Regional gravity field modelling with radial basis functions

Tobias Wittwer

Regional gravity field modelling
with radial basis functions

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op woensdag 9 december 2009 om 12.30 uur

door Tobias Florian WITTWER

Diplom-Ingenieur für Geodäsie und Geoinformatik
Universität Stuttgart, Duitsland

geboren te Ostfildern/Ruit, Duitsland

Dit proefschrift is goedgekeurd door de promotor:
Prof. Dr.-Ing. habil. R.A.P. Klees

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. Dr.-Ing. habil. R.A.P. Klees	Technische Universiteit Delft, promotor
Prof. dr. ir. R.F. Hanssen	Technische Universiteit Delft
Prof. ir. B.A.C. Ambrosius	Technische Universiteit Delft
Prof. Dr.-Ing. habil. J. Kusche	Universität Bonn
Prof. W. Featherstone	Curtin University of Technology
Prof. Dr. W. Keller	Universität Stuttgart
Prof. Dr.-Ing. habil. B. Heck	Universität Karlsruhe

Acknowledgements

Contents

Summary

Regional gravity field modelling with radial basis functions

Terrestrial gravimetry, airborne gravimetry, and the recent dedicated satellite gravity missions Challenging Minisatellite Payload (CHAMP), Gravity Recovery and Climate Experiment (GRACE), and Gravity and Ocean Circulation Explorer (GOCE) provide us with high-quality, high-resolution gravity data, which are used in many application areas such as

1. the computation of global static gravity fields, in support of precise orbit determination of many Earth observation satellites;

2. the quantification and interpretation of mass transport in the Earth system such as the shrinking of ice sheets, the shifting of ocean currents, and water storage variations;

3. the computation of high resolution regional and local gravity fields in support of height system realisation and the modelling of reservoirs and geophysical features.

Traditionally, for each data set (satellite, airborne, terrestrial) dedicated data processing schemes have been developed using different estimation principles, parametrisations, etc. The optimal combination of different data sets would benefit of a methodology that can be used for any type of data. Elements of this methodology comprise a uniform parametrisation, estimation principle, data weighting scheme, regularisation, and error propagation.

In the framework of this thesis, such a methodology is developed. It uses radial basis functions (RBFs) as parametrisation. They have parameters that allow us to tune their approximation properties as function of the data coverage and distribution and the signal variations. This makes them equally well suited for global and local parametrisation. Moreover, there exists an analytical relationship between a spherical harmonic representation and a radial basis function representation, which allows the latter to be transformed into the former, without any approximation error. Among others, this has the advantage that one can make use of existing processing tools, such as spectral analysis.

Although radial basis functions are not new in gravity field modelling, there are many important issues which have not yet been addressed or require further research. The

main research question underlying this thesis is: "Are radial basis functions a suitable parametrisation for global and regional models of the mean and time-variable gravity field, and if so, how do they perform compared with spherical harmonic solutions?" Directly related to this is the question: "Are there situations where radial basis functions models outperform spherical harmonic solutions?" The answer to both questions is positive as will be shown in this thesis.

There are two important aspects that determine the quality of a gravity field model based on radial basis functions: 1) the spatial distribution of the radial basis functions, i.e. the basis function network design, and 2) the choice of the bandwidths of the radial basis functions. For both problems, semi-automatic algorithms have been developed. Data-adaptive network design and local refinement avoid respectively over- and under-parametrisation by fine-tuning the basis function network based on the data. The basis function bandwidth is determined by optimising the fit to the data including control data.

The computation of regional gravity fields constitutes a considerable numerical workload, especially since the methodology presented here does not use an iterative normal equation solver (e.g., the preconditioned conjugate gradient method). Instead, a Cholesky solver is used, which requires the assembly of the complete normal equation system. For this purpose the program is numerically optimised and fully parallelised for hybrid high performance computer architectures. This guarantees optimal performance on all types of parallel computers and handles the memory requirements.

The modelling of satellite data with radial basis functions is investigated using real data of the GRACE satellites collected over the period 2003-2006. An optimal Wiener filter has been developed for radial basis functions in line with the optimal Wiener filter approach previously developed at DEOS for spherical harmonic representations. Monthly GRACE gravity models computed using radial basis function are compared to spherical harmonic models, and validated using independent data provided by the Ice Cloud and Land Elevation Satellite (ICESat), radar altimetry satellites, and the global hydrological model PCR-GLOBWB. Two applications were considered: 1) mass variations over Greenland and Antarctica and 2) water storage variations in river basins. The results show that the radial basis function approach yields solutions that are of at least the same quality as global models using spherical harmonics. There is evidence that radial basis functions may provide better spatial resolution and more realistic amplitudes in particular in high-latitude areas. For instance, it will be shown that radial basis function solutions detected signal that could not be seen in spherical harmonic solutions.

Two test areas are used for regional gravity field modelling using real terrestrial data: An area in the northeastern USA and a larger area in eastern Canada. The results show that the data-adaptivity and local refinement algorithms developed in the framework of this thesis provide good solutions of constant quality regardless of the initially chosen grid spacing. The models are compared to the official regional geoid models GEOID03 and CGG05, respectively. In both cases, rms errors of several centimetres remain, which are attributed to different input data and processing strategies.

The combination of satellite and terrestrial data is tested using simulated global and regional data sets. It is shown that a joint inversion of the two data sets yields combined solutions which are significantly better than a solution using the traditional remove-restore approach. The addition of satellite data with the corresponding stochastic model

compensates the reduced quality of the terrestrial data at long wavelengths.

The examples show that the regional modelling methodology presented here is a very flexible approach that can be applied to all types of gravity data and data distributions, regardless of application, data source, and area size. The quality of the solutions is at least equal to the solutions developed for the stand-alone inversion of individual data sets, while radial basis functions offer numerical benefits. As a result, this approach is already used for marine geoid modelling, and recommended for the modelling of airborne gravity data and data of the GOCE satellite, and for the joint inversion of satellite, airborne and ground-based gravity data.

Samenvatting

Regionale zwaartekrachtmodellering met radiale basisfuncties

Terrestrische gravimetrie, vliegtuiggravimetrie en de recente specifieke satelliet-zwaartekrachtmissies Challenging Minisatellite Payload (CHAMP), Gravity Recovery and Climate Experiment (GRACE) en Gravity and Ocean Circulation Explorer (GOCE) leveren data van hoge kwaliteit en hoge resolutie, die voor veel toepassingen gebruikt worden, zoals:

1. De berekening van mondiale statische zwaartekrachtvelden, ter ondersteuning van de nauwkeurige baanbepaling van vele aardobservatiesatellieten;

2. Het kwantificeren en de interpretatie van massatransport in het systeem aarde, zoals het afsmelten van de ijskappen, het verplaatsen van oceaanstromingen en variaties in wateropslag;

3. De berekening van regionale en lokale zwaartekrachtvelden met hoge resolutie, ter ondersteuning van de realisatie van hoogtesystemen en de modellering van reservoirs en geofysische kenmerken.

Voor elk data set (satelliet, vliegtuig, terrestrisch) zijn specifieke verwerkingsschema's ontwikkeld die gebruik maken van verschillende schattingsprincipes en parametrisaties etc. De optimale combinatie van verschillende data sets zou profiteren van een methodiek die voor elke data set gebruikt kan worden. Onderdelen van een dergelijke methodiek bevatten een uniforme parametrisatie, schattingsprincipe, methode van data weging en foutvoortplanting.

In het kader van dit proefschrift wordt een dergelijke methodiek ontwikkeld. Zij gebruikt radiale basisfuncties (RBFs) als parametrisatie.

Radiale basisfuncties hebben parameters die het mogelijk maken hun approximatie-eigenschappen aan te passen als functie van de bedekking en verdeling van de waarnemingen en de signaalvariatie. Dit maakt ze geschikt voor zowel mondiale als ook voor regionale parametrisaties. Verder bestaat een analytische relatie tussen een sferisch harmonische representatie en een representatie in radiale basisfuncties, die het mogelijk maakt de laatste, zonder approximatiefouten, in de eerste te transformeren. Dit heeft

onder andere het voordeel dat men gebruik kan maken van bestaande tools, zoals spectraalanalyse.

Ondanks het feit dat radiale basisfuncties niet nieuw zijn in de zwaartekrachtmodellering, zijn er vele belangrijke aspecten die tot nu toe niet behandeld zijn of die dieper onderzocht moeten worden. De hoofdonderzoeksvraag die aan de basis van dit proefschrift staat is: "Zijn radiale basisfuncties een geschikte parametrisatie voor mondiale en regionale modellen van het gemiddelde en tijdvariabele zwaartekrachtveld, en zo ja, hoe presteren zij vergeleken met sferisch harmonische oplossingen?" Direct gerelateerd hieraan is deze vraag: "Zijn er situaties waarin modellen die als parametrisatie radiale basisfunctie gebruiken, beter presteren dan sferisch harmonische oplossingen?" Het antwoord op beide vragen is positief, zoals in dit proefschrift aangetoond zal worden.

Er zijn twee belangrijke aspecten die de kwaliteit van een zwaartekrachtsveldmodel gebaseerd op radiale basisfuncties bepalen: 1) het ontwerp van het netwerk van basisfuncties 2) de keuze van de bandbreedtes van de radiale basisfuncties. Voor beide problemen werden semiautomatische algoritmen ontwikkeld. Een data-adaptief netwerkontwerp en locale verfijning voorkomen respectievelijk over- en onderparametrisatie door het netwerk van basisfuncties aan te passen op de verdeling van de waarnemingen. De bandbreedte van de basisfuncties wordt bepaald door optimalisatie van de data fit inclusief controledata.

De berekening van regionale zwaartekrachtvelden vereist een behoorlijke numerieke werklast Dit wordt versterkt doordat de hier gepresenteerde methode geen gebruik maakt van een iteratieve normaalvergelijkings-oplosser (bijv. de preconditioned conjugate gradient methode). In plaats daarvan wordt een Cholesky oplosser gebruikt, die de opbouw van het gehele systeem van normaalvergelijkingen vereist. Voor dit doel is het programma numeriek geoptimaliseerd en geheel geparallelliseerd voor hybride high-performance computer architecturen. Dit waarborgt optimale prestaties op alle soorten parallelle computers en voldoet aan de geheugeneisen.

Het modelleren van satellietdata met behulp van radiale basisfuncties wordt onderzocht met gebruik van echte data van de GRACE satellieten, verzameld in de periode 2003-2006. Een optimale Wiener filter werd ontwikkeld voor radiale basisfuncties in overeenstemming met de optimale Wiener filter aanpak eerder ontwikkeld bij DEOS voor sferisch harmonische representaties. Maandelijkse GRACE oplossingen, berekent met radiale basisfuncties, worden vergeleken met sferisch harmonische modellen, en gevalideerd met gebruik van externe data geleverd door de Ice Cloud and Land Elevation Satelliet (ICESat), radar-altimetrie satellieten en een mondiaal hydrologisch model. Twee toepassingen werden beschouwd: 1) massavariaties boven Groenland en Antarctica en 2) wateropslagvariates in rivierbasins. De resultaten tonen aan dat de aanpak met radiale basisfuncties oplossingen levert die van tenminste dezelfde kwaliteit zijn als mondiale modellen berekent met behulp van sferisch harmonischen. Er zijn indicaties dat radiale basisfuncties mogelijk een betere spatiële resolutie en realistischere amplitudes leveren, vooral in gebieden dicht bij de polen. Wij zullen bijv. aantonen dat de oplossingen met radiale basisfuncties een signaal gedetecteerd hebben wat niet zichtbaar is in een sferisch harmonische oplossing.

Twee testgebieden worden gebruikt voor regionale zwaartekrachtmodellering met echte terrestrische data: Een gebied in het noordoosten van de VS en een groter gebied in

het oosten van Canada. De resultaten tonen aan dat de algoritmen voor data-adaptief netwerkontwerp en lokale verfijning goede oplossingen leveren, onafhankelijk van de in eerste instantie gekozen gridafstand. De modellen worden vergeleken met de officiële geoïdes GEOID03 en CGG05. In beide gevallen was het kwadratisch gemiddelde van de afwijkingen meerdere centimetres, die aan verschillende data en verwerkingsmethodieken toegekend worden.

De combinatie van satelliet- en terrestrische data wordt getest met gesimuleerde mondiale en regionale data sets. Het wordt aangetoond dat een gecombineerde inversie van de twee data sets oplossingen levert die significant beter zijn dan een oplossing die gebruik maakt van de traditionele remove-restore aanpak. De toevoeging van satellietdata met het bijbehorende stochastische model compenseert de lagere kwaliteit van de terrestrische data in de lange golflengtes.

De voorbeelden laten zien dat de hier gepresenteerde methodiek voor regionale modellering een zeer flexibele aanpak is die voor alle soorten zwaartekrachtdata en dataverdelingen toegepast kan worden, ongeacht de toepassing, databron en gebiedsgrootte. De kwaliteit van de oplossingen is tenminste gelijk aan de oplossingen ontwikkeld voor stand-alone inversies van de individuele data sets, terwijl radiale basisfuncties numerieke voordelen bieden. Als resultaat hiervan wordt deze aanpak al gebruikt voor het modelleren van een marine geoïde, en wordt hij aanbevolen voor het modelleren van vliegtuiggravimetriedata en data van de GOCE satelliet, en voor de gecombineerde inversie van satelliet-, vliegtuig- en terrestrische zwaartekrachtdata.

Zusammenfassung

Regionale Schwerefeldmodellierung mit radialen Basisfunktionen

Terrestrische Gravimetrie, Flugzeuggravimetrie und die aktuellen Schwerefelds-Satellitenmissionen Challenging Minisatellite Payload (CHAMP), Gravity Recovery and Climate Experiment (GRACE) und Gravity and Ocean Circulation Explorer (GOCE) liefern uns Schweredaten von hoher Qualität und hoher Auflösung, die für viele Anwendungen benützt werden, zum Beispiel

1. die Berechnung von globalen statischen Schwerefeldern zur Unterstützung der genauen Bahnberechnung vieler Erdbeobachtungssatelliten;

2. die Quantifizierung und Interpretation von Massentransport im System Erde, wie das Schrumpfen von Eisfeldern, die Verschiebung von Ozeanströmungen, und Variationen von Wasserspeichern;

3. die Berechnung hochauflösender regionaler und lokaler Schwerefelder zur Realisierung von Höhensystemen und zur Modellierung von Rohstoffspeichern und geophysischen Merkmalen.

Traditionell wurde für jeden Datensatz (Satelliet, Flugzeug, terrestrisch) spezifische Prozessierungsmethoden entwickelt, unter Verwendungen verschiedener Schätzprinzipen, Parametrisierungen etc. Die optimale Kombination verschiedener Datensätze würde von einer Methodik profitieren, die mit allen Datentypen verwendet werden kann. Teile solch einer Methodik sind eine einheitliche Parametrisierung, Schätzprinzip, Datengewichtung, Regularisierung und Fehlerfortpflanzung.

Im Rahmen dieser Dissertation wird eine derartige Methodik entwickelt. Sie verwendet radiale Basisfunktionen (RBFs) als Parametrisierung.

Diese kennen Parameter die es möglich machen ihre Approximationseigenschaften auf Datenbedeckung, Datenverteilung und Signalvariationen abzustimmen. Dadurch sind sie in gleicher Weise geeignet für globale und lokale Parametrisierungen. Weiterhin besteht eine analytische Beziehung zwischen einer Darstellung in sphärisch Harmonischen und einer Darstellung in radialen Basisfunktionen, die es möglich macht letztere in erstere ohne Näherungsfehler zu transformieren. Unter anderem hat dies den Vorteil, dass existierende Prozessierungstools, z.B. Spektralanalyse, benützt werden können.

Obwohl radiale Basisfunktionen nicht neu sind in der Schwerefeldmodellierung, gibt es viele wichtige Themen die bisher nicht untersucht wurden oder weitere Forschung benötigen. Die Hauptfrage an der Basis dieser Dissertation ist: "Sind radiale Basisfunktionen eine geeignete Parametrisierung für globale und regionale Modelle des mittleren und zeitvariablen Schwerefelds, und wenn ja, wie gut funktionieren sie verglichen mit Lösungen in sphärisch Harmonischen?" Direkt verwandt hieran ist ist die Frage: "Gibt es Situationen in denen Modelle in radialen Basisfunktionen besser sind als Lösungen in sphärisch Harmonischen?" Die Antwort auf beide Fragen ist positiv, was in dieser Dissertation gezeigt wird.

Es gibt zwei wichtige Aspekte die die Qualität eines Schwerefeldmodells in radialen Basisfunktionen bestimmen: 1) die räumliche Verteilung der radialen Basisfunktionen, der Netzwerkentwurf und 2) die Wahl der Bandbreiten der radialen Basisfunktionen. Für beide Probleme wurden semiautomatische Algorithmen entwickelt. Datenadaptiver Netzwerkentwurf und lokale Verfeinerung verhindern Über- und Unterparametrisierung durch Feinabstimmung des Basisfunktionsnetzwerks auf Basis der Daten. Die Bandbreite der Basisfunktionen wird durch den Datenfit, einschließlich Kontrolldaten, bestimmt.

Die Berechnung regionaler Schwerefelder formt eine nennenswerte numerische Last, unter anderem weil die hier vorgestellte Methodik keinen iterativen Löser (z.B. die vorkonditionierte konjugierte Gradientenmethode) benützt. Stattdessen wird ein Cholesky-Löser verwendet, der das Aufstellen des kompletten Normalgleichungssystems erfordert. Für diesen Zweck ist das Programm numerisch optimiert und vollständig für hybride Hochleistungsrechnerarchitekturen parallelisiert. Dies garantiert optimale Leistung auf allen Typen Hochleistungsrechner und ausreichend Hauptspeicher.

Die Modellierung von Satellitendaten wird untersucht an Hand von echten Daten der GRACE Satelliten, die im Zeitraum 2003-2006 gesammelt wurden. Ein optimaler Wiener Filter wurde für die radialen Basisfunktionen entwickelt, übereinstimmend mit dem vorher bei DEOS für die Darstellung in sphärisch Harmonischen entwickelten Filter. Monatliche GRACE Lösungen werden verglichen mit Modellen in sphärisch Harmonischen, und validiert unter Verwendung von unabhängigen Daten vom Ice Cloud and Land Elevation Satellite (ICESat), Radaraltimetriesatelliten und einem globalen hydrologischen Modell. Zwei Anwendungen werden betrachtet: 1) Massenveränderungen über Grönland und der Antarktis und 2) Wasserspeicherschwankungen in Flussbecken. Die Ergebnisse zeigen, dass der Ansatz mit radialen Basisfunktionen mindestens die gleiche Qualität liefert wie globale Modelle mit Verwendung von sphärisch Harmonischen. Es gibt Hinweise, dass die radialen Basisfunktionen bessere räumliche Auflösung und realistischere Signalamplituden liefern, besonders in Regionen hoher geographischer Breite. Zum Beispiel wird gezeigt, dass die radialen Basisfunktionen Signal detektiert haben das in den Lösungen mit sphärisch Harmonischen nicht zu sehen ist.

Zwei Testgebiete werden für die regionale Schwerefeldmodellierung mit echten terrestrischen Daten verwendet: Ein Gebiet im Nordosten der USA und ein grösseres Gebiet im Osten von Kanada. Die Ergebnisse zeigen, dass die Algorithmen zur Datenadaption und lokalen Verfeinerung, die im Rahmen dieser Dissertation entwickelt wurden, unabhängig vom anfänglich gewählten Rasterabstand gute Lösungen von konstanter Qualität liefern. Die Modelle werden mit den offiziellen Geoiden GEOID03 und CGG05 verglichen. In beiden Fällen blieben gemittelte Fehler von mehreren Zentimetern, die

unterschiedlichen Eingangsdaten und Prozessierungsmethoden zugerechnet werden.

Die Kombination von Satellitendaten und terrestrischen Daten wird mit simulierten globalen und regionalen Datensätzen getestet. Es wird gezeigt, dass die kombinierte Inversion beider Datensätze kombinierte Lösungen liefert, die signifikant besser sind als eine Lösung die den traditionellen remove-restore Ansatz verwendet. Das Zufügen von Satellitendaten und des entsprechenden stochastischen Modells kompensiert die reduzierte Genauigkeit der terrestrischen Daten bei langen Wellenlängen.

Die Beispiele zeigen, dass die hier vorgestellte Methode zur regionalen Schwerefeldmodellierung ein flexibler Ansatz ist, der für alle Typen Schweredaten und Datenverteilungen verwendet werden kann, unabhängig von Anwendung, Datenquelle und Gebietsgröße. Die Qualität der Lösungen ist mindestens vergleichbar zu den Lösungen aus den Methoden zur eigenständigen Inversion des jeweiligen Datensatzes, wobei radiale Basisfunktionen numerische Vorteile bieten. Daraus resultiert, dass dieser Ansatz bereits für die Modellierung des Marineschwerefeldes verwendet wird, und empfohlen wird für die Modellierung von Flugzeuggravimetriedaten, Daten des GOCE Satelliten, und die gemeinsame Inversion von Satelliten-, Flugzeug- und terrestrischen Schweredaten.

Nomenclature

Acronyms

CHAMP   Challenging Minisatellite Payload
CG   Conjugate Gradient method
CRF   Celestial Reference Frame
CSR   Center for Space Research, University of Texas
DAND   Data-Adaptive Network Design
DEOS   Delft Institute for Earth Observation and Space Systems
DMT-1   DEOS Mass Transport model 1
EWH   Equivalent Water Height
FW   Frequency-dependent data Weighting
GCV   Generalised Cross Validation
GIA   Glacial Isostatic Adjustment
GOCE   Gravity field and steady-state Ocean Circulation Explorer
GPS   Global Positioning System
GRACE   Gravity field and Climate Explorer
ICESat   Ice, Cloud, and land Elevation Satellite
LoS   Line of Sight
LR   Local Refinement
MC   Monte Carlo simulation
PCCG   Pre-Conditioned Conjugate Gradient method
PGR   Post-Glacial Rebound
RBF   Radial Basis Function
SH   Spherical Harmonics
SHA   Spherical Harmonic Analysis
SHS   Spherical Harmonic Synthesis
TRF   Terrestrial Reference Frame
VCE   Variance Component Estimation

Notation and symbols

Scalars

$\bar{c}_{lm}$	spherical harmonic coefficient
G	gravitational constant
l	degree of spherical harmonics
m	order of spherical harmonics
M	mass of the Earth
n	order of a basis function
N	geoid undulation
r	spherical radius
r_j	partial redundancy
R	radius of the Bjerhammar sphere
R_E	mean Earth radius
t	time
T	disturbing potential
α	regularisation parameter
α_i	radial basis function coefficient
γ	grid level
δg	gravity disturbance
Δg	gravity anomaly
λ	spherical longitude
φ	spherical latitude
ψ_l	Legendre coefficient
ρ	inter-satellite range
σ^2	variance
σ_j^2	variance component
σ_l^2	degree variance
ϑ	spherical co-latitude
ζ	height anomaly

Functions

$P_l(x)$	Legendre polynomial
$\bar{P}_{lm}(x)$	fully normalised associated Legendre function
$\Psi_i(x)$	radial basis function

Matrices and vectors

A	design matrix
b	right-hand-side vector
$\mathbf{C}_x$	covariance matrix of vector **x**
D	signal covariance matrix
e	vector of residuals
$\mathbf{e}_i$	unit vector
I	identity matrix
N	normal matrix
P	weight matrix
r	inter-satellite range vector
r̈	inter-satellite acceleration vector
R	regularisation matrix
x	vector of unknown parameters
x̂	least-squares estimate
y	vector of observations
z	vector of random numbers

Operators and mathematical notations

$\mathbf{A}^{-1}$	inverse of **A**
$\mathbf{A}^{T}$	transpose of **A**
$E\{\cdot\}$	expectation
$D\{\cdot\}$	dispersion
$|\cdot|$	absolute value
$\|\cdot\|$	norm
$\hat{\cdot}$	normalised vector
$\bar{\cdot}$	mean value

Chapter 1

Introduction

1.1 Background

Geodesy (from Greek $\gamma\varepsilon\omega\delta\alpha\iota\sigma\iota\alpha$, "division of the Earth") is traditionally the science of measuring the size and shape of the Earth. This includes measuring the Earth's geometry as well as the gravity field and its variation in time. The Earth's gravity field is often represented by the geoid, a surface of equal gravity potential (figure 1.1). The geoid coincides with the mean surface of the sea at rest.
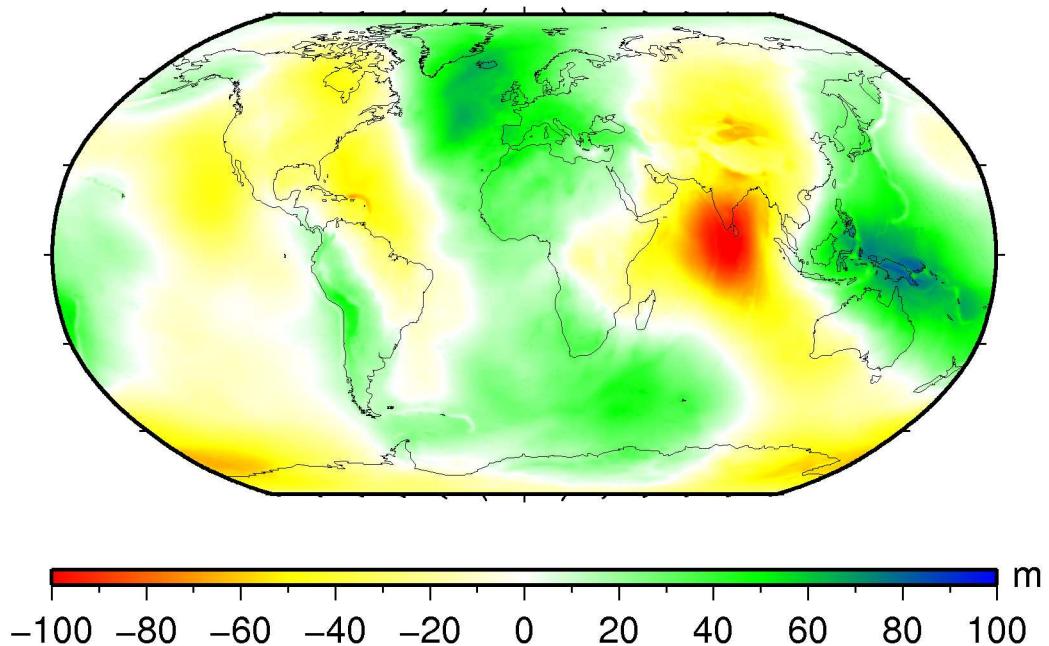
Figure 1.1: Geoid heights computed from EIGEN-GL04C up to degree 360 (Foerste et al., 2008).

Knowledge of the shape of the geoid is of major importance for many applications (Rummel et al., 2002). The geoid serves as a height reference surface; with the knowledge of the geoid, ellipsoidal heights measured by GPS can be transformed into gravity-related

heights and augment traditional levelling, a time-intensive technique with increasing error budget over long distances. Since the geoid represents a resting mean sea surface, it is required for all applications that need to determine dynamic sea surface topography from altimetry measurements. Precise orbit determination of satellites, particularly for low-orbiting satellites, also requires accurate gravity field models. The static gravity field reflects the internal structure of the Earth and can be used to constrain dynamic models of processes in the Earth's interior, and can help in the exploration of mineral and hydrocarbon reservoirs.

