	7.71	1.63	0.48	17

These conclusions can also be seen in Table 6.3 by comparing the diagonal-only EGM96 covariance matrix with the fully-populated version. Again, for both test network cases, overly-optimistic variance factors resulted from ignoring the correlations, which is often

encountered in practice when implementing the contribution of the global geopotential models due to the very large matrices involved. The results show that the off-diagonal elements should not be considered insignificant and efforts should be made, when possible, to include all of the available CV information, especially for high precision applications.

6.6 Case Study IV - Role of the type of parametric model

All variance component estimation procedures outlined in chapter 5, pre-suppose that *no biases* or *systematic effects* are present in the data. Any unmodelled effects may propagate into the estimated variances and give unreliable results (Persson, 1981; Koch, 1999; Rao and Kleffe, 1988). This case study was designed to determine the role of the type of the parametric model, if any, on the final estimated variance components. In particular, it was suspected that the inappropriateness of a parametric model would be revealed through the VCE procedure, which is influenced by unmodelled systematic effects. A recent study using simulated data was conducted to detect systematic errors through VCE, and initial effects were evident (Nafisi, 2003).

Six different parametric models were selected for the investigations in this case study, namely the classic 4-parameter transformation model, described in chapter 3 and repeated below for completeness

$$p_c(\varphi, \lambda) = x_0 + x_1 \cos \varphi \cos \lambda + x_2 \cos \varphi \sin \lambda + x_4 \sin \varphi \quad (6.9)$$

four simple nested polynomial regression models from first to fourth order as given by

$$p_{nm}(\varphi, \lambda) = \sum_{m=0}^4 \sum_{n=0}^4 x_{nm} (\varphi - \varphi_0)^n (\lambda - \lambda_0)^m \quad (6.10)$$

where φ, λ are the horizontal coordinates of the benchmarks and φ_o, λ_o represent the midpoint of the network. In addition to the models given by Eqs. (6.9) and (6.10), a constant bias (i.e., the first term, x_o , in Eq. 6.9) was also tested for the purposes of discussion. The computed estimated variance components corresponding to each model type for the Swiss and southern BC/AB test networks are shown in Tables 6.8 and 6.9, respectively.

Table 6.8: Estimated variance components for the Swiss test network using various parametric models

Parametric Model Type	$\hat{\sigma}_h^2$	$\hat{\sigma}_H^2$	$\hat{\sigma}_N^2$
constant bias	2.67	10.45	0.99
4-param model	2.82	5.06	1.01
1 st n=m=1	2.83	4.80	1.01
2 nd n=m=2	2.94	4.50	1.05
3 rd n=m=3	3.08	3.96	0.97
4 th n=m=4	divergence, negative estimates		

Table 6.9: Estimated variance components for the southern BC/AB test network using various parametric models

Parametric Model Type	$\hat{\sigma}_h^2$	$\hat{\sigma}_H^2$	$\hat{\sigma}_N^2$
constant bias	7.16	16.59	0.12
4-param model	7.20	3.97	0.28
1 st n=m=1	6.92	4.72	0.26
2 nd n=m=2	7.85	3.58	0.28
3 rd n=m=3	divergence, negative estimates		
4 th n=m=4	divergence, negative estimates		

From the results in the tables, it is clear that the type of parametric model used in the combined least-squares adjustment affects the estimated values for the variance components. In both networks, the variance factor for the orthometric heights was influenced the most by the change in corrector surface, resulting in differences of several centimetres in the estimated variance factors from the use of a constant bias and the other parametric models.

In all cases, the number of iterations required to converge to the final estimated components also varied depending on the type of model used. For instance, in the Swiss test network, 99 iterations were needed for the case of the 4-parameter model, whereas 180 iterations were required for the third-order bivariate polynomial and over 250 iterations for the estimates obtained after a constant fit.

Predictably, a suitable solution was not achievable in all cases leading to unreasonable estimates, divergence and/or negative values. This may indicate that an inadequate model was used for the systematic effects resulting in 'residual' biases that corrupt the performance of the VCE method. This is the suspected cause for the unreasonably large estimates of $\hat{\sigma}_H^2$ when only a constant bias is applied to remove all systematic effects; a model that is certainly inadequate for both test networks (see results in chapter 4). Another possibility is numerical instabilities caused by over-parameterization, which occurred when a complete fourth-order model was used in the Swiss test network and with both the third and fourth-order models in the southern BC/AB test network.

In any case, these first results are revealing as they hint towards a means for identifying the inappropriateness of the tested corrector surface. This latter comment is stated with prudence, as it should be independently verified from additional studies. Therefore, the same parametric models were implemented for the same data and tested using the cross-validation process described in section 3.5.2 and statistical testing methods described in sections 3.5.1 and 3.5.3. These results showed that for the Swiss test network, indeed the

failure of the fourth-order model could be deduced to numerical instabilities caused by attempting to fit a model of higher frequency attributes than is visible from the data. In fact, for the fourth-order model the condition number is 5.62×10^{16} , which provides a first inference into the numerical stability incurred by employing this parametric model. Furthermore, statistical and cross-validation tests showed that this model is not the best candidate for the region.

The southern BC/AB test network is perhaps more sensitive to the choice of the incorporated parametric model as the data is sparsely distributed, and therefore both the third and fourth-order bivariate polynomial models caused wild oscillations and negative variance components throughout the iteration process, which do not converge. The results from section 4.2 also verify the inadequacy of these models for absorbing the systematic errors and datum inconsistencies in this test network area.

6.7 Summary

The implementation of VCE techniques for the optimal combined adjustment of heterogeneous height data, namely ellipsoidal, orthometric and geoid height data was described. Specifically, the iterated MINQUE and AUE procedures were used and found to work well for testing and scaling the supplied CV matrices for the height coordinate provided by GPS post-processing software packages, calibrating geoid error models (regional and global), and assessing/evaluating the accuracy of the orthometric heights from results of national or regional adjustments.

Through four numerical case studies with real data, a number of key issues aimed at improving the combined adjustment for height related applications, were studied in detail. Firstly, it was determined that by using iterative procedures, there was no effect of the changing a-priori variance factor on the final estimated variance component values (globally best estimator). Secondly, by applying a constrained algorithm, namely, IAUE,

provisions can be made for estimating non-negative variance components that yield very similar results to the rigorous IMINQUE method. Thirdly, the effect of overly-optimistic covariance matrices obtained by neglecting correlations for observations of the same type was validated.

Last, but not least, an interesting relationship between the type of parametric model incorporated for dealing with the systematic effects and the VCE process was discovered. In particular, the divergence of the VCE solution or the computation of negative variance components provided some insight into the selected parametric model effectiveness. Results were more revealing when compared to independent evaluations of the parameter model performance using the testing procedure described in chapter 3 for identical data sets.

All of the tests were conducted using two distinct networks in terms of data characteristics, quality, distribution, and coverage area. Challenges encountered in working with this data were described throughout, providing some insight into what can be expected in practical applications. However, it is important to continue testing with other data sets. In general, it was found that optimistic error models are initially obtained, which can be re-scaled through the implementation of VCE algorithms, leading to more realistic accuracy measures. This is essential for determining the relative contribution of each of the height data types to the overall error budget. Furthermore, in these studies only three variance components ($\sigma_h^2, \sigma_H^2, \sigma_N^2$) were estimated, which is reasonable based on the information that is currently available. However, the estimation procedure can easily be expanded to include other height sources and their associated CV information, such as sea surface topography (σ_{SST}^2) and land uplift (σ_U^2). The prospect and incorporation of such information will be discussed in more detail in chapter 7.

Chapter 7

Practical Considerations for Modernizing Vertical Control

7.1 Introduction

An overview of the complete procedure developed for the optimal combination of ellipsoidal, orthometric and geoid height data is provided in Figure 7.1. As mentioned previously, the reasons for combining the different types of height data are innumerable, with the most prevalent outlined in section 2.4. Thus far, the problem has been approached from the perspective of dealing with the systematic and random error sources inherent in each of the height data types and their combination. Ultimately, the main motivation for this research is embedded in the need to introduce *modern* tools and techniques in the establishment of vertical control. In this chapter, some of the practical implementation issues involved with establishing vertical control using modern techniques are discussed. In particular, the process of GNSS-levelling is demonstrated using a test network in the province of Alberta, Canada. The second part of this chapter involves the integrated optimal network adjustment of terrestrial benchmarks and tide gauge stations typically situated along the coast, in harbours, estuaries, etc. Specifically, the incorporation of sea surface topography (SST) values and the corresponding variance-covariance information in the optimal adjustment of the heterogeneous height data is discussed. Modifications to the practical algorithms for variance component estimation are also provided that may be used for future 'calibration' of SST error models.

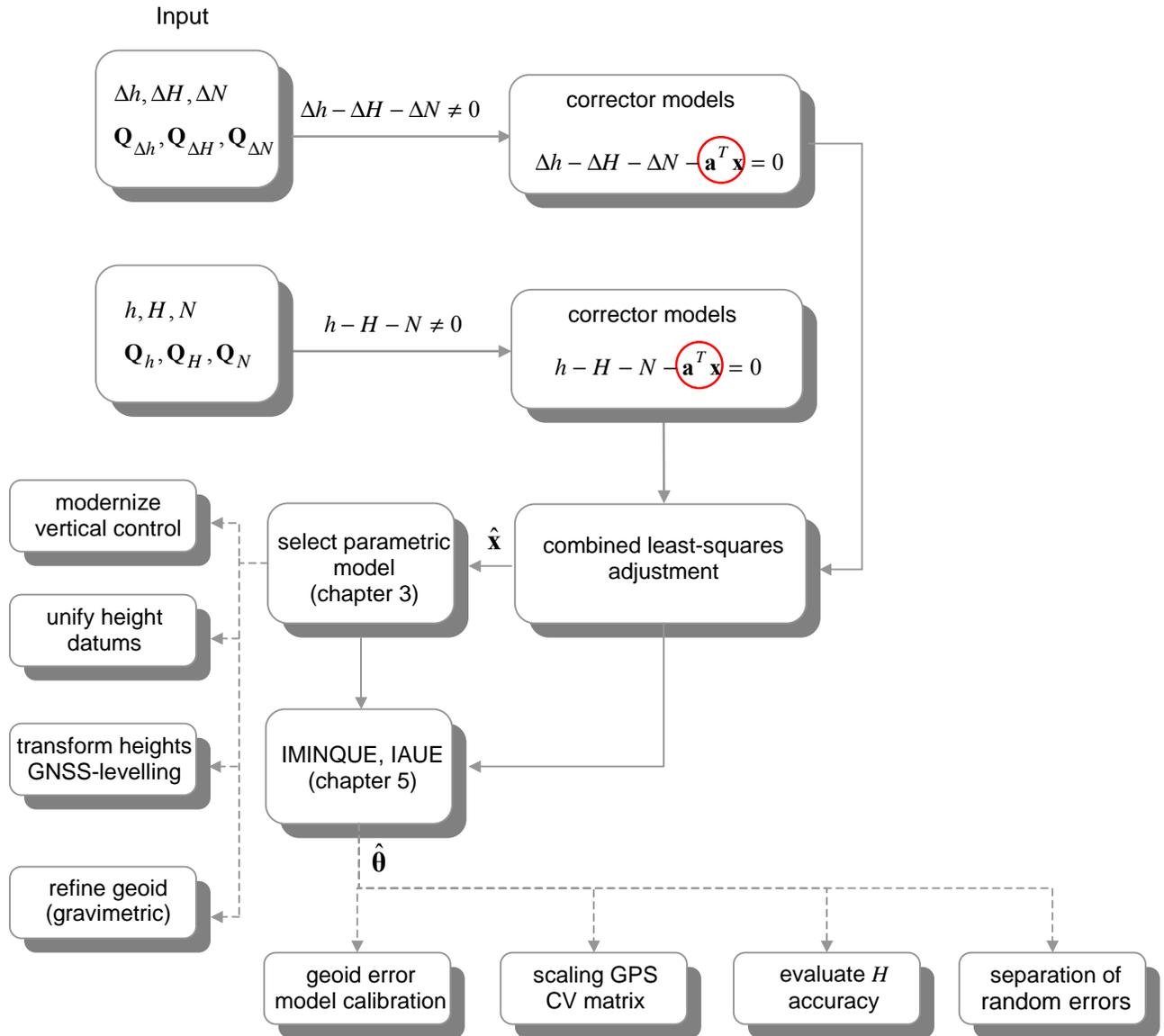


Figure 7.1: Optimal combination of ellipsoidal, orthometric and geoid height data

7.2 GNSS-levelling: What are the issues?

The main problems that may be encountered by users when trying to implement GNSS-levelling using the procedure described in Figure 7.1, in practice, include:

- obtaining approximate initial covariance matrices for the ellipsoidal, orthometric and geoid height data (i.e., $\mathbf{Q}_{\Delta h}, \mathbf{Q}_{\Delta H}, \mathbf{Q}_{\Delta N}$) for a specified region
- determining the accuracy contribution of the corrector surface model parameters, $\mathbf{C}_{\hat{\mathbf{x}}}$ (computed using Eq. 3.21)
- determining the achievable accuracy of GNSS-levelling for a new baseline

Although this study does not make use of actual height values, the height accuracy information that is used and the test network configuration are simulated to mirror realistic conditions. The selected test network consists of a subset of the GPS control benchmarks in the southwestern part of Canada (314 points in Alberta and British Columbia) covering $49^\circ \leq \varphi \leq 55^\circ$ and $-120^\circ \leq \lambda \leq -110^\circ$, which translates to an approximate network coverage area of $667 \text{ km} \times 685 \text{ km}$ as depicted in Figure 7.2 (the solid line is the political boundary for southern Alberta).

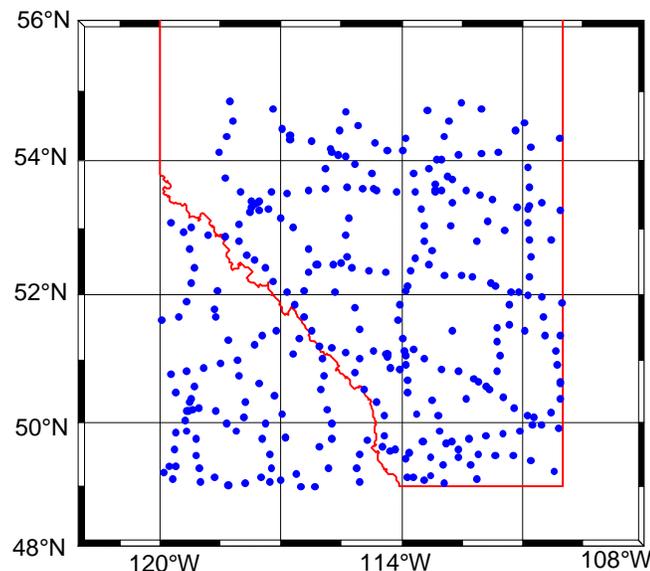


Figure 7.2: Test network area and distribution of the GPS control benchmarks

Full CV matrices describing the accuracy of the orthometric and the ellipsoidal heights at the test network points were obtained through separate simulative least-squares adjustments. These separate adjustments used the measuring accuracy of GPS and spirit-levelling as input, according to the standard formulation

$$\sigma(mm) = b \cdot \sqrt{d} \quad (7.1)$$

where σ is the standard deviation of the observed height differences in millimeters, b is the height difference accuracy based on published standards for GPS and levelling, and d is the baseline length in kilometres (Kearsley *et al.*, 1993).