The Earth's gravity field is not constant in time. It changes as result of mass redistribution, ocean and atmospheric circulation, water flux between terrestrial water storages, ice melting, basin discharges into the oceans, and convective flow inside the Earth's mantle are mass transport processes that constantly redistribute mass in the atmosphere, on the Earth's surface, and inside the Earth. All these processes interact in various ways on differing time and spatial scales. This makes it difficult to develop realistic models for accurate predictions. The new gravity satellite missions and resulting gravity field models make it possible for the first time to monitor these mass transport processes on a global scale and with high temporal resolution. This allows us to investigate the causes and build a basis for the prediction of future changes. These insights are essential for the understanding of the global water cycle, and accurate predictions of the future effects of climate change.

1.2   Motivation

1.2.1   Regional modelling from satellite data

Satellites provide global or almost-global (in non-polar orbits) data coverage. Spherical harmonics, as global basis function, are usually used for modelling gravity fields from satellite data. While past satellite missions only allowed for the computation of models of the static gravity field, the recent GRACE (Gravity field and Climate Experiment) mission has made it possible to produce gravity field solutions for each month or even shorter time spans. This makes it possible for the first time to monitor mass variations worldwide (Tapley et al., 2004).

The two largest mass variation signals that can be detected by GRACE are the melting of polar ice caps and water storage variations. Both signals are analysed at the large basin scale, not globally.

GRACE does not provide the same data quality everywhere on the globe. Figure 1.2 shows the locations of GRACE A over the Northern Hemisphere in February 2006, at a 5-second interval (the sampling used for data collection). At higher latitudes, orbits are more tightly spaced. Whether this increased data sampling can be exploited to produce higher-quality gravity fields at higher latitudes than around the equator needs to be investigated.

Spherical harmonic (SH) solutions that try to capture all signal at higher latitudes will overparametrise at lower latitudes. This may lead to numerical instabilities and require

regularisation. Noise at lower latitudes will be modelled and have to be treated by smoothing, which introduces a bias by reducing the signal. A regional solution can be fine-tuned to the specific region of interest.
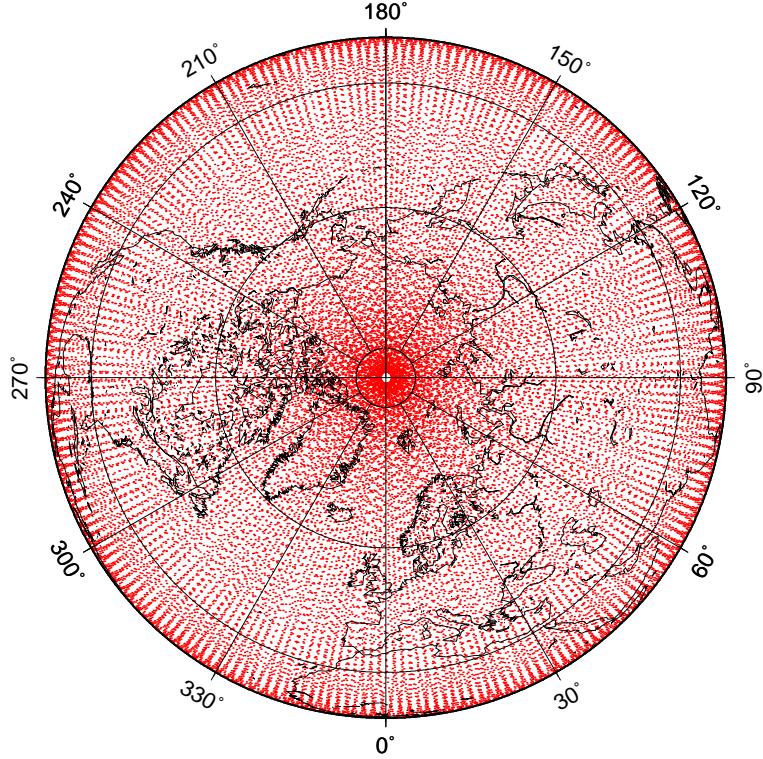
Figure 1.2: Orbit of GRACE A, February 2006, at 5s interval.

The GOCE mission, launched in March 2009, aims at generating data that makes it possible to compute accurate static gravity fields with at least twice the spatial resolution obtainable from GRACE. This increased resolution requires a similar increase in the number of gravity field parameters and leads to a corresponding increase in numerical complexity. Not everywhere will the same resolution be required - the gravity field is much smoother over the parts of the oceans, with strong gravity signals from trenches and sea mounts. This motivates the use of regional modelling and refinement techniques.

Using spherical harmonics for regional solutions, the numerical complexity will be much higher compared to a regional parametrisation. Achieving a resolution of 0.5 degrees requires a spherical harmonic representation up to degree $l_{max} = 360$, more than 130,000 spherical harmonic coefficients. The number of radial basis functions (RBFs) required to achieve the same resolution depends on the area size. For instance, 400 functions are sufficient for an area of 10 by 10 degrees.

1.2.2 Regional modelling from terrestrial data

Terrestrial data are collected for comparatively small regions, i.e. the territory of a country. Compared to satellite data, the data are spaced much tighter, and a gravity field of much higher resolution can be computed from them. Terrestrial data contain signals with

shortest wavelengths of a few kilometres, which would result in a maximum spherical harmonic degree necessary for modelling that is impossible to handle in a least-squares adjustment.

Terrestrial data are also inhomogeneously distributed. Data gaps may exist e.g. due to the inaccessibility of an area. Areas of rugged terrain (mountains) feature stronger signal gradients and require more tightly-spaced functions for a good approximation than areas of smooth topography. This requires adaptation of the basis function network and the basis functions' approximation characteristics to the data's spatial distribution and spectral signal content, something that can be achieved with RBFs.

1.2.3   Combined modelling of satellite and terrestrial data

Satellite data are accurate at long wavelengths, but limited in their spatial resolution. The resolution that can be obtained from terrestrial data is much higher, but the quality at long wavelengths is not as good as the quality of satellite data. A combination of both data sources is required for the computation of accurate high-resolution regional gravity fields.

Up to now, combination of these two data sources has almost exclusively been achieved by a remove-restore approach using Stokes's integral (Stokes, 1849; Heiskanen and Moritz, 1967). A global field is subtracted from the terrestrial data before the gravity field estimation to remove long-wavelength signal and reduce edge effects. This does however not take the stochasticity of either data set into account. This motivates research into better combination strategies.

1.2.4   Radial basis functions

While other techniques for regional gravity field modelling exist, e.g. techniques based on Stokes's integral (Stokes, 1849) and least-squares collocation (Moritz, 1980), a choice has been made to use RBFs as gravity field representation.

Both the spatial distribution and the approximation characteristics of RBFs can be adjusted. This makes it possible to use them for all kinds of data sets. Inhomogeneous data distributions and varying spectral data content can be addressed by carefully choosing the locations of the basis functions and their bandwidth.

An approach using RBFs offers numerical advantages. The number of basis functions required depends only the resolution that can be achieved with the data set and on the size of the problem area. The number of observations does not directly influence the number of basis functions.

Radial basis functions can be related to spherical harmonics. This makes it possible to transform a gravity field solution computed with RBFs into SH coefficients. Existing processing tools and techniques for spherical harmonics can thus still be used.

Overall, RBFs are a flexible presentation that can be used for all types of gravity field modelling. For this reason they have been chosen for the methodology described here.

1.3 Prior research on radial basis functions

The use of RBFs in gravity field modelling is not new. Over the years, several attempts to find a methodology have been made. So far, no "standard" methodology has been established, and none of the methodologies tested so far have covered the whole spectrum of gravity field modelling.

The use of point masses as an alternative representation to spherical harmonics can be traced back to Weightman (1965). The use of several spherical layers of point masses for gravity anomaly recovery has been described in Reilly and Herbrechtsmeier (1978). An approach for automatic positioning of point masses was developed by Barthelmes (1986). Point masses were used by Vermeer (1984, 1989, 1990) for regional gravity field modelling in Finland and the Baltic. With exceptions, such as Antunes et al. (2003), point masses have now been replaced by other types of RBFs.

Least-squares collocation (Moritz, 1980) uses the auto-covariance function of the data as kernel function. It has long been used for gravity field modelling (Forsberg and Kenyon, 1994; Marchenko et al., 2001). Challenges are the proper computation of the covariance function (Alberts, 2009) and the associated numerical complexity.

At Kaiserslautern University, Germany, the use of RBFs and spherical wavelets has been investigated for a long time, mostly focusing on multiscale representations. Recent applications to gravity field modelling where the computation of the multi-scale geopotential model SWITCH-03 from CHAMP data (Fengler et al., 2003), and the modelling of regional and temporal gravity variations from GRACE (Fengler et al., 2007). In the first case, so called $\mathscr{H}$- splines and the Abel-Poisson kernel were used, while the latter used cubic polynomials (CuP) as scaling functions.

Another approach to multi-scale wavelet methodology has been developed at DGFI, Munich, Germany. The methodology uses wavelets with Shannon and Blackman scaling functions. Recently, the method has been applied to regional computations of potential field from a combination of CHAMP and terrestrial data, as well as temporal variations from GRACE (Schmidt et al., 2007).

Marchenko (1998) and Marchenko et al. (2001) have focused on the use of radial multipoles for local gravity field modelling. The parameters of the radial multipoles are determined by the signal covariance function. This approach requires a very small number of basis functions in order to obtain a good solution, but finding the multipole parameters is very time consuming. It has been applied to airborne gravity data modelling (Marchenko et al., 2001), and compared to the approach presented here (Klees et al., 2005).

At Bonn University, Germany, a methodology using so-called harmonic splines as kernels in an otherwise conventional RBF approach has been developed. The kernels are fine-tuned to the signal characteristics. The methodology has been employed for a variety of applications, such as regional refinement of global satellite-derived gravity fields, and the combination of GRACE and altimetry data. A detailed description of the application of this methodology to the local refinement of static GRACE gravity models can be found in Eicker (2008).

1.4   Research objectives

The goal of the research presented here is the development of a unified methodology for regional gravity field modelling using RBFs. Unified implies that there is no limitation to a specific data source or application, instead the aim is the development of a methodology suitable for modelling of satellite data, terrestrial data, or a combination of both. It should be possible to include other data types, such as airborne gravity measurements. Radial basis functions are used because of their flexibility, adaptability, and relation to spherical harmonics. The unknown RBF coefficients are estimated by least squares.

To achieve this goal, it has been further subdivided into certain important aspects that have to be investigated.

Network design

To avoid numerical instabilities due to overparametrisation. and poor data approximation due to underparametrisation, the number and location of the RBFs needs to be governed by the data distribution. Network design involves 1) choosing a suitable grid type and grid spacing for basis function placement, and 2) algorithms for locally adapting and refining the RBF network to observation density and signal content (section 2.6).

Bandwidth estimation

An RBF's bandwidth describes its approximation characteristics. In order to achieve a good solution, an optimal bandwidth needs to be used for each basis function on a given network. An algorithm for estimation of the optimal basis function bandwidth out of the data or a reference model will be described in section 2.5.2.

Numerical efficiency

Regional gravity field modelling with RBFs routinely deals with hundreds of thousands of observations and tens of thousands of unknown gravity field parameters. These parameters are estimated by least-squares. The resulting numerical effort for the construction and solution of the linear equation system requires 1) an efficient numerical implementation and 2) parallelisation (chapter 4).

Modelling of satellite data

The main application of the methodology presented here is the computation of regional gravity fields using data collected by the GRACE satellite mission. The functional model and optimal filtering approach developed at the Delft Institute for Earth Observation and Space Systems (DEOS) are to be used. Network design and bandwidth estimation as they relate to satellite data are investigated. The approach is to be validated by comparing the results obtained with the RBF approach to other GRACE models and external data, using real-world applications, in chapter 5.

Modelling of terrestrial data

The second application is the computation of local gravity fields using terrestrial data. This requires proper bandwidth estimation and a network that is adapted to the heterogeneous data distribution, for which the data-adaptation and local refinement algorithms are used. The methodology is to be tested using real data sets and compared to official gravity fields for the computation areas (chapter 6).

Combined modelling of satellite and terrestrial data

Finally, strategies for the combination of satellite and terrestrial data beyond the remove-restore approach are to be investigated. This includes the suitable choice of RBF distribution and bandwidth, and the incorporation of the stochastic properties of both data sets. The strategies are tested in a controlled simulated environment in chapter 7.

1.5   Outline of thesis

This thesis is outlined as follows. Chapter 2 describes different types of RBFs, their behaviour in the spectral and the spatial domain, and how they relate to spherical harmonics. An overview of the use of RBFs for gravity field modelling is given. The three main choices that need to be made for gravity field modelling are explained: kernel choice, RBF centre location, and bandwidth. Chapter 2 includes a short description of an approach to multi-scale modelling.

The mathematical model and estimation principle used here are explained in chapter 3. Functional and stochastic model are described. Least-squares estimation, regularisation and solution strategies are mentioned, including the use of variance component estimation for data weighting and regularisation parameter estimation.

Chapter 4 addresses numerical issues. These include numerical optimisation and a fast synthesis scheme for the computation of the design matrix. The focus is on the description of the parallelisation of the program used here. The programming techniques used for parallelisation for hybrid high performance computer architectures are explained, and the resulting speed increase is shown.

Results of modelling data collected by the GRACE satellite mission with RBFs are given in chapter 5. The functional and stochastic model used are described. Included is a detailed explanation of the application of the optimal filter approach developed at DEOS to RBF solutions. The issues of network design and bandwidth as they apply to GRACE modelling are also covered. The results focus on two applications: the recovery of ice mass loss in Greenland and Antarctica, and the monitoring of hydrological signals. The former includes a comparison with ICESat, while the latter uses a hydrological model and satellite radar altimetry for evaluation.

Local gravity field modelling, using terrestrial data as input, is covered in chapter 6. This once again includes the description of functional model, network design, and bandwidth

choice. Results are shown for two test areas and data sets: A 10° x 10° area in the northeastern USA, and a larger area in eastern Canada. In both cases, a comparison to the official geoid models for these areas was made.

Chapter 7 describes strategies for the combination of satellite and terrestrial data for the joint computation of an accurate high-resolution regional gravity field model. A global and a regional setting using simulated data are used for the investigation of the possible combination strategies.

This thesis concludes with chapter 8, in which the research presented here is briefly summarised. Recommendations are given concerning problems that still need to be investigated and possible future research topics and related applications.

Chapter 2

Radial basis functions

This chapter introduces RBFs and discusses issued that are important for their use for gravity field modelling. First, the properties of RBFs both in the spectral and spatial domain are described and the relation to spherical harmonics is established.

Three attributes of RBFs need to be addressed in the context of gravity field modelling. These items, kernel choice (section 2.5.1), bandwidth selection (section 2.5.2), and RBF centre location (section 2.6) are described in detail. The chapter concludes with a link to multi-scale modelling.

2.1   Gravity field representations

2.1.1   Spherical harmonics

The disturbing potential is usually represented using spherical harmonics, essentially a Fourier expansion on the sphere:

$$
T\left(\mathbf{x}\right) = \frac{GM}{R_E} \sum_{l=0}^{\infty} \left(\frac{R_E}{|\mathbf{x}|}\right)^{l+1} \sum_{m=-l}^{l} \bar{c}_{lm} \bar{P}_{lm}\left(\sin\varphi\right). \tag{2.1}
$$

where $\mathbf{x}$ is a point with the spherical coordinates λ, φ, r. $\bar{P}_{lm}\left(\sin\varphi\right)$ are the fully normalised associated Legendre functions of degree l and order m. G is the gravitational constant, M is the mass of the Earth and R_E its mean equatorial radius. $\bar{c}_{lm}$ are the fully normalised spherical harmonic geopotential coefficients.

Spherical harmonics are global basis functions, as is evident from figure 2.1, which shows a spherical harmonic function of degree 10 and order 5. Spherical harmonics provide, as the name implies, a harmonic set of functions to globally describe a certain field, in our case the gravity potential. A full spherical harmonic representation of degree l_{max} makes use of $\left(l_{max}\right)^2 + 1$ functions.

Due to their global nature, spherical harmonics are a natural choice for global methods, such as computing global gravity fields from satellite data. The spectral representation
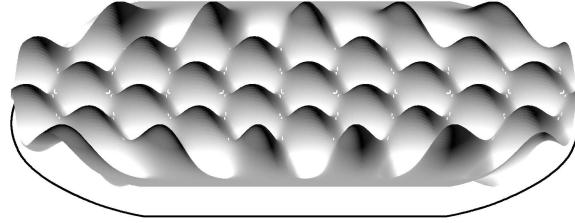
Figure 2.1: Spherical harmonic of degree 10 and order 5 in the spatial domain.
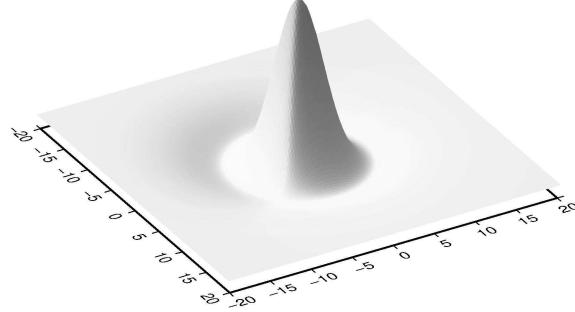
Figure 2.2: A radial basis function in the spatial domain.

makes it easy to apply filters in the frequency domain and to analyse the spectral be-
haviour of the signal. Translation into the spatial domain is accomplished by spherical
harmonic synthesis (SHS).

2.1.2   Radial basis functions

Radial basis functions, also named spherical basis functions (e.g. Narcowich and Ward
(1996)), spherical radial basis functions, localising basis functions, are radial-symmetric
functions which are localising in space (figure 2.2), i.e. most of their energy is confined
to a local area. They have either global or local support. All basis functions described
here have global support and do not differ from spherical harmonics in that respect.
Wendland (1995, 2005) has proposed locally supported basis functions which are often
used in science and engineering. This leads to a reduced numerical complexity due
to sparse linear equations systems. They may however not be used for gravity field
modelling because the trace of a harmonic function on the sphere can never have local
support.

σ_R denotes the surface of the sphere of radius R, which is located completely inside the
topographic masses (Bjerhammar sphere), $\sigma_R = \left\{ (x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = R^2 \right\}$. Int
σ_R denotes the interior and Ext σ_R denotes the exterior. Two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, $\mathbf{y} \neq 0$,
with $\mathbf{x} = (x_1, x_2, x_3)^T \in$ Ext σ_R and $\mathbf{y} = (y_1, y_2, y_3)^T \in$ Int σ_R with unit vectors $\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|}$,
$\hat{\mathbf{y}} = \frac{\mathbf{y}}{|\mathbf{y}|}$ are considered. The exterior RBF at centre location $\mathbf{y}_i$ evaluated at $\mathbf{x}$ is defined
through

$$\Psi_i(\mathbf{x}, \mathbf{y}_i) = \sum_{l=0}^{\infty} \psi_l \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} P_l \left(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i \right), \tag{2.2}$$

with ψ_l being the Legendre coefficients of the basis function kernel, and P_l the Legendre polynomial of degree l. The kernel determines both the spectral and spatial characteristics of the basis function. The exterior RBF is a harmonic function in Ext σ_R. It is restricted to the sphere rotationally symmetric around the $\hat{\mathbf{y}}$ axis.

A function T harmonic outside a sphere of radius R can be written as an infinite sum of RBFs according to

$$T(\mathbf{x}) = \frac{GM}{R_E} \sum_{i=1}^{\infty} \alpha_i \Psi_i(\mathbf{x}, \mathbf{y}_i), \tag{2.3}$$

where α_i are the basis function coefficients.

In practice, only a finite number of RBFs is being used to represent the disturbing potential T. The unknown RBF coefficients α_i are estimated by least-squares as described in chapter 3.

2.2 RBF types and behaviour in the spectral domain

The behavior of the basis functions in the spectral domain is determined by the choice of the basis function kernel and its Legendre coefficients ψ_l. Many different kernels have been employed for RBFs. This section presents some of the most popular kernels used so far in gravity field modelling and their spectrum.

Point mass kernel

The point mass kernel has long been used in gravity field modelling (Heikkinen, 1981; Vermeer, 1984, 1989, 1990; Barthelmes, 1986). Its analytical representation is

$$\Psi(\mathbf{x}, \mathbf{y}) = \frac{1}{|\mathbf{x} - \mathbf{y}|}. \tag{2.4}$$

When using the expansion of the reciprocal distance in spherical harmonics, it can easily be shown that the point mass kernel's Legendre coefficients are (Klees et al., 2007)

$$\psi_l = \frac{|\mathbf{y}|^l}{R^{l+1}}. \tag{2.5}$$

The spectral behaviour of the kernel is only dependent on $|\mathbf{y}|$, the radial distance of the basis function from the centre of the Bjerhammar sphere, which is equivalent to

$$d = R - |\mathbf{y}|, \tag{2.6}$$

its depth below the Bjerhammar sphere. Figure 2.3 shows the resulting spectrum for three different depths. It can be seen that the point mass kernel acts as low-pass filter. Placing the basis functions at shallower depths increases the kernel's sensitivity for higher degrees of spherical harmonics.
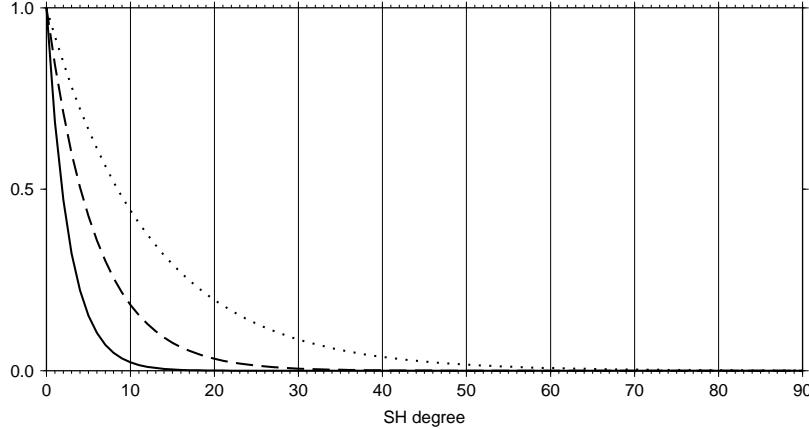
Figure 2.3: Normalised spectrum of point mass kernel of depth 2000 km (solid), 1000 km (dashed), and 500 km (dotted).

Radial multipole

The radial multipole has been used by Marchenko et al. (2001) for regional gravity field modelling. Its analytical representation is

$$\Psi(\mathbf{x}, \mathbf{y}) = \frac{1}{n!} \left(\frac{\partial}{\partial |\mathbf{y}|} \right)^n \frac{1}{|\mathbf{x} - \mathbf{y}|}. \tag{2.7}$$

Its Legendre coefficients are given as (Klees et al., 2007)

$$\psi_l = \binom{l}{n} \frac{|\mathbf{y}|^{l-n}}{4\pi R^2}. \tag{2.8}$$

There are two different parameters that determine the shape of the radial multipole: The radial distance $|\mathbf{y}|$ and related depth, and the order n. The resulting spectra for three different depths and order $n = 3$ are shown in figure 2.4. The radial multipole is not a truly band-limited function (the Legendre coefficients are never 0), but it has the characteristics of a band-pass filter. The shallower the depth, the higher the frequencies covered (and the wider the spectrum).

The spectra for a fixed depth of 1000 km and various orders n are shown in figure 2.6.

Poisson kernel

The analytical representation of the Poisson kernel (Klees et al., 2008a) is given as

$$\Psi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi R} \frac{|\mathbf{x}|^2 - |\mathbf{y}|^2}{|\mathbf{x} - \mathbf{y}|^3}. \tag{2.9}$$

The corresponding Legendre coefficients are (Klees et al., 2007)

$$\psi_l = \frac{(2l + 1)}{4\pi R^3} |\mathbf{y}|^l. \tag{2.10}$$
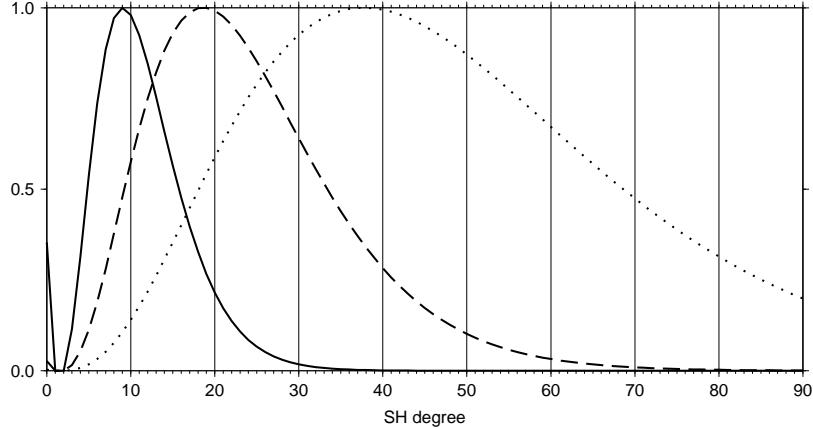
Figure 2.4: Normalised spectrum of radial multipole of order 3 and depth 2000 km (solid), 1000 km (dashed), and 500 km (dotted).
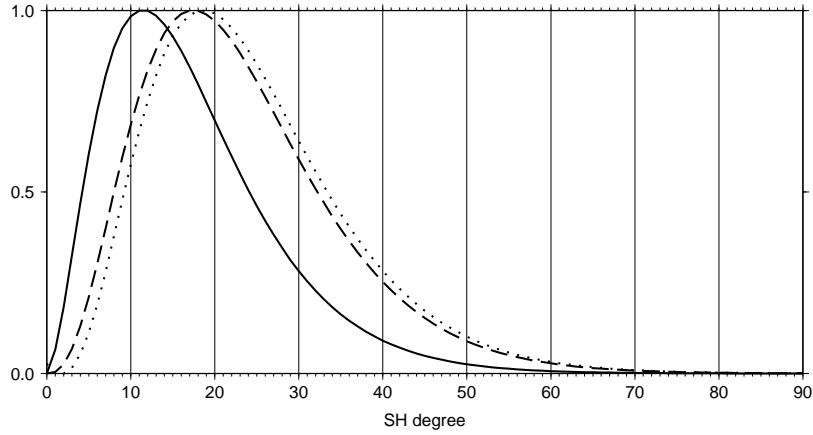
Figure 2.5: Normalised spectrum of radial multipole of depth 1000 km and order 1 (solid), order 2 (dashed), and order 3 (dotted).

The resulting spectrum (figure 2.6) also exhibits band-pass characteristics. At the same depth, the maximum of the Poisson kernel is shifted more towards the lower degrees compared to the radial multipole of order 3.