7.2.1 Computing the accuracy of the levelling data

The multi-data integrated adjustment that was described in section 3.1 and used to extract the accuracy of the corrector surface parameters, requires the relative accuracy $\mathbf{C}_{\Delta H}$ of the levelling/orthometric heights in the control network (see Eq. 3.21 and consider its relative form). Such a covariance matrix was computed using a separate minimally constrained least-squares adjustment on the levelling network (see Figure 7.3), according to the equation

$$\mathbf{C}_{\Delta H} = \mathbf{A}_{net} \mathbf{C}_H \mathbf{A}_{net}^T \quad (7.2)$$

where \mathbf{A}_{net} is a design matrix (composed of -1 , 1 and 0 only) corresponding to the baseline configuration of the multi-data network adjustment (see Figure 7.9), and \mathbf{C}_H is the covariance matrix for the orthometric heights at all points in the test network. Actually, the latter is the quantity that is obtained through the separate adjustment of the levelling network, as follows:

$$\mathbf{C}_H = \left(\mathbf{A}_{lev}^T \mathbf{P}_{lev} \mathbf{A}_{lev} \right)^{-1} \quad (7.3)$$

where \mathbf{A}_{lev} is a design matrix (composed also of -1 , 1 and 0) corresponding to the baseline configuration of the levelling network adjustment (see Figure 7.3).

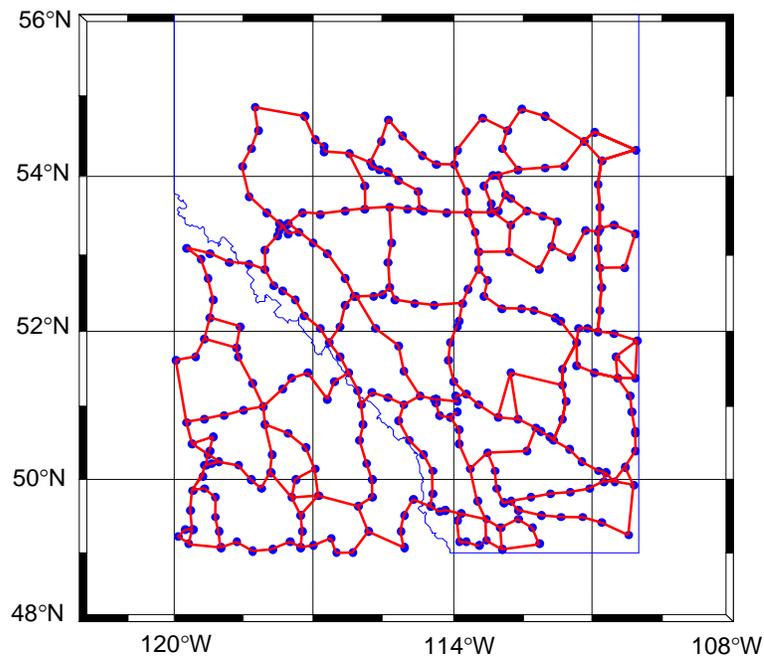


Figure 7.3: Baseline configuration for the levelling network

\mathbf{P}_{lev} is a diagonal weight matrix that takes into account the measuring accuracy of each levelling baseline in the test network, i.e.

$$\mathbf{P}_{lev} = \begin{pmatrix} 1/\sigma_{\Delta H_1}^2 & 0 & \dots & \dots & 0 \\ 0 & \ddots & 0 & 0 & \vdots \\ \vdots & 0 & 1/\sigma_{\Delta H_i}^2 & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots \end{pmatrix} \quad (7.4)$$

Three different orders of accuracy were used for assigning the a-priori values $\sigma_{\Delta H}$ for each levelling baseline in the weight matrix \mathbf{P}_{lev} , namely $0.7mm\sqrt{d(km)}$, $1.3mm\sqrt{d(km)}$ and $2mm\sqrt{d(km)}$, referring to first, second and third order, respectively (see Figure 7.4). National standards for the accuracy of vertical control vary depending on the country (Kearsley *et al.*, 1993; Ollikainen, 1997; van Onselen, 1997). In our case, the U.S. standards were implemented for both levelling and GPS, as they were readily available (National Geodetic Survey, 1994). For the case of levelling, larger baselines ($d > 80km$) usually constitute part of a national levelling campaign and adhere to first order levelling standards, followed by denser regional levelling campaigns ($30km < d \leq 80km$) of second order, and finally local levelling lines ($d \leq 30km$) which are of third order accuracy.

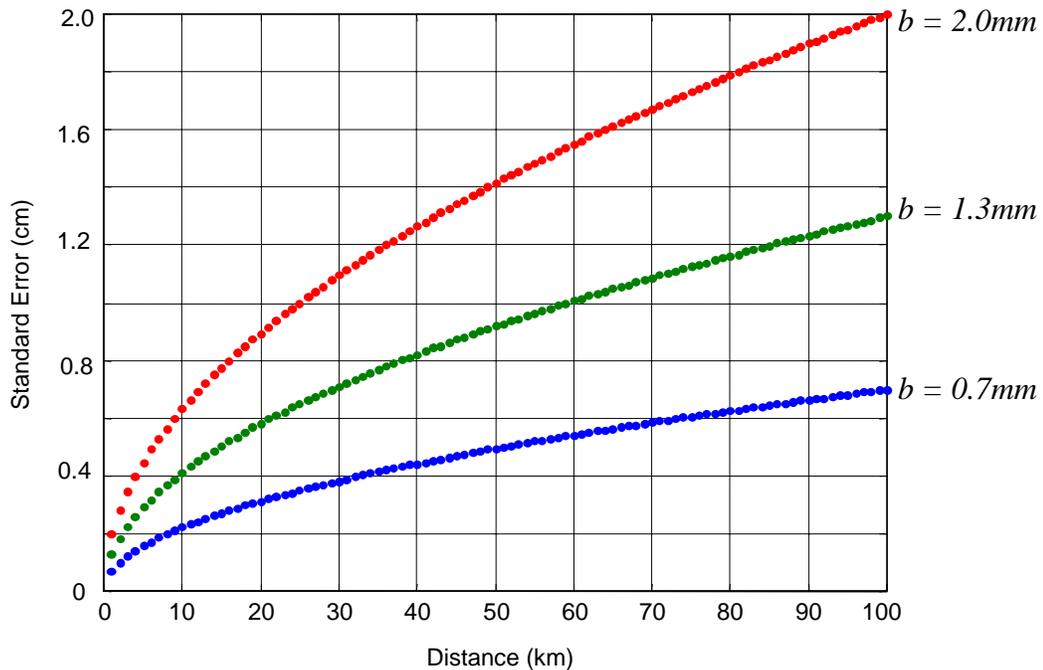


Figure 7.4: Relative measurement accuracy for the levelling baselines

7.2.2 Computing the accuracy of the GPS data

A similar procedure to the one described in the previous section was followed in order to obtain the relative accuracy $\mathbf{C}_{\Delta h}$ of the ellipsoidal/geometrical heights in the control test network, which is also required as input into Eq. (3.21) for the multi-data adjustment. The GPS network is shown in Figure 7.5 and the final equation is:

$$\mathbf{C}_{\Delta h} = \mathbf{A}_{net} \mathbf{C}_h \mathbf{A}_{net}^T \quad (7.5)$$

where \mathbf{A}_{net} is the same design matrix that was used in Eq. (7.2) and it corresponds to the baseline configuration of the multi-data network adjustment (see Figure 7.9), and \mathbf{C}_h is the covariance matrix for the ellipsoidal heights at all points of the test network. The latter was obtained through a separate simulative (minimally constrained) least-squares adjustment for the GPS network, according to the formula

$$\mathbf{C}_h = \left(\mathbf{A}_{GPS}^T \mathbf{P}_{GPS} \mathbf{A}_{GPS} \right)^{-1} \quad (7.6)$$

where \mathbf{A}_{GPS} is a design matrix (composed only of -1 , 1 and 0) corresponding to the baseline configuration of the GPS network adjustment (see Figure 7.5). It is evident by comparing the two network configurations, shown in Figures 7.3 and 7.5, that the levelling network has a weaker geometry than the GPS network due to the stringent line-of-sight restrictions of spirit-levelling (Ollikainen, 1997). This results in the GPS network having more observations between the network points and a stronger geometry overall. Finally, \mathbf{P}_{GPS} is a diagonal weight matrix that takes into account the measuring accuracy of the vertical component for each GPS baseline in the test network, i.e.

$$\mathbf{P}_{GPS} = \begin{pmatrix} 1/\sigma_{\Delta h_1}^2 & 0 & \dots & \dots & 0 \\ 0 & \ddots & 0 & 0 & \vdots \\ \vdots & 0 & 1/\sigma_{\Delta h_i}^2 & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots \end{pmatrix} \quad (7.7)$$

The accuracy for observed GPS height differences also degrades as the baseline length increases, mainly due to the spatial decorrelation of atmospheric errors that affect the GPS observables (see, e.g., discussion in section 2.3 and Fotopoulos, 2000). In this analysis, ten different orders for the relative ellipsoidal height accuracy are used as defined by the U.S. standards (National Geodetic Survey, 1994). The a-priori values $\sigma_{\Delta h}$ in the weight matrix \mathbf{P}_{GPS} were assigned based on the length of each GPS baseline in the test network. For every 10 km increase in baseline length, the value of $\sigma_{\Delta h} = b\sqrt{d}$ changed according to $b = (0.5, 0.7, 1, 1.3, 2, 3, 6, 15, 30, 60)$ mm (e.g., for $d = 55$ km, $b = 3$ mm), as shown in Figure 7.6.

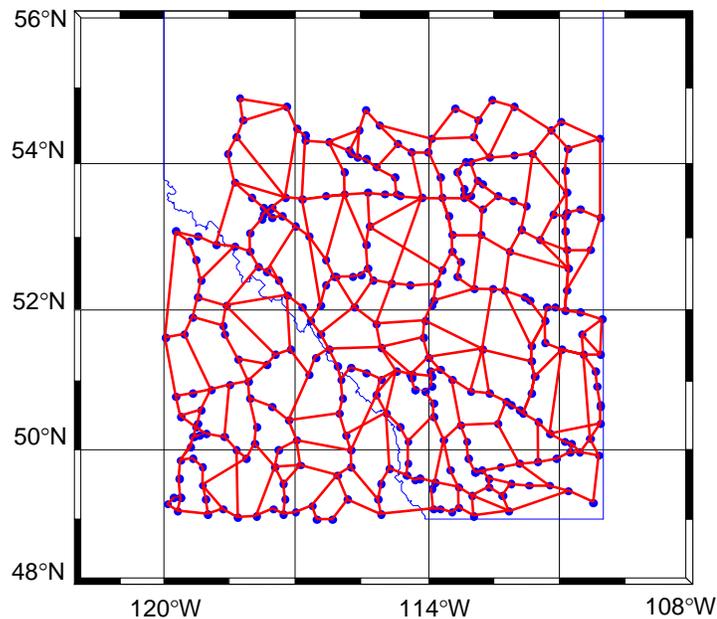


Figure 7.5: Baseline configuration for the GPS network

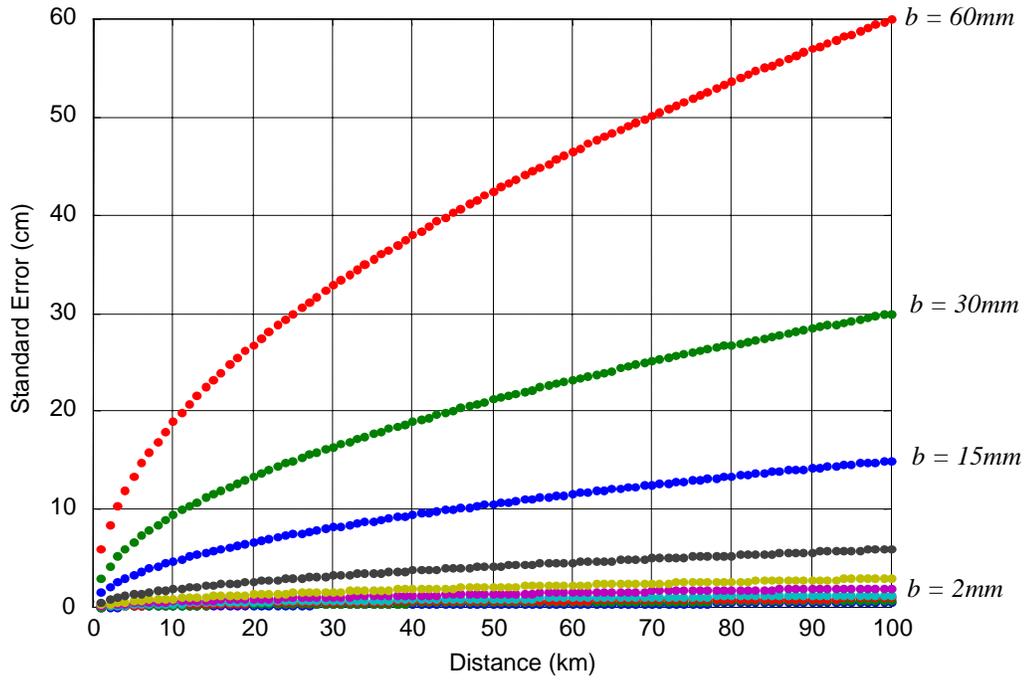


Figure 7.6: Relative measurement accuracy for the height component of GPS baselines

7.2.3 Computing the accuracy of the geoid height differences

The input accuracy for the geoid undulation differences in the adjustment of the multi-data test network was approximated by a diagonal CV matrix $\mathbf{C}_{\Delta N}$, as follows:

$$\mathbf{C}_{\Delta N} = \begin{pmatrix} \sigma_{\Delta N_1}^2 & 0 & \dots & \dots & 0 \\ 0 & \ddots & 0 & 0 & \vdots \\ \vdots & 0 & \sigma_{\Delta N_i}^2 & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\Delta N_m}^2 \end{pmatrix} \quad (7.8)$$

where the subscript m denotes the number of baselines in the multi-data test network configuration (Figure 7.9). Its diagonal values $\sigma_{\Delta N}^2$ were chosen based on the performance evaluation of the GSD95 Canadian geoid model in Alberta and British Columbia (Véronneau, 1997). Specifically, baselines located in the flat or rolling hills area of the network were assigned a relative geoid accuracy of $1\text{ cm} + 4\text{ mm}/\text{km}$, and those in more mountainous regions were assigned a lower accuracy of $1\text{ cm} + 10\text{ mm}/\text{km}$, as shown in Figure 7.7. The mountainous regions within the network coverage area were identified by terrain changes shown on a regional digital elevation map for western Canada (Alberta and British Columbia). The digital elevation data used for this case was part of the Canadian Digital Elevation Data (CDED) products (3"×3" resolution), which are distributed by the Centre of Topographic Information of Natural Resources Canada (www.geod.nrcan.gc.ca). Therefore, the relative geoid accuracy depended both on the spatial separation of the points as well as on the geographical location of the baseline.

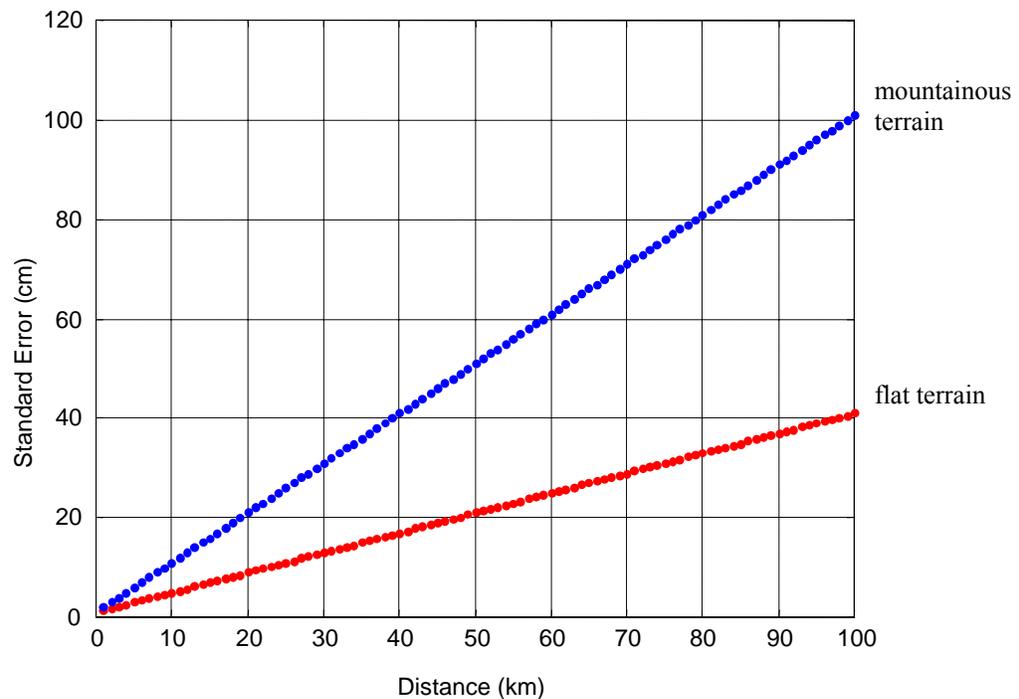


Figure 7.7: Relative geoidal height accuracy

It should be noted that a full covariance matrix $\mathbf{C}_{\Delta N}$ for the internal accuracy of the gravimetric geoid model could be theoretically computed by properly propagating the errors in the local gravity and height data and the geopotential harmonic coefficients into the final geoid solution, as described in section 6.3.2 (for formulas, see Sideris and Schwarz, 1987; Li and Sideris, 1994). The process for obtaining a full CV matrix for the internal geoid accuracy is considerably complex (mainly due to the lack of adequate information on the accuracy of the input data in the geoid solution). An indication of the significance of this assumption is evident from the results in section 6.4. However, for the tests conducted herein, a diagonal CV matrix for the relative accuracy of the geoidal undulation signal is considered to be sufficient. Furthermore, all the software programs that were developed for this study were designed to facilitate the inclusion of a fully-populated error covariance matrix $\mathbf{C}_{\Delta N}$, should it be available.

7.2.4 Achievable accuracy of GNSS-levelling for a new baseline

Given the theoretical relationship among the three types of height data and the incorporation of an appropriate corrector surface model, the orthometric height difference for a new baseline (k, l) (i.e., not belonging in the original test network) as obtained from relative GPS/geoid levelling is given by

$$\Delta H_{kl} = \Delta h_{kl} - \Delta N_{kl} - (\mathbf{a}_l^T - \mathbf{a}_k^T) \hat{\mathbf{x}} \quad (7.9)$$

where the estimated parameters $\hat{\mathbf{x}}$ are obtained from the adjustment of the multi-data observations in the control network (see chapter 3). In this study, we are interested in the achievable *accuracy* of the orthometric height difference ΔH_{kl} . By simply applying variance-covariance propagation to Eq. (7.9), the accuracy of GPS/geoid levelling can be obtained according to the following formula:

$$\sigma_{\Delta H_{kl}}^2 = \sigma_{\Delta h_{kl}}^2 + \sigma_{\Delta N_{kl}}^2 + (\mathbf{a}_l^T - \mathbf{a}_k^T) \mathbf{C}_{\hat{\mathbf{x}}} (\mathbf{a}_l - \mathbf{a}_k) \quad (7.10)$$

where $\sigma_{\Delta h_{kl}}^2$ and $\sigma_{\Delta N_{kl}}^2$ represent the relative accuracy of the ellipsoidal and geoidal heights of the new baseline, and $\mathbf{C}_{\hat{\mathbf{x}}}$ is the CV matrix of the estimated parameters of the corrector surface model, as shown in Eq. (3.21). A description of the procedure followed to obtain input accuracy levels used for the numerical tests was given in the previous sections.

In this section, investigations on how accurately orthometric height differences can be determined from relative GNSS-levelling are conducted. More specifically, numerical investigations analyze the impact of the GNSS and geoid data accuracy, and the accuracy of the corrector surface model parameters, on the determination of orthometric height differences. A test network of spirit levelled GPS control points, situated in the western part of Canada, is used as a basis for all of the numerical investigations. In this network, various simulative adjustments of GPS/levelling/geoid data are performed in order to determine the CV matrix for the estimated parameters of the corrector surface model. Then, an evaluation of the achievable accuracy of GPS-derived orthometric heights for new baselines (not included in the original test network) is made. A number of different scenarios are studied by varying the:

- a) configuration of the multi-data control network (number of baselines, density of observations),
- b) location of the new baseline with respect to the test network,
- c) length of the new baseline, and
- d) type of the corrector surface model.

A summary of the various test network configurations that were used is provided in Table 7.1. The three network configurations labeled *dense*, *mixed* and *sparse* (see Figure 7.9) refer to the multi-data (GPS/levelling/geoid) integrated adjustment used to compute the a-

posteriori accuracy $C_{\hat{x}}$ of the parameters for the corrector surface model. The different configurations for the multi-data test network were selected in order to assess the effect of varying control network geometry on the accuracy of relative GPS/geoid levelling for a new baseline that was not included in the original control network. The other two configurations for the test network that appear in Table 7.1 (i.e. levelling, GPS) correspond to the individual auxiliary adjustments that were performed in order to assess the relative accuracy of the levelling and the GPS heights (see sections 7.2.1 and 7.2.2).

Table 7.1: Statistics of various test network configurations
(d denotes baseline length in km)

Network	# obs	$\mu(d)$	$\sigma(d)$	max (d)
Levelling	390	27.9	13.5	95.2
GPS	472	36.9	25.4	164.3
Dense	514	41.1	33.3	438.2
Mixed	425	32.3	20.3	128.1
Sparse	408	30.6	18.9	128.1

Figure 7.8 provides an illustrative view of the location of the new baseline(s) with respect to the control test network. For our numerical tests, the length and the location of the new baseline, as well as the test network configuration, will all be varied. In Table 7.2, the results of the achievable accuracy for relative GPS/geoid levelling, according to the procedure described in the previous sections (see Eq. 7.10), are shown for baselines of varying lengths, from a minimum of 10 km to a maximum of 100 km.

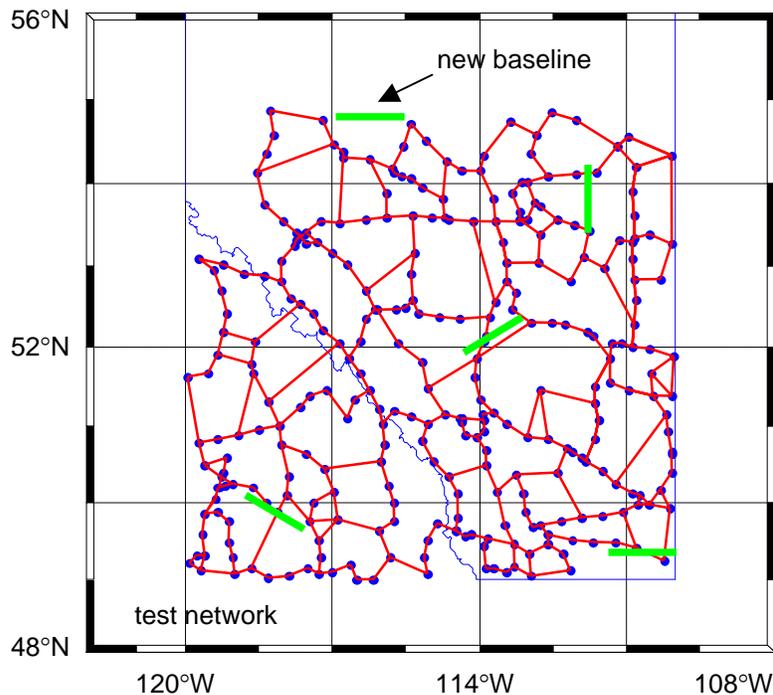


Figure 7.8: Example of the locations of newly established baselines within the test network area

These results correspond to the averaged accuracy levels for baselines tested within and on the edge of the borders of the test networks (dense, mixed and sparse). The type of the parametric model used was also varied according to the relative forms of Eqs. (3.31), (3.33) and a seven-parameter third-order polynomial of the form in Eq. (3.28). However, by examining the values in Table 7.2 it becomes evident that the input accuracy of the geoidal height differences ΔN_{kl} and the GPS height differences Δh_{kl} overshadow any contribution from the corrector surface model, surpassing the contribution of this third term in Eq. (7.10) by several orders of magnitude. This is the reason for not distinguishing in Table 7.2 between results from the different network configurations and parametric models used, as the results did not vary significantly. Hence, the final numerical values given in the fourth column of Table 7.2 correspond to the mean accuracy level over the three corrector surface models and the three multi-data network configurations.

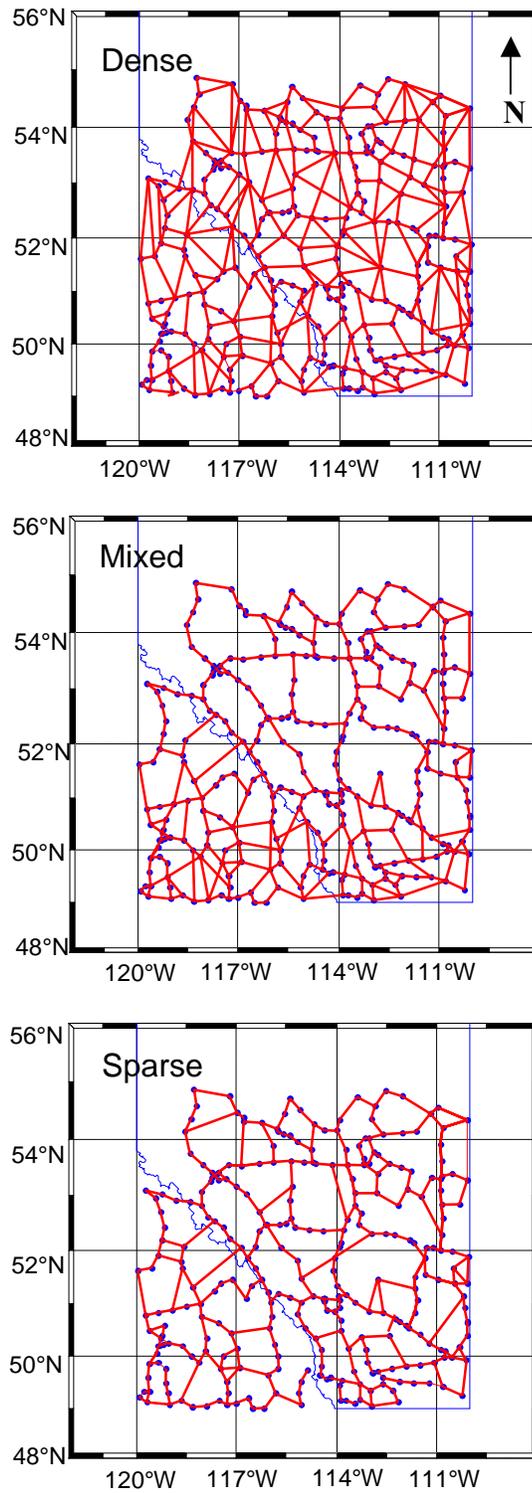


Figure 7.9: Test network configurations for the multi-data adjustment

Table 7.2: Standard error of relative height components (in cm) as a function of baseline length

d (km)	$\sigma_{\Delta h_{kl}}$	$\sigma_{\Delta N_{kl}}$	$\sigma_{\Delta H_{kl}}^*$
10	4.7	5.0	6.9
20	6.7	9.0	11.2
30	8.2	13.0	15.4
40	9.5	17.0	19.5
50	10.6	21.0	23.5
60	11.6	25.0	27.6
70	12.6	29.0	31.6
80	13.4	33.0	35.6
90	14.2	37.0	39.6
100	15.0	41.0	43.7

* accuracy of GPS/geoid levelling, computed from Eq. (7.10)

Since the internal relative accuracy of the GPS and geoid heights (second and third column values in Table 7.2) is independent of the type of the parametric corrector surface used and the test network configuration, it is more illuminating to base the discussion of the results on the third term in Eq. (7.10). This *third term* contains the contribution of the adjusted parameters of a pre-specified model (corrector surface) to the final GPS/geoid levelling accuracy, and is plotted for a number of different cases in Figure 7.10. In this figure, each error profile/line is labeled based on an abbreviation, which states the test network configuration ('d', 's', 'm' for dense, sparse and mixed, respectively), followed by a 3 or 7 for the relative model types corresponding to Eqs. (3.31) and (3.28), respectively. The results for the 4-parameter model (relative form of Eq. 3.33) were also computed, however they were very similar to the 3-parameter model values and therefore omitted to avoid cluttering the graph.

The results in Figure 7.10 depict the values obtained when using the dense, mixed and sparse test network configurations for the multi-data adjustment. The figure also shows

that, in all cases, the 7-parameter model deteriorates the final ΔH_{kl} accuracy by approximately 30%, compared to the 3-parameter model. This may be due to the fact that over-parameterization of the systematic effects and datum inconsistencies in the corrector surface model can lead to large correlation values in the $C_{\hat{x}}$ matrix. However, it should be noted that all these differences are at the sub-cm level. At this point it can be said that the errors in the estimated parameters of the corrector surface model are not a major contributing factor to the overall accuracy of GPS/geoid levelling. The tests actually showed that relative GPS/geoid levelling can result in sub-decimetres accuracy for a 10 km baseline, regardless of the parametric model chosen. The main factor that contributes to the achievable accuracy is the baseline length, as it comes into play through the a-priori observation accuracy of the GPS and geoid height data. When the baseline length is increased up to 100 km, the relative GPS/geoid levelling accuracy degrades to over 40 cm. By changing the relative accuracy of the GPS and geoid heights at the newly established baseline, these results will change accordingly.

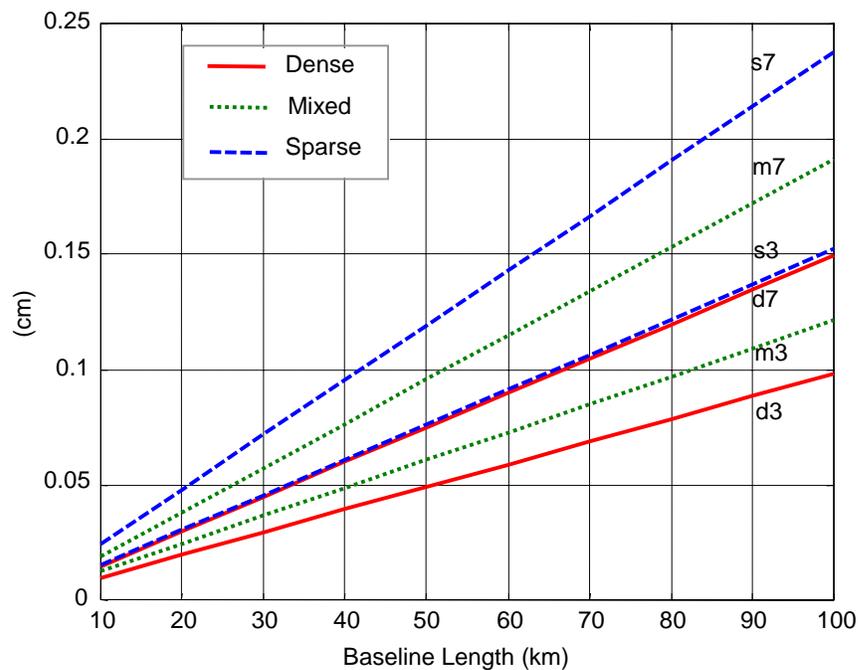


Figure 7.10: Accuracy contribution of the corrector surface model to GPS-levelling

It should be noted that, although the *accuracy* of the parametric model does not significantly affect the final results, its incorporation in the height determination procedure is absolutely essential for obtaining high accuracy in GPS-levelling applications, since the magnitude of the datum inconsistencies and systematic distortions in the original control height data can reach up to several metres (see results in chapter 4).