Poisson wavelet

The Poisson "wavelet" was introduced by Holschneider et al. (2003). A slightly modified (removal of terms that are eliminated by normalisation) but equivalent definition used in (Klees et al., 2008a) is

$$\psi_l = \frac{2l+1}{4\pi R^3} l^n |\mathbf{y}|^l, \tag{2.11}$$

with the radial distance $|\mathbf{y}|$ and the order n as free parameters. The spectrum for $n = 3$ and three different depths is given in figure 2.7. It is quite similar to the spectrum of the radial multipole of the same order and depth. With $n = 0$, the Poisson kernel is obtained. The Poisson wavelet also has band-pass filter characteristics.
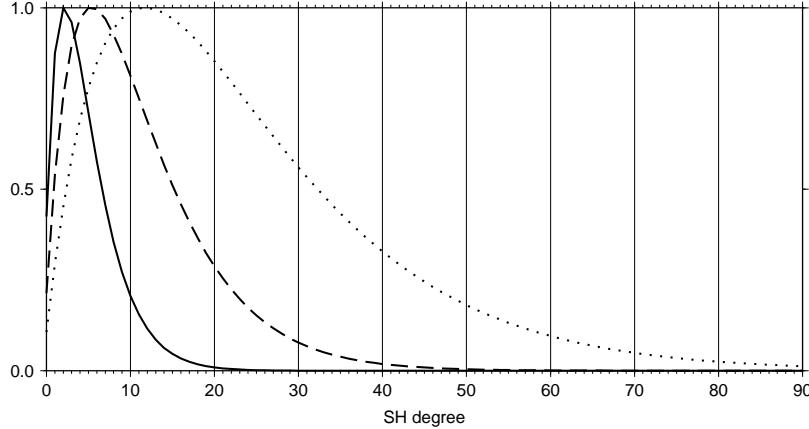
Figure 2.6: Normalised spectrum of Poisson kernel of depth 2000 km (solid), 1000 km (dashed), and 500 km (dotted).
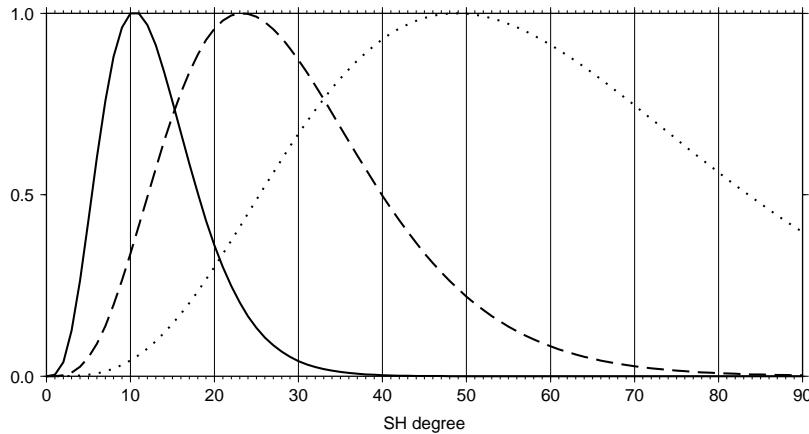
Figure 2.7: Normalised spectrum of Poisson wavelets of order 3 and depth 2000 km (solid), 1000 km (dashed), and 500 km (dotted).

The spectrum for a fixed depth of 1000 km and three different orders n is shown in figure 2.8. Changing the order n has a similar effect to changing the depth, but higher-order Poisson wavelets exhibit stronger gradients than lower-order Poisson wavelets of similar bandwidth (compare the order 3, depth 2000 km curve in figure 2.7 to the order 1, depth 1000 km curve in figure 2.8).

The corresponding analytical representation has been derived as (Klees et al., 2007)

$$\Psi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi R^2} \left(2\chi_{n+1} + \chi_n\right),$$
$$\chi_n = \left(|\mathbf{y}| \frac{\partial}{\partial |\mathbf{y}|}\right)^n \frac{1}{|\mathbf{x} - \mathbf{y}|}. \tag{2.12}$$

Since the Poisson wavelet is the type of kernel used for all computations described here, the analytical kernel will be covered in more detail.

Klees et al. (2007) have shown that the functions χ_n can be computed recursively by

$$\chi_n = 2|\mathbf{y}|^{n+1} b_{n+1} + \sum_{i=1}^{n} \beta_{n,i} |\mathbf{y}|^i b_i \tag{2.13}$$
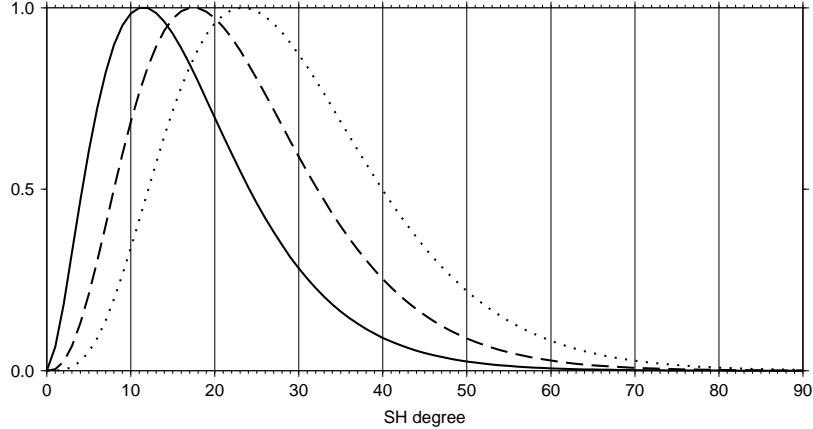
Figure 2.8: Normalised spectrum of Poisson wavelets of depth 1000 km and order 1 (solid), 2 (dashed), and 3 (dotted).

for $n \geq 1$. For $n = 0$,

$$\chi_0 = 2\,|\mathbf{y}|\,b_1 + b_0 \tag{2.14}$$

holds.

The b_n have been derived as (Klees et al., 2007)

$$b_n = (2n - 1)\,|\mathbf{x} - \mathbf{y}|\,b_1 b_{n-1} - (n-1)^2 b_0^2 b_{n-2}, \tag{2.15}$$

which can be recursively computed with

$$b_0 = \frac{1}{|\mathbf{x} - \mathbf{y}|}, \; b_1 = \frac{|\mathbf{y}| - |\mathbf{x}|\cos\vartheta}{|\mathbf{x} - \mathbf{y}|^3}, \tag{2.16}$$

where ϑ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

The factors $\beta_{3,i}$, $i = 1, 2, 3$ are $\beta_{3,1} = 3$, $\beta_{3,2} = 17$, and $\beta_{3,3} = 13$.

Blackman wavelets

The Blackman wavelets have been used for multi-scale global and regional gravity field modelling, e.g. Schmidt et al. (2007). They are defined as

$$\psi_l = \begin{cases} 1 - \left(0.42 - 0.5\cos\frac{\pi l}{l_1} + 0.08\cos\frac{2\pi l}{l_1}\right) & l = l_1 \ldots 2l_1 - 1 \\ 1 & l = 2l_1 \ldots l_2 - 1 \\ 0.42 - 0.5\cos\frac{\pi l}{l_2} + 0.08\cos\frac{2\pi l}{l_2} & l = l_2 \ldots 2l_2 - 1 \\ 0 & \text{elsewhere} \end{cases}, \tag{2.17}$$

with two free parameters l_1 and l_2, which determine the bandwidth. Between degrees l_1 and l_2, the Blackman wavelets are larger than 0. They are equal to 1 between $2l_1$ and

$l_2 - 1$. They gradually increase between l_1 and $2l_1$, and decrease between l_2 and $2l_2 - 1$. Hence, the Blackman wavelets are a bandlimited function with smooth cut-off (figure 2.9). Sharp cut-offs are undesirable, as they result in oscillations in the spatial domain (ringing or Gibbs effect).
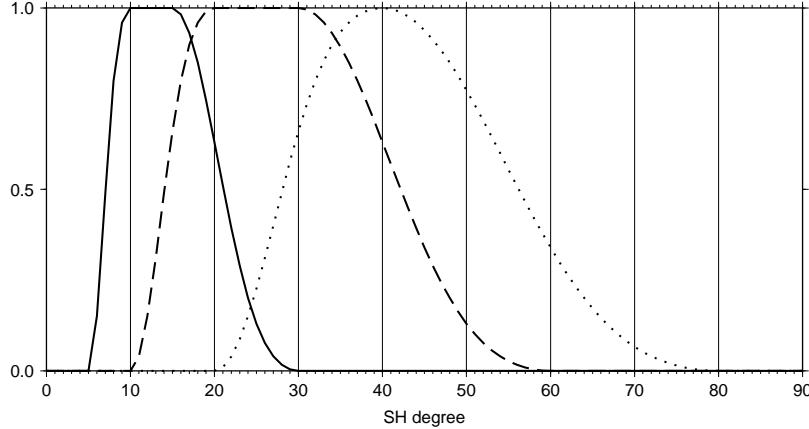
Figure 2.9: Normalised spectrum of Blackman wavelets with $l_1 = 5$, $l_2 = 15$ (solid), $l_1 = 10$, $l_2 = 30$ (dashed), and $l_1 = 20$, $l_2 = 60$ (dotted).

Harmonic splines

The Harmonic splines, as used by Eicker (2008), use the Legendre coefficients

$$\psi_l = \sigma_l^2, \tag{2.18}$$

where σ_l^2 are the SH degree variances of the gravity field to be modelled. Since the true degree variances are not known, the expected frequency behaviour has to be used. This can be derived from an existing gravity field model, or using an existing model of the spectral behaviour such as Kaula's rule of thumb:

$$\psi_l = (2l+1)\frac{10^{10}}{l^4}. \tag{2.19}$$

Kaula's rule of thumb is shown in figure 2.10. When a reference model of degree l_{max} is subtracted from the input data, the Legendre coefficients are modified to ensure that they match the expected signal behaviour. This is achieved either by also subtracting the reference model's signal degree variances from the coefficients,

$$\psi_l = \begin{cases} \sigma_l^2 - \sigma_{l,Ref}^2, & l \leq l_{max} \\ \sigma_l^2, & l > l_{max} \end{cases}, \tag{2.20}$$

or by replacing the degree variances in the part covered by the reference model by the reference model's error degree variances,

$$\psi_l = \begin{cases} \sigma_{e,l,Ref}^2, & l \leq l_{max} \\ \sigma_l^2, & l > l_{max} \end{cases}. \tag{2.21}$$
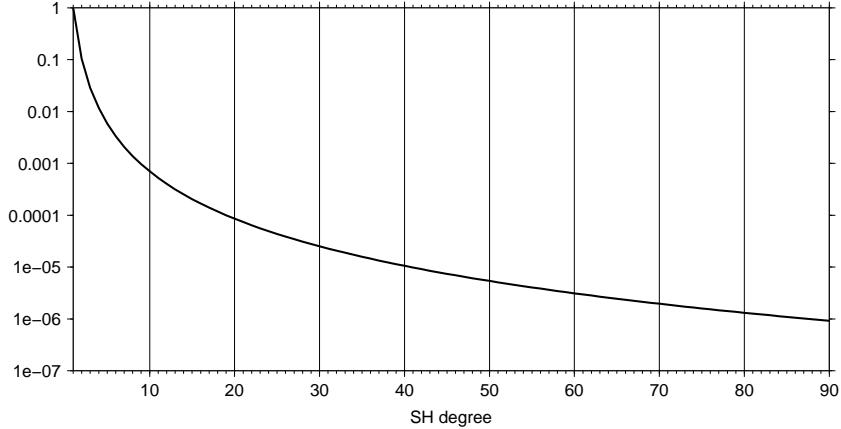
Figure 2.10: Spectrum of a harmonic spline when Kaula's rule of thumb is used.

2.3   Behaviour in the spatial domain

The behaviour of RBFs in the spatial domain is determined by the choice of the kernel and its Legendre coefficents ψ_l. The more power at the lower degrees the wider the basis function is. Figure 2.11 shows Poisson wavelets of order and at depths of 2000km, 1000 km, and 500 km in the space domain, the associated spectrum can be seen in figure 2.7.
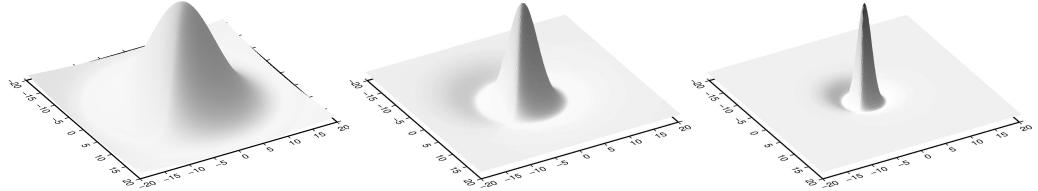
Figure 2.11: Poisson wavelet of order 3 and of depth 2000km (left), 1000km (centre), and 500km (right) in the spatial domain.

The spectrum plots already showed that different kernels behave differently in the frequency domain. This behaviour propagates into the space domain, as is evident from figure 2.12, which shows the six kernels described in section 2.2 in the space domain. We can detect the characteristics that were already visible in the spectrum. At the same depth of 500 km, the point mass kernel and the Poisson kernel concentrate more energy at lower frequencies than the radial multipoles and the Poisson wavelets of order 3. As a result, the point mass and Poisson kernel have a much wider support in the spatial domain. The radial multipole's and Poisson wavelet's similar spectra lead to similar characteristics in space, too. It should be noted that the point mass kernel and Poisson kernel are always positive.

The Blackman wavelet's band-pass characteristics and relatively sharp cut-off in the frequency domain causes strong oscillations in the spatial domain. These oscillations can be reduced by spacing the parameters l_1 and l_2 further apart. The Harmonic spline's behaviour in space is governed by the chosen degree variance model.

When computing RBFs using their series expansion, it is important not to truncate the summation too early. Figure 2.13 shows the spatial plots of Poisson wavelets with a
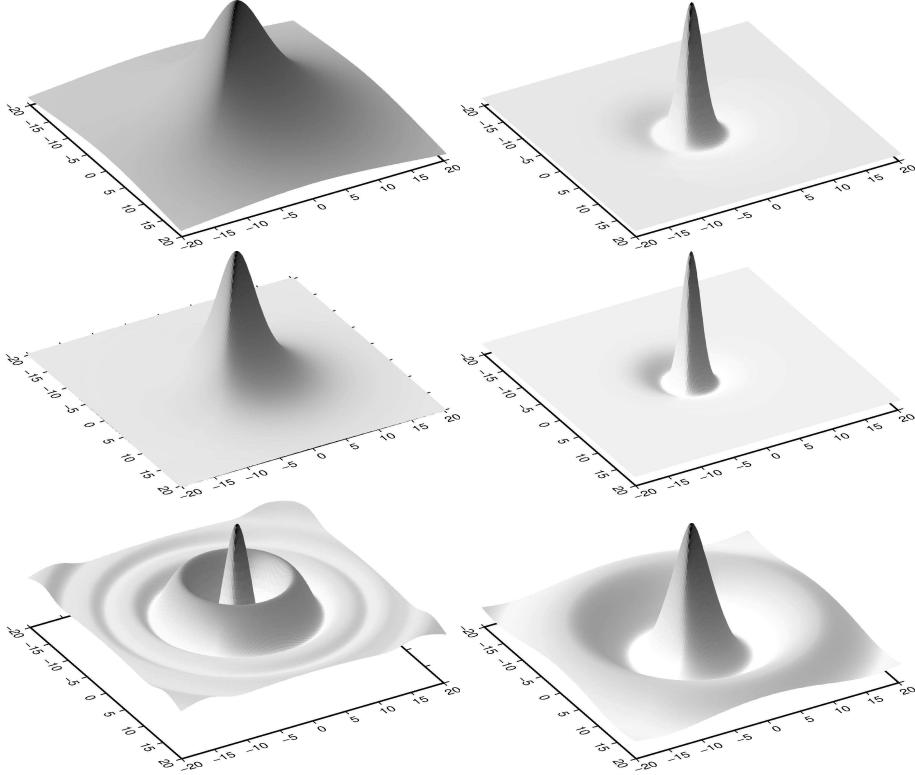
Figure 2.12: Kernels in the space domain. From left to right and top to bottom: Point mass kernel at 500 km depth; radial multipole of order 3 at 500 km depth; Poisson kernel at 500 km depth; Poisson wavelet of order 3 at 500 km depth; Blackman wavelet, $l_1 = 20$, $l_2 = 60$; harmonic spline using Kaula's rule of thumb between degree 20 and 200 ($\psi_l = 0$ for $l < 20$ and $l > 200$).

depth of 500 km. The series expansion has been truncated at degree 300 and degree 30, respectively. The basis function truncated at degree 30 is not only much wider, but also oscillates quite strongly. This will lead to numerical instabilities. Where the summation can be truncated depends on the bandwidth of the basis function - the wider the bandwidth, the earlier the series can be truncated.

For shallow basis functions (small bandwidth), the series will have to be evaluated to a very high degree, with considerable numerical effort. One way to avoid this is the use of analytical expressions for the basis functions (see section 2.2). If no analytical expression is available, a fast synthesis scheme can be implemented (see section 4.2).

2.4   Relation of RBFs to a spherical harmonic representation

Suppose we have a representation of a harmonic function by a finite set of RBFs and want to convert it into an equivalent spherical harmonic representation. A potential field is represented by RBFs according to
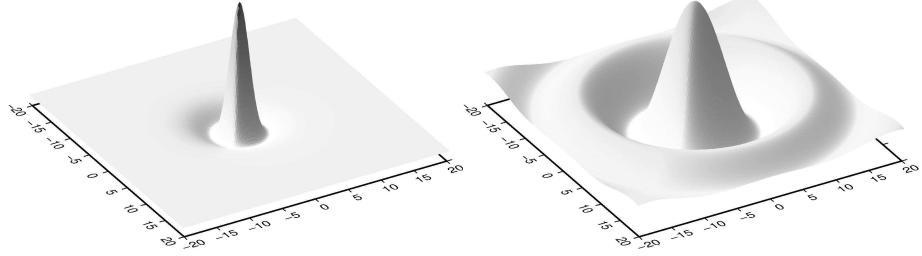
Figure 2.13: Poisson wavelet of depth 500 km truncated at degree 300 (left) and 30 (right) in the spatial domain.

$$T(\mathbf{x}) = \frac{GM}{R_E} \sum_{i=1}^{n} \alpha_i \Psi_i (\mathbf{x}, \mathbf{y}_i), \qquad (2.22)$$

where α_i is the coefficient of basis function i, and Ψ_i as in eq. (2.2).

Using the addition theorem for Legendre functions (e.g. Heiskanen and Moritz, 1967),

$$P_l \left(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i \right) = \frac{4\pi}{2l+1} \sum_{m=0}^{l} \bar{P}_{lm} (\hat{\mathbf{x}}) \bar{P}_{lm} (\hat{\mathbf{y}}_i), \qquad (2.23)$$

eq. (2.22) can be written as

$$
\begin{aligned}
T(\mathbf{x}) &= \frac{GM}{R_E} \sum_i \alpha_i \sum_l \psi_l \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} \frac{4\pi}{2l+1} \sum_m \bar{P}_{lm} (\hat{\mathbf{x}}) \bar{P}_{lm} (\hat{\mathbf{y}}_i) \\
&= \frac{GM}{R_E} \sum_l \sum_m \psi_l \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} \frac{4\pi}{2l+1} \sum_i \alpha_i \bar{P}_{lm} (\hat{\mathbf{y}}_i) \bar{P}_{lm} (\hat{\mathbf{x}}).
\end{aligned}
\qquad (2.24)
$$

A representation in spherical harmonics is given by

$$T(\mathbf{x}) = \frac{GM}{R_E} \sum_{l=0}^{n} \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} \sum_{m=-l}^{l} \bar{c}_{lm} \bar{P}_{lm} (\hat{\mathbf{x}}). \qquad (2.25)$$

When comparing eq. (2.25) to eq. (2.24), it can be seen that

$$\bar{c}_{lm} = \psi_l \frac{4\pi}{2l+1} \sum_{i=1}^{n} \alpha_i \bar{P}_{lm} (\hat{\mathbf{y}}_i). \qquad (2.26)$$

Eq. (2.26) allows the computation of spherical harmonic coefficients out of a given representation in RBFs. Eq. (2.26) can also be written in matrix-vector form as

$$\mathbf{x}_{SH} = \mathbf{A}\mathbf{x}_{RBF}, \qquad (2.27)$$

where

$$\mathbf{x}_{SH} = \begin{pmatrix} \bar{c}_{00} \\ \vdots \\ \bar{c}_{l_{max}l_{max}} \end{pmatrix}, \tag{2.28}$$

$$\mathbf{x}_{RBF} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \tag{2.29}$$

$$\mathbf{A} = \begin{pmatrix} \psi_0 \bar{P}_{00}(\mathbf{y}_1) & \cdots & \psi_0 \bar{P}_{00}(\mathbf{y}_n) \\ \vdots & \ddots & \vdots \\ \psi_{l_{max}} \frac{4\pi}{2l_{max}+1} \bar{P}_{l_{max}l_{max}}(\mathbf{y}_1) & \cdots & \psi_{l_{max}} \frac{4\pi}{2l_{max}+1} \bar{P}_{l_{max}l_{max}}(\mathbf{y}_n) \end{pmatrix}. \tag{2.30}$$

If $\mathbf{C}_{\mathbf{x}_{RBF}}$ is the covariance matrix of the estimated RBF coefficients, error propagation is then achieved by computing

$$\mathbf{C}_{\mathbf{x}_{SH}} = \mathbf{A}\mathbf{C}_{\mathbf{x}_{RBF}}\mathbf{A}^T, \tag{2.31}$$

where $\mathbf{C}_{\mathbf{x}_{SH}}$ is the resulting covariance matrix of the spherical harmonic coefficients.

The ability to transform an RBF solution into an equivalent set of spherical harmonic coefficients is useful for two applications:

1. It allows for the computation of the power spectrum of a potential field in RBFs and spectral analysis.

2. Since spherical harmonics are the most-used representation, many post-processing tools, such as spectral analysis or spectral filters, expect SH coefficients as input. The transformation makes it possible to use these tools. However, a very high maximum degree l may be necessary to get an accurate representation of $\mathbf{x}_{RBF}$ in terms of spherical harmonics.

2.5   Choice of RBF characteristics

For gravity field modelling, three choices with regard to the RBFs used need to be made in order to achieve a good approximation quality.

1. The choice of the kernel. Section 2.2 has introduced a number of kernels that might be used for gravity field modelling. Of course, in practice there may be many more.

2. The bandwidth of the RBFs. We have seen that the choice of bandwidth of the RBFs affects their spectral and spatial characteristics and thus their approximation properties. These must be chosen appropriately in order to obtain a good potential field approximation.

3. The location of the basis function centres (network design). Unlike spherical harmonics, which have global support, RBFs are localising in space. As a result, the locations of the RBF centres need to be selected appropriately.

2.5.1 Choice of the kernel

The first choice that needs to be made when RBFs are used in gravity field modelling is the kernel that is to be used. Section 2.2 has listed a number of possible kernels for regional gravity field modelling using RBFs and their associated spectral and spatial characteristics. A comparison conducted using the point mass kernel, the radial multipole, the Poisson kernel, and the Poisson wavelet has shown that, for the test data set, all four kernels yield similar results if the bandwidth is chosen correctly (Tenzer and Klees, 2008). This means that if a bandwidth selection scheme as described in section 2.5.2 is implemented, any one of the kernels can be used. Additionally, the four kernels used in the comparison have a desirable analytical expression, which can significantly reduce the computational load. Whether there are other RBFs that are better suited for potential field modelling has not been investigated.

The Blackman wavelets are easily fine-tuned to a certain spectral bands, but oscillate strongly in space due to their sharp frequency cut-off. They are probably best used in a wavelet analysis. The Harmonic spline approach of directly using the expected spectrum as kernel is quite elegant, but requires knowledge about the signal when designing the kernel. A disadvantage is that no analytical expression for the kernel is available.

All computations presented here were done using Poisson wavelets of order three as kernel for the basis functions. This is a somewhat arbitrary choice from the class of RBFs with known analytical expressions.

2.5.2 Bandwidth selection

The bandwidth of an RBF affects its approximation characteristics. RBFs with large bandwidths cover a lower-frequency part of the spectrum, while small bandwidths are required for modelling high-frequency signal.

There is no single definition for a basis function's bandwidth. Narcowich and Ward (1996) introduced a basis function's variance as measure for the bandwidth. Here, the bandwidth is considered to be the basis function's correlation length, $\vartheta_{0.5}$. The bandwidth is defined as the spherical angle where the basis function is half its maximum value, and is determined by the type of basis function chosen and the associated parameters. In the case of the Poisson wavelets of order 3, the bandwidth is determined by the radial distance $|\mathbf{y}|$ and thus the depth $d = R - |\mathbf{y}|$.

It is of utmost importance that the bandwidth is chosen properly. Otherwise, the basis functions will not have the optimal approximation characteristics. Too narrow bandwidths might lead to a good approximation of the data and a well-conditioned normal equation system, but a poor quality of the solution between the data points. A bandwidth that is chosen too large will result in an ill-conditioned normal equation system,

may lead to a solution that is too smooth, and might result in leakage since the basis functions cannot represent the spectral content contained in the signal.

Choosing the bandwidth is a problem of model identification, hence many methods can be used. Possible methods are computing the signal's covariance function and modifying the basis functions' parameters to fit it (Marchenko et al., 2001), or using the expected signal spectrum as spectrum for the basis functions (Eicker, 2008). As part of the methodology presented here, a different approach has been chosen: The bandwidth is selected by minimising a functional computed out of the data used in the approximation.

The simplest functional is the root mean square (rms) of the least-squares residuals and thus the best fit to the data:

$$f_{rms} = \sqrt{\frac{1}{n} \|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|^2}. \tag{2.32}$$

Use of the residual rms has drawbacks. It can be made arbitrarily small by increasing the number of RBFs. It is minimised when all observations are approximated exactly, thus interpolated. This includes the modelling of all noise contained in the observations. Nevertheless, if used with care and with a number of basis functions significantly smaller than the number of observations, minimising the observation residual rms can yield a smooth, good solution.

The over-fit can be avoided by using generalised cross validation, based on the leave-one-out idea (e.g. Picard and Cook , 1984). Solutions $\hat{\mathbf{x}}_k$ are computed leaving out a different observation y_k in each estimation, and predicting the observation $\hat{y}_k$ using the solution $\hat{\mathbf{x}}_k$. A well-selected bandwidth would lead to small errors in the predicted observations. General Cross Validation (GCV, Golub et al. (1979)), aims at minimising the functional (Kusche and Klees, 2002)

$$\begin{aligned} f_{GCV} &= \frac{n\|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}\|^2}{(n - u + \alpha\Im)^2} \\ \Im &= \text{trace}\left(\mathbf{L}\left(\mathbf{A}^T\mathbf{P}\mathbf{A} + \alpha\mathbf{R}\right)^{-1}\mathbf{L}^T\right) \end{aligned} \tag{2.33}$$

where $\mathbf{L}$ is the lower triangular matrix resulting from the Cholesky decomposition of the regularisation matrix $\mathbf{R}$. If more observations are available than necessary for the estimation (oversampling), some observations can be left out and then be used as criterion for minimising the rms.

Sometimes, information about the signal to be recovered is available, for example when computing a new gravity field model for an area where an older model or reference model is available. A smooth solution can then be obtained by minimising the difference between the reference and the new model, instead of the data rms or GCV. This method has to be used with caution, as the new model will inevitably be biased towards the reference model.

A smooth solution is obtained by placing the basis functions deep rather than shallow. There is a limit to how deep the basis functions can be placed. When placed too deep, they overlap too much, resulting in an ill-conditioned normal matrix. Looking at the

condition number of the normal matrix can help in estimating optimal depths. This approach most however used with caution as it may lead to sub-optimal results, but has been used successfully when modelling satellite data.