7.2.5 Tests in northern Canada

The general idea described herein is to use the accuracy information of the different height data types that is available from existing vertical control in densely surveyed areas (such as southwestern Canada) and propagate that information for determining the accuracy of the orthometric height difference for a newly established baseline in an area with no (or very limited) vertical control, such as northern Canada. The test network used is shown in the boxed area of Figure 7.11 and contains 252 vertical control points (a subset of the same network used in the previous studies, Figure 7.2). The procedure described in the previous sections was used to determine the achievable $\sigma_{\Delta H}$ for a new baseline as it moves further north - *away from* - the test network. The covariance information for the levelling and GPS data was approximated in the same manner as described in sections 7.2.1 and 7.2.2, respectively.

A diagonal covariance matrix for the geoid heights in the region of interest were obtained from a $1^\circ \times 1^\circ$ world-wide grid of the commission errors of EGM96, as computed from the accuracy of the spherical harmonic coefficients up to degree and order 70 and was used to interpolate σ_N of each benchmark in the control test network. For this particular test network area, the $1^\circ \times 1^\circ$ grid resolution of the geoidal undulation errors was deemed sufficient as the gravity coverage employed for this area in the computation of EGM96 was reasonably dense and homogeneously spaced (Lemoine *et al.*, 1998). By using this global geoid error model and bilinearly interpolating the grid of commission errors for the points in the test network, a diagonal error CV matrix of the absolute geoid height

errors \mathbf{C}_N was obtained. The computation of the relative geoid height accuracy $\mathbf{C}_{\Delta N}$, required as input into Eq. (3.21), was obtained by propagating the absolute height accuracy as follows:

$$\mathbf{C}_{\Delta N} = \mathbf{A}_{net} \mathbf{C}_N \mathbf{A}_{net}^T \quad (7.11)$$

where \mathbf{A}_{net} is the same design matrix as mentioned above, corresponding to the baseline configuration of the multi-data network adjustment. This results in a fully-populated form of $\mathbf{C}_{\Delta N}$.

From the results in section 7.2.4, it is evident that the major error contributor of the three height types to the final result was $\mathbf{C}_{\Delta N}$. By varying the type of corrector surface model used in the adjustment, the achievable accuracy also changed. For instance, for a 50 km baseline located in the north ($\varphi = 60^\circ N$, $\lambda = 116^\circ W$) with input accuracies of $0.6 \cdot \sqrt{d(km)}$ and 23 cm for GPS and geoid height differences, respectively, the resultant relative orthometric height accuracy was 48.3 cm with the 3-parameter model (relative form of Eq. 3.31) and 52.4 cm with the 6-parameter model (relative form of Eq. 3.34, with the sixth term removed). Similarly, for a 100 km baseline the achievable accuracy degraded by ~ 13 cm with the 6-parameter model as compared to the 3-parameter model. It is evident from these results that the more parameters in the model, the more amplified its error contribution is (see Table 7.3). This is interesting, as we are not evaluating the performance/fitting of different parametric models (no actual height data is used), rather the focus is placed on how the random errors flow through the model and affect the accuracy of the final value.

Table 7.3 provides a summary of some of the numerical results. The first column refers to the approximate latitude of a 40 km baseline ($\lambda \sim 112^\circ W$). The latitude varies as the newly established baseline is moved northward with respect to the test network area (see

Figure 7.11). The column labeled 3^{rd} term refers to the accuracy contribution of the corrector surface model (3^{rd} term in Eq. 7.10). There are three main groups of results, where *Full CVs* refers to fully-populated covariance matrices for the three height types, *Diagonal CVs* refers to diagonal covariance matrices for the three height types, and *Full & Diagonal* refers to fully-populated CV matrices for levelling and GPS heights and a diagonal covariance matrix for the geoid heights. All results are based on fixed input accuracies of $\sigma_{\Delta N} = 15$ cm and $\sigma_{\Delta h} = 0.15 \cdot \sqrt{d(km)}$ cm for Eq. (7.10).

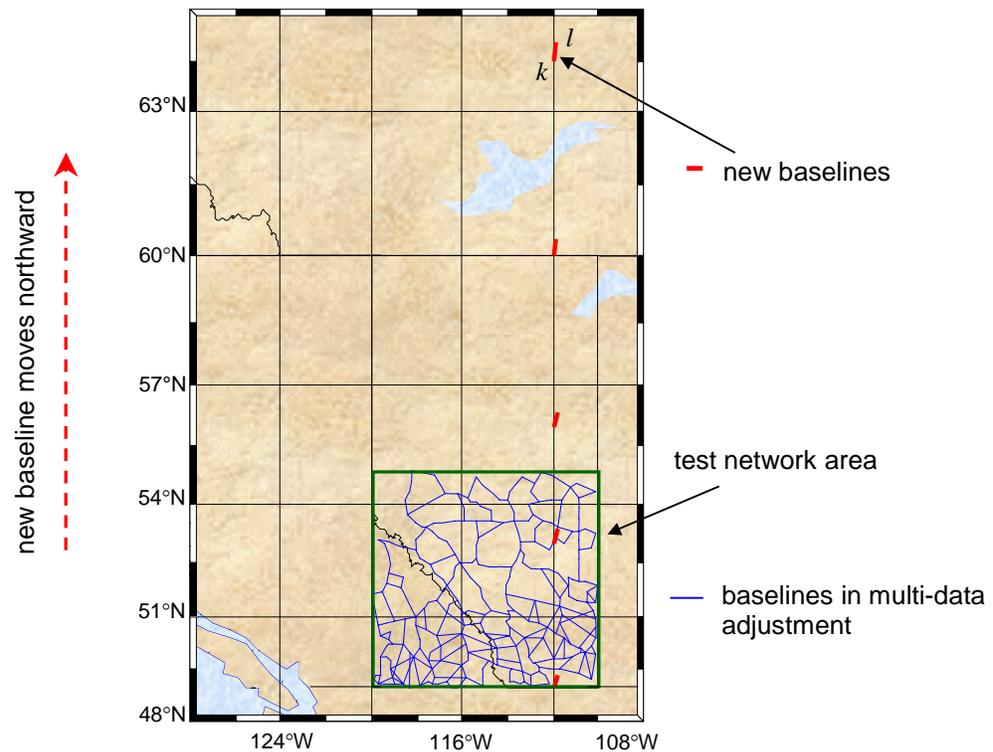


Figure 7.11: Test network coverage area and locations of newly established baselines in the north

The results show that the achievable $\sigma_{\Delta H}$ for the new baseline was worse as it moved farther north from the control test network. This was mainly due to the increased error

contribution of the corrector surface model parameters. Perhaps the most interesting result was the difference between using fully-populated and (approximate) diagonal CV matrices for Δh , ΔH , and ΔN as input into Eq. (3.21). In studies where the newly established baseline was located *within* the test network coverage area, it was found that there was no difference between using fully-populated versus (more approximate) diagonal CV matrices. However, in cases where the baseline is moved farther north (away from) the test network area, differences up to several tens of centimetres resulted. This result is quite significant as it indicates that approximate versions of the CV matrices should *not* be used, for new baselines situated *away from* the test network, in order to take advantage of the highest accuracy that GPS-based levelling provides.

Table 7.3: Results of baselines moving northward from the test network area (units: cm)

<i>3 - parameter corrector surface model</i>						
φ	Full CVs		Diagonal CVs		Full and Diagonal CVs	
	$\sigma_{\Delta H}$	3 rd term	$\sigma_{\Delta H}$	3 rd term	$\sigma_{\Delta H}$	3 rd term
49°	17.8	0.1	18.0	2.7	18.0	2.7
53°	17.8	0.6	18.0	3.6	18.1	3.6
56°	17.8	1.1	18.8	6.0	18.8	6.0
60°	17.8	1.8	20.2	9.6	20.2	9.6
64°	17.9	2.5	22.2	13.3	22.2	13.3
<i>6 - parameter corrector surface model</i>						
49°	17.8	1.4	18.3	4.6	18.3	4.6
53°	17.8	0.8	18.3	4.5	18.3	4.5
56°	18.5	5.1	24.4	16.8	24.4	16.8
60°	24.2	16.4	54.5	51.6	54.6	51.6
64°	36.9	32.4	102.2	100.6	102.3	100.8

A flowchart outlining the major steps of the software program that was designed to compute the orthometric height difference accuracy via GNSS-levelling is provided in Figure 7.12.

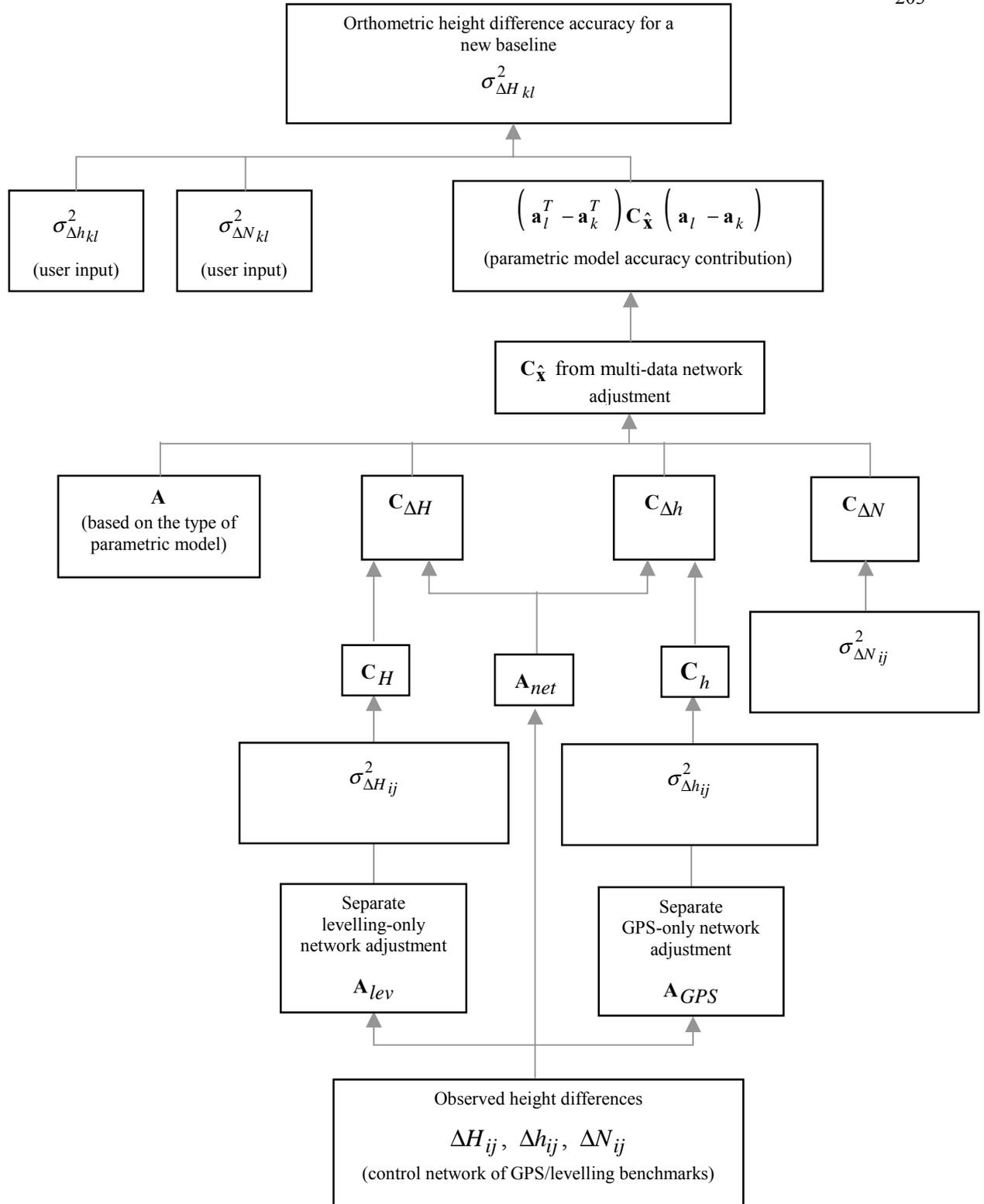


Figure 7.12: Flowchart outlining main components of software program designed to compute the accuracy of orthometric height differences via GNSS-levelling

7.3 Combining heterogeneous heights at sea

One of the major sources of distortion in vertical control networks is caused by neglecting sea surface topography at tide gauge stations. Often, the orthometric height is fixed to zero at these stations, without applying proper corrections for the deviation of the mean sea surface from the equipotential surface represented by the geoid. Over the past couple of decades, numerous advances have been made in modelling the low-to-medium frequency components of the sea surface, particularly through satellite altimetry measurements, which provide accurate and global uniform coverage (Andersen and Knudsen, 2000). The principle satellite altimetric measurements and the relationship between different height reference surfaces is illustrated in Figure 7.13. Over the open seas, altimetric-derived stationary SST (the equivalent of GNSS-levelling at sea) is theoretically given by the following relationship (Gruber and Steigenberger, 2000):

$$SST = h_{alt} - N_m \quad (7.12)$$

where the ellipsoidal heights with respect to the geocentric reference ellipsoid, h_{alt} , are derived from satellite altimetry measurements along the satellite tracks and the geoid heights, N_m , are interpolated at track points from a marine geoid (or global geopotential model). In the past, the utility of this approach has been hindered by uncertainties in the global gravity field models and errors in the altimetric measurements (radial orbit error, atmospheric effects, tides, electromagnetic bias and measurement noise). Most of the errors affecting the altimetric ranges can be modelled and corrected leaving only residual effects at the cm-level (Cheney *et al.*, 1994). Thus, the dominant error source affecting the SST computed via Eq. (7.12) is due to the global gravity field model, which is expected to greatly improve, as discussed below.

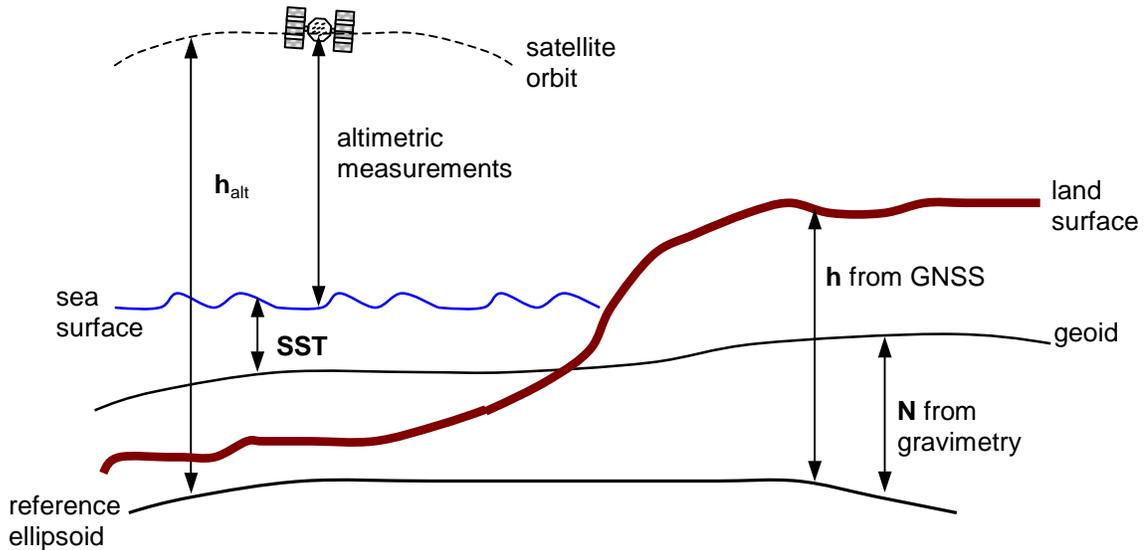


Figure 7.13: Satellite altimetry measurements and the relationship between various height reference surfaces

The proposed scheme involves the combined adjustment of terrestrial GPS-on-benchmark data and GPS-on-tide gauge data typically located in coastal areas, harbours, estuaries, and/or river mouths. This requires that accurate SST values are approximated at tide gauge locations. Although the performance of satellite altimetry is very good over open seas (i.e., 2 to 4 cm for TOPEX/POSEIDON), the measurement accuracy significantly deteriorates along the coast, over the shelf (which varies from tens to hundreds of kilometres off the coasts), in shallow depths, and at fresh water inflows (Hipkin, 2000). Unfortunately it is along this ocean/land boundary where the altimetric and global gravity field information is most critical for vertical positioning as it is where the tide gauges are situated.