2.6 RBF network design

RBF network design addresses the issues of how many RBFs are to be used in the estimation process and their horizontal position. There are generally two possible choices for the network design:

1. Placing the RBFs on a grid, with grid type and spacing chosen suitably.

2. Scattered RBF positioning, guided by the location of the observations (data-adaptive).

Grids are a straightforward approach. Grid type and grid parameters (spacing, level) determine the positions of the basis functions. However, type and parameters need to be chosen carefully to avoid under- and overparametrisation. Grids are well-suited for use with homogeneous data distributions. They can also be used for heterogeneous data distributions to get an initial solution that may then be refined in a two-step approach (Klees et al., 2008a). Data-adaptivity may be required to avoid overparametrisation in the initial step.

Scattered RBF positioning does not make use of grids, but optimises the RBF centre locations based on the characteristics of the data. Least-squares collocation (Moritz, 1980) places one basis function under each observation. While this makes the RBF centre location easy, it will lead to an unnecessarily high numerical complexity and stability issues for an RBF solution in the case of signal oversampling by the data. Additionally, it may lead to modelling of noise and not yield the desired smooth solution. In practice, RBFs are not placed below every observation. There exists however no objective criterion for leaving out RBFs, the decision is up to the experience of the user.

Barthelmes (1986) designed an algorithm for use with point masses. It solves a non-linear optimisation problem with four parameters per point mass, three 3D position coordinates and one parameter to fix the magnitude. The associated numerical complexity might be the reason that this approach has not gained wider acceptance.

An alternative approach has been developed by Marchenko (1998); Marchenko et al. (2001) for the use with radial multipoles. The sequential multipole algorithm places basis functions under the largest (residual) observations, and at the same time determines order and depth of the multipole. The latter two parameters are fixed using the gravity anomaly covariance function in the vicinity of the data point. It has been proven that this approach leads to a significantly reduced number of basis functions compared to simpler approaches (Klees et al., 2005) while maintaining a high-quality solution. The network optimisation procedure is however quite complex and time-consuming.

Here, another route has been taken. An adaptive approach is used, but to reduce the numerical complexity each solution is first computed using a grid of basis functions, with

all basis functions located on a sphere inside the Bjerhammar sphere. Individual grid points may be omitted in a data-adaptive manner in order to ensure a stable solution. If the resulting approximation quality is not considered sufficient, the RBF network is locally refined in a second step and a joint solution then computed. This leads to good approximations with a relatively small number of basis functions, while avoiding a complicated optimisation procedure.

2.6.1 Grids

When using grids for the placement of basis functions, two choices need to be made:

1. The type of the grid. Mentioned here are equi-angular grids, grids based on the subdivision of an Icosahedron, and Reuter grids. Many other grid types exist and are viable choices (Eicker, 2008).

2. The grid spacing. The grid spacing needs to be chosen in such a way that the spectral data content can be recovered with RBFs of suitable bandwidth. Signals with large gradients and high-frequency content requires dense grids, while coarse grids are sufficient for smooth signals.

A natural grid choice is a grid with equiangular spacing in longitudinal (λ) and latitudinal (φ) direction, respectively. An example of such a grid is shown in figure 2.14.
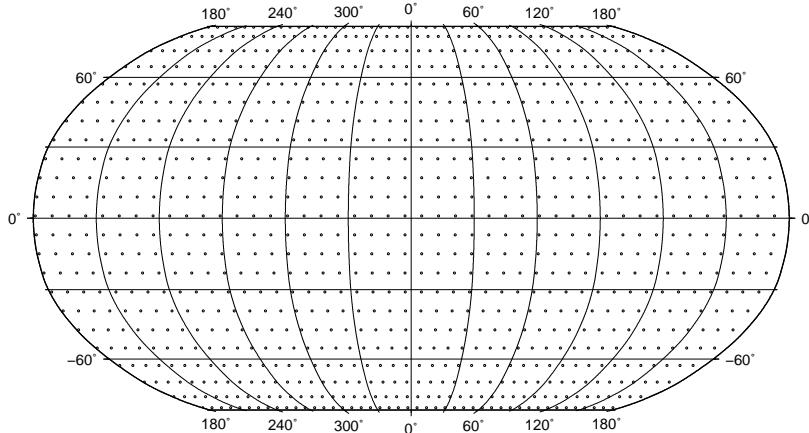
Figure 2.14: Global equiangular grid.

An advantage of such a grid is that it can easily be tailored to a specific purpose by choosing the proper spacing in both directions. A major disadvantage is visible in figure 2.15. At higher latitudes, the grid becomes increasingly dense in longitudinal direction, leading to overparametrisation and thus numerical instabilities. Use of an equi-angular grid should thus be avoided for high-latitude areas.

One solution to this problem is using a grid derived from subdividing a polyhedron (Saff and Kuijlaars, 1997; Freeden, 1999). A very popular choice is the icosahedron, a polyhedron with 20 faces. The number of faces increases four-fold with each subdivision, so a level γ icosahedron grid has

$$n = 20 \cdot 4^{\gamma-1} \tag{2.34}$$
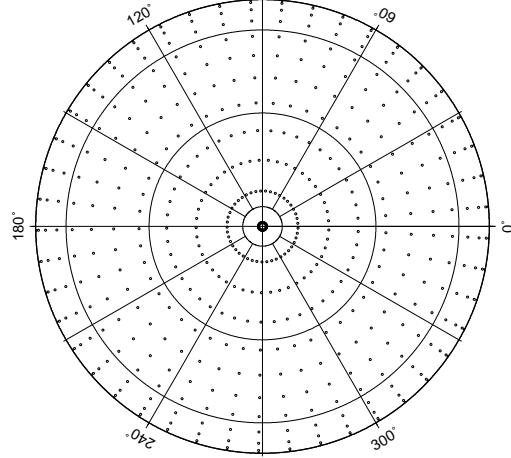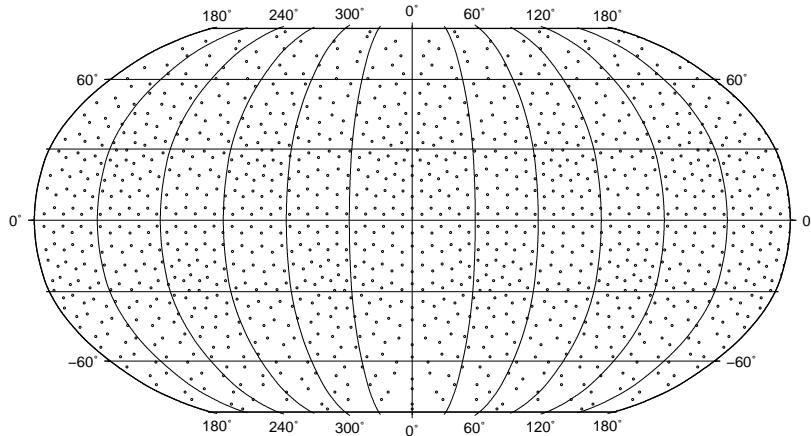
Figure 2.15: Equiangular grid at high latitudes.

points. Figures 2.16 and 2.17 show icosahedron grids of level $\gamma = 4$. The grid has been computed with DOME, available from `http://www.senecass.com/software.html`.

Figure 2.16: Global icosahedron grid of level $\gamma = 4$.

The icosahedron grid shows an almost equidistant point distribution at high latitudes. A disadvantage is the four-fold increase in the number of points from level to level. This can make it difficult to find the correct level for a given application, as slight in- or decreases in the number of grid points and thus the resolution are not possible.

An interesting alternative that does not have this disadvantage is the Reuter grid (Reuter, 1982). It was originally designed for numerical integration, approximation and interpolation on the sphere with harmonic splines. A Reuter grid of level γ is constructed in the following way:

$$
\begin{array}{ll}
\text{(i)} & \vartheta_0 = 0,\ \lambda_{01} = 0\,(\text{North Pole}) \\
\text{(ii)} & \Delta\vartheta = \pi/\gamma \\
\text{(iii)} & \vartheta_i = i\Delta\vartheta,\ 1 \le i \le \gamma - 1 \\
\text{(iv)} & \gamma_i = 2\pi/\arccos\left(\left(\cos\Delta\vartheta - \cos^2\vartheta_i\right)/\sin^2\vartheta_i\right) \\
\text{(v)} & \lambda_{ij} = \left(j - \tfrac{1}{2}\right)\left(2\pi/\gamma_i\right),\ 1 \le j \le \gamma_i \\
\text{(vi)} & \vartheta_\gamma = \pi,\ \lambda_{\gamma 1} = 0\,(\text{South Pole}).
\end{array}
$$
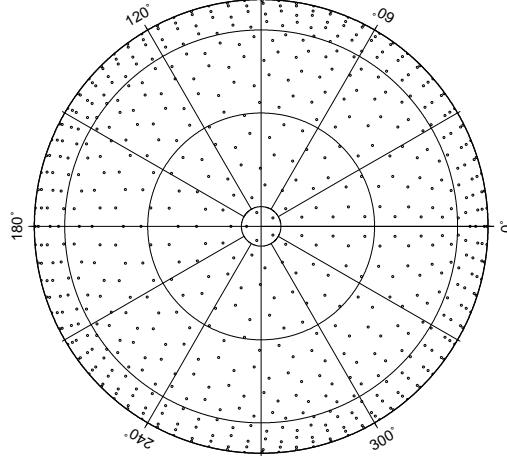
Figure 2.17: Icosahedron grid of level $\gamma = 4$ at high latitudes.

A level γ Reuter grid has a number of grid points which can be estimated by
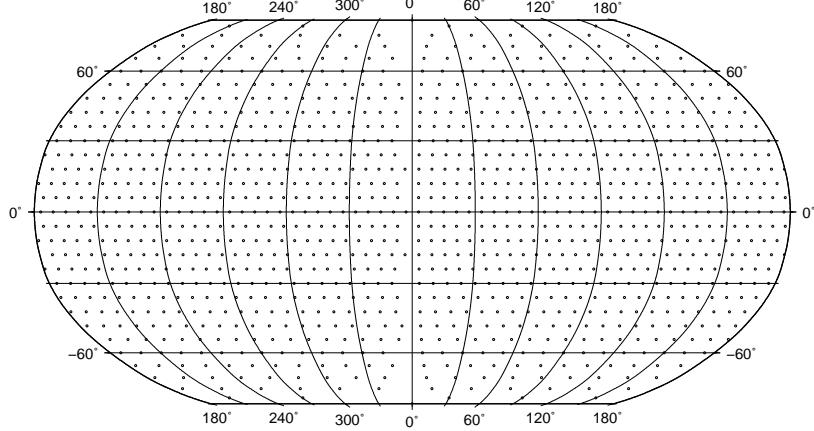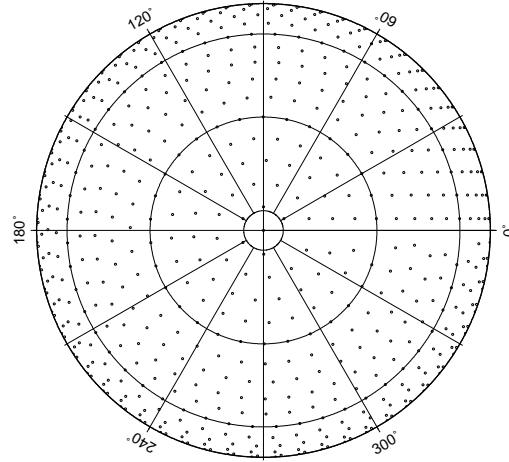
$$n \approx 2 + \frac{4}{\pi}\gamma^2. \qquad (2.35)$$

Figures 2.18 and 2.19 show a Reuter grid of level $\gamma = 30$. Like the icosahedron grid, it provides an almost equidistant point distribution. The advantage of the Reuter grid over the icosahedron grid is that the number of points on a Reuter grid increases more slowly, with γ^2 instead of 4^γ, making it easier to find an optimal point distribution for a given application. Furthermore, as will be shown in chapter 5, the Reuter grid level γ is approximately equal to the maximum spherical harmonic degree l_{max} that can be modelled with a RBF representation using this grid.

Most computations presented here make use of the Reuter grid for the mentioned reasons. It has to be noted that Reuter grids have one drawback: They exhibit gaps near longitude 0° at certain degrees of latitude, depending on the chosen grid level γ. They are evident in figure 2.18, e.g. at 30 and 60 degrees of latitude. These gaps can cause small errors due to unmodelled signal. It is suggested to reduce these errors by rotating the grid in a way that the gaps fall in areas of little signal. This can be achieved by rotating the grid by 180 degrees of longitude, thus placing the gaps in the Pacific ocean.

2.6.2 Adaptation to data

With the methodology presented here, an initial computation always places all basis functions on a grid at the same depth. In the case of heterogeneous data distributions, which are often encountered when dealing with terrestrial gravity data, this grid will not be optimal. Even if the grid spacing is chosen carefully, there may be areas with very sparse data coverage. Placing an excessive number of basis functions in such an area will lead to overparametrisation and an unstable normal equation system. This can either be addressed by regularisation (section 3.3), or by adapting the grid to the data distribution (data-adaptive network design, DAND) (Klees et al., 2008a).

Figure 2.18: Global Reuter grid of level $\gamma = 30$.

Figure 2.19: Reuter grid of level $\gamma = 30$ at high latitudes.

Here, the data adaption is achieved by using only RBFs that have $q \geq 1$ observations within their influence radius (Klees and Wittwer, 2005). q is usually set to 1, but should be larger in the case of significant oversampling of the signal.

The influence radius is a function of the correlation length of the basis function:

$$R_I = \alpha \cdot \vartheta_{0.5}, \tag{2.36}$$

where $\alpha \in \mathbb{R}^+$ an arbitrary factor. α is usually set to one, but sometimes smaller to avoid overparametrisation. Additionally, it is possible to only consider observations with a (residual) signal amplitude exceeding a certain threshold, e.g. σ. This way, the modelling of noise can be reduced.

2.6.3 Local refinement

In gravity field recovery, one has to deal with signal that varies strongly depending on the geographic location. The signal is quite smooth over the oceans and in areas with

little changing topography (such as plains), while there are large variations (gradients) in mountaineous areas or at plate boundaries. A set of RBFs distributed on a coarse grid will usually not be able to model all signal in areas with large variations if the grid is designed in such a way not to lead to over-parametrisation in the areas with smooth signal, and may lead to leakage that may corrupt the solution. It is thus necessary to locally refine the coarse grid solution in order to recover the full signal.

Such a local refinement algorithm should place additional basis functions in areas of large signal variations, where the coarse grid solution alone is not sufficient. The goal is to capture more usable signal while at the same time avoiding overparametrisation, numerical instabilities, and the modelling of noise. The implementation used here is shown in figure 2.20. The starting point is a coarse grid solution. The algorithm performs the following steps:

1. The observation with the largest residual is chosen as candidate location for a local refinement basis function.

2. The optimal depth of the new basis function is determined by minimising the residual when estimating the basis function's coefficient using observations in the vicinity of the data point.

3. The contribution of the RBF to the observations is subtracted in the vicinity.

This process continues with the next largest residual, until no residual above a chosen threshold can be found. The threshold should depend on the signal noise level, to avoid modelling of noise instead of signal, while making sure that as much signal as possible is being recovered. Additionally, the mean value of all observations within the vicinity of the considered observation should exceed some threshold, to avoid the modelling of outliers. It is also important that the distance between the basis function exceeds a certain minimum, otherwise numerical instabilities will occur.

A full readjustment is then performed, using the basis functions on the coarse grid, as well as all basis functions found during the local refinement, with their respective depths, to compute the final solution (Klees et al., 2008a).

A variety of parameters have to be chosen in the local refinement algorithm:

1. The threshold used for selection of observations.

2. The threshold for the mean value of all observations in the vicinity (vicinity threshold). This threshold avoids placement of basis functions under isolated large residuals, which are likely to be outliers.

3. The radius of the vicinity in which observations are used for estimating the basis functions (selection vicinity).

4. The radius of the vicinity in which the contribution of the local refinement basis function is subtracted (update vicinity).

```
              ┌─────────────┐
              │ coarse grid │
              │ adjustment  │
              └──────┬──────┘
                     │    ◄──────────────────────┐
                     ▼                            │
              ┌─────────────┐                     │
              │ find largest│                     │
              │  residual   │                     │
              └──────┬──────┘                     │
                     │                            │
                     ▼                            │
          No      ╱─────────╲                     │
       ◄─────────╱ residual  ╲                    │
       │         ╲ above      ╱                   │
       │          ╲threshold?╱                    │
       │           ╲───┬────╱                     │
       │               │ Yes                      │
       │               ▼                          │
       │        ┌─────────────┐                   │
       │        │ determine   │                   │
       ▼        │   depth     │                   │
┌──────────────┐└──────┬──────┘                   │
│ readjust with│       │                          │
│all basis     │       ▼                          │
│functions     │┌─────────────┐                   │
└──────────────┘│  compute    │                   │
                │ coefficient │                   │
                └──────┬──────┘                   │
                       │                          │
                       ▼                          │
                ┌─────────────┐                   │
                │  subtract   │                   │
                │contribution │                   │
                └──────┬──────┘                   │
                       │                          │
                       └──────────────────────────┘
```
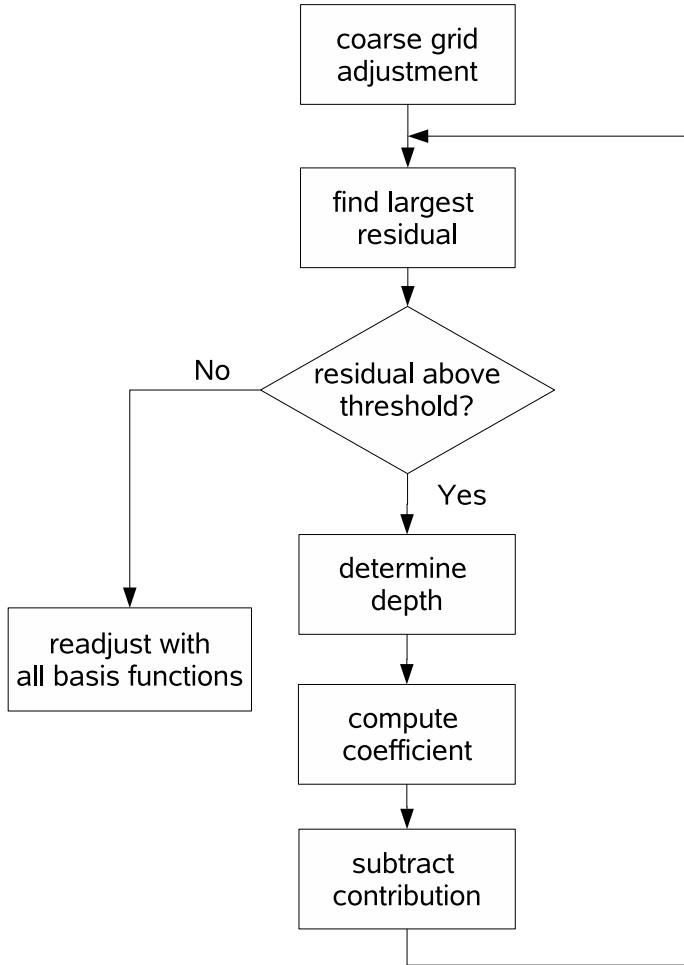
Figure 2.20: flowchart of the local refinement algorithm

5. The minimum distance between basis functions. This prevents numerical instabilities.

It is difficult to give definitive numbers for these parameters, as the optimal values will depend on the test case. From many numerical tests involving different data sets, it can be said that

1. the threshold should be related to the expected noise in the data, with 3σ being a good starting value;

2. the vicinity threshold is very difficult to select, some experimentation is required to find a good value, which is usually quite small (threshold $< \sigma$);

3. the observation selection vicinity should be chosen in such a manner that several, but not too many, observations are chosen;

4. the update vicinity should be larger than the selection vicinity, with three times as large being a good starting value;

5. the minimum RBF distance should be around half the coarse grid spacing; if a tighter spacing delivers a better solution, it might be advisable to increase the number of coarse grid RBFs by using a tighter grid spacing.

2.7 Multi-scale modelling

2.7.1 Introduction

Multi-scale modelling, sometimes called multiresolution modelling, describes a method where the signal is decomposed into different spectral bands, or scales. The signal is represented by a smoothed version and several detail levels. The motivation behind this is the feasibility to compute the different scales independently from different input data with different spectral behaviour, such as satellite and terrestrial data. The ability to look at different spectral bands also aims at easier identification of certain geophysical features.

Wavelets have successfully been used for multi-scale gravity modelling (Fengler et al., 2007; Schmidt et al., 2007). Wavelet methods estimate the full signal using some scaling function (which is in fact a radial basis function), and then compute the individual detail scales using filter operations in a top-down fashion. In the context of the research described here, a different method working bottom-up was developed and analysed.

2.7.2 Methodology

The idea behind a bottom-up algorithm is that only parts of a region will have high-frequency signal, while others will feature a rather smooth signal. With heterogeneous data sets, some areas will have sparse data coverage, making it impossible to use a tight spacing of basis functions. By starting at a coarse scale, the long wavelengths contained in the data will be modelled. Subtracting them from the original input data should leave only high-frequency signal remaining, which will then be modelled by the next higher scale, with basis functions placed shallower on a denser grid. An algorithm working in such a manner may be able to completely model the signal, while using less basis functions than a single-scale algorithm when dealing with data of heterogeneous distribution and/or spectral content.

Let $\hat{\mathbf{x}}_s$ denote the least-squares estimate of the single-scale solution

$$\mathbf{y} + \mathbf{e}_s = \mathbf{A}_s \mathbf{x}_s \tag{2.37}$$

which models the potential field (in vector notation)

$$\mathbf{T} = \hat{\mathbf{x}}_s^T \mathbf{\Psi}_s. \tag{2.38}$$

using the vector of single-scale basis functions $\mathbf{\Psi}_s$. Let $\hat{\mathbf{x}}_j$ denote a solution at level j, $j = j_{min}, ..., j_{max}$ with

$$\mathbf{y} + \mathbf{e}_j = \mathbf{A}_j \mathbf{x}_j. \tag{2.39}$$

A flowchart showing the multi-scale algorithm is shown in figure 2.21. The minimum and maximum scale j_{min} and j_{max} have to be chosen, depending on the expected spectral content of the available data. Estimation starts with scale j_{min}. The scale j is chosen as icosahedron grid level and thus determines the spacing of the basis functions. For each level, the data-adaptive algorithm is applied, as well as the optimal bandwidth/depth estimation. Once the optimal depth has been found, the solution $\hat{\mathbf{x}}_j$ is computed and its contribution is subtracted from the observations:

$$\mathbf{y} + \mathbf{e}_{j+1} - \mathbf{A}_j \hat{\mathbf{x}}_j = \mathbf{A}_{j+1} \mathbf{x}_{j+1}. \tag{2.40}$$

The observation equations for level i are thus

$$\mathbf{y}_n = \mathbf{y} - \sum_{j=j_{min}}^{i-1} \mathbf{A}_j \mathbf{x}_j = \mathbf{A}_i \mathbf{x}_i, \ i \in j_{min}, ..., j_{max}. \tag{2.41}$$

The scale is then increased by one, and the process starts again for the new scale and solution $\hat{\mathbf{x}}_{j+1}$. The highest level should be chosen in such a manner that no usable signal remains after the estimation of the highest level. The total solution is then

$$\mathbf{T} = \sum_{j=j_{min}}^{j_{max}} \hat{\mathbf{x}}_j^T \Psi_j. \tag{2.42}$$

It is hoped that $\|\mathbf{T}_s - \mathbf{T}\|$ is small and that the number of basis functions for all levels j is smaller than the number of basis functions used for the single-scale solution $\hat{\mathbf{x}}_s$.

Note that the methodology does not make use of the local refinement algorithm as described in section 2.6.3. The combination of the bottom-up approach with data-adaptivity and thresholding essentially leads to local refinement, with the exception that all "refinement" basis functions of one scale are located at the same depth. Unlike with the local refinement method, there is no joint estimation of all basis functions. Each scale is estimated independently, and no readjustment is performed.

2.7.3 Filtering

While initial results using this methodology looked promising (Klees and Wittwer, 2007), further research showed that a multi-scale representation computed in such a manner was of lower quality than a single-scale solution. Testing showed that high-frequency signal that cannot be modelled by the lowest scale j_{min} leaked into and corrupted the estimated coarse-scale RBF coefficients $\hat{\mathbf{x}}_{j_{min}}$. Hence, the residuals $\mathbf{y} - \mathbf{A}_{j_{min}} \hat{\mathbf{x}}_{j_{min}}$ contain a long-wavelength error that can not be modelled by higher levels. To deal with the leakage problem, it is necessary to low-pass filter the input data for the lower scales.

Applying a filter $\mathbf{F}_i$ for level i leads to the observation equations
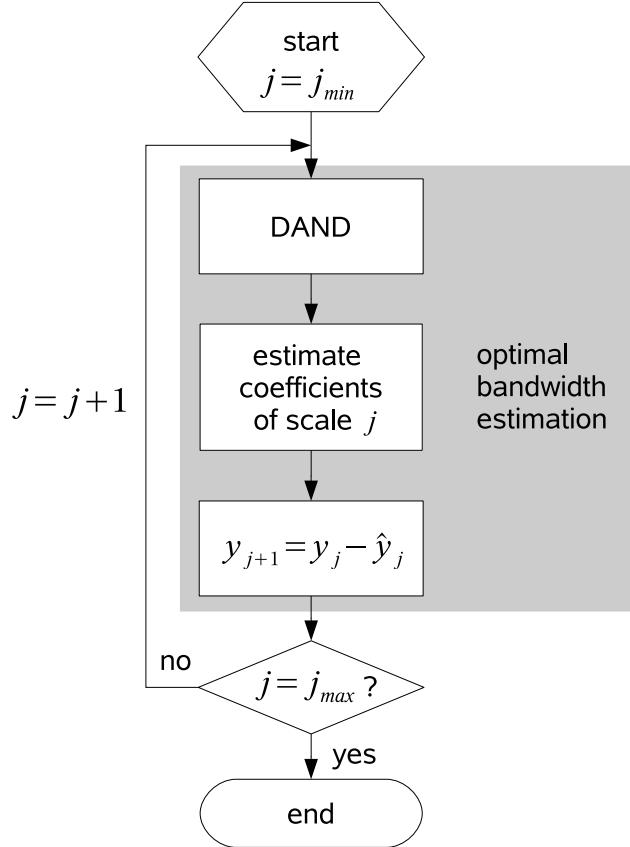
Figure 2.21: flowchart of multi-scale algorithm

$$\mathbf{F}_i(\mathbf{y} - \sum_{j=j_{min}}^{i-1} \mathbf{A}_j\mathbf{x}_j) = \mathbf{A}_i\mathbf{x}_i,\, i \in j_{min},...,j_{max} \ \text{and}\, \mathbf{F}_{j_{max}} = \mathbf{I}. \tag{2.43}$$

The observation equations for a two-scale solution are thus given by

$$\begin{aligned}\mathbf{F}_1\mathbf{y}_1 &= \mathbf{A}_1\mathbf{x}_1 \\ \mathbf{y}_2 = \mathbf{y}_1 - \mathbf{A}_1\mathbf{x}_1 &= \mathbf{A}_2\mathbf{x}_2\end{aligned}, \tag{2.44}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are estimated by least-squares. The total solution is then

$$\mathbf{T} = \hat{\mathbf{x}}_1\Psi_1 + \hat{\mathbf{x}}_2\Psi_2. \tag{2.45}$$

$\mathbf{F}_i$ must be adapted to the basis function vector Ψ_i of the same level, i.e. the half-width of the filter must correspond to the correlation length of the basis functions.

Filtering and filter parameter choice is a well-known, common problem. In order to investigate the effectiveness of filtering in the context of the multi-scale methodology described here, several possible filters were examined. A test was done using noise-free data computed from EIGEN-GL04C (Foerste et al., 2008) up to to degree 70. Gravity

disturbances were synthesised on a grid with 2 degree spacing, resulting in 16,200 observations. Using these data, the following single-scale and multi-scale solutions were computed:

1. A single-scale solution using Poisson wavelets at a depth of 800km on a level 6 icosahedron grid, resulting in 12,400 RBFs.

2. Multi-scale solutions using Poisson wavelets at two depths: 5,100 RBFs on a level 5 icosahedron grid at a depth of 1,600km, and 12,400 RBFs on a level 6 icosahedron grid at a depth of 800km. Various low-pass filters were applied to the input data for the level 5 solution.

While a single-scale solution resulted in an rms error of 0.45 mm compared to a true signal computed from the same set of SH coefficients, a solution using RBFs on a level 5 and 6 icosahedron grid and the methodology described in the previous section yielded an rms error of 5.5 cm. The correlation length of the RBFs on the level 5 grid was 500 km. The input data for the level 5 estimation was subsequently filtered with

1. a Gauss filter with a correlation length of 500 km, resulting in an rms error of 0.42 mm;

2. a Blackman filter with the filter parameter $p = 25$ (correlation length 496 km), resulting in an rms error of 0.35 mm.

In both cases, the correlation length had been chosen equal to the correlation length of the level 5 basis functions. Previous testing had shown that this yielded the smallest error rms.

The cumulative degree errors for the single-scale as well as the two filtered solutions are shown in figure 2.22. The single-scale solution used 12,400 RBFs, the multi-scale solution required an additional 5,100 or 41% more basis functions.

A multi-scale solution constructed in such a way can thus deliver the same quality as a single-scale solution. Unfortunately, the goal of reducing the number of basis functions required could not be achieved with the chosen test case, global gravity recovery. Instead, the number of basis functions required for a multi-scale solution is 30-40% larger than with a single-scale solution. Because of this, the multi-scale approach was abandoned and all solutions shown in later chapters are single-scale solutions. Nevertheless, the methodology described here is a viable approach if a multi-scale representation is desired. A multi-scale solution using two scales to model EIGEN-GL04C up to degree 70 is shown in figure 2.23.
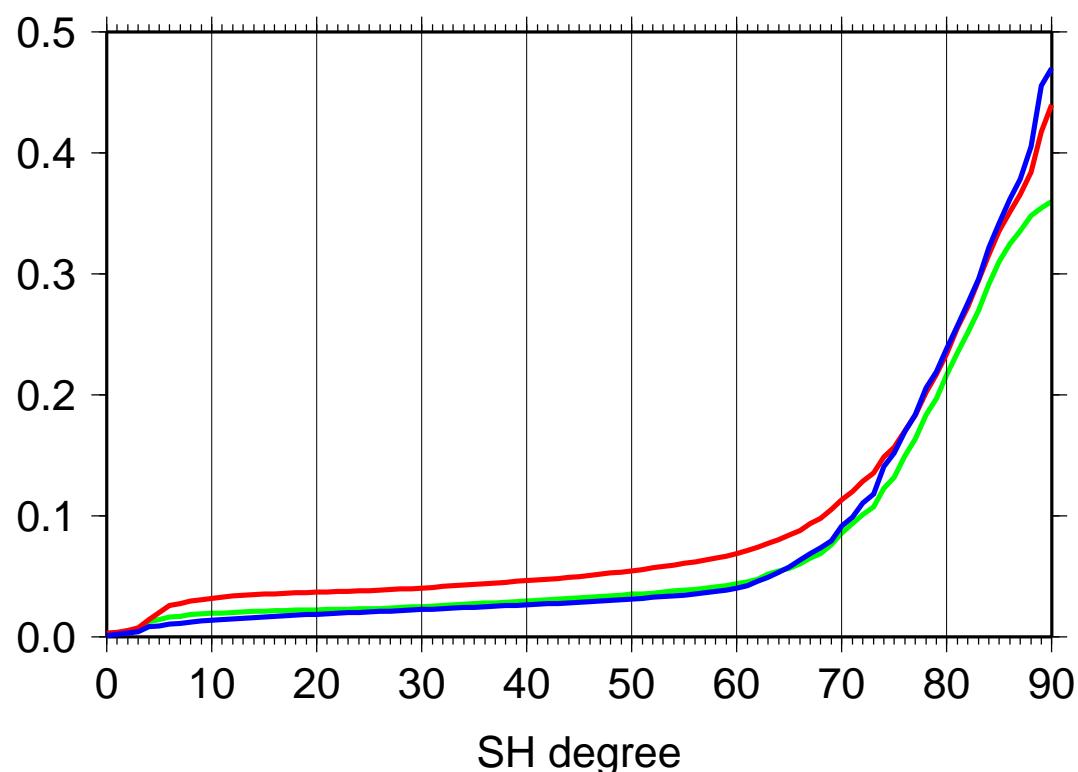
Figure 2.22: Cumulative errors [mm] for single-scale solution (blue), Blackman-filtered solution (green), and Gauss-filtered solution (red).
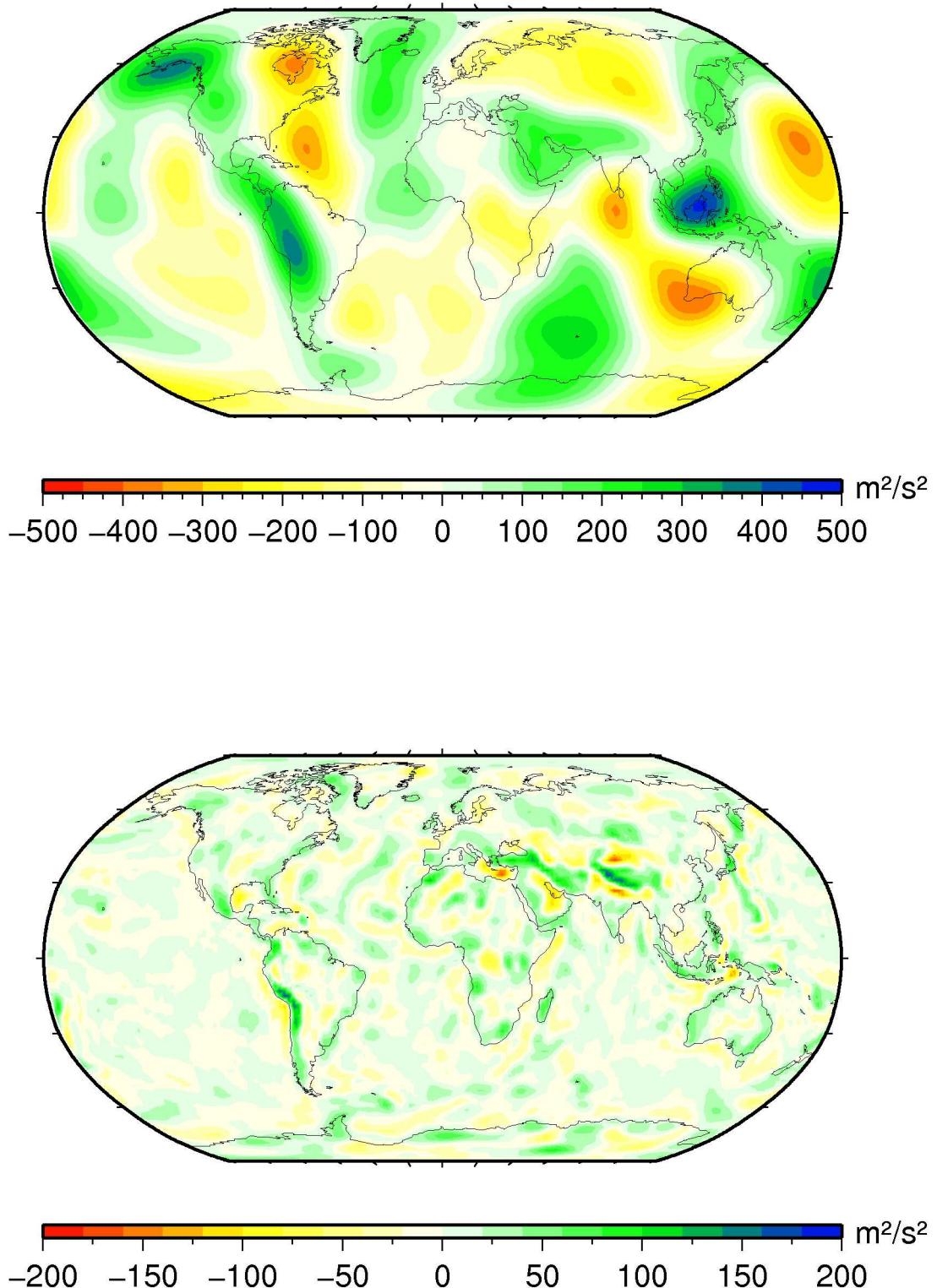
Figure 2.23: Multi-scale solution of disturbing potential: Scale 5 (top) and detail signal at scale 6 (bottom).

Chapter 3

Mathematical model and estimation principle

This chapter describes the mathematical model and the estimation principle used for computing gravity field models with RBFs. Functional model, stochastic model, estimation principle, regularisation, data weighting, and solution strategies are all addressed.

3.1 Functional model

The functional model for gravity field modelling with RBFs (cf. eq. (2.3)) may be written as standard Gauss-Markov model

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{3.1}$$

where $\mathbf{y}$ is the $n \times 1$ vector of observations, $\mathbf{e}$ is the $n \times 1$ vector of residuals, and $\mathbf{x}$ is the $u \times 1$ vector of unknown coefficients α_i. The $n \times u$ design matrix $\mathbf{A}$ describes the functional relations between $\mathbf{x}$ and $\mathbf{y}$.

3.2 Stochastic model

The stochastic properties of the residuals are described by the stochastic model

$$E\{\mathbf{e}\} = 0,\ E\{\mathbf{e}\mathbf{e}^T\} = D\{\mathbf{e}\} = \mathbf{C}, \tag{3.2}$$

where $E\{\cdot\}$ denotes the expectation operator, and $D\{\cdot\}$ the dispersion operator. The noise in the observations is described by the variance-covariance matrix of $\mathbf{e}$, which is denoted $\mathbf{C}$.

3.3 Least-squares estimation and regularisation

The coefficients of the RBFs are estimated using least-squares. The quadratic functional

$$\Phi(\mathbf{x}) = \mathbf{e}^T \mathbf{C}^{-1} \mathbf{e} \tag{3.3}$$

is minimised, where $\mathbf{e}$ is the vector of least-squares residuals and $\mathbf{C}$ is the covariance matrix of the observations. The minimum is obtained by computing

$$\hat{\mathbf{x}} = \mathbf{N}^{-1} \mathbf{b}, \tag{3.4}$$

with the normal matrix

$$\mathbf{N} = \sum_{j=1}^{J} \frac{1}{\sigma_j^2} \mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j \tag{3.5}$$

and the right-hand side vector

$$\mathbf{b} = \sum_{j=1}^{J} \frac{1}{\sigma_j^2} \mathbf{A}_j^T \mathbf{P}_j \mathbf{y}_j \tag{3.6}$$

with $\mathbf{A}_j$ the design matrix of observation group j, the weight matrix $\mathbf{P}_j = \frac{1}{\sigma_j^2} \mathbf{C}_j^{-1}$, $\mathbf{C}_j$ being the covariance matrix of observation group j with variance factor σ_j^2, and the observation vector $\mathbf{y}_j$.

The least-squares estimates of the observation vector and the vector of residuals are given as

$$\hat{\mathbf{y}} = \mathbf{A}\hat{\mathbf{x}}, \tag{3.7}$$

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}}. \tag{3.8}$$

It is possible that the normal equation matrix $\mathbf{N}$ is ill-conditioned. Then, some sort of stabilisation is needed to make the computation of a solution possible. This can be achieved by Tikhonov regularisation

$$\hat{\mathbf{x}} = (\mathbf{N} + \alpha \mathbf{R})^{-1} \mathbf{b}, \tag{3.9}$$

where $\mathbf{R}$ is the regularisation matrix and α is the regularisation parameter. In this thesis $\mathbf{R} = \mathbf{I}$ is used (zero-order Tikhonov regularisation). The choice of $\mathbf{R}$ affects the quality of the solution (Kusche and Klees, 2002; Ditmar et al., 2003b), although some studies (e.g. Ilk (1993)) have concluded that the choice of $\mathbf{R}$ is of secondary importance. In the research presented here, regularisation is only needed and used with terrestrial data (chapter 6).

There are several methods for choosing the regularisation parameter α (e.g. Ditmar et al., 2003b). It can be chosen manually by looking at the condition number of the normal equation matrix $\mathbf{N}$ or the smoothness of the solution, which will yield a regularisation parameter as small as possible for a stable normal equation system. More desirable is an "automatic" optimal estimation, which can be achieved by generalised cross-validation or variance component estimation (see section 3.5).

3.4 Solution strategies

For regional gravity field modelling, we often have to deal with a large number of observations and unknown parameters. As a result, the numerical cost to compute $\mathbf{N}$ and its inverse are very large. Two strategies can be used to handle the numerical complexity. In addition, a parallel implementation is required to deal with memory requirements and to reduce computing times. Parallelisation is treated in section 4.3.

3.4.1 Cholesky factorisation

Since the normal matrix $\mathbf{N}$ (or $\mathbf{N} + \alpha \mathbf{R}$ in the case that $\mathbf{N}$ is ill-conditioned) is a symmetric positive-definite matrix with real entries, it can be decomposed as (e.g. Golub and van Loan, 1996)

$$\mathbf{N} = \mathbf{G}\mathbf{G}^T, \tag{3.10}$$

where $\mathbf{G}$ is a lower triangular matrix with positive diagonal entries, and $\mathbf{G}^T$ denotes the conjugate transpose of $\mathbf{G}$. The decomposition in eq. (3.10) is called Cholesky factorisation. It can be used for the numerical solution of a linear system of equations $\mathbf{N}\mathbf{x} = \mathbf{b}$ by first computing the Cholesky decomposition, then solving the system $\mathbf{G}\mathbf{p} = \mathbf{b}$ for $\mathbf{p}$, and finally solving $\mathbf{GTx} = \mathbf{p}$ to obtain the solution $\mathbf{x}$.

The assembly of $\mathbf{N}$ requires the computation of the dense matrix-matrix product $\mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}$, a numerically intensive operation. Luckily, efficient parallel routines for both the computation of $\mathbf{N}$ and the Cholesky decomposition in eq. (3.10) exist.

3.4.2 Conjugate gradients

The system of normal equations can be solved using the conjugate gradients (CG) method (Hestenes and Stiefel, 1952), an iterative solver. The expensive assembly of $\mathbf{N}$ is not required for the CG method. Keeping the whole design matrix $\mathbf{A}$ in memory can also be avoided, which drastically lowers the memory requirement.

In order to achieve an acceptable convergence rate, the so-called pre-conditioned conjugate gradients (PCCG) method needs to be used. The convergence rate is strongly affected by the choice of the pre-conditioner $\mathbf{P}$. While a suitable preconditioner is easily found for spherical harmonics, resulting in estimations that converge within a few iterations (Wittwer, 2006), no effort was made to find a suitable preconditioner for RBFs.

Therefore, all computations presented here were performed with a parallel solver making use of Cholesky factorisation. Furthermore, assembly of the normal matrix $\mathbf{N}$ makes it easy to compute the covariance matrix $\mathbf{C} = \mathbf{N}^{-1}$ of the estimated coefficients.

3.5   Variance component estimation

3.5.1   Normal equations

If more than one observation group is used, each observation group j has to be introduced with its proper weighting factor. The normal equations are

$$\left(\sum_{j=1}^{J} \frac{1}{\sigma_j^2} \mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j \right) \mathbf{x} = \sum_{j=1}^{J} \frac{1}{\sigma_j^2} \mathbf{A}_j^T \mathbf{P}_j \mathbf{y}_j, \tag{3.11}$$

or, in the case of regularisation,

$$\left(\left(\sum_{j=1}^{J} \frac{1}{\sigma_j^2} \mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j \right) + \alpha \mathbf{R} \right) \mathbf{x} = \sum_{j=1}^{J} \frac{1}{\sigma_j^2} \mathbf{A}_j^T \mathbf{P}_j \mathbf{y}_j, \tag{3.12}$$

with σ_j^2 being the variance factors of the observation groups. The regularisation parameter α can be interpreted as inverse variance factor of the observation group $\mathbf{e}_{J+1} + \mathbf{y}_{J+1} = \mathbf{x}$ where $D\{\mathbf{e}_{J+1}\} = \frac{1}{\alpha}\mathbf{R}^{-1} = \sigma_x^2 \mathbf{R}^{-1}$. Then, α can be interpreted as the inverse variance factor of the unknown parameters.

3.5.2   Variance component estimation

The variance factors of the observation groups are estimated using variance component estimation (VCE). If regularisation is applied, VCE can also be used to estimate the regularisation parameter, i.e. the variance factor of the unknown parameters. The variance factors for the observation groups are, according to Kusche (2003),

$$\hat{\sigma}_j^2 = \frac{\hat{\mathbf{e}}_j^T \mathbf{P}_j \hat{\mathbf{e}}_j}{r_j}, \tag{3.13}$$

with the residuals

$$\hat{\mathbf{e}}_j = \mathbf{A}_j \hat{\mathbf{x}} - \mathbf{y}_j \tag{3.14}$$

and the partial redundancies

$$r_j = n_j - \frac{1}{\sigma_j^2} \mathrm{tr}\left(\mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j \mathbf{N}^{-1} \right), \tag{3.15}$$

where n_j denotes the number of observations in observation group j.

The variance factor for the unknown parameters is, in the case of $\mathbf{R} = \mathbf{I}$, obtained from

$$\hat{\sigma}_x^2 = \frac{\hat{\mathbf{x}}^T \hat{\mathbf{x}}}{r_x}, \tag{3.16}$$

with

$$r_x = u - \mathrm{tr}\left(\frac{1}{\sigma_x^2}\mathbf{N}^{-1}\right), \tag{3.17}$$

and u being the number of unknown parameters (Koch and Kusche, 2002).

3.5.3 Stochastic trace estimation

The computation of the partial redundancies r_i, as defined in eq. (3.15), is difficult, as the inverse of the normal equation matrix $\mathbf{N}^{-1}$ is required, which may not be available if an iterative method is used to compute the least-squares solution $\hat{\mathbf{x}}$, and time-consuming, because the matrix product $\mathbf{A}_j^T\mathbf{P}_j\mathbf{A}_j\mathbf{N}^{-1}$ needs to be computed. The solution of this problem is the use of Monte-Carlo methods for stochastic trace estimation.

According to Hutchinson (1990),

$$E(\mathbf{z}^T\mathbf{B}\mathbf{z}) = \mathrm{tr}\,\mathbf{B}, \tag{3.18}$$

holds, where $\mathbf{B}$ denotes a symmetric $n \times n$ matrix and $\mathbf{z}$ an $n \times 1$ vector of n independent samples from a random variable Z with $E(Z) = 0$ and $D(Z) = \mathbf{I}$. If Z is a discrete random variable which takes with probability $1/2$ the values -1 and +1, then $\mathbf{z}^T\mathbf{B}\mathbf{z}$ is an unbiased estimator of $\mathrm{tr}\,\mathbf{B}$ with minimum variance. For large-scale problems, it is usually sufficient to employ only one realisation of $\mathbf{z}$ (Kusche and Klees, 2002).

Partial redundancies, unweighted case

In the case of $\mathbf{P}_j = \mathbf{I}$, eq. (3.15) may be rewritten as

$$r_j = n_j - \frac{1}{\sigma_j^2}\mathrm{tr}\left(\mathbf{A}_j\mathbf{N}^{-1}\mathbf{A}_j^T\right). \tag{3.19}$$

Inserting eq. (3.18) yields

$$r_j = n_j - \frac{1}{\sigma_j^2}\mathbf{z}^T\mathbf{A}_j\mathbf{N}^{-1}\mathbf{A}_j^T\mathbf{z}. \tag{3.20}$$

Computing the product $\mathbf{N}^{-1}\mathbf{A}_j^T\mathbf{z}$ can be replaced by solving the linear equation system

$$\beta_j = \mathbf{N}\xi_j, \; \beta_j = \mathbf{A}_j^T \mathbf{z}. \tag{3.21}$$

This results in the following equation for the partial redundancies:

$$r_j = n_j - \frac{1}{\sigma_j^2}\mathbf{z}^T \mathbf{A}_j \hat{\xi}_j \tag{3.22}$$

Partial redundancies, weighted case

If $\mathbf{P}_j \neq \mathbf{I}$, a rearrangement as in eq. (3.19) is not possible. Instead, applying eq. (3.18) directly to eq. (3.15) yields

$$r_j = n_j - \frac{1}{\sigma_j^2}\mathbf{z}^T \mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j \mathbf{N}^{-1} \mathbf{z}. \tag{3.23}$$

The further computation is similar to the unweighted case, with the computation of $\mathbf{N}^{-1}\mathbf{z}$ being replaced by solving the linear equation system $\beta_j = \mathbf{N}\xi_j$. The partial redundancies are then computed as

$$r_j = n_j - \frac{1}{\sigma_j^2}\mathbf{z}^T \mathbf{A}_j^T \mathbf{P}_j \mathbf{A}_j \hat{\xi}_j. \tag{3.24}$$

Partial redundancies, unknown parameters

For the variance factor of the unknown parameters, the equations are somewhat simpler. Inserting eq. (3.18) into eq. (3.17) yields

$$r_x = u - \frac{1}{\sigma_x^2}\mathbf{z}^T \mathbf{N}^{-1} \mathbf{z} = u - \frac{1}{\sigma_x^2}\mathbf{z}^T \hat{\xi}, \tag{3.25}$$

where $\hat{\xi} = \mathbf{N}^{-1}\mathbf{z}$. The variance factor can then be computed according to eq. (3.16).

Chapter 4

Numerical aspects

The computations performed in the context of the research presented here use large numbers of observations and unknown coefficients. This leads to significant numerical costs in terms of runtime and memory. Three areas of numerical complexity can be identified:

- The computation of the design matrix $\mathbf{A}$. This consists of many small operations that are memory intensive. A numerically optimised implementation is necessary. Interpolation schemes (see section 4.2) allow for a fast computation of the design matrix elements even when a series expansion of the RBFs needs to be used.

- The largest computational effort is required for the computation of the normal equation matrix by dense matrix-matrix multiplication, $\mathbf{N} = \mathbf{A}^T \mathbf{P} \mathbf{A}$. Luckily, this is a well-understood operation for which efficient parallel routines exist.

- The solution of the normal equation system is relatively quicker to compute, but can also benefit from parallel routines.

To get an idea of the numerical complexity, GRACE is taken as an example. When computing an optimally filtered GRACE four-year time series, this requires the assembly and solution of 48 monthly solutions. Each monthly solution typically uses 500,000 observations to estimate 10,000 unknown coefficients. It would require approximately 4 hours of computation time on a single 3 GHz CPU if memory constraints were not an issue. In fact, since optimal filtering requires iteration (see section 5.3), this time would be required not 48 times for a 4-year time series, but 480 times in the case of 10 iterations, for a total of 1,920 hours or 80 days. It is obvious that a numerically efficient and parallel implementation is necessary in order to perform the computations within an acceptable time frame.

This chapter starts with some numerical optimisation issues. Section 4.2 describes a fast synthesis scheme that can be used to accelerate the computation of the design matrix. The actual parallelisation of the program used here is described in section 4.3. The chapter ends with a summary.

4.1   Numerical optimisation

4.1.1   Constant expressions in "do"-loops

When dealing with spectral representations of RBFs, frequently series expansions like

$$\beta(\mathbf{x}) = \sum_{l=0}^{l_{max}} (2l+1) \left(\frac{R}{|\mathbf{x}|}\right)^l \alpha(\mathbf{x}, l), \tag{4.1}$$

have to be evaluated, often up to a very high maximum degree l_{max}, e.g. when computing the series expansion of an RBF (see section 2.3). Usually, the series has to be computed several times with slightly different factors $\alpha(\mathbf{x}, l)$, the rest of the expression remaining constant. The computation of the expression $\left(\frac{R}{|\mathbf{x}|}\right)^l$ alone might use as many as $\frac{l_{max}^2 + l_{max} + 1}{2}$ floating point operations, due to the recursive nature of the power computation.

The computation of several series can be accelerated significantly if the constant expression $(2l+1)\left(\frac{R}{|\mathbf{x}|}\right)^l$ is computed beforehand for each degree l and stored in a vector. Even if the series is computed only once, the costly exponentiation can be avoided by computing $\left(\frac{R}{|\mathbf{x}|}\right)^l$ recursively as $\left(\frac{R}{|\mathbf{x}|}\right)^{l-1}\left(\frac{R}{|\mathbf{x}|}\right)$. Only about l_{max} floating point operations are then required.

Smaller expressions, like $\frac{R}{|\mathbf{x}|}$, that are frequently reused will often be optimised by the compiler, by computing them only once and storing them. Since there is no guarantee that this is done, based on the complexity of the expression and the availability of registers, it is advisable to do this manually (by computing them once and storing them in a variable) if a significant reduction in computational workload is expected. It will require additional storage, but today storage is usually a smaller problem than computing power.

4.1.2   Computation of the design matrix

A typical example of the repeated evaluation of the same series is the computation of the design matrix $\mathbf{A}$. For n observations and u basis functions, the series

$$\Psi_{i,j}(\mathbf{x}_j, \mathbf{y}_i) = \sum_{l=0}^{l_{max}} \psi_{l,i} \left(\frac{R}{|\mathbf{x}_j|}\right)^{l+1} P_l(\hat{\mathbf{x}}_j^T \hat{\mathbf{y}}_i), \tag{4.2}$$

has to be evaluated $n \cdot u$ times, where $\left(\frac{R}{|\mathbf{x}|}\right)^{l+1}$ is the same in each series belonging to the same observation j. $\psi_{l,i}$ is the same for all basis functions i with the same parameters, such as depth. This means that the series depends only on the angle between observation $\mathbf{x}_j$ and basis function $\mathbf{y}_i$. By computing the constant expressions $\psi_{l,i} \left(\frac{R}{|\mathbf{x}_j|}\right)^{l+1}$ only

once for each observation, and only the term $P_l\left(\hat{\mathbf{x}}_j^T\hat{\mathbf{y}}_i\right)$ for each pair of observation and basis function, a significant acceleration can be achieved.

When using analytical expressions, the resulting speed improvement will be much smaller. Still, computing constant expressions only once per observation will accelerate the computations.