Improvements in global SST models (including in coastal areas) are expected with the increase in accuracy of the satellite-only global geopotential models from the new and upcoming dedicated LEO satellite gravity missions. Already, with preliminary results from the CHAMP and more recently the GRACE data, dramatic improvements in the

long wavelength features of the SST models have been observed (Gruber and Steigenberger, 2000). The meso-scale features are also expected to be enhanced with the availability of the GOCE mission data. Furthermore, the accuracy of oceanographic methods for SST will also improve from the assimilation of the improved gravity field solutions into the estimation scheme. Therefore, it is quite certain that SST models will rely heavily on the altimetric method in the future. Along the coasts, it is most likely that this information will be combined with oceanographic methods for determining SST, such as steric levelling and global ocean circulation models (Hipkin, 2000).

In view of the significant improvements in sea surface topography determination made in the past decade and the expected improvement in global gravity field models in the near future, it is appropriate to consider the incorporation of SST into establishing vertical control. The purpose of this section is to develop a consistent procedure for incorporating the mean SST values into the combined height network adjustment problem described throughout this thesis. VCE can play a vital role in improving/calibrating SST covariance information (from altimetric, oceanographic or assimilated schemes) at tide gauge stations and surrounding areas and is also described herein.

To facilitate the discussion, a typical vertical control network consisting of terrestrial inland stations (GPS-on-benchmarks) and stations on the coasts at tide gauges (GPS-on-tide gauges) is depicted in Figure 7.14. In this example, two tide gauges for the island of Tasmania form the perimeter of a vertical control network consisting of 48 stations inland. The two main issues that arise for the proper incorporation of SST information into the optimal heterogeneous height network adjustment algorithm include:

- modelling the systematic errors and datum discrepancies among the height data types (h, SST, H, N) using a corrector surface, and
- separation of random errors for estimating variance components for each height type.

It should be noted that in the algorithms provided below, the ellipsoidal height information at the tide gauges is typically determined through GPS measurements and the SST values may be interpolated from oceanographic, altimetric or a combined model at the tide gauge stations.

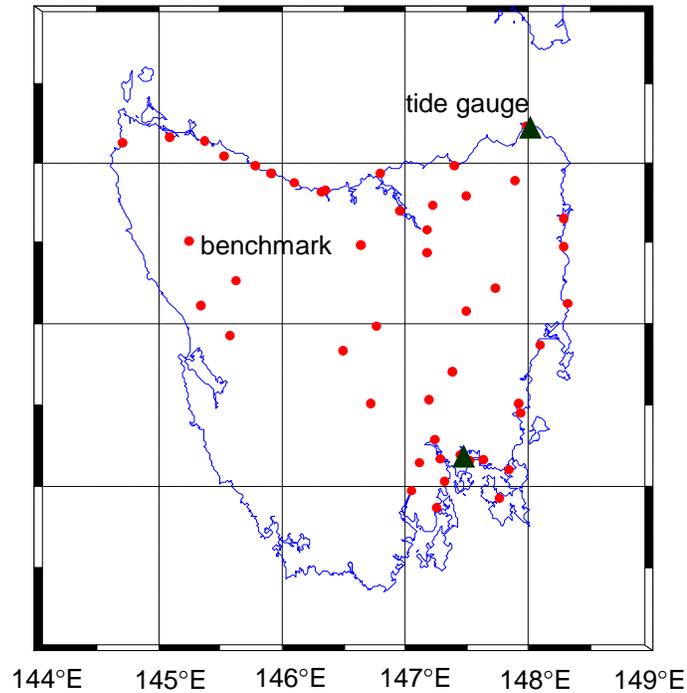


Figure 7.14: Example of a mixed network of terrestrial benchmarks and tide gauge stations

The new observation equation model for this 'mixed' adjustment is described as follows:

$$\ell_i = h_i - \tilde{H}_i - N_i = \mathbf{a}_i^T \mathbf{x} + v_i \quad (7.13)$$

where \tilde{H}_i for terrestrial points refers to the orthometric height $\tilde{H}_i = H_i$, as described thus far. At the tide gauge stations the observation equation model is modified to

accommodate SST values such that $\tilde{H}_i = SST_i$. The new vector of observations consists of two types of observations as expressed by

$$\ell = \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix} = \begin{bmatrix} h_i - H_i - N_i \\ h_i - SST_i - N_i \end{bmatrix} \quad (7.14)$$

where the terrestrial points are contained in the $m_1 \times 1$ subvector ℓ_1 and the observations at the tide gauges stations comprise the $m_2 \times 1$ subvector ℓ_2 . In virtually all practical situations, there will be fewer tide gauge stations than benchmarks on land in the network, $m_2 < m_1$ (as is the case in Figure 7.14).

The vector of random errors originally described by Eq. (3.13), is modified to include the new observation type as follows:

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_h^T & \mathbf{v}_{\tilde{H}}^T & \mathbf{v}_N^T \end{bmatrix}^T \quad (7.15)$$

where $\mathbf{v}_{(\cdot)}$ is an $(m_1 + m_2) \times 1$ vector of random errors for each of the h, H (SST), N data types. The corresponding covariance matrix is described by the general formulation in Eq. (3.11c).

The solution for the unknown parametric model coefficients follows in the same manner as described in section 3.1, with the observation vector replaced with Eq. (7.14) and is given by

$$\hat{\mathbf{x}} = \left[\mathbf{A}^T (\mathbf{C}_h + \mathbf{C}_{\tilde{H}} + \mathbf{C}_N)^{-1} \mathbf{A} \right]^{-1} \mathbf{A}^T (\mathbf{C}_h + \mathbf{C}_{\tilde{H}} + \mathbf{C}_N)^{-1} \ell \quad (7.16)$$

It should be noted from the above equation that the corrector surface parameters are computed and applied to all stations (terrestrial and tide gauge) and therefore remains common throughout.

Finally, the combined adjusted residuals from the adjustment are given by

$$\mathbf{B}\hat{\mathbf{v}} = \hat{\mathbf{v}}_h - \hat{\mathbf{v}}_{\tilde{H}} - \hat{\mathbf{v}}_N \quad (7.17)$$

where we can explicitly solve for the **separate** adjusted residuals, according to height data type as was done in chapter 3. The values corresponding to the ellipsoidal and the geoid heights are equivalent to Eq. (3.19a) and Eq. (3.19c), respectively. The separate adjusted residuals corresponding to \tilde{H} are computed as follows:

$$\hat{\mathbf{v}}_{\tilde{H}} = \begin{bmatrix} \hat{\mathbf{v}}_H \\ \hat{\mathbf{v}}_{SST} \end{bmatrix} = - \begin{bmatrix} \mathbf{C}_H & 0 \\ 0 & \mathbf{C}_{SST} \end{bmatrix} \left\{ \begin{bmatrix} \mathbf{C}_H & 0 \\ 0 & \mathbf{C}_{SST} \end{bmatrix} + \mathbf{C}_h + \mathbf{C}_N \right\}^{-1} \mathbf{M} \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix} \quad (7.18)$$

where $\mathbf{C}_{SST} = E \left\{ \mathbf{v}_{SST} \mathbf{v}_{SST}^T \right\}$ is the covariance matrix for the SST data at the tide gauge stations in the network.

The new block-diagonal covariance matrix for the disjunctive observations can be formulated as follows:

$$\mathbf{C}_\ell = \begin{bmatrix} \mathbf{C}_{h_1} + \mathbf{C}_{H_1} + \mathbf{C}_{N_1} & 0 \\ 0 & \mathbf{C}_{h_2} + \mathbf{C}_{SST_2} + \mathbf{C}_{N_2} \end{bmatrix} \quad (7.19)$$

where the subscript, '1', refers to points on land and '2' refers to tide gauge stations. The individual CV matrices are defined as follows:

$$\begin{aligned}
\mathbf{C}_{h_1} &= \sigma_h^2 \mathbf{Q}_{h_1}, \mathbf{C}_{h_2} = \sigma_h^2 \mathbf{Q}_{h_2} \\
\mathbf{C}_{H_1} &= \sigma_H^2 \mathbf{Q}_{H_1} \\
\mathbf{C}_{N_1} &= \sigma_N^2 \mathbf{Q}_{N_1}, \mathbf{C}_{N_2} = \sigma_N^2 \mathbf{Q}_{N_2} \\
\mathbf{C}_{SST_2} &= \sigma_{SST}^2 \mathbf{Q}_{SST_2}
\end{aligned} \tag{7.20}$$

where $\mathbf{Q}_{h_1}, \mathbf{Q}_{H_1}, \mathbf{Q}_{N_1}$ are the positive-definite cofactor matrices for the ellipsoidal, orthometric and geoid heights, respectively, on land. \mathbf{Q}_{h_2} and \mathbf{Q}_{N_2} are the positive-definite cofactor matrices for the ellipsoidal and geoid height data, respectively, at the tide gauge stations.

The utility of the MINQUE procedure described in chapter 5 is demonstrated through the effective incorporation of a fourth variance component, σ_{SST}^2 , for the sea surface topography values at the tide gauge stations in the network. We start with the general form of the stochastic model for the observations given by

$$\mathbf{C}_\ell = \sum_{i=1}^k \theta_i \mathbf{T}_i \tag{7.21}$$

where $\boldsymbol{\theta} = [\sigma_h^2 \quad \sigma_H^2 \quad \sigma_N^2 \quad \sigma_{SST}^2]^T$ contains the four unknown variance components and the \mathbf{T}_i matrix for each height data type is explicitly stated by

$$\mathbf{T}_h = \begin{bmatrix} \mathbf{Q}_{h_1} & \mathbf{Q}_{h_{12}} \\ \mathbf{Q}_{h_{21}} & \mathbf{Q}_{h_2} \end{bmatrix} \tag{7.22a}$$

$$\mathbf{T}_H = \begin{bmatrix} \mathbf{Q}_{H_1} & 0 \\ 0 & 0 \end{bmatrix} \quad (7.22b)$$

$$\mathbf{T}_h = \begin{bmatrix} \mathbf{Q}_{N_1} & \mathbf{Q}_{N_{12}} \\ \mathbf{Q}_{N_{21}} & \mathbf{Q}_{N_2} \end{bmatrix} \quad (7.22c)$$

$$\mathbf{T}_{SST} = \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{Q}_{SST_2} \end{bmatrix} \quad (7.22d)$$

Since a common adjustment is performed for all of the ellipsoidal and the geoid heights (regardless if they are situated at terrestrial benchmarks or at tide gauge stations), the cross-covariance information between the two sets of points for the ellipsoidal and geoid heights is included and denoted by $\mathbf{Q}_{h_{12}} = \mathbf{Q}_{h_{21}}^T$ and $\mathbf{Q}_{N_{12}} = \mathbf{Q}_{N_{21}}^T$, respectively. This also results in a common variance factor to be estimated for each of the ellipsoidal and geoid heights. The positive-definite cofactor matrix for the sea surface topography values is given by \mathbf{Q}_{SST_2} . The decomposition of the stochastic model through the use of the matrices in Eq. (7.21) allows for unbalanced data to be used for the heterogeneous height data types and for more than one variance component to be estimated for the group of heights collectively denoted by \tilde{H} .

Substituting the above equations into the general rigorous MINQUE formulation, the four unknown variance components are estimated by Eq. (5.24), where

$$\mathbf{S} = \begin{bmatrix} tr(\mathbf{RT}_h \mathbf{RT}_h) & tr(\mathbf{RT}_h \mathbf{RT}_H) & tr(\mathbf{RT}_h \mathbf{RT}_N) & tr(\mathbf{RT}_h \mathbf{RT}_{SST}) \\ tr(\mathbf{RT}_H \mathbf{RT}_h) & tr(\mathbf{RT}_H \mathbf{RT}_H) & tr(\mathbf{RT}_H \mathbf{RT}_N) & tr(\mathbf{RT}_H \mathbf{RT}_{SST}) \\ tr(\mathbf{RT}_N \mathbf{RT}_h) & tr(\mathbf{RT}_N \mathbf{RT}_H) & tr(\mathbf{RT}_N \mathbf{RT}_N) & tr(\mathbf{RT}_N \mathbf{RT}_{SST}) \\ tr(\mathbf{RT}_{SST} \mathbf{RT}_h) & tr(\mathbf{RT}_{SST} \mathbf{RT}_H) & tr(\mathbf{RT}_{SST} \mathbf{RT}_N) & tr(\mathbf{RT}_{SST} \mathbf{RT}_{SST}) \end{bmatrix}$$

$$\mathbf{q} = \begin{bmatrix} \ell^T \mathbf{R}^T \mathbf{T}_h \mathbf{R} \ell \\ \ell^T \mathbf{R}^T \mathbf{T}_H \mathbf{R} \ell \\ \ell^T \mathbf{R}^T \mathbf{T}_N \mathbf{R} \ell \\ \ell^T \mathbf{R}^T \mathbf{T}_{SST} \mathbf{R} \ell \end{bmatrix} \quad (7.23)$$

The matrix \mathbf{R} is computed from Eq. (5.26) by substituting the appropriate form of the covariance matrix for the observations as given in Eq. (7.19). The above formulation is applied iteratively as explained in section 5.6.

7.4 Other practical issues

The discussion above essentially showed how to incorporate an additional height data type (that is not necessarily available at each point) into the mixed adjustment and VCE algorithms. Additional modifications can also be made to include the vertical displacement in height values (at both GPS-on-benchmark and GPS-on-tide gauge stations) resulting from crustal movement due to post-glacial rebound or land subsidence. These values may be interpolated from a land uplift model derived using geophysical parameters (Peltier, 1999). The rates of vertical displacements can also be computed from repeated precise levelling and analysis of tide gauge time series and more recently GPS information has been integrated as part of post-glacial rebound studies (Mäkinen, 2000; Mäkinen *et al.*, 2000). The estimated accuracy of the land uplift may be determined from standard errors of computed land uplift differences obtained from the adjustment of precise levelling loops (Ollikainen, 1997). In this case, VCE could be applied to test the standard error values. Explicit formulas for this type of data are not provided herein as they are very similar to the inclusion of the SST information and can be derived along the same lines. The limiting factor in all of these studies is data availability or rather lack of data and obtaining initial CV matrices for the height data in a particular region. However, in practice these effects are visible over time and future improvements in models should allow for the proper numerical incorporation of uplift values for the vertical component.

Disseminating parameters to users

There are several implementation issues that must be considered for a practical 'modernized' approach to establishing vertical control, including (i) how to make model parameters available to users, (ii) to which region should the parameters refer, (iii) how frequently should the model parameters be updated, and (iv) how to make height data quality indicators available to users. Currently, it is common survey practice to use geoid models embedded in GPS receivers and apply the relationship given in Eqs. (1.1) or (1.2), to obtain orthometric heights (or differences) using GPS/geoid levelling. However, as demonstrated from the studies in this thesis, the contribution of the corrector surface is significant and should not be neglected. The experience gained from wide area differential GPS can be applied in this area in order to identify possible pitfalls and needs of the geodetic surveying community. In general, there are three possible alternatives for making *corrector surface parameters* available to users, namely (Fotopoulos, 2000):

- a) provide one set of parameters for an entire network coverage area (i.e. country)
- b) provide multiple sets of parameters for pre-determined 'regions' within the network coverage area
- c) provide both a set of corrections for the entire network (as in (a) above) and multiple sub-sets of parameters for the smaller sub-regions (as in (b) above).

The appropriate choice will depend on the size of the network coverage area and the density of existing vertical control. The advantage of having sub-sets is the ability to model more localized error sources, as demonstrated in the Australian test network in section 4.3. In all cases, the parameters are usually provided in the form of a grid and interpolated at the point(s) of interest (Véronneau *et al.*, 2002). Regardless of the method chosen, it is important that the geographical boundaries used to compute the model parameters are also made available to the user so that they are knowledgeable of the valid coverage area and to avoid extrapolation, which may lead to less accurate results (as evidenced in section 7.2.5 for the northern Canada results).