4.1.3 Normalisation of coordinates

Normalised coordinates of basis functions and observations are frequently required during the assembly of the design matrix, as well as during the data-adaptive network design and local refinement steps. They are used for computing the angle between observation j and basis function i by means of the scalar product, $\hat{\mathbf{x}}_j^T\hat{\mathbf{y}}_i$ with $\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|}$ and $\hat{\mathbf{y}} = \frac{\mathbf{y}}{|\mathbf{y}|}$. In the computation of the design matrix, this angle has to be computed $n \cdot u$ times. Computing a normalised coordinate is performed by

$$\frac{\mathbf{x}}{|\mathbf{x}|} = \frac{\mathbf{x}}{\sqrt{x_1^2 + x_2^2 + x_3^2}}, \tag{4.3}$$

which requires three multiplications, two additions, three divisions, and a very costly square root operation. A significant acceleration can thus be achieved by computing the normalised coordinates just once and storing them.

4.1.4 Normalisation of basis functions

The values attained by the basis functions can be very large or very small. In the interest of numerical stability, it is better to work with normalised RBFs:

$$\Psi_{norm} = \frac{\Psi_i}{\|\Psi_i\|}, \tag{4.4}$$

A suitable normalisation factor $\|\Psi_i\|$ can be computed by

$$\|\Psi_i\| = \Psi\left(r = r_{mean}, \vartheta = 0\right), \tag{4.5}$$

i.e. the basis function with spherical angle $\vartheta = 0$ between basis function and observation and the observation radius r equal to the mean observation radius. It is advisable to compute the normalisation factor only once per program run and store it.

4.2 Fast synthesis

It has been shown in section 2.3 that it is necessary to evaluate the series expansion of the basis functions to a sufficiently high degree. Depending on the maximum degree (which is governed by the depth of the basis function, see section 2.3), this poses a
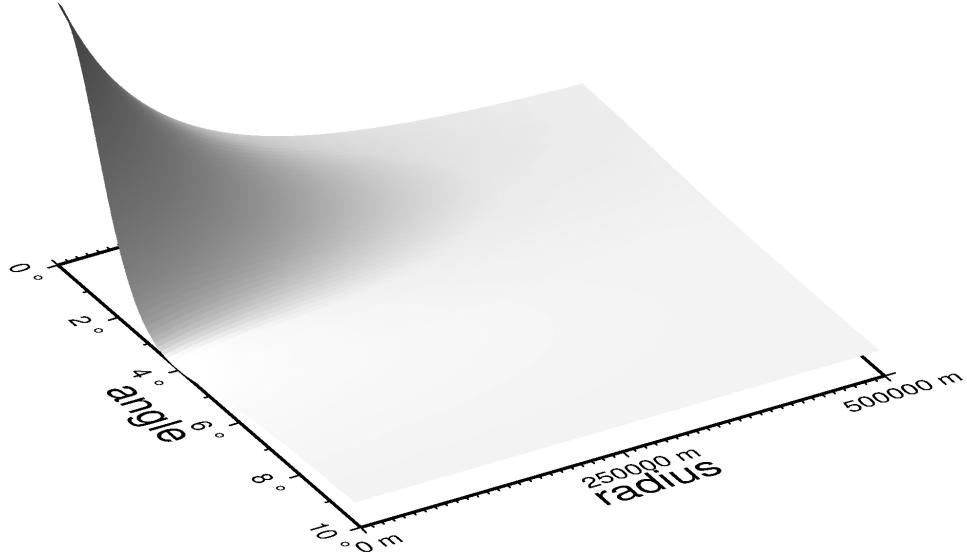
Figure 4.1: Basis function of 500km depth for various angles ϑ and radii r.

considerable numerical effort, as the series will have to be evaluated for every pair of basis functions and observations, even if the numerical optimisations in section 4.1 are used. If no analytical expression for the basis functions is available, a fast synthesis scheme allows for the quick computation of the basis functions. This scheme is based on the idea developed for the fast synthesis for spherical harmonics by Ditmar and Klees (2002).

The basis function Ψ is a function of two parameters: The spherical angle ϑ between the centre of the basis function and the observation, and the radial distance r of the observation from the center of the Earth, assuming that all basis functions are located on the same sphere. Since the function of ϑ and r is quite smooth (fig. 4.1), it is possible to interpolate between values to compute the basis function with sufficient accuracy.

As a first step (the synthesis step), the basis functions will have to be computed at a sufficiently dense grid, which will be used for interpolation. Grid spacings of $1 \cdot 10^{-4}$ rad at 16 different radii have been proven to be sufficient in all test cases.

The second step is interpolating the desired values using 2-dimensional Overhauser splines (Ditmar and Klees, 2002). For the interpolation, 16 nodes are required: 4 nodes at 4 different altitudes each. With these 16 nodes, values can be interpolated for the area between the 4 centre points (the shaded area in fig. 4.2).

The interpolated value is calculated by

$$S(u,v) = \sum_{i=1}^{16} w_i(u,v) f_i, \tag{4.6}$$

where f_i is the function value at point i. The 16 two-dimensional shape functions w_i are the product of four one-dimensional shape functions $b_j(u)$ and $b_k(v)$ with $t \in [0,1]$:
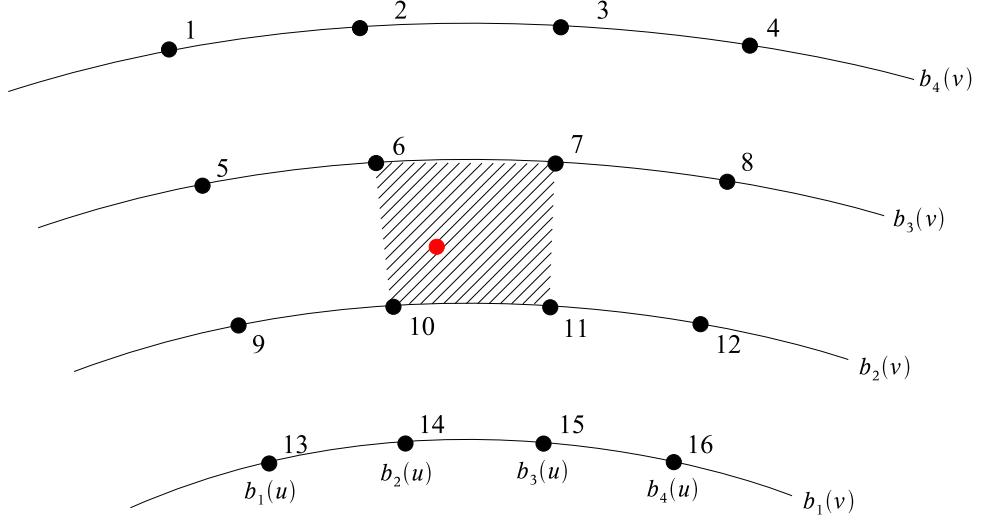
Figure 4.2: Node setup for Overhauser spline interpolation.

$$
\begin{aligned}
b_1(t) &= -\tfrac{1}{2}t(1-2t+t^2) \\
b_2(t) &= 1 - \tfrac{5}{2}t^2 + \tfrac{3}{2}t^3 \\
b_3(t) &= \tfrac{1}{2}t(1+4t-3t^2) \\
b_4(t) &= -\tfrac{1}{2}t^2(1-z)
\end{aligned}
\tag{4.7}
$$

The w_i are then given by $w_1(u,v) = b_1(u)b_4(v)$, $w_{10}(u,v) = b_2(u)b_2(v)$ etc. (see fig. 4.2). The parameters u, v are defined as follows:

$$
u = \frac{\vartheta - \vartheta_2}{\Delta\vartheta}, \ v = \frac{r - r_2}{\Delta r}
\tag{4.8}
$$

Eq. 4.7 and 4.8 only hold for a grid of synthesised points that is equidistant both in ϑ and in r direction. As all points need to be in the centre area, with two synthesised points each to the left, the right, the top, and the bottom, the shells need to be constructed such that all data points are between the second and the second-to-last spherical shell.

4.3 Parallelisation

4.3.1 Problem description

The estimation of a gravity field model requires the assembly and solution of the normal equations $\mathbf{Nx} = \mathbf{b}$, where $\mathbf{N} = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$ or $\mathbf{N} = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{A} + \alpha\mathbf{R}$ and $\mathbf{b} = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{y}$.

Three main areas of computational effort have previously been identified:

1. the computation of the design matrix $\mathbf{A}$,

2. the dense matrix-matrix multiplication $\mathbf{N} = \mathbf{A}^T\mathbf{C}^{-1}\mathbf{A}$,

3. the inversion $\mathbf{N}^{-1}$, which is usually replaced by solving the linear equation system $\mathbf{N}\hat{\mathbf{x}} = \mathbf{b}$.

The method described here involves other areas, of considerable computational effort, such as data-adaptive network design (section 2.6.2) and local refinement (section 2.6.3).

The subject of parallelisation can be treated here only in a superficial manner. A much more detailed description is given in Wittwer (2006).

4.3.2   Parallel computer architectures

For the purpose of parallel programming, we need to differentiate between two kinds of parallel computer architectures: shared memory and distributed memory. Figure 4.3 gives a schematic view of a shared memory parallel computer. All CPUs are connected to all memory banks by some sort of interconnect. If all CPUs have equal access to all memory, we speak of SMP (symmetric multiprocessing). If the access to some memory banks is slower than to others, we call this NUMA (non-unified memory access).

Shared memory systems are easy to program, as all CPUs can access the whole memory. The distribution of data is not of importance, good performance can simply be reached by distributing the workload in a clever way. The downside to shared memory systems is their limited maximum size, as the interconnect becomes increasingly complex with larger numbers of CPUs and memory banks. This puts a practical limit on the maximum number of CPUs. These days, true SMP systems are rarely found with more than 16 CPUs.
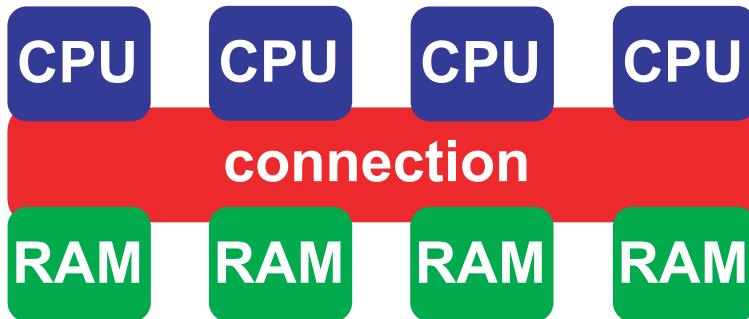

Figure 4.3: Schematic of a shared memory parallel computer.

A schematic of a distributed memory system is shown in figure 4.4. Each CPU has its own memory. The CPUs are connected by some sort of network. Since the connection between the CPUs can have a considerably lower bandwidth and latency than the connection between CPUs and memory, this interconnect can be much simpler than the interconnect in shared memory systems. In PC-based clusters, even simple (and cheap) Ethernet is used. As a result, distributed memory systems can be much larger than shared memory systems. Indeed, systems with tens of thousands of processors are not uncommon.

While distributed memory computers are easier to expand, they are more difficult to use. Each CPU has only access to its own memory, so the access of other memory and communication between CPUs has to be done by explicitly exchanging data. This is called message passing. It is thus necessary to distribute the data in a smart way, as data exchange over the interconnect is slow. Communication should be avoided as much as possible.
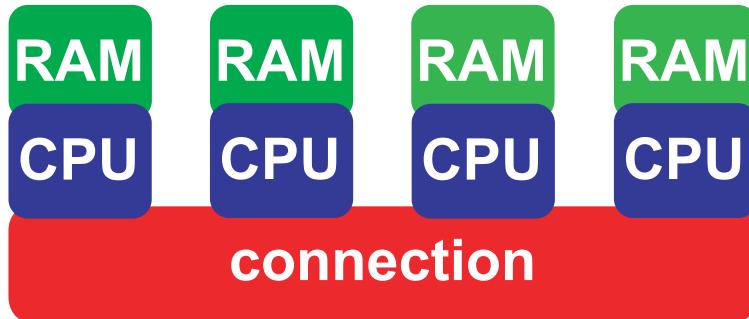

Figure 4.4: Schematic of a distributed memory parallel computer.

Today, most high performance computers are of a hybrid type. They are usually constructed as a cluster of SMP nodes. Each node forms its own shared memory system, which in turn communicates with other nodes over a communication network in the manner of a distributed memory system.

4.3.3 Parallelisation for shared memory computers

Shared-memory parallelisation usually makes use of *multi-threading*. Computations are split up in several threads, which are executed by different CPUs. The easiest way to write multi-threaded programs is the use of OpenMP (OpenMP, 2002) for loop parallelisation. OpenMP needs to be supported by the compiler.

Design matrix computation

The design matrix $\mathbf{A}$ contains $n \times u$ elements for n observations and u unknown parameters. The entries depend on the location of observation ($\mathbf{x}$) and basis function ($\mathbf{y}$), as well as the observation's type and the associated functional model. It is thus natural to arrange the design matrix computation as loop over the observations and basis functions:

```
do i=1,n (loop over observations)
do j=1,u (loop over basis functions)
```
$\mathbf{A}_{ij} = \Phi(\mathbf{x}_i, \mathbf{y}_j, type_i)$
```
end do
end do
```

Since $\mathbf{x}$ and observation type are constant for one line of the design matrix, and a number of expressions evaluated in the computation are constant if all basis functions are located at the same depth, it is numerically more efficient to compute all elements for one line at once (see 4.1.2):

```
do i=1,n
```
$A_{i:} = \Phi(x_i, y_{:}, type_i)$
```
end do
```

Different lines of the design matrix can be computed independently from each other. This makes distributing the observations over the threads a natural choice for parallelisation. It can be achieved with a single OpenMP statement:

```
!$OMP PARALLEL DO
do i=1,n
```
$A_{i:} = \Phi(x_i, y_{:}, type_i)$
```
end do
!$OMP END PARALLEL DO
```

Distribution of the loop iterations i is done automatically by the operating system's scheduler. The number of threads used can be set with the `OPENMP_NUM_THREADS` environment variable. It should be set to the number of CPUs/cores of the parallel computer system.

Computation of normal equation matrix and solution of normal equations

The computation of the normal equation matrix $\mathbf{N} = \mathbf{A}^T\mathbf{A}$ or $\mathbf{N} = \mathbf{A}^T\mathbf{P}\mathbf{A}$ is performed using the DSYRK or DGEMM routines of BLAS (basic linear algebra subprograms, Lawson et al. (1979)). BLAS contains routines for vector-vector (level 1), matrix-vector (level 2) and matrix-matrix (level 3) operations. It is the standard for these linear algebra operations and available in many different implementations. High performance implementations, such as the Goto BLAS (Goto and van de Geijn, 2008), AMD's ACML (AMD Core Math Library) and Intel's MKL (Math Kernel Library) are multi-threaded. Using one of these libraries will automatically lead to a significant acceleration on multi-CPU systems. As a side effect, these libraries are also optimised and much faster than the BLAS reference implementation, even on single-CPU systems.

Solving the linear equation system $\mathbf{N}\hat{\mathbf{x}} = \mathbf{b}$ is done using the DPOSV routine of LA-PACK (linear algebra package) (Anderson et al., 1999). LAPACK defines a set of standards for routines such as solving linear equation systems and eigenvalue computation. Like BLAS, several implementations are available. It makes use of BLAS routines, and the performance is directly affected by the performance of the BLAS library used. Thus, using an optimised multi-threaded BLAS library like the abovementioned Goto BLAS, or the AMD ACML and Intel MKL (which contain, besides BLAS, also LAPACK functionality) results in the desired acceleration.

4.3.4   Parallelisation for distributed memory computers

While parallelisation for shared memory computers is comparatively easy, due to OpenMP and multi-threaded numeric libraries, distributed memory parallelisation is much more of a challenge. Explicit communication (message passing) between the processes is required. Parallelising the design matrix computation in this way is straightforward (by distributing the computations of rows over the processes), but distributing matrices and

doing the resulting distributed computations in an efficient manner is difficult. Luckily, for linear algebra, the distributed memory library ScaLAPACK (Scalable LAPACK) (Blackford et al., 1997) can do this.

ScaLAPACK

ScaLAPACK is a distributed memory implementation of a subset of the LAPACK functionality. It also contains the PBLAS (parallel BLAS), a distributed memory version of BLAS. Both make use of BLACS (basic linear algebra communication subroutines), a wrapper for the actual MPI (message passing interface) or PVM (parallel virtual machine) communication subroutines. MPI and PVM are two standards for communication on distributed memory systems. ScaLAPACK is designed in a way that programs using BLAS and LAPACK can be modified for distributed memory systems with reasonable effort. Much of the message passing is hidden from the user, and he does not have to deal with the intricate details of matrix distribution and communication.

Vectors and matrices are distributed among the notes in a block-cyclic manner. Figure 4.5 shows how a matrix is distributed on a 4x4 process grid.

p_0	p_1	p_2	p_3	p_0	p_1	p_2	p_3	p_0	$\cdots$
p_4	p_5	p_6	p_7	p_4	p_5	p_6	p_7	p_4	$\cdots$
p_8	p_9	p_{10}	p_{11}	p_8	p_9	p_{10}	p_{11}	p_8	$\cdots$
p_{12}	p_{13}	p_{14}	p_{15}	p_{12}	p_{13}	p_{14}	p_{15}	p_{12}	$\cdots$
p_0	p_1	p_2	p_3	p_0	p_1	p_2	p_3	p_0	$\cdots$
p_4	p_5	p_6	p_7	p_4	p_5	p_6	p_7	p_4	$\cdots$
p_8	p_9	p_{10}	p_{11}	p_8	p_9	p_{10}	p_{11}	p_8	$\cdots$
p_{12}	p_{13}	p_{14}	p_{15}	p_{12}	p_{13}	p_{14}	p_{15}	p_{12}	$\cdots$
p_0	p_1	p_2	p_3	p_0	p_1	p_2	p_3	p_3	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Figure 4.5: Distribution of matrix on 4x4 process grid.

Design matrix computation

The computation of the design matrix remains mostly unchanged with ScaLAPACK, except that each process only computes the matrix elements for its parts of the distributed matrix. Assuming process p has r rows and c columns of **A**, and the routine `idxl2g` is used to get the global indices n and m for the local indices i and j, the computation looks like:

```
do i=1,r
n = idxl2g(p,i)
do j=1,c
m = idxl2g(p,j)
```
$\mathbf{A}_{ij} = \Phi(\mathbf{x}_n, \mathbf{y}_m, type_n)$
```
end do
end do
```

Computation of normal equation matrix and solution of normal equations

The computation of the normal equation matrix and solving the linear equation system is changed little from the shared memory version shown above. The BLAS routines DSYRK and DGEMM are replaced by the PDSYRK and PDGEMM from PBLAS. These routines take care of all communication required - the user does have not to worry about an efficient implementation of a distributed matrix-matrix multiplication.

For solving, the PDPOSV routine is used. It replaces the DPOSV routine of LAPACK. All PBLAS and ScaLAPACK routines make use of BLAS routines, so ScaLAPACK performance is governed by the underlying BLAS library, as well as the speed of the communication library (usually an MPI implementation) and network used.

4.3.5   Hybrid parallelisation

The combination of the methods described for shared and distributed memory parallelisation is known as hybrid parallelisation. Since most of today's high performance computers are of a hybrid nature (interconnected shared memory nodes), hybrid parallelisation offers a number of benefits over parallelisation only for distributed memory, i.e. without the use of OpenMP.

- Multi-threading on shared memory nodes does not require the explicit communication by message passing involved in distributed memory parallelisation, and is thus potentially faster. Message passing is only required between different nodes, not CPUs on the same node.

- With distributed memory parallelisation, there is one process per CPU. Each process may need a full set of certain data, such as observations. When processing a year of GRACE data (about 6 million observation epochs, stored in a 320 byte structure per epoch), this amounts to a total 1.8 GB. Using four processes on a four CPU node would thus already require 7.2 GB for observation storage, leaving insufficient memory available on a node equipped with 8 GB of RAM (as used in the research presented here). Using only one multi-threaded process per node severely cuts down memory use, while still allowing the efficient utilisation of all four CPUs.

- In the case of relatively small numbers of unknowns, as is often the case with regional models, distributing the unknowns over all the processes requires very small block sizes. Small block sizes result in inefficient communication. Using fewer processes by employing multi-threading makes it possible to use larger blocks, while still being able to distribute them over all nodes.

Since hybrid programming offers a number of benefits, and shared memory parallelisation is much easier than distributed memory parallelisation (by means of OpenMP and multi-threaded libraries), it is advisable to combine the two in order to get a flexible program that can achieve high performance on different system architectures.

4.3.6 Results of parallelisation

System description

The computations presented here were performed on two high performance computing systems. *Cleopatra* is the DEOS Physical and Space Geodesy group's computing cluster. It consists of 33 compute nodes, each equipped with two dual-core AMD Opteron 280 processors clocked at 2.4 GHz, and 8 GB of RAM. The nodes are interconnected by a fast Infiniband network. Total theoretical peak performance is 633 GFLOPS.

Huygens is the Netherland's national supercomputer. It has 104 nodes, each with 16 dual-core IBM Power 6 processors (4.7 GHz) and 128 GB or 256 GB RAM. It also uses Infiniband as node interconnect. The total theoretical peak performance is 60 TFLOPS. Both systems use the Linux operating system.

Memory requirement

A parallel computer is required not only to shorten program runtimes, but also to fulfill the memory requirements resulting from using a Cholesky solver, which requires assembling both the design matrix $\mathbf{A}$ and the normal equation matrix $\mathbf{N}$. Table 4.1 shows the memory requirements for three typical test cases.

problem	# of observations	# of unknowns	size of $\mathbf{A}$	size of $\mathbf{N}$
monthly GRACE solution	500,000	1,500	5.6 GB	17 MB
yearly GRACE solution	6,000,000	5,000	223.5 GB	191 MB
typical terrestrial data set	300,000	10,000	22.4 GB	763 MB

Table 4.1: Memory requirements for various problem sizes.

It is obvious that the storage of the design matrix $\mathbf{A}$ is the bigger challenge. Even with 10,000 unknowns, the size of $\mathbf{N}$ is less than a Gigabyte. For a one-year global GRACE solution, keeping $\mathbf{A}$ in memory can be difficult. It would require the full memory of 28 nodes of *Cleopatra*, not taking into account memory required for other data. Subtracting 2.5 GB from the available memory per node (the observations alone require 1.8 GB), 41 nodes would be required, whereas only 33 are available. Two solutions to this problem are possible:

1. Build $\mathbf{A}$ in blocks (e.g. only for one month), and compute $\mathbf{N} = \sum \mathbf{A}_i^T \mathbf{A}_i / \mathbf{N} = \sum \mathbf{A}_i^T \mathbf{P}_i \mathbf{A}_i$. This requires the recomputation of $\mathbf{A}$ if it is required more than once.

2. Use an out-of-core approach, where parts of the matrix are stored on hard disks.

Since the out-of-core approach was more difficult to program, and it is questionable whether the slow disk access would provide a speed benefit over the quite quick (re-) computation of $\mathbf{A}$, the block-wise build was implemented.

number of processes	setup of $\mathbf{A}$	$\mathbf{N} = \mathbf{A}^T\mathbf{A}$	solving	total runtime
4	200	101	1	316
8	107	59	2	175
16	49	53	1	120
32	28	47	1	97
64	17	42	4	91

Table 4.2: Runtimes [s] *Cleopatra*, GRACE monthly solution.

Monthly GRACE solution

To show the benefits of parallelisation, a monthly global GRACE solution (April 2004) was computed using various numbers of processes, both on *Cleopatra* and *Huygens*. 518,400 observations and 1,543 RBFs were used. The time required for the setup of the design matrix $\mathbf{A}$, the matrix-matrix multiplication $\mathbf{N} = \mathbf{A}^T\mathbf{A}$, the solving of the linear equation system, and the total program runtime are used to judge the effect of parallelisation.

Table 4.2 shows the runtimes for *Cleopatra*. The time required for the setup of $\mathbf{A}$ scales linearly from 4 to 16 processes. Beyond that, the improvement is a little smaller, but still very good. The matrix-matrix multiplication does not benefit significantly from more than 8 processes due to the small size of $\mathbf{N}$. Unlike the setup of $\mathbf{A}$, more communication is required when more processes are used, resulting in the noticed performance dropoff. The same is true for the solution of the system of linear equations - it is a relatively cheap operation and does not benefit from more processes, but also does not contribute significantly to the total program runtime. The total runtime is governed by the first two operations. As a result, a significant runtime improvement going from 4 to 8 processes can be seen, with smaller improvements beyond that.

The runtimes for the same test on *Huygens* are listed in table 4.3. The time required for the setup of $\mathbf{A}$ scales almost linearly up to 64 processes. Once again, the benefit for the matrix-matrix multiplication is smaller, with almost no benefit for the solution of the system of linear equations. The total runtime improvement is once again a combination of the two, with a almost linear improvement going from 4 to 8 processes, and less beyond that.

Comparing the results to *Cleopatra*, we see that

- The setup of $\mathbf{A}$ requires about the same amount of time on *Huygens* as on *Cleopatra*, while the matrix matrix multiplication is faster on *Huygens*, a consequence of the processors used;

- The matrix-matrix multiplication on *Huygens* benefits more from more processes than on *Cleopatra*, an indication that *Huygens* offers faster inter-process communication.

number of processes	setup of $\mathbf{A}$	$\mathbf{N} = \mathbf{A}^T\mathbf{A}$	solving	total runtime
4	172	48	1	237
8	87	26	1	135
16	46	16	1	83
32	26	12	1	60
64	14	14	1	50

Table 4.3: Runtimes [s] on *Huygens*, GRACE monthly solution.

number of processes	setup of $\mathbf{A}$	$\mathbf{N} = \mathbf{A}^T\mathbf{A}$	solving	total runtime
8	180	459	14	742
16	83	247	6	387
32	44	171	4	258
64	22	157	7	220

Table 4.4: Runtimes *Cleopatra*, terrestrial data, in seconds.

Terrestrial data

Another test was done using a terrestrial data set. The terrestrial data requires more RBFs, resulting in a higher number of unknown parameters compared to the above-mentioned GRACE case, which leads to a greater importance of the performance of the matrix-matrix multiplication. The test was done with the set of Canadian gravity data (320,000 observations) described in section 6.4.2, and 6447 basis functions.

The runtimes for *Cleopatra* are given in table 4.4. The results are very similar to the previous test case. The runtimes for the setup of $\mathbf{A}$ improve almost linearly with an increasing number of processes, while the improvement concerning the matrix-matrix multiplication is less pronounced, and best going from 8 to 16 processes. Having four times the number of unknowns leads to longer runtimes for the linear equation solver, with an associated larger benefit from using more processes. Combined, the total runtime scales better than in the previous case, with of course the biggest improvement going from 4 to 8 processes.

Table 4.5 lists the results for *Huygens*. The results are comparable to the previous test case, with almost linear scaling of the $\mathbf{A}$ setup, a little less improvement for the matrix-matrix multiplication, and little to no improvement of the solver runtime. The total runtime drops significantly, even when going from 32 to 64 processes.

number of processes	setup of $\mathbf{A}$	$\mathbf{N} = \mathbf{A}^T\mathbf{A}$	solving	total runtime
8	166	232	7	442
16	84	119	3	229
32	49	79	2	145
64	27	54	2	94

Table 4.5: Runtimes *Huygens*, terrestrial data, in seconds.

4.4   Summary and conclusions

This chapter has discussed the most important numerical issues associated with the software developed within the research presented here. A number of numerical optimisations have been applied to decrease the required program runtime. Most important of all, parallelisation was employed to decrease the runtimes to acceptable levels.