In terms of frequency of corrector surface model updates, the values are not required in (near) real-time as in the case for corrections in wide area differential GPS applications. Rather, the model parameters should only be updated depending on new network adjustments, provision of additional height data and computation of new geoid models. Therefore, it is important that the corrector surface parameters (in the form of a grid) are provided with an epoch to which the height data refer. In this manner valid comparisons between GPS/geoid levelling derived heights referring to different epochs can be conducted.

A final comment should be made on the availability of *quality* indicators for height data values. It is suggested that full covariance matrices should be provided from software processing packages or at least some indication of the expected accuracy. Users should be aware that these values are usually overly-optimistic (as shown in chapter 6) and VCE can be applied in secondary adjustments in order to 'improve' the CV information.

7.5 Summary

Several practical implementation issues for establishing vertical control using modern techniques were discussed. Specifically, a means for obtaining approximate initial covariance matrices for the ellipsoidal, orthometric and geoid height data was described. This process was tested with a simulated test network in northern Canada and found to be a realistic alternative to obtaining covariance matrices from separate adjustments. The combined height adjustment was performed using relative height differences as opposed to absolute height values, which has been the main focus thus far. Investigations into the baseline configuration and network geometry were also conducted in order to establish a link between the performance of the corrector surface model as a function of network geometry. The actual contribution of the specified parametric model to the overall GNSS-levelling procedure was also studied and found to vary.

Modifications to the mixed height network adjustment algorithm, provided in section 3.1, were described in order to incorporate data from both terrestrial benchmarks and tide gauge stations. Following the same procedure, additional height information can be included (such as land uplift models) as they become available. An important aspect was the modified VCE algorithm, which allows for the calibration of the SST and uplift CV information. Finally, some of the practical issues involved with providing users of vertical control with the appropriate corrector surface model parameters were described.

Chapter 8

Conclusions and Recommendations

In this final chapter the major conclusions and recommendations that can be drawn from this thesis work are outlined. Our starting point was the seemingly simple geometrical relationship between ellipsoidal, orthometric and geoid heights shown in chapter 1. Recognizing the profound impact that satellite-based measurement systems have had on the practice of geodesy and surveying, a detailed deconstruction of the major issues affecting the optimal combination of these heterogeneous height types was conducted. The following discussion provides a summary of the key findings regarding three main areas that were identified as outstanding issues, namely modelling systematic errors and datum inconsistencies, separation of random errors and estimation of variance components for each height type, and practical considerations for modernizing vertical control systems. It should be mentioned that several numerical results for specific test networks (Switzerland, parts of Canada and Australia) are included in the main text and only the major points will be repeated here.

8.1 Conclusions

A general methodology for testing/assessing candidate parametric model performance over a vertical network of co-located GPS-levelling benchmarks was developed. The procedure consists of five major modules including, empirical precision and cross-

validation tests, statistical measures of goodness of fit, indicators of numerical stability and testing of individual parameter significance. The following lists the main conclusions regarding this part of the study:

- Numerical results showed that incorrect conclusions could be drawn regarding the choice of the optimal corrector surface model if the selection is solely based on the computed statistics from $|N_{grav} - N_{GPS/levelling}|$. Instead, the complete procedure developed in chapter 3 should be used which provides additional information regarding the numerical stability and performance of the model when used to predict height values at 'new' points (as in GPS-levelling).
- Two alternative formulations for testing the significance of parameters of a corrector surface model were presented. The key advantage of the implemented formulation was the ability to scrutinize the significance of individual (or several) parameters and eliminate them *without* the need to repeat the combined least-squares height adjustment.
- High correlation among model parameters was found to distort the results of the statistical testing process. Gram-Schmidt orthonormalization was implemented and shown to be a viable method for providing a *new* set of uncorrelated parameters. Subsequent evaluation of parameter significance resulted in a reduction in the number of vital terms, which was undetectable before orthonormalization.
- A significant improvement in modelling the systematic discrepancies between gravimetrically-derived geoid heights and a nation-wide set of GPS-levelling control points was demonstrated through the use of a mosaic of different parametric models tailored to smaller sub-regions. This regional parametric modelling approach was proven to provide better results than using the 'best' national fit for the Australian vertical control test network of GPS-levelling benchmarks as described in chapter 4.
- Various types of parametric models were tested including, polynomials of various degrees and order and base functions (MRE) and similarity-based transformation models. In general, the dangers of over-parameterization were described and the

importance of comparing the numerical stability of each model over the region of interest was demonstrated. Such models should particularly be avoided if the parametric model is used to predict values for independent/new points exclusive of the original network adjustment, as shown by the results of the cross-validation tests.

Another major contribution of this research work is the detailed variance component estimation studies related to the common adjustment of the ellipsoidal, orthometric and geoid height data types. In particular, the iterative minimum norm quadratic unbiased estimation algorithm using the separated height residuals (according to type) was determined as an appropriate method for this problem. The following conclusions can be drawn from the extensive numerical tests on the Swiss and Canadian test networks:

- The iterative form of the MINQUE method was shown to be a globally best estimator providing consistent results independent of the a-priori covariance matrix values. This led to the 'validation' of the empirical (and somewhat arbitrary) scaling of the a-posteriori covariance matrix for the GPS-derived ellipsoidal heights. For example, in the case of the Swiss test network, an empirically derived scaling factor of 10^2 for \mathbf{Q}_h was validated through VCE. Therefore, more realistic measures for the accuracy of the ellipsoidal height data can be obtained using VCE as compared to the typically optimistic a-priori CV matrices supplied by GPS post-processing software.
- Global and local geoid error models were calibrated via VCE. In all cases, the over-bearing contribution of the global geopotential model (EGM96) to the total error budget was clear. Results revealed the necessary down-scaling of the provided/computed covariance matrices (\mathbf{Q}_N) for the given test areas, for both the global and the regional geoid models. Overall, it was found that the use of a diagonal CV matrix for the geoid height errors gives a very poor representation of the actual contribution of geoid uncertainty to the total error budget.
- The effectiveness of evaluating the orthometric height accuracy from national adjustments (in the Swiss case) and regional adjustments (for the Canadian test

- network) of levelling data was shown. Overall, the VCE results verified that the orthometric heights were the most accurate of the three height data types, as expected, however the corresponding a-priori covariance matrices (\mathbf{Q}_H) were found to be optimistic in all cases.
- The estimation of variance components for each height type was conducted using two methods, namely, the IMINQUE and the IAUE method. Numerical studies showed that both methods provide very similar results for the estimated variance components (differences at the mm-level). It was also observed that the IAUE approach was on average 50% more computationally efficient than the rigorous IMINQUE method. Therefore, the IAUE method should be used when computational efficiency is an issue.
 - The effect of data redundancy on the estimation of variance components was investigated through several numerical tests. In general, a positive-valued numerical solution for all variance components was obtained unless the number of observations was significantly reduced (assuming the number of unknown parameters sought remains constant). The effect of changing the network geometry on the admissibility of the estimated variance components was exemplified, to some end, by the results of the Canadian test network. In this case, the removal of observations (GPS-levelling benchmarks) from the already sparse network resulted in the estimation of some negative variance components during the initial stages of the iterative process. This in turn resulted in a slower rate of convergence.
 - Overly optimistic values for the variance factors result when correlations among heights of the same type are neglected (i.e., diagonal-only CV matrices). The number of iterations required for a solution also dramatically increases (i.e. from 99 to 152 iterations in the Swiss test network case) when diagonal-only CV matrices are used. The greatest impact on the final estimated variance components was obtained when a diagonal CV matrix for the EGM96 model was used compared to the fully-populated version. In this case, the use of an a-priori diagonal CV matrix resulted in extremely optimistic values for the Swiss network. For the Canadian test network, a positive-

valued variance component was not obtainable when a diagonal a-priori CV matrix corresponding to EGM96 was used. Thus, correlations should not be neglected, especially when a global geopotential model is used.

- An interesting relation between the type of parametric model and the final estimated variance components was observed. It was found that the use of an *inappropriate* parametric model in the common adjustment of the height data might lead to problems in estimating variance components (i.e., divergence or negative-values). The model assessment procedure described in chapter 3 was used to verify the inadequacy of these models for absorbing the systematic errors and datum inconsistencies in the test network area. Therefore, the intrinsic connection between the systematic effects and datum inconsistencies and the VCE process (which presupposes the absence of biases and no systematic effects in the data) is evident. According to the results shown herein, this can be extended to the possible *detection* of inappropriate parametric models (not the selection of a proper model).

Several other issues were also studied that provided insight into the practical problems encountered when implementing GPS-levelling. The following interesting conclusions were made:

- A method for evaluating the achievable accuracy of relative GPS-levelling that incorporates the quality of the ellipsoidal and geoid height data and the contribution of the parametric model (which is typically neglected in practice) was presented. This approach does not require any actual observations and can therefore be used in the pre-design stages of surveys.
- The accuracy contribution of the parametric model was found to be insignificant in general compared to the accuracy of the relative geoid and ellipsoidal height data. However, the incorporation of the parametric model is absolutely essential in the common height adjustment as the magnitude of the datum inconsistencies and systematic errors in the original control network can reach several metres.

- A practical problem encountered by 'secondary' users of height data is the lack of covariance information. A method for generating realistic covariance matrices for the $\Delta h, \Delta H, \Delta N$ data when they are not readily available was described and tested.
- Tests conducted using GPS-levelling in remote areas with few vertical control points showed that as the distance between the newly established baseline and the original control network increased, the significance of using fully-populated versus diagonal CV matrices for the three height types became evident. Therefore, it is recommended that fully-populated forms of CV matrices for the height types be used in the evaluation of the accuracy of orthometric heights computed via GPS-levelling.
- Modifications to the combined adjustment scheme (described in chapter 3) and the IMINQUE algorithm (described in chapter 5) to accommodate additional height information were provided. In particular, the integrated optimal adjustment of terrestrial benchmarks and tide gauge stations was described. The application of the MINQUE algorithm for the case of unbalanced data where sea surface topography information is available at tide gauge stations was also presented. A similar approach can be applied to include additional height information expected in the near future such as land uplift models.
- Practical options for disseminating corrector surface parameters to GPS-levelling users were described. It is suggested that fully-populated covariance matrices for the height data are also made available to practitioners so that the optimal adjustment procedure summarized in chapter 7 can be implemented.

8.2 Recommendations for future work

The following is a list of some of the areas recommended for future work:

- Parametric models composed of orthogonal base functions should be testing using real data and evaluated using cross-validation at independent control points after insignificant terms have been identified and deleted from the analytical model.

- The least-squares adjustment scheme presented in this work focused on using either absolute height values (h, H, N) at all points in the test network or relative height differences ($\Delta h, \Delta H, \Delta N$) between all points in the test network. It would be worthwhile to re-formulate the problem using a mixture of observed input values including relative height differences and absolute values, which is a closer representation of real databases.
- The variance component estimation studies conducted herein should be extended to consider estimating covariance components as well. This complicates the problem somewhat through the introduction of more unknown parameters and reduces the data redundancy. The viability of a numerical solution using available a-priori height data CV matrices should be tested.
- The correlation between the local gravimetric geoid heights and the levelled orthometric heights at the GPS-levelling benchmarks was ignored for the purpose of this study. The effect of such a correlation, which exists in the case where the gravimetric geoid solution has incorporated height information from the levelling network, may merit further investigation.
- The adjusted residuals were separated according to height type and used to estimate a single variance component for each group. It may be worthwhile, especially in larger disconnected networks, to estimate more than one variance component for each height type.
- Given the upcoming satellite-only global gravity field models expected from the LEO missions for dedicated gravity field research, it is suggested that the VCE procedures described herein are implemented to calibrate the new global geoid error models. The results should be compared to those obtained from previous models such as EGM96.
- Temporal variations of height datums are very important and must be considered in view of the goal for a 1-cm accurate geoid. This involves the incorporation of time variable geophysical, geodynamic and oceanographic models to account for processes such as sea level change, post-glacial rebound, plate subduction and plate movement,

land subsidence, etc. in the geoid model. Essentially we engage in a bootstrap procedure where data from different time periods is used to test the validity of these time variable models while this data is also integrated to improve the models themselves.

Given the current state of technologies with emerging global navigation satellite systems (i.e., GALILEO) and modernized existing systems (GPS) as well as new LEO missions for dedicated gravity field research (i.e., CHAMP, GRACE, GOCE), it is expected that a new era in geodesy is upon us offering many challenges and promising resolutions. The author hopes that this research will be a small contribution to the changing world of geodesy and surveying in view of the vertical datum problem.

References

- Adam J, Augath W, Brouwer F, Engelharat G, Gurtner W, Harsson G, Ihde J, Ineichen D, Lang H, Luthardt J, Sacher M, Schlüter W, Springer T, Wöppelmann G (2000): Status and development of the European height systems. *International Association of Geodesy Symposia*, vol. 121, Schwarz KP (Ed.), Springer-Verlag. *Proceedings of the International Association of Geodesy General Assembly, Geodesy and Beyond 2000 - The Challenges of the First Decade*, Birmingham, UK, July 19-30, pp. 55-60.
- Allister NA, Featherstone WE (2001): Estimation of Helmert orthometric heights using digital barcode levelling, observed gravity and topographic mass-density data over part of the Darling Scarp, Western Australia, *Geomatics Research Australasia*, vol. 75, pp. 25-52.
- Andersen OB and Knudsen P (2000): The role of satellite altimetry in gravity field modelling in coastal areas. *Physics and Chemistry of the Earth*, vol. 25, no. 1, pp. 17-24.
- Baker TF, Curtis DT, Dodson AH (1995): Ocean tide loading and GPS. *GPS World*, March, pp. 54-59.
- Balasubramania N (1994): Definition and Realization of a Global vertical Datum. Ohio State University, Department of Geodetic Science and Surveying, Report No. 427, April 1994.
- Beutler G, Baueršima I, Gurtner W, Rothacher M, Schildknecht T, and Geiger A (1987): Atmospheric refraction and other important biases in GPS carrier phase observations. University of New South Wales, School of Surveying, Monograph 12.
- Bin AY, Khoon TY, and Soon VKH (1995): A geometric geoid model for Singapore. Technical Proceedings of the 5th South East Asian and 35th Australian Surveyors Congress, *Network into the 21st Century*, July 16-20, Singapore, vol. 1, pp. 317-382.
- Bomford G (1971): *Geodesy*. Oxford University Press, New York.
- Bouman J (1998): Quality of regularization methods. DEOS Report no. 98.2, Delft University Press, The Netherlands.
- Braasch, MS (1996): Chapter 14: Multipath Effects. In *Global Positioning System: Theory and Applications*, Volume I. Eds. Parkinson, B. and Spilker, J., American Institute of Aeronautics and Astronautics, Inc., pp. 547-566.

- Burša M, Kenyon S, Kouba J, Raděj K, Vatrt V, Vojtišková M, and Šimek J (2001): World height system specified by geopotential at tide gauge stations. *International Association of Geodesy Symposia*, vol. 124, Drewes H, Dodson A, Fortes LPS, Sánchez L and Sandoval P (Eds.), Springer-Verlag. *Proceedings of the Vertical Reference Systems Symposium*, Cartagena, Colombia, Feb. 20-23, pp. 291-296.
- Carroll JD and Green PE (1976): *Mathematical tools for applied multivariate analysis*. Academic Press.
- Caspary WF (1987): *Concepts of network and deformation analysis*. The University of New South Wales, School of Geomatic Engineering, Monograph 11, pp. 97-111.
- Chen YQ, Chrzanowski A, and Kavouras M (1990): Assessment of Observations Using Minimum Norm Quadratic Unbiased Estimation (MINQUE), *CISM Journal of ACSGS*, Ottawa, Canada, vol. 44, pp. 39 – 46.
- Cheney B, Miller L, Agreen R, Doyle N and Lillibridge J (1994): TOPEX/POSEIDON: The 2-cm Solution. *Journal of Geophysical Research*, 99 (C12), pp. 24555-24564.
- Colombo OL (1980): *A World Vertical Network*. Ohio State University, Department of Geodetic Science, Report No. 296.
- Craymer MR, Penney R, Lapelle E (1997): *CBN Status Report: Processing, adjustment and integration*. Presented to the Canadian Geodetic Reference System Committee, Ottawa, April 23.
- Crocetto N, Gatti M, and Russo P (2000): Simplified formulae for the BIQUE estimation of variance components in disjunctive observation groups. *Journal of Geodesy*, 74, pp. 447-457.
- Davis PJ (1963): *Interpolation and Approximation*. Dover Publications, Inc., New York.
- Davis RE, Foote FS, Anderson JM, Mikhail EM (1981): *Surveying theory and practice*. Sixth edition, McGraw-Hill Inc.
- de Bruijne A, Haagmans R, and de Min E (1997): *A preliminary North Sea Geoid Model GEONZ97*. Project report, Directoraat-Generaal Rijkswaterstaat, Meetkundige Dienst.
- de Min E (1990): *On the Computation of the Geoid and its Errors: With Special Attention to the Inner Zone*. *MSc. Thesis*, Faculty of Geodesy, Technical University of Delft.

- Denker H, Torge W, Wenzel G, Ihde J and Schirmer U (2000): Investigation of different methods for the combination of gravity and GPS/levelling data. *International Association of Geodesy Symposia*, vol. 121, Schwarz KP (Ed.), Springer-Verlag. *Proceedings of the International Association of Geodesy General Assembly, Geodesy Beyond 2000 - The Challenges of the First Decade Symposium*, Birmingham, UK, July 19-30, pp. 137-142.
- Dermanis A and Rossikopoulos D (1991): Statistical Inference in Integrated Geodesy. Presented at the *IUGG XXth General Assembly, International Association of Geodesy*, Vienna, August 11-24.
- Dinter G, Illner M, and Jäger R (2001): A general concept for the integration of GPS-heights into standard height systems comprising the quality analysis and the refinement of geoid models. Presented at the *International Association of Geodesy Symposium on Vertical Reference Systems*, Cartagena, Colombia, Feb. 20-23
- Dodson AH (1995): GPS for height determination. *Survey Review*, vol. 33, no. 256, pp. 66-76.
- Duquenne H, Jiang Z, and Lemarie C (1995): Geoid determination and levelling by GPS: Some experiments on a test network. *International Association of Geodesy Symposia*, vol. 113, Springer-Verlag. *Proceedings of the Gravity and Geoid Symposium*, Graz, Austria, pp. 559-568.
- Ekholm S (1998): Determination of Greenland surface topography from satellite altimetry and other elevation data. National Survey and Cadastre – Denmark, *Technical Report No. 9*.
- Ekman M (1989): Impacts of geodynamic phenomena on systems for height and gravity. *Bulletin Géodésique*, vol. 63, pp. 281-296.
- Elósegui P, Davis JL, Jaldehag TK, Johansson JM, Niell AE and Shapiro II (1995): Geodesy using the Global Positioning System: The effects of signal scattering on estimates of site position. *Journal of Geophysical Research*, vol. 100, no. B7, pp. 9921-9934.
- El-Rabbany AE (1994): The Effect of Physical Correlations on the Ambiguity Resolution and Accuracy Estimation in GPS Differential Positioning. *PhD Thesis*, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada.
- Engelis T, Rapp RH, and Bock Y (1985): Measuring orthometric height differences with GPS and gravity data. *Manuscripta Geodaetica*, vol. 10, pp. 187-194.

- ESA (1999): The four candidate Earth explorer missions - Gravity field and steady-state ocean circulation. European Space Agency, ESA SP-1233 (1), 217 pp.
- ESA (2002): GALILEO: The European Programme for Global Navigation Services. European Space Agency, BR-186, ESA Publications Division, The Netherlands.
- Featherstone W (1998): Do we need a gravimetric geoid or a model of the Australian height datum to transform GPS Heights in Australia? *The Australian Surveyor*, vol. 43, no. 4, pp. 273-280.
- Featherstone W (2000): Refinement of gravimetric geoid using GPS and levelling data. *Journal of Surveying Engineering*, vol. 126, no. 2, May 2000, pp. 27-56.
- Featherstone W and Guo W (2001): Evaluations of the Precision of the AUSGeoid98 Versus AUSGeoid93 Using GPS and Australian Height Datum Data. *Geomatics Research Australasia*, no. 74, pp. 75-102.
- Featherstone WE, Kirby JF, Kearsley AHW, Gilliland JR, Johnston J, Steed R, Forsberg R and Sideris MG (2001): The AUSGeoid98 geoid model of Australia: data treatment, computations and comparisons with GPS/levelling data. *Journal of Geodesy*, vol. 75, pp. 313-330.
- Featherstone WE, Kirby JF, Holmes SA, Fotopoulos G and Goos JM (2002): Recent research towards an improved geoid model for Australia. Presented at the *Joint AURISA and Institute of Surveyors Conference*, Adelaide, Australia, Nov. 25-30.
- Forsberg R (1994): Terrain Effects in Geoid Computations, In International School for the Determination of the Geoid, Lecture Notes, Milan, Italy, Oct. 10-15, pp. 101-134.
- Forsberg R and Madsen F (1990): High precision geoid heights for GPS levelling. Proceedings of *GPS'90/SPG'90*, pp. 1060-1074.
- Forsberg R (1998): Geoid Tayloring to GPS - with Example of a 1-cm Geoid of Denmark. Proceedings of the *2nd Continental Workshop on the Geoid in Europe*, Budapest, Hungary, March 10-14, Vermeer M (Ed.). Report of Finnish Geodetic Institute, 98:4, pp. 191-198.
- Forsberg R and Featherstone WE (1998): Geoids and cap sizes. International Association of Geodesy Symposia, vol. 119, Forsberg R, Feissel M and Dietrich R (Eds.), Springer-Verlag. *Proceedings of the Geodesy on the Move Symposium*, Rio de Janeiro, pp. 194-200.

- Fotopoulos G (2000): Parameterization of DGPS carrier phase errors over a regional network of reference stations. *MSc. Thesis*, University of Calgary, Department of Geomatics Engineering, Report No. 20142.
- Fotopoulos G and Sideris MG (2003): On the estimation of variance components using GPS, geoid and levelling data. Presented at the *Canadian Geophysical Union Annual meeting: Challenges and opportunities for geophysics in Canada*, Banff, Canada, May 10-14.
- Fotopoulos G, Kotsakis C, and Sideris MG (2001a): How accurately can we determine orthometric height differences from GPS and geoid data? *Journal of Surveying Engineering*, vol. 129, no. 1, pp. 1-10.
- Fotopoulos G, Kotsakis C, and Sideris MG (2001b): Determination of the Achievable Accuracy of Relative GPS/Geoid Levelling in Northern Canada. Presented at the *International Association of Geodesy 2001 Scientific Assembly*, Budapest, Hungary, September 2-8.
- Fotopoulos G, Featherstone WE and Sideris MG (2002): Fitting a gravimetric geoid model to the Australian height datum via GPS data. *Proceedings of the 3rd Meeting of the International Gravity and Geoid Commission*, Tziavos (Ed.), Thessaloniki, Greece, Aug. 26-30, pp. 173-178.
- Gareau RM (1986): History of Precise Levelling in Canada and the North American Vertical Datum Readjustment. *MSc. Thesis*, University of Calgary, Department of Geomatics Engineering, Report No. 20018.
- Geodetic Survey Division (1992): Guidelines and specifications for GPS surveys. Release 2.1, Geodetic Survey Division, Canada Centre for Surveying, Surveys, Mapping and Remote Sensing Sector.
- Geodetic Survey Division (1998): Providing Access to the Canadian Spatial Reference System (CSRS) 1998. McGregor-Sauvé, S. and Scott, D., Geodetic Survey Division, Geomatics Canada (www.geod.nrcan.gc.ca).
- Grafarend EW (1984): Variance-covariance component estimation of Helmert type in the Gauss-Helmert model. *ZfV*, vol. 109, pp. 34-44.
- Grafarend EW (1985): Variance-covariance component estimation theoretical results and geodetic applications. *Statistics and Decisions*, Supplement Issue No. 2, pp. 407-441.
- Grafarend EW and Ardalan AA (1997): W_0 : an estimate in the Finnish Height Datum N60, epoch 1993.4, from twenty-five GPS points of the Baltic Sea Level Project. *Journal of Geodesy*, vol. 71, pp. 673-679.

- Grodecki J (1997): Estimation of Variance-Covariance Components for Geodetic Observations and Implications on Deformation Trend Analysis. *Ph.D. dissertation*, Department of Geodesy and Geomatics Engineering Technical Report No. 186, University of New Brunswick, Fredericton, New Brunswick, Canada.
- Groten E and Müller T (1991): Combined Sea Surface Determination Based on Various Geometric and Dynamic Techniques. *Marine Geodesy*, vol. 14, pp. 185-195.
- Gruber T and Steigenberger P (2002): Impact of new gravity field missions for sea surface topography determination. *Proceedings of the 3rd Meeting of the International Gravity and Geoid Commission, GG2002*, Tziavos (Ed.), Thessaloniki, Greece, Aug. 26-30, pp. 320-325.
- Haagmans R, de Bruijne A, and de Min E (1998): A procedure for combining gravimetric geoid models and independent geoid data, with an example in the North Sea region. *DEOS Progress Letters 98.1*, Delft University Press.
- Haagmans R, de Min E, and van Gelderen M (1993): Fast evaluation of convolution integrals on the sphere using 1D FFT, and a comparison with existing methods for Stokes' integral. *Manuscripta Geodaetica*, vol. 18, pp. 227-241.
- Hartung J (1981): Non-negative minimum biased invariant estimation in variance component models. *The Annals of Statistics*, vol. 9, Issue 2, pp. 278-292.
- Heck B (1990): An evaluation of some systematic error sources affecting terrestrial gravity anomalies. *Bulletin Géodésique*, vol. 64, pp. 88-108.
- Heck B and Rummel R (1990): Strategies for solving the vertical datum problem using terrestrial and satellite geodetic data. *International Association of Geodesy Symposia*, vol. 104, Sünkel H and Baker T (Eds.), Springer-Verlag, pp. 116-128.
- Heiskanen WA and Moritz H (1967): *Physical Geodesy*. W.H. Freeman and Company San Francisco.
- Helmert FR (1924): *Die Ausgleichungsrechnung nach der Methode der Kleinsten Quadratic*, 3. AUFL., Leipzig/Berlin.
- Hofmann-Wellenhof B, Licktenegger H and Collins J (1992): *GPS Theory and Practice*, Third revised edition, Springer-Verlag.
- Hipkin R (2000): Modelling the geoid and sea surface topography in coastal areas. *Physics and Chemistry of the Earth*, vol. 25, no. 1, pp. 9-16.

- Horn SD and Horn RA (1975): Comparison of estimators of heteroscedastic variances in linear models. *Journal of the American Statistical Association*, vol. 70, Issue 352, pp. 872-879.
- Jäger R (1999): State of the art and present developments of a general concept for GPS-based height determination. *Proceedings of Geodesy Surveying in the Future, The Importance of Heights*, International Federation of Surveyors, Gävle, Sweden, March 15-17, pp. 161-174.
- Jekeli C (1979): Global accuracy estimates of point and mean undulation differences obtained from gravity disturbances, gravity anomalies and potential coefficients. Ohio State University, Department of Geodetic Science, Report No. 288.
- Jekeli C (2000): Heights, the Geopotential, and Vertical Datums. Ohio State University, Geodetic Science and Surveying, Department of Civil and Environmental Engineering and Geodetic Science, Report No. 459.
- Jiang Z and Duquenne H (1995): Optimal fitting of gravity geoid to GPS levelling points. Latest Developments in the Computation of Regional Geoids, *Proceedings of Session G4, EGS XX General Assembly*, Vermeer M (Ed.), Hamburg, Germany, April 3-7, pp. 33-42.
- Jiang Z and Duquenne H (1996): On the combined adjustment of a gravimetrically determined geoid and GPS levelling stations. *Journal of Geodesy*, vol. 70, pp. 505-514.
- Johnston GM and Luton GC (2001): GPS and the Australian Height Datum, *Proceedings of the 5th International Symposium on Satellite Navigation Technology and Applications*, Kubik K, Rizos C, Featherstone WE (Eds.), Canberra, Australia, July, [CD-ROM].
- Kaplan D (1996): *Understanding GPS Principles and Applications*. Artech House Publishers, Boston.
- Kearsley AHW, Ahmad Z, and Chan A (1993): National Height Datums, Levelling, GPS Heights and Geoids. *Australian Journal of Geodesy, Photogrammetry, and Surveying*, no. 59, pp. 53-88.
- Kelly KM (1991): Weight estimation in geodetic levelling using variance components derived from analysis of variance. *Master of Applied Science Thesis*, University of Toronto, Department of Civil Engineering.

- Klees R and van Gelderen M (1997): The Fiction of the Geoid. Presented at the *Staring Symposium, Royal Academy of Sciences*, Amsterdam, Netherlands, Oct. 12, 1997. Also in Delft Institute for Earth-Oriented Space Research (DEOS) Progress Letters, 97.1, pp. 17-26.
- Koch KR (1986): Maximum likelihood estimate of variance components. *Bulletin Géodésique*, vol. 60, no. 4, pp. 329-338.
- Koch KR (1987): Bayesian inference for variance components. *Manuscripta Geodaetica*, vol. 12, pp. 309-313.
- Koch KR (1988): Bayesian statistics for variance components with informative and noninformative priors. *Manuscripta Geodaetica*, vol. 13, pp. 370-373.
- Koch KR (1999): Parameter Estimation and Hypothesis Testing in Linear Models, Second edition, Springer-Verlag, Berlin Heidelberg, Germany.
- Koch KR and Kusche J (2002): Regularization of geopotential determination from satellite data by variance components. *Journal of Geodesy*, vol. 76, pp. 259-268.
- Kotsakis C and Sideris MG (1999): On the adjustment of combined GPS/levelling/geoid networks. *Journal of Geodesy*, vol. 73, no. 8, pp. 412-421.
- Kotsakis C, Fotopoulos G, Sideris MG (2001): Optimal fitting of gravimetric geoid undulations to GPS/levelling data using an extended similarity transformation model. Presented at the 27th Annual meeting joint with the 58th Eastern snow conference of the Canadian Geophysical Union, Ottawa, Canada, May 14-17.
- Kubik K (1970): The estimation of weights of measured quantities within the method of least squares. *Bulletin Géodésique*, vol. 95, pp. 21-40.
- Kusche J (2003): A Monte-Carlo technique for weight estimation in satellite geodesy. *Journal of Geodesy*, 76, pp. 641-652.
- Lancaster P and Šalkauskas K (1986): Curve and Surface Fitting: An Introduction. Academic Press Inc., London, England.
- Lehmann R (2000): Altimetry-gravimetry problems with free vertical datum. *Journal of Geodesy*, vol. 74, pp. 327-334.