Hybrid parallelisation, using OpenMP, optimised numerical libraries, and ScaLAPACK has been employed, resulting in a program that delivers good performance on shared memory, distributed memory, and hybrid high performance computer architectures. Even larger computations, involving hundreds of thousands of observations and thousands of unknown basis functions, can be performed in a matter of minutes on these machines.

Chapter 5

Gravity field modelling from satellite data

Global gravity field modelling relies on satellites, as only satellite missions can provide global data coverage. Satellite geodesy was born in 1957 with the launch of Sputnik I. Measurements of Sputnik's orbit allowed the computation of the flattening of the Earth. Other missions like PAGEOS (passive geodynamic satellite) and later LAGEOS (laser geodynamic satellite) made it possible to determine the low SH coefficients of the Earth's gravity field, using optical and laser measurements to determine the satellite orbits. An improvement was provided by radar altimetry missions such as Geosat. Geosat was launched in 1986, and Geosat and ERS-1 data were used for estimation of the EGM96 gravity field model (Lemoine et al., 1998). Altimetry provides high resolution, but is of course limited to measurements over sea.

A major step forward was the CHAMP (challenging minisatellite payload) satellite, launched in 2000. A flight altitude of approx. 450 km, GPS-based orbit determination and an onboard accelerometer made it possible to use CHAMP to compute static gravity fields complete to spherical harmonic degree 70 (Reigber et al., 2002; Han et al., 2003; Gerlach et al., 2003; Ditmar et al., 2006).

Currently, research focuses on the GRACE (gravity recovery and climate experiment) satellite mission. GRACE is a formation-flying mission consisting of two identical satellites in the same near-polar orbit spaced approx. 220 km apart at 500 km altitude (fig. 5.1). The satellites were launched in 2002. Besides being equipped with GPS receivers and accelerometers, the inter-satellite distance is continuously measured with a microwave (k-band) link, down to micrometer accuracy. This yields a significant improvement in terms of resolution and accuracy of the derived gravity field models (Tapley et al., 2004).

Static GRACE gravity fields have been computed by GFZ Potsdam, both GRACE-only (Reigber et al., 2005) and combined fields that also made use of terrestrial data (Foerste et al., 2008,b), and by CSR (Tapley et al., 2005). Static GRACE-only gravity fields have been computed up to degree 180 (Mayer-Gürr, 2007). GRACE had an even higher impact on modelling and understanding the time-variability of gravity. Monthly GRACE solutions are computed by CSR (Bettadpur, 2007), GFZ (Flechtner, 2007), and
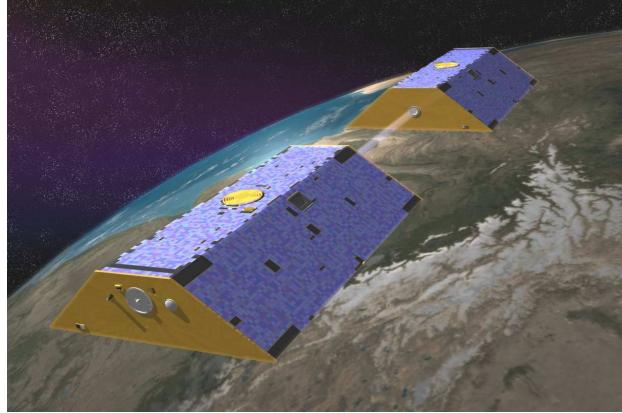
Figure 5.1: Artist's impression of the GRACE satellites. Source: NASA

JPL (Watkins and Yuan, 2007). DEOS optimally filtered monthly GRACE gravity fields are computed up to degree 120 (Liu, 2008; Klees et al., 2008b). Computing 10-day solutions is possible and regularly done by CNES, albeit at lower resolution (Biancale et al., 2007).

Global satellite-based gravity field modelling will make another step forward with the GOCE (gravity field and steady-state ocean circulation explorer) mission. An advanced spacecraft, flying at only 250 km altitude and using an ion engine to counter atmospheric drag, it will use a highly accurate gradiometer to measure gravity gradients. The measurements are expected to make it possible to accurately compute static gravity models up to degree 200 and beyond. GOCE was launched in March 2009.

The methodology described here can be applied to all satellite gravity missions. The results presented here focuses exclusively on the GRACE mission, as it provides the best data currently available. GRACE makes it possible for the first time to compute monthly high-resolution gravity fields that clearly show mass variations in the Earth system, which yield new insights into mass transport processes at a global scale. The computations presented here will also show monthly solutions exclusively, and use these monthly solutions to derive mass transport quantities.

The chapter starts with a description of the functional model used for gravity recovery from GRACE. Section 5.2 describes the stochastic model that was used for some of the computations. The extension of the optimal filtering algorithm developed at DEOS to RBF solutions is explained in section 5.3. Network design and bandwidth estimation are addressed in sections 5.4 and 5.5, respectively.

Results of GRACE gravity field modelling are presented in section 5.6. Two main applications of GRACE monthly models are used as representative examples:

1. The estimation of ice mass loss in Greenland and Antarctica, including a comparison of GRACE and ICESat results.

2. The estimation of water storage variations at river basin scale. Time-series for several river basins are shown. Comparisons are made between various monthly GRACE models as well as between GRACE and the PCR-GLOBWB hydrological model. A comparison to radar altimetry is made for Lake Victoria in Africa.
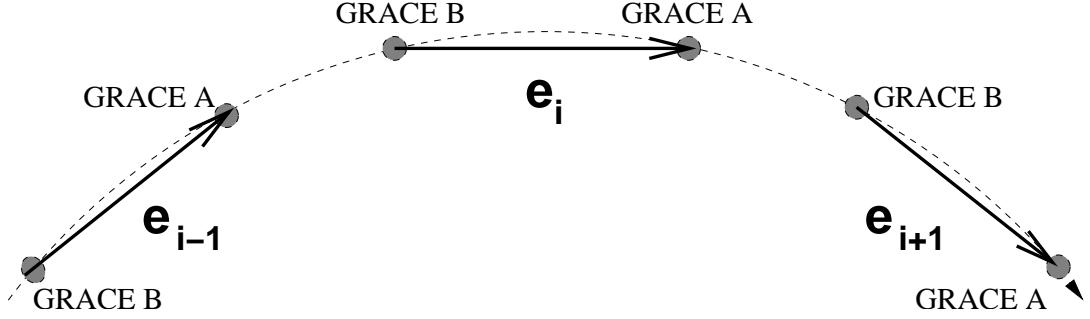
Figure 5.2: Unit vectors of LoS directions at three successive epochs.

5.1 Functional model

5.1.1 Three-point range combination approach

All GRACE gravity field models were computed using the DEOS 3-point range combination approach. This allows the direct comparison of SH and RBF solutions that use identical processing except for the parametrisation. The approach is sketched here. For more details, the reader should refer to Liu (2008).

GRACE provides biased inter-satellite ranges ρ using the k-band link. These ranges can be related to the gravity potential gradient. Using a three-point scheme, the inter-satellite range vectors of three epochs, $\mathbf{r}_{i-1}$, $\mathbf{r}_i$, and $\mathbf{r}_{i+1}$ can be used to compute the average 3D inter-satellite acceleration:

$$\ddot{\mathbf{r}}_i = \frac{\mathbf{r}_{i-1} - 2\mathbf{r}_i + \mathbf{r}_{i+1}}{(\Delta t)^2}, \tag{5.1}$$

where Δt is the sampling interval. Since $\mathbf{r}_k = \mathbf{e}_k \cdot \rho_k$, with $k = i-1, i, i+1$ and $\mathbf{e}_k$ the line of sight (LoS) unit vectors at epoch k,

$$\mathbf{e}_k = \frac{\mathbf{r}_k}{\rho_k}. \tag{5.2}$$

Eq. (5.1) can be written as

$$\ddot{\mathbf{r}}_i = \frac{\mathbf{e}_{i-1} \cdot \rho_{i-1} - 2\mathbf{e}_i \cdot \rho_i + \mathbf{e}_{i+1} \cdot \rho_{i+1}}{(\Delta t)^2}. \tag{5.3}$$

By projecting both sides of eq. (5.3) onto the unit vector $\mathbf{e}_i$, we obtain

$$\mathbf{e}_i \cdot \ddot{\mathbf{r}}_i = \frac{\xi_{i-1} \cdot \rho_{i-1} - 2\rho_i + \xi_{i+1} \cdot \rho_{i+1}}{(\Delta t)^2}, \tag{5.4}$$

where $\xi_{i-1} = \mathbf{e}_i \cdot \mathbf{e}_{i-1}$ and $\xi_{i+1} = \mathbf{e}_i \cdot \mathbf{e}_{i+1}$, see figure 5.2.

In the actual computation, residual quantities are used. They follow from eq. (5.3) by linearisation:

$$\delta\ddot{\mathbf{r}}_i = \ddot{\mathbf{r}}_i - \ddot{\mathbf{r}}_c$$

$$\delta\rho_i = \rho_i - \rho_c \tag{5.5}$$

$$\mathbf{e}_{i,c} \cdot \delta\ddot{\mathbf{r}}_i = \frac{\xi_{i-1,c} \cdot \delta\rho_{i-1} - 2\delta\rho_i + \xi_{i+1,c} \cdot \delta\rho_{i+1}}{(\Delta t^2)}, \tag{5.6}$$

with $\ddot{\mathbf{r}}_c$, ρ_c, $\mathbf{e}_{i,c}$, $\xi_{i-1,c}$, and $\xi_{i+1,c}$ being reference quantities computed from precise reference orbits. $\mathbf{e}_k$ and ρ_k cannot be obtained from the k-band measurements, but need to be computed from orbits instead. The use of small residual quantities reduces errors in the functional model. The static EIGEN-GL04C model (Foerste et al., 2008) up to degree 150 has been been used as background gravity field model for the computation of the reference orbits.

5.1.2   Residual accelerations

Analytical expressions for the radial base functions used here have been derived as

$$\Psi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi R^2} \left(2\chi_{n+1} + \chi_n\right),$$
$$\chi_n = \left(|\mathbf{y}| \frac{\partial}{\partial |\mathbf{y}|}\right)^n \frac{1}{|\mathbf{x} - \mathbf{y}|}. \tag{5.7}$$

The χ_n can be computed by

$$\chi_n = 2 |\mathbf{y}|^{n+1} b_{n+1} + \sum_{i=1}^{n} \beta_{n,i} |\mathbf{y}|^i b_i \tag{5.8}$$

for $n \geq 1$.

The b_n have been derived as

$$b_n = (2n - 1) |\mathbf{x} - \mathbf{y}| b_1 b_{n-1} - (n - 1)^2 b_0^2 b_{n-2}, \tag{5.9}$$

which can be computed recursively with

$$b_0 = \frac{1}{|\mathbf{x} - \mathbf{y}|}, \, b_1 = \frac{|\mathbf{y}| - |\mathbf{x}| \cos\gamma}{|\mathbf{x} - \mathbf{y}|^3}, \tag{5.10}$$

where γ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

The residual accelerations in eq. (5.6) are related to the gravity potential gradient. The gradient is defined as

$$grad\, T = \frac{\partial T}{\partial x_1} + \frac{\partial T}{\partial x_2} + \frac{\partial T}{\partial x_3}. \tag{5.11}$$

To compute the gradient of the RBF in eq. (5.7), we need the partial derivatives of the coefficients b_n with respect to the RBFs. Applying the product rule to eq. (5.9), we find

$$
\begin{aligned}
\frac{\partial b_n}{\partial x_i} = {} & (2n-1)\frac{\partial |\mathbf{x}-\mathbf{y}|}{\partial x_i}b_1 b_{n-1} + (2n-1)|\mathbf{x}-\mathbf{y}|\left(\frac{\partial b_1}{\partial x_i}b_{n-1} + b_1\frac{\partial b_{n-1}}{\partial x_i}\right) \\
& -(n-1)^2\left(2b_0\frac{\partial b_0}{\partial x_i}b_{n-2} + b_0^2\frac{\partial b_{n-2}}{\partial x_i}\right).
\end{aligned}
\tag{5.12}
$$

The partial derivatives $\frac{\partial b_n}{\partial x_i}$, $\mathbf{x} = (x_1, x_2, x_3)$ are required:

$$
\frac{\partial b_0}{\partial x_i} = -\frac{x_i - y_i}{|\mathbf{x}-\mathbf{y}|^3}
\tag{5.13}
$$

and

$$
\frac{\partial b_1}{\partial x_i} = \frac{3(x_i - y_i)|\mathbf{y}|}{|\mathbf{x}-\mathbf{y}|^5} + \frac{1}{|\mathbf{y}|}\left(\frac{y_i}{|\mathbf{x}-\mathbf{y}|^3} - 3\frac{\mathbf{x}^T\mathbf{y}(x_i - y_i)}{|\mathbf{x}-\mathbf{y}|^5}\right),
\tag{5.14}
$$

as well as

$$
\frac{\partial |\mathbf{x}-\mathbf{y}|}{\partial x_i} = \frac{x_i - y_i}{|\mathbf{x}-\mathbf{y}|}.
\tag{5.15}
$$

Using eq. (5.7) to (5.15), the point-wise design matrices $\mathbf{A}_{pw,A}$ for satellite A and $\mathbf{A}_{pw,B}$ for satellite B are computed. The design matrices computed in this manner are defined in the terrestrial reference frame (TRF) of $\mathbf{x}$ and $\mathbf{y}$. The observation in eq. (5.6) is the averaged residual acceleration in the LoS direction between the two satellites. The final design matrix is

$$
\mathbf{A} = \mathbf{R}_{ll}\left(\mathbf{A}_{pw,B} - \mathbf{A}_{pw,A}\right),
\tag{5.16}
$$

with $\mathbf{R}_{ll}$ being the transformation matrix that links the point-wise accelerations in the TRF to the averaged observations in the LoS frame. The following transformation steps are required:

1. A rotation from the TRF to the celestial reference frame (CRF).

2. Three-point averaging, since averaged accelerations are used.

3. A rotation from the CRF to the LoS between the satellites.

$\mathbf{R}_{ll}$ can be thus written as

$$
\mathbf{R}_{ll} = \mathbf{R}^{(C \to LoS.x)}\mathbf{E}\mathbf{R}^{(T \to C)},
\tag{5.17}
$$

where $\mathbf{R}^{(T \to C)}$ is the rotation matrix from TRF to CRF, and $\mathbf{R}^{(C \to LoS.x)}$ is the matrix that projects a 3D-vector defined in the CRF onto the x-axis of the LoS vector, which is directed along the line-of-sight (LoS) (Liu, 2008). $\mathbf{E}$ is the averaging filter that computes the averaged three-point accelerations (Ditmar and van Eck, 2004).

5.1.3   Equivalent water heights

Spatial plots of GRACE monthly gravity models and associated mass balances are usually not provided in terms of potential, but in terms of equivalent water height (EWH), also known as equivalent water layer thickness (Wahr et al., 1998).

A representation in terms of EWH is computed from a spherical harmonic expansion by

$$\delta h_w(\mathbf{x}) = R_E \frac{\rho_e}{3\rho_w} \sum_{l=0}^{\infty} \sum_{m=0}^{l} \frac{2l+1}{1+k_l'} \bar{c}_{lm} \bar{Y}_{lm}(\mathbf{x}), \tag{5.18}$$

where R_E is the Earth's mean radius, ρ_e is the Earth's mean density (5,500 $\frac{kg}{m^3}$), ρ_w is the density of water (1,000 $\frac{kg}{m^3}$), and k_l' are the load Love numbers. Inserting eq. (2.26) into eq. (5.18) and using eq. (2.23), we find

$$\delta h_w(\mathbf{x}) = R_E \frac{\rho_e}{3\rho_w} \sum_{l=0}^{\infty} \frac{2l+1}{1+k_l'} \psi_l \sum_{i=1}^{n} \alpha_i P_l(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i). \tag{5.19}$$

It is not possible to use the analytical representations for the basis functions, as only a spectral representation of the functional model for equivalent water heights can easily be used. Luckily, the necessary maximum summation degree l is quite low. Load Love numbers up to degree 1024 were used (Petrov and Boy, 2004), which is sufficient for the deeply based RBFs used for satellite gravity field modelling. If runtimes become an issue, an interpolation scheme as described in section 4.2 can be used.

5.1.4   Trend and signal amplitude estimation

This chapter uses not only single solutions, but also trend and yearly signal amplitude estimates from time series of solutions. The functional model for the trend and annual signal amplitude estimate is

$$\delta h(t) = a + bt + c\sin 2\pi t + d\cos 2\pi t, \tag{5.20}$$

where $\delta h(t)$ is the mass variation at point t in time (measured in years), a is the bias, b is the linear yearly trend, and $\sqrt{c^2 + d^2}$ is the amplitude of the annual signal.

Trends and signal amplitudes can be estimated for basin averages, but also per point. The latter yields a spatial map of linear mass changes and yearly mass variations.

5.2   Stochastic model

It is assumed that the noise in GRACE range observations is white and Gaussian. Hence, the noise in the three-point range combinations is coloured, i.e. frequency dependent.

Most of the noise energy is concentrated at the high frequencies. The frequency behaviour of the noise can be approximated by a model (Ditmar et al., 2007), which approximates the square root of the noise power spectral density $u(f)$:

$$\sqrt{u(f)} = \frac{2\sigma}{(\Delta t)^{\frac{3}{2}}} \left(1 - \cos(2\pi f \Delta t) + \left(\frac{\Delta t}{\tau} \right)^2 \right), \tag{5.21}$$

where σ is the standard deviation of the k-band ranges, Δt is the sampling rate, and τ is the filter halfwidth. For the computation of the DEOS global models, the values used were $\sigma = 40 \, \mu m$, $\Delta t = 5 \, s$, and $\tau = 30 \, s$ (Liu, 2008).

It has been shown by Liu (2008) that in the case of the 3-point range combination approach, use of the stochastic model in eq. (5.21), instead of the assumption of white noise, does not lead to significant differences up to spherical harmonic degree 40, and small differences (about 5%) up to degree 120. Nevertheless, application of the stochastic model may be desirable or necessary to get realistic estimates for the noise covariance matrix $\mathbf{C}_{xx} = \mathbf{N}^{-1}$ and the a-posteriori variance $s^2 = \frac{\mathbf{e}^T \mathbf{P} \mathbf{e}}{n-u}$, especially in the context of optimal filtering, as discussed in section 5.3.3.

Figure 5.3 shows the differences between solutions computed with and without this frequency-dependent weighting. A monthly solution with and without frequency-dependent weighting can show significant differences in certain areas. Point differences are below one cm for yearly signal amplitudes and yearly trends. It can be concluded that frequency-dependent weighting will have little to no impact when computing trends and yearly signal amplitudes, but may significantly affect monthly solutions. When computing basin averages, differences will be even smaller as can be seen in section 5.6.

Using frequency-dependent weighting leads to a considerable increase in the required computation time, so it is best disabled for large computations that need to be made quickly.

5.3 Optimal filtering

5.3.1 Introduction

As will be seen in section 5.4, GRACE solutions are strongly affected by noise. Due to the GRACE orbit and its resulting high sensitivity to unaccounted for effects in the north-south directions, noise appears as north-south "stripes". Methods to reduce this noise are the truncation of the spherical harmonic expansion at a very low degree, simple Gaussian smoothing of the GRACE models (Wahr et al., 1998), or "de-striping" techniques, which are often combined with Gaussian smoothing (Swenson and Wahr, 2006).

None of these methods are entirely satisfactory, leading to decreases in spatial resolution and loss of signal. Another approach is the use of so-called "optimal" filters, which require information about the signal. Such an optimal filter has been developed in Klees et al. (2008b). The optimal filter is a Wiener filter, i.e. it minimises the global mean of the mean square difference between the unfiltered and filtered signal. It is both anisotropic
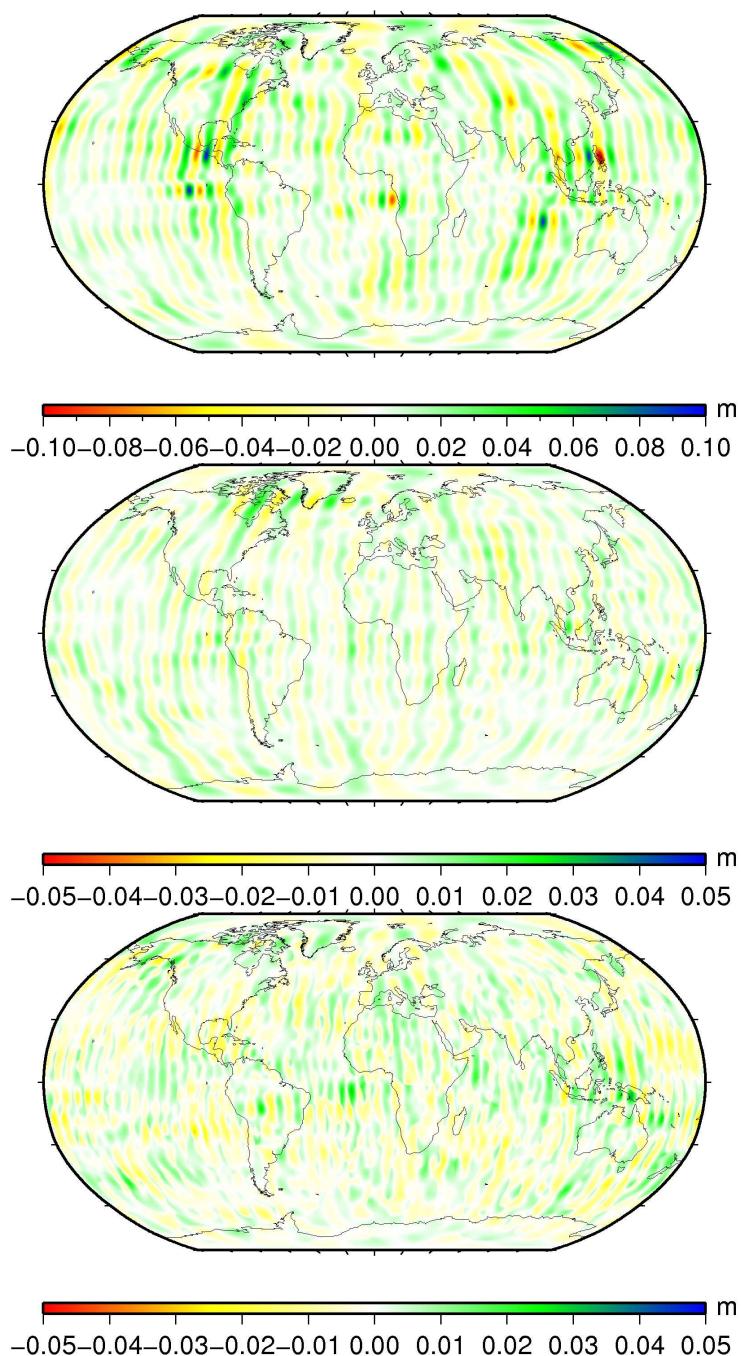
Figure 5.3: Differences between level 35 solutions (RBFs at 1600km depth) with and without frequency-dependent weighting; April 2005 (top), yearly trend (centre), and yearly signal amplitude (bottom), in EWH. Note different colorbar scales.

and non-symmetric and thus capable of removing noise artifacts of arbitrary shape, such as stripes.

The filter can either be applied as post-processing filter or as regularisation-type filter. In the latter case, application of the filter is done by

$$\hat{\mathbf{x}} = \left(\mathbf{N} + \mathbf{D}^{-1} \right)^{-1} \mathbf{b}, \tag{5.22}$$

where $\mathbf{D}$ is the signal covariance matrix, $\mathbf{N}$ is the normal matrix, and $\mathbf{b}$ is the right-hand-side vector. The inverse signal covariance matrix is thus applied as regularisation matrix. This means that

- strong regularisation, resulting in a smooth field, will be applied in areas of small signal to noise ratio;

- weak regularisation will be applied in areas of large signal to noise ratio.

5.3.2 Signal covariance matrix computation

An important factor governing the performance of a filter constructed as in eq. (5.22) is the signal covariance matrix which is used as regularisation matrix. It would be possible to construct the signal covariance matrix based on an a-priori model, such as a global hydrological model, but such a matrix would be biased towards this model. Instead, the DEOS approach computes the signal covariance matrix in an iterative way using no a-priori model (Klees et al., 2008b).

The signal covariance matrix $\mathbf{D}_f$ is defined in space, in units of EWH. Variances of points $\mathbf{y}_i$ on an equidistant Reuter grid (section 2.6.1) are computed using n realisations t (monthly solutions)

$$\sigma_f^2 \left(\mathbf{y}_i \right) = \frac{1}{n-1} \sum_{t=1}^{n} \left(\mathbf{f}_t \left(\mathbf{y}_i \right) - \bar{\mathbf{f}} \left(\mathbf{y}_i \right) \right)^2 \tag{5.23}$$

with the signal in EWH

$$\mathbf{f}_t = \mathbf{B} \hat{\mathbf{x}}_t. \tag{5.24}$$

Signal correlations at the grid points are omitted, leading to a diagonal matrix $\mathbf{D}_f$, which needs to be transformed into the RBF domain. Since $\mathbf{B}$ cannot be inverted (it is not square), the pseudoinverse

$$\Gamma = \left(\mathbf{B}^T \mathbf{B} \right)^{-1} \mathbf{B}^T \tag{5.25}$$

is used. The transformation is then computed by

$$\mathbf{D}_\alpha = \Gamma \mathbf{D}_f \Gamma^T. \tag{5.26}$$

The iterative computation of the signal covariance matrix $\mathbf{D}_\alpha$ in the RBF basis is shown in figure 5.4.

1. An initial signal covariance matrix $\mathbf{D}_f$ is constructed using some arbitrary starting values, e.g. a signal variance of 25 cm^2 at each grid point.

2. At the start of each iteration, the signal covariance matrix $\mathbf{D}_f$ is transformed into the domain of the RBFs (RBF analysis).

3. Regularised solutions $\hat{\mathbf{x}}_t$ are computed for the whole available GRACE time series.

4. The resulting signal in terms of EWH is computed on the same equidistant Reuter grid as used in step 1, by $\mathbf{f}_t = \mathbf{B}\hat{\mathbf{x}}_t$ (RBF synthesis).

5. The signal variance in each point is computed (eq. 5.23) and used to populate $\mathbf{D}_f$ for the next iteration.

6. The computation is stopped as soon as some stopping criterion, such as sufficiently small difference (e.g. <3 cm) between the signal covariance matrices of two iterations, is reached. In all computations, not more than ten iterations were sufficient.