- Lemoine FG, Kenyon SC, Factor JK, Trimmer RG, Pavlis NK, Chinn DS, Cox CM, Klosko SM, Luthcke SB, Torrence MH, Wang YM, Williamson RG, Pavlis EC, Rapp RH, and Olson TR (1998): The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) geopotential model EGM96. NASA Technical Publication -1998 - 206861, July 1998.
- Li Y (2000): Airborne gravimetry for geoid determination. *PhD Thesis*, University of Calgary, Department of Geomatics Engineering, Report No. 20141.
- Li Y and Sideris MG (1994): Minimization and estimation of geoid undulation errors. *Bulletin Géodésique*, vol. 68, no. 4, pp. 201-219.
- Lucas JR and Dillinger WH (1998): MINQUE for block diagonal bordered systems such as those encountered in VLBI data analysis. *Journal of Geodesy*, vol. 72, pp. 343-349.
- Mainville A, Forsberg R and Sideris MG (1992): Global Positioning System Testing of Geoids Computed from Geopotential Models and Local Gravity Data. *Journal of Geophysical Research*, vol. 97, no. B7, pp. 11,137-11,147.
- Mainville A, Véronneau M, Forsberg R and Sideris MG (1994): A Comparison of Terrain Reduction Methods in Rough Topography. *Presented at the Joint Meeting of the International Gravity Commission and the International Geoid Commission*, Graz, Austria.
- Mäkinen J (2000): Geodynamical studies using gravimetry and levelling. *Academic dissertation in Geophysics*, Faculty of Science, University of Helsinki, Helsinki, Finland.
- Mäkinen J, Koivula H, Poutanen M and Saaranen V (2000): Contemporary postglacial rebound in Finland: comparison of GPS results with repeated precise levelling and tide gauges. *Presented at the International Association of Geodesy, Gravity, Geoid and Geodynamics Symposium*, Banff, Canada, July 31-Aug. 5.
- Marana B and Sanso F (1996): On the sensitivity of the least squares solution to errors in the covariance of the observations. *Reports on Surveying and Geodesy*. Università Degli Studi di Bologna, Edizioni Nautilus.
- Marti U (2002): Modelling of differences of height systems in Switzerland. *Proceedings of the 3rd Meeting of the International Gravity and Geoid Commission*, Tziavos (Ed.), Thessaloniki, Greece, Aug. 26-30, pp. 378-388.
- Marti U, Schlatter A, Brockmann E and Wiget A (2000): The way to a consistent national height system for Switzerland. *International Association of Geodesy Symposia, Vistas for Geodesy in the New Millennium*, vol. 125, Adam and Schwarz (Eds.), pp. 90-95.

- Melchior P (1978): *The Tides of the Planet Earth*. Pergamon Press, Oxford.
- Mikhail EM (1976): *Observations and Least-Squares*. Thomas Y. Crowell Company, Inc., New York.
- Milbert DG (1995): Improvement of a high resolution geoid height model in the United States by GPS height on NAVD88 benchmarks. *International Association of Geodesy Bulletin d'Information, New Geoids in the World*, Balmino and Sanso (Eds.), IGeS Bulletin no. 4, pp. 13-36
- Mitchum GT (2000): An Improved Calibration of Satellite Altimetric Heights Using Tide Gauge Sea Levels with Adjustment for Land Motion. *Marine Geodesy*, vol. 23, pp. 145-166.
- Moritz H (1980): *Advanced Physical Geodesy*. Herbert Wichmann Verlag, Karlsruhe, Abacus Press, Tunbridge Wells, Kent.
- Naeije MC, Haagmans RHN, Schrama EJO, Scharroo R and Wakker KF (1998): Ocean Circulation, Height Systems, and Gravity Field Studies Based on TOPEX/POSEIDON Altimetry. *AVISO Altimetry Newsletter*, April 1998, No. 6.
- Nafisi V (2003): Detection of systematic errors by variance components. *Survey Review*, vol. 37, no. 288, pp. 155-161.
- National Geodetic Survey (1994): *Input Formats and Specifications of the National Geodetic Survey Data Base: Volume II, Vertical Control Data*. Systems Development Division, National Geodetic Survey (NGS), NOAA, Silver Spring, MD.
- Ollikainen M (1997): *Determination of Orthometric Heights Using GPS Levelling*. *Publication of the Finnish Geodetic Institute*, No. 123.
- Ou Ziqiang (1989): Estimation of variance and covariance components. *Bulletin Géodésique*, vol. 63, pp. 139-148.
- Ou Ziqiang (1991): Approximative Bayes estimation for variance components. *Manuscripta Geodaetica*, vol. 16, pp. 168-172.
- Pan M and Sjöberg LE (1998): Unification of vertical datums by GPS and gravimetric geoid models with application to Fennoscandia. *Journal of Geodesy*, vol. 72, pp. 64-70.
- Papoulis A (1990): *Probability and Statistics*. Prentice-Hall, Inc.
- Parkinson B and Spilker J (1996a): *Global Positioning System: Theory and Applications*, Volume I. American Institute of Aeronautics and Astronautics, Inc., Washington, DC.

- Parkinson B and Spilker J (1996b): *Global Positioning System: Theory and Applications*, Volume II. American Institute of Aeronautics and Astronautics, Inc., Washington, DC.
- Pavlis NK (1988): Modeling and estimation of low degree geopotential model from terrestrial gravity data. Ohio State University, Department of Geodetic Science and Surveying, Report 386.
- Pavlis NK and Kenyon SC (2002): Analysis of surface gravity and satellite altimetry data for their combination with CHAMP and GRACE information. *Proceedings of the 3rd Meeting of the International Gravity and Geoid Commission, GG2002*, Tziavos (Ed.), Thessaloniki, Greece, Aug. 26-30, pp. 249-261.
- Peltier WR (1999): Global sea level rise and glacial isostatic adjustment. *Global and Planetary Change*, vol. 20, pp. 93-123.
- Persson CG (1981): On the estimation of variance components in linear models - with special reference to geodetic applications. Royal Institute of Technology, Division of Geodesy, Stockholm.
- Poutanen M, Vermeer M and Mäikinen J (1996): The permanent tide in GPS positions. *Journal of Geodesy*, vol. 70, pp. 499-504.
- Pukelsheim F (1981): On the existence of unbiased non-negative estimates of variance covariance components. *The Annals of Statistics*, vol. 9, Issue 2, pp. 293-299.
- Rao CR (1970): Estimation of Heterogeneous Variances in Linear Models. *Journal of American Statistical Association*, vol. 65, pp. 161 – 172.
- Rao CR (1971): Estimation of Variance Components - MINQUE Theory. *Journal of Multivariate Statistics*, vol.1, pp.257-275.
- Rao CR and Kleffe J (1988): Estimation of Variance Components and Applications. North-Holland *Series in Statistics and Probability*, vol.3.
- Rapp RH and Rummel R (1975): Methods for the Computation of Detailed Geoids and their Accuracy. Ohio State University, Department of Geodetic Science and Surveying, Report No.233.
- Roelse A, Granger HW, Graham JW (1971): The adjustment of the Australian levelling survey 1970 – 1971, *Technical Report 12*, Division of National Mapping, Canberra, 81 pp.

- Roman D and Smith DA (2000): Recent investigations toward achieving a one-centimeter geoid. *International Association of Geodesy Symposia*, vol. 123, Springer-Verlag, Sideris MG (Ed.), *Gravity, Geoid and Geodynamics 2000 Symposium*, Banff, Canada, July 31 - Aug. 4, pp. 285-289.
- Rothacher M (2001): Estimation of station heights with GPS. *International Association of Geodesy Symposia*, vol. 124, Springer-Verlag, Drewes H, Dodson A, Fortes LPS, Sanchez L and Sandoval P (Eds.), *Vertical Reference Systems Symposium*, Cartagena, Colombia, Feb. 20-23, pp. 81-90.
- Rummel R (2000): Global Unification of Height Systems and GOCE. *International Association of Geodesy Symposia*, vol. 123, Springer-Verlag, Sideris MG (Ed.), *Gravity, Geoid and Geodynamics 2000 Symposium*, Banff, Canada, July 31-Aug. 4, pp. 13-20.
- Rummel R and Teunissen P (1989): Height datum definition, height datum connection and the role of the geodetic boundary value problem. *Bulletin Géodésique*, vol. 62, pp. 477 - 498.
- Sahin M, Cross PA, and Sellers PC (1992): Variance Component Estimation Applied to Satellite Laser Ranging, *Bulletin Géodésique*, vol. 66, pp. 284 – 295.
- Santerre R (1991): Impact of GPS satellite sky distribution. *Manuscripta Geodaetica*, no. 16, pp. 28-53.
- Satirapod C, Wang J and Rizos C (2002): A simplified MINQUE procedure for the estimation of variance-covariance components of GPS observables. *Survey Review*, vol. 36, no. 286, pp. 582-590.
- Schaffrin B and Iz HB (2002): BLIMPBE and its Geodetic Applications. *International Association of Geodesy Symposia*, vol. 125, Adam J and Schwarz KP (Eds.), Springer-Verlag. *Vistas for Geodesy in the New Millennium Symposium*, Budapest, Hungary, Sept. 2-7, pp. 377-381.
- Schwarz KP, Sideris MG, Forsberg R (1987): Orthometric heights without levelling. *Journal of Surveying Engineering*, vol. 113, no. 1, pp. 28-40.
- Seeber G (1993): *Satellite Geodesy: Foundations, Methods and Applications*. Walter de Gruyter, New York.
- Sen A and Srivastava M (1990): *Regression Analysis: Theory, Methods and Applications*. Springer Texts in Statistics, Springer, New York.

- Searle SR, Casella G and McCulloch CE (1992): Variance Components. Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, Inc., New York.
- She BB (1993): A PC-Based Unified Geoid for Canada. *MSc. Thesis*, University of Calgary, Department of Geomatics Engineering, Report No. 20051.
- Shretha R, Nazir A, Dewitt B, and Smith S (1993): Surface Interpolation Techniques to Convert GPS Ellipsoid Heights to Elevations. *Surveying and Land Information Systems*, vol. 53, no. 2, pp. 133-144.
- Shum CK, Woodworth PL, Andersen OB, Egbert E, Francis O, King C, Klosko S, Le Provost C, Li X, Molines JM, Parke M, Ray R, Schlax M, Stammer D, Thierney C, Vincent P, and Wunsch C (1997): Accuracy assessment of recent ocean tide models. *Journal of Geophysical Research*, vol. 102, C11, pp. 173-194.
- Sideris MG (1994): Chapter 4: Regional geoid determination, In *Geophysical Interpretation of the geoid*, Vaniček P and Christou N (Eds.), CRC Press inc.
- Sideris MG, Mainville A, Forsberg R (1992): Geoid testing using GPS and levelling (or GPS testing using levelling and the geoid?). *Australian Journal of Geodesy, Photogrammetry and Surveying*, no. 57, pp. 62-77.
- Sideris MG and Schwarz KP (1987): Improvement of Medium and Short Wavelength Features of Geopotential Solutions by Local Gravity Data. *Bolletino di Geodesia e Scienze Affini*, no. 3, pp. 207-221.
- Sideris MG and Forsberg R (1991): Review of geoid prediction methods in mountainous regions. *International Association of Geodesy Symposia*, vol. 106, Springer-Verlag. Proceedings of the *Determination of the geoid – Present and future Symposium*, Milan, Italy, June 11-13, pp. 51-62.
- Sideris MG and Li Y (1992): Improved geoid determination for levelling by GPS. Presented at the 6th *International Geodetic Symposium on Satellite Positioning*, Columbus, Ohio, March 17-20.
- Sjöberg L (1983): Unbiased estimation of variance-covariance components in condition adjustment with unknowns - a MINQUE approach. *ZfV*, vol. 108, no. 9, pp. 382-387.
- Sjöberg L (1984): Non-negative Variance Component Estimation in the Gauss-Helmert Adjustment Model. *Manuscripta Geodaetica*, vol. 9, pp.247-280.
- Sjöberg L (1995): The best quadratic minimum biased non-negative definite estimator for an additive two variance component model. *Manuscripta Geodaetica*, vol. 20, pp. 139-144.

- Smith A (1999): Application of Satellite Altimetry for Global Ocean Tide Modelling. *Ph.D. Thesis*, Delft University Press, The Netherlands.
- Soler T (1998): A compedium of transformation formulas useful in GPS work. *Journal of Geodesy*, vol. 72, pp. 482-490.
- Soler T and Marshall J (2002): Rigorous transformation of variance-covariance matrices of GPS-derived coordinates and velocities. *GPS Solutions*, vol. 6, pp. 76-90.
- Strang G and Borre K (1997): Linear algebra, geodesy and GPS. Wellesley-Cambridge Press, Wellesley, USA.
- Strang van Hees G (1986): Precision of the geoid, computed from terrestrial gravity measurements. *Manuscripta Geodaetica*, vol. 11, pp. 1-14.
- Tiberius C and Kenselaar F (2003): Variance component estimation and precise GPS positioning: case study. *Journal of Surveying Engineering*, vol. 129, no. 1, pp. 11-18.
- Tscherning C (2001): Computation of spherical harmonic coefficients and their error estimates using least-squares collocation. *Journal of Geodesy*, vol. 75, pp. 12-18.
- Tscherning C, Arabelos D and Strykowski G (2000) The 1-cm geoid after GOCE. *International Association of Geodesy Symposia*, vol. 123, Sideris MG (Ed.), *Gravity, Geoid and Geodynamics 2000 Symposium*, Banff, Canada, July 31 - Aug. 4, pp. 267-270.
- Tscherning C, Radwin A, Tealeb A, Mahmoud S, El-Monum, A, Hassan R, El-Syaed, I and Saker K (2001): Local geoid determination combining gravity disturbances and GPS/levelling: a case study in the Lake Nasser area, Aswan, Egypt. *Journal of Geodesy*, vol. 75, pp. 343-348.
- Van Dam TM, Blewitt B, Heflin MB (1994): Atmospheric pressure loading effects on GPS coordinate determination. *Journal of Geophysical Research*, vol. 99, pp. 23939-23950.
- van Onselen K (1997): Quality Investigation of Vertical Datum Connection. Delft University of Technology, DEOS Report No. 97.3.
- Vaniček P (1991): Vertical Datum and NAVD88. *Surveying and Land Information Systems*, vol. 51, no. 2, pp. 83-86.
- Vaniček P, Castle RO, Balazs EI (1980): Geodetic leveling and its applications. *Reviews of Geophysics and Space Physics*, vol. 18, no. 2, pp. 505-524.

- Vaniček P and Krakiwsky EJ (1986): *Geodesy the Concepts*. North-Holland, Amsterdam.
- Vaniček P and Christou NT (1994): *Geoid and its Geophysical Interpretations*. CRC Press Inc.
- Véronneau M (1997): The GSD95 geoid model for Canada. *International Association of Geodesy Symposia*, vol. 117, Springer-Verlag. Proceedings of the *Gravity, Geoid and Marine Geodesy Symposium*, Tokyo, Japan, Sept. 30 - Oct. 5, pp. 573-580.
- Véronneau M (2002): The Canadian Gravimetric Geoid Model of 2000 (CGG2000), Internal Report, Geodetic Survey Division, Earth Sciences Sector, Natural Resources Canada, Ottawa, Canada.
- Véronneau M, Mainville A, Craymer M (2002): The GPS height Transformation (v2.0): An Ellipsoidal- CGVD28 Height Transformation for Use With GPS in Canada. Internal Report, Geodetic Survey Division, Earth Sciences Sector, Natural Resources Canada, Ottawa, Canada.
- Wang J (1999): Modelling and quality control for precise GPS and GLONASS satellite positioning. *PhD Thesis*, Curtin University of Technology, School of Spatial Sciences.
- Wesolowsky GO (1976): *Multiple regression and analysis of variance*. John Wiley and Sons, New York.
- Wichiencharoen C (1982): *The Indirect Effect on the Computation of Geoidal Undulations*. Department of Geodetic Science and Surveying, Ohio State University, Report No.336, Columbus, Ohio.
- Wolf H (1978): The Helmert block method, its origin and development. *Proceedings of the Second International Symposium on Problems Related to the Redefinition of the North American Geodetic Networks*, Arlington, VA, April 24-28, pp. 319-326.
- Yu Z (1992): A generalization theory of estimation of variance-covariance components. *Manuscripta Geodaetica*, vol. 17, pp. 295-301.
- Yu Z (1996): A universal formula of maximum likelihood estimation of variance-covariance components. *Journal of Geodesy*, vol. 70, pp. 233-240.
- Zilkoski DB, Richards JH, and Young GM (1992): Special Report: Results of the General Adjustment of the North American Vertical Datum of 1988. *Surveying and Land Information Systems*, vol. 52, no. 3, pp. 133-149.