It has been shown by Klees et al. (2008b) that the spectrum of the signal computed from the signal covariance matrix $\mathbf{D}_f$ has almost constant signal degree variances. This was overcome by dividing the degree variances by the spherical harmonic degree l^2, which can be written as the application of an additional scaling matrix $\mathbf{S}$:

$$\mathbf{D}_{SH} = \mathbf{S}\Gamma\mathbf{D}_f\Gamma^T\mathbf{S}. \tag{5.27}$$
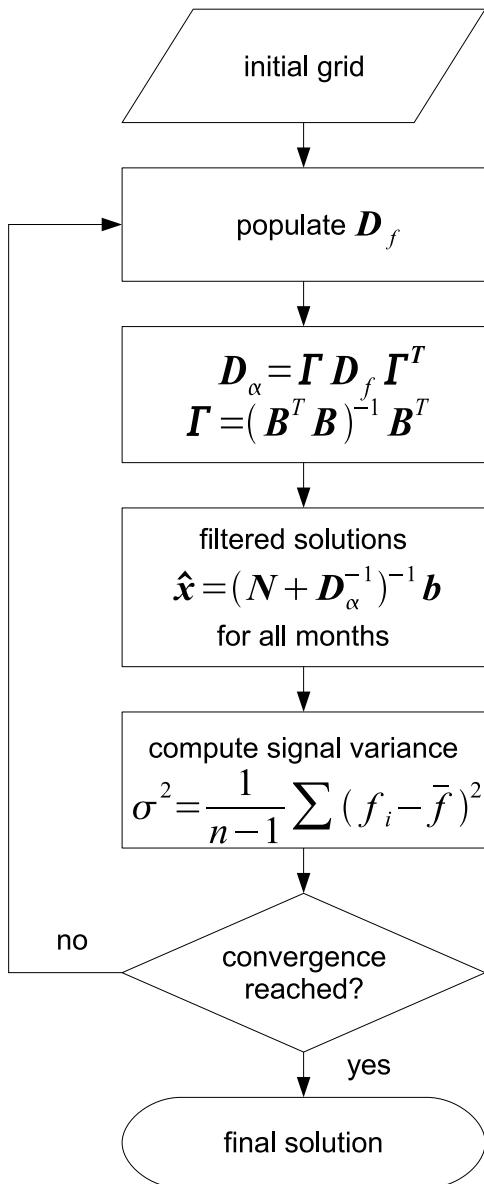
$\mathbf{S}$ is diagonal with diagonal elements $\frac{1}{l}$. Applying spectral scaling in this way is however not possible with RBFs. For the RBFs, the scaling has to be applied in the functional model for EWHs (eq. 5.28) when computing the design matrix $\mathbf{B}$:

$$\delta h_{w,scaled}(\mathbf{x}) = R_E \frac{\rho_e}{3\rho_w} \sum_{l=0}^{\infty} l\frac{2l+1}{1+k_l'}\psi_l \sum_{i=1}^{n} \alpha_i P_l(\hat{\mathbf{x}}^T\hat{\mathbf{y}}_i). \tag{5.28}$$

Note that scaling is performed with l instead of $\frac{1}{l}$, since Γ is the inverse of $\mathbf{B}$.

Figure 5.5 shows the signal amplitudes (the square root of the signal variances in $\mathbf{D}_f$) in the space domain after 10 iterations computed using a level 90 RBF solution. Large signal variations are detected in areas such as the Amazon river basin, Alaska, Greenland, and parts of Antarctica. These are due to seasonal hydrological signals or ice melting. Two spots are visible in the ocean around Indonesia: Mass displacement due to the Sumatra Earthquake in December of 2004 (Chen et al., 2007).

The diagonal elements of the regularisation matrix $\mathbf{D}_\alpha^{-1}$ can be plotted in space, at the location of their respective RBF. Such a plot the signal covariance matrix in figure 5.5 is shown in figure 5.6. It shows an inverted image compared to figure 5.5. Strong

Figure 5.4: Flowchart of iterative D_f computation
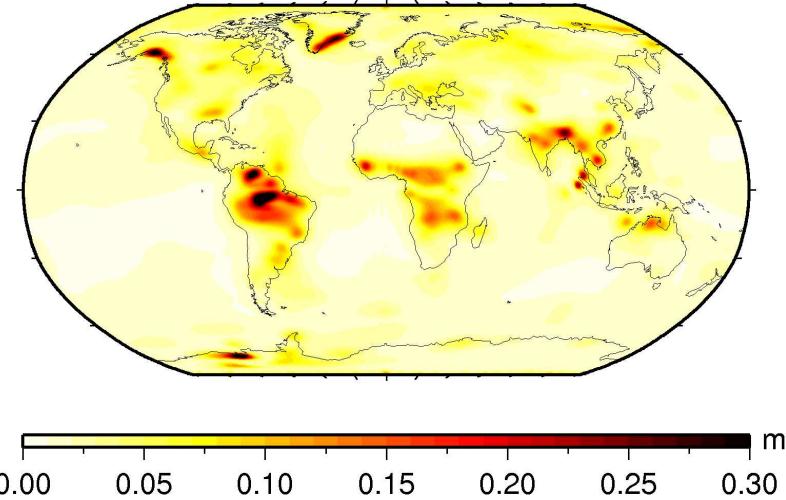
Figure 5.5: Signal amplitudes in units of EWH. These are the square root of the signal variances in $\mathbf{D}_f$.
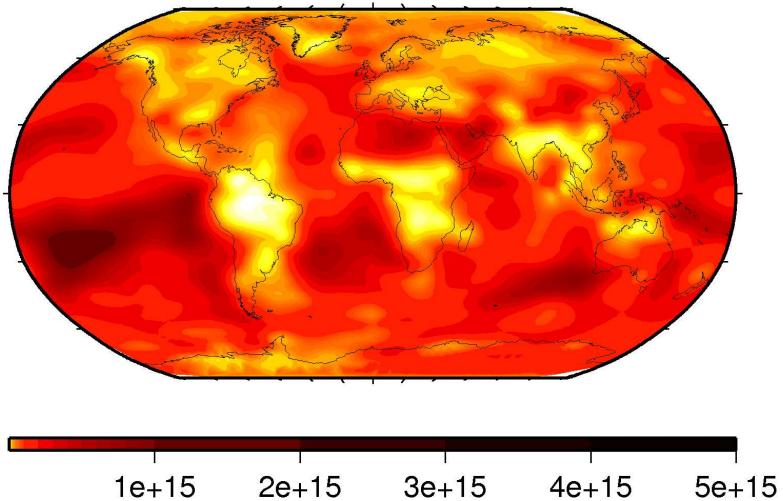

Figure 5.6: Diagonal elements of the regularisation matrix $\mathbf{D}_\alpha^{-1}$ in the space domain.

regularisation is applied in regions of small signal to noise ratio, and weak regularisation in areas that have been identified as area of large signal to noise ratio.

The iterative computation of the optimally filtered solutions constitutes a considerable numerical effort, because

- the design matrix $\mathbf{A}$ requires the spectral representation of the basis functions in eq. (5.28);

- the computation of Γ and $\mathbf{D}_\alpha$ requires several matrix-matrix multiplications;

- the normal matrix $\mathbf{N}$ has to be built for each month and iteration, if it is not stored.

A major acceleration can be achieved when using the following scheme:

1. Normal matrices **N** and right-hand side vectors **b** are computed once for every month, using no regularisation, and stored on disk. Each process stores only its part of **N** and **b**, preferably on a fast (local) disk.

2. The transformation matrix Γ is computed only once at the beginning of the iterative computation, and kept in memory.

3. The design matrix **A**, required for synthesis of the signal on the spatial grid, is also computed only once and kept in memory.

4. The signal covariance matrix $\mathbf{D}_\alpha$ is computed once at the beginning of each iteration.

5. Each regularised monthly solution is then computed by reading **N** and **b**, applying $\mathbf{D}_\alpha$, and solving $(\mathbf{N} + \mathbf{D}_\alpha)\hat{\mathbf{x}} = \mathbf{b}$, which is a very fast operation, i.e. takes only seconds.

In this manner, computing one iteration takes (for a typical parametrisation, depending on the number of processors used) only a few minutes instead of several hours. The time-consuming operation of building design matrices for EWH is only executed once, and matrix-matrix multiplications are necessary only once per iteration. For each monthly solution, operations are limited to reading a very small part of a matrix from disk, adding $\mathbf{D}_\alpha$, and solving the regularised normal equations.

5.3.3 Noise level estimation

The inverse signal covariance matrix is applied as regularisation matrix (eq. 5.22). In order to get correct results, proper relative scaling of normal equation matrix **N** and regularisation matrix $\mathbf{D}_\alpha^{-1}$ is required. Since $\mathbf{N} = \mathbf{A}^T \frac{1}{\sigma^2}\mathbf{P}\mathbf{A}$, this means that the correct weight matrix **P** and the proper noise level σ of the observations needs to be taken into account. Too large values of σ will result in too much smoothing and the loss of signal; too small values of σ result in insufficient smoothing and the modelling of noise.

When frequency-dependent observation weighting (see 5.2) is used, the noise level is set using the σ of the k-band ranges in eq. (5.21). Initially, $\sigma = 40\,\mu m$ as with the DMT-1 model (Klees et al., 2008b) was used. Comparing the a-priori noise level of the observations to the noise level obtained from the post-fit residuals, a value of $\hat{\sigma} = 18\,\mu m$ was obtained and used for all computations. Monthly solution with $\sigma = 40\,\mu m$, $\sigma = 18\,\mu m$, and $\sigma = 13\,\mu m$) are shown in figure 5.7. As can be expected, solutions computed with a smaller σ show more noise, but also more signal and apparently improved spatial resolution.

Comparing a-priori and a-posteriori noise levels is not possible when no frequency-dependent weighting is used, thus when white noise is assumed. The computed a-posteriori noise levels are then dominated by high-frequency noise and would lead to much too large values of σ. Instead, other methods of determining the noise level are required:
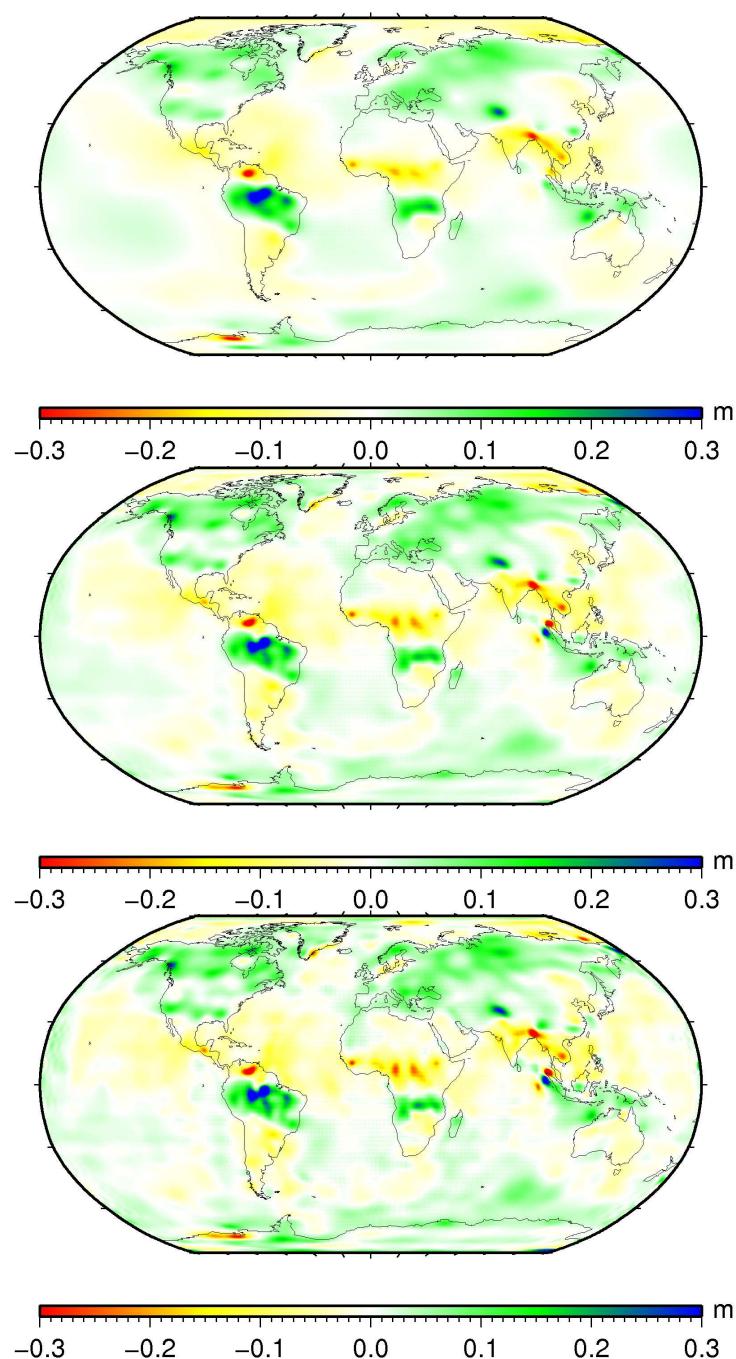
Figure 5.7: Level 90 @ 700 km monthly solutions for April 2005 with frequency-dependent weighting, in EWH; k-band range noise-level was set to 40 μm (top), 18 μm (centre), and 13 μm (bottom).

An unscientific approach is "guesstimating" the proper noise level by looking at plots of solutions and judging, based on the noisiness of the plots, whether the chosen noise level was correct. A better method is the validation by some external information. Used here as described in Stolk (2009) was the signal above the no-precipitation zone in eastern Antarctica . The noise level was chosen in such a manner that the resulting signal above this signal-free area corresponded to the noise level that can be expected from GRACE (Liu et al., 2009).

Working with and without frequency-dependent weighting yields different results. Both methods use different ways of computing σ and result in different covariance matrices **C**. Figure 5.8 shows the differences in terms of a monthly solution, yearly trends, and yearly signal amplitudes. Significant differences do exist, but are restricted to certain geographic areas. It is difficult to say which solution is "better" without taking a closer look, which will be done in section 5.6.

5.4 RBF network design

5.4.1 Grid choice

Satellite-derived observations contain mostly low-wavelength information. It it thus not necessary to use a very dense grid of basis functions to recover this information. Since satellite methods deliver essentially global data coverage, we often want to model the gravity field globally. There are also regional applications in high-latitude areas (e.g. ice balance estimation for Antarctica and Greenland). Both of these require a uniform basis function distribution over the whole globe. A grid with increasingly dense spacing at high latitudes has to be avoided to prevent overparametrisation and numerical instabilites. Possible grid candidates were discussed in section 2.6. All computations presented in this chapter make use of the Reuter grid.

Unfiltered solutions

The residual GRACE quantities used as observations are contaminated by strong high-frequency noise. This severely limits the amount of usable information that can be retrieved from an unfiltered monthly solution. From numerical experiments, it can be said that a Reuter grid of level 35 is very suitable for the computation of monthly models without filtering. Higher-level grids model too much noise. This is visible in figure 5.9. The same observations were used for both solutions. The first solution was computed using basis function on a level 35 grid at a depth of 1,600 km, while the second solution used a level 45 grid and a depth of 1,100 km. The depths were found using the algorithm described in section 2.5.2. The higher-level solution contains no usable extra information, but is distorted by stripes, which is an indication that too much noise is modelled.

This behaviour also holds for trends (fig. 5.10) and yearly signal amplitudes (fig. 5.11) computed from a four-year time series (2003-2006). The level 45 solution shows little to
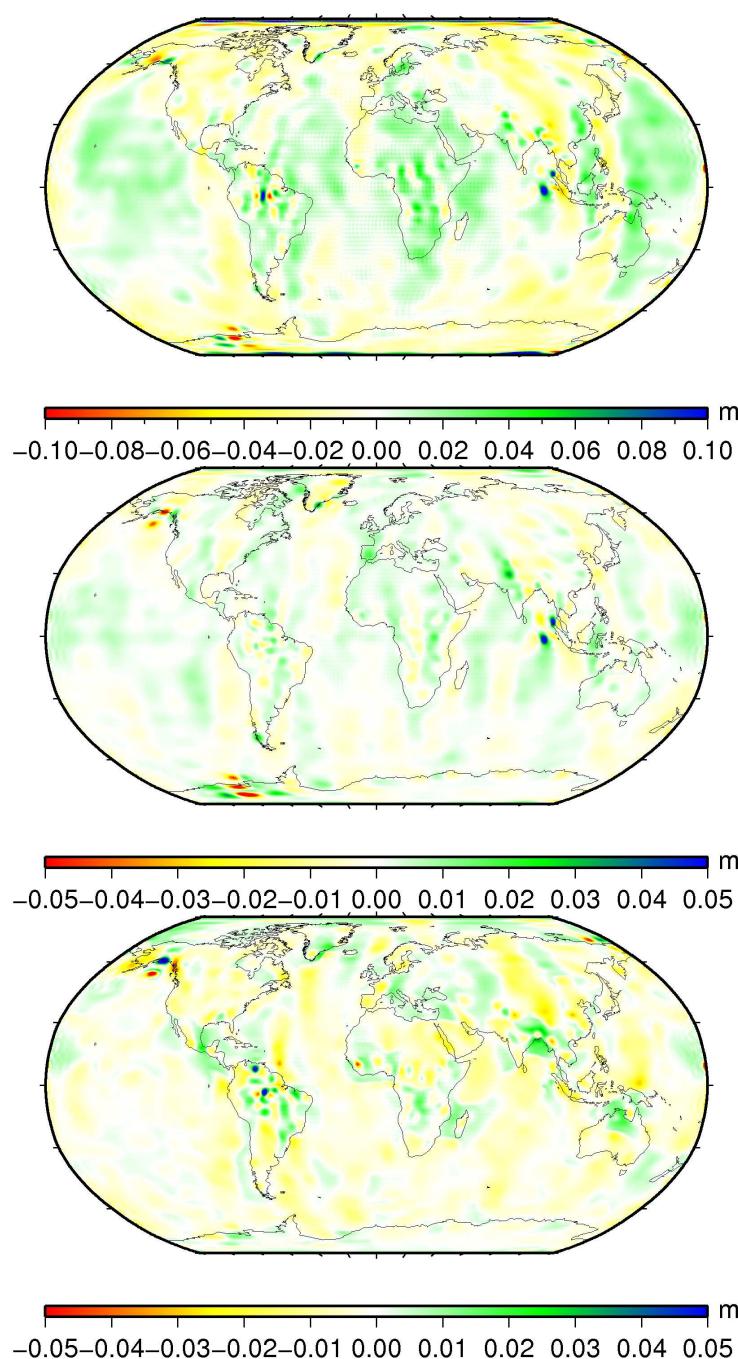
Figure 5.8: Differences between level 90 @ 700 km solutions with and without frequency-dependent weighting, in EWH; April 2005 (top), yearly trend (centre), and yearly signal amplitude (bottom).
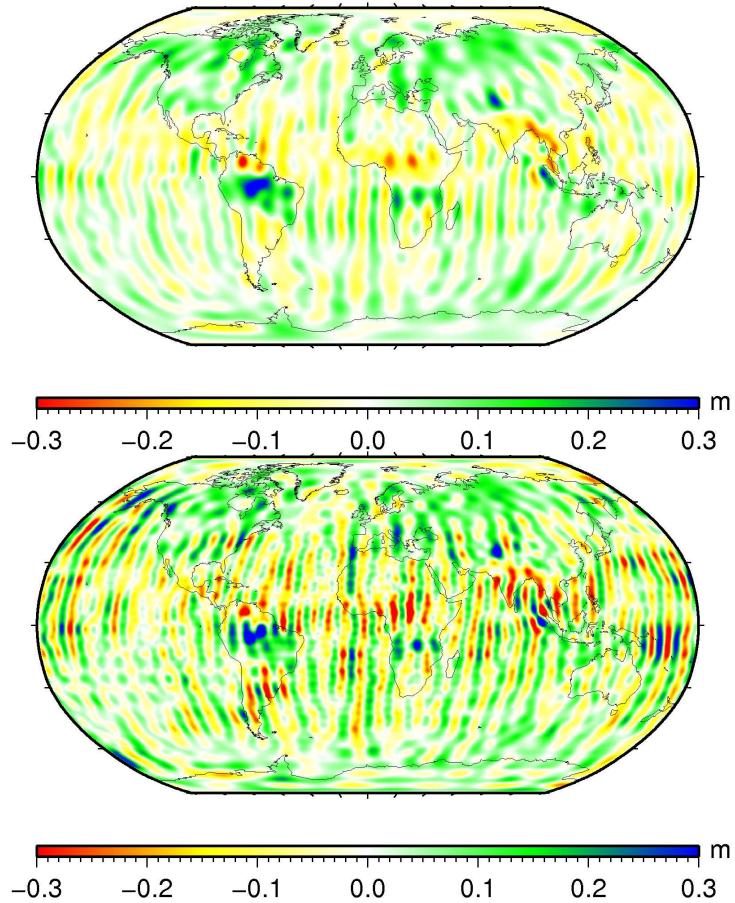
Figure 5.9: Unfiltered monthly solutions, level 35 @ 1,600 km (top) and level 45 @ 1,100 km (bottom), in EWH.

no extra information, just more noise. It is visible that errors decrease at higher latitudes. Grid levels higher than 35 may thus be of interest when focusing on high-latitude areas such as Greenland and Antarctica.
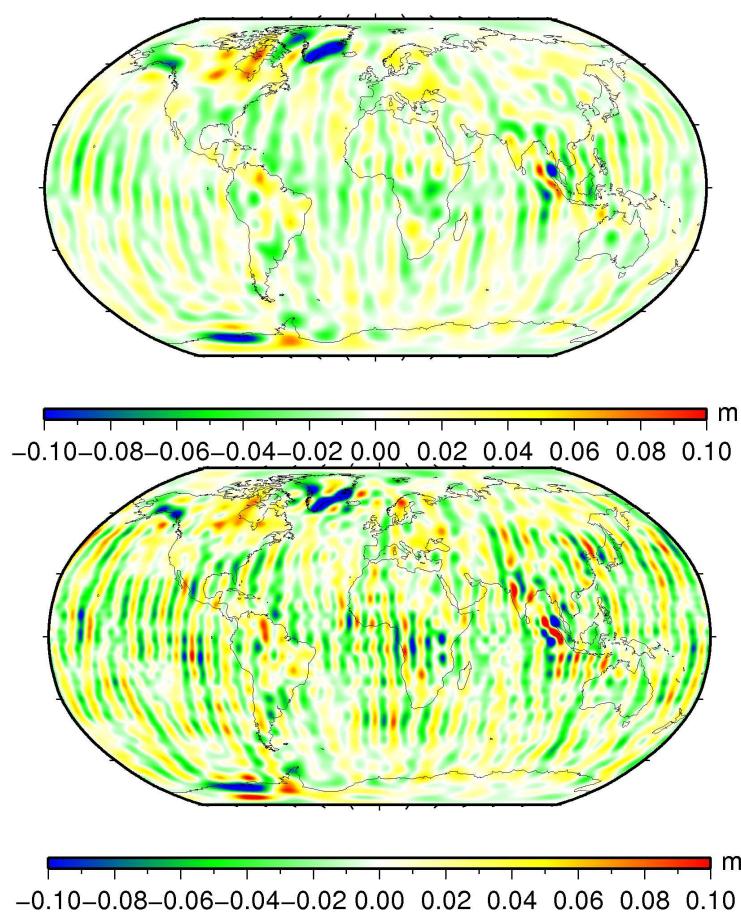
Figure 5.10: Yearly trends (2003-2006) computed from level 35 @ 1,600 km solution (top) and level 45 @ 1,100 km solution (bottom), in EWH.
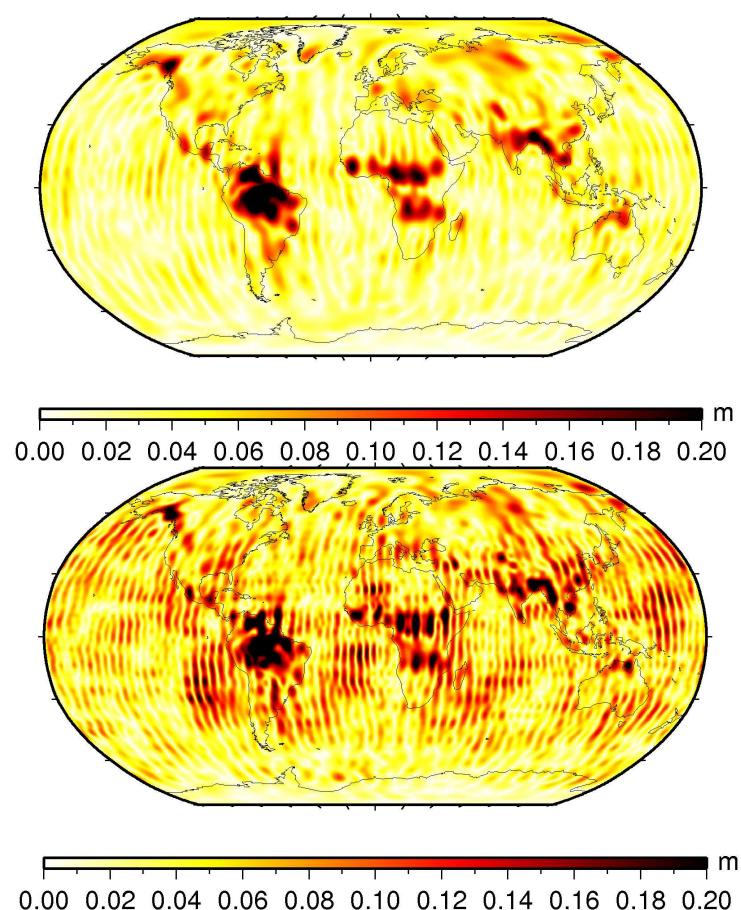
Figure 5.11: Yearly signal amplitudes (2003-2006) computed from level 35 @ 1,600 km solution (top) and level 45 @ 1,100 km solution (bottom), in EWH.

Dependence on area of interest - unfiltered solutions

The GRACE ground tracks are spaced tighter at higher latitudes. This leads to an increased observation density and usually an improved signal-to-noise ratio. This allows for increased spatial resolution, but requires an adapted grid spacing and related bandwidth. Hence, the optimal grid spacing depends on the area of interest, and the latitudes covered by the area of interest.

To demonstrate the effect of various grid spacings and associated optimal bandwidths, figure 5.12 shows nine unfiltered gravity fields solutions for Greenland computed with Reuter grid levels from 20 to 60. The results are shown as yearly mass variations, computed as difference between the October 2003 and October 2006 solution. If the grid level is chosen too small (i.e. underparametrisation), spatial resolution is lost and signal is smeared out. If the grid level is chosen too high (i.e. overparametrisation), too much noise is modelled. The level 45 solution seems to be the solution with the best trade-off between spatial resolution and noise level.

The same results (yearly mass variation as difference between October 2003 and 2006) for Antarctica are shown in figure 5.13. Here, the noise is disturbing at lower grid levels already, and the level 35 solution seems to be the highest one with an acceptable noise level.

By contrast, smaller grid levels are required when computing solutions for equatorial regions. Monthly solutions for Africa in October 2006, using grid levels from 20 to 60, are shown in figure 5.14. The solutions from level 45 onward are dominated by stripes. The level 25 solution already has a stripe propagating into the Sahara desert, a region where little to no rainfall occurs, and no signal should thus be present. Higher levels show more noise, but also allow for better localisation of some features. The level 35 solution has signal over Madagascar, which is smeared out into the ocean in the lower-resolution solutions.

Further insight can be gained by comparing the regional solution to a global spherical harmonic solution. Figure 5.15 shows the level 20 and 25 solutions next to the DEOS global solution, which has been destriped and smoothed with a 400 km Gaussian filter. The SH solution is even smoother than the level 20 solution, so the level 20 solution might be considered the optimal solution here when trying to match the SH solution.

It can be summarised that level 20 to 30 and the related RBF depths yield reasonable solutions at equatorial latitudes when no optimal filtering is used. A higher resolution in certain areas (level 35, Madagascar) is offset by more noise. Higher spatial resolution can indeed be achieved at higher latitudes, where level 35 to 45 are valid choices.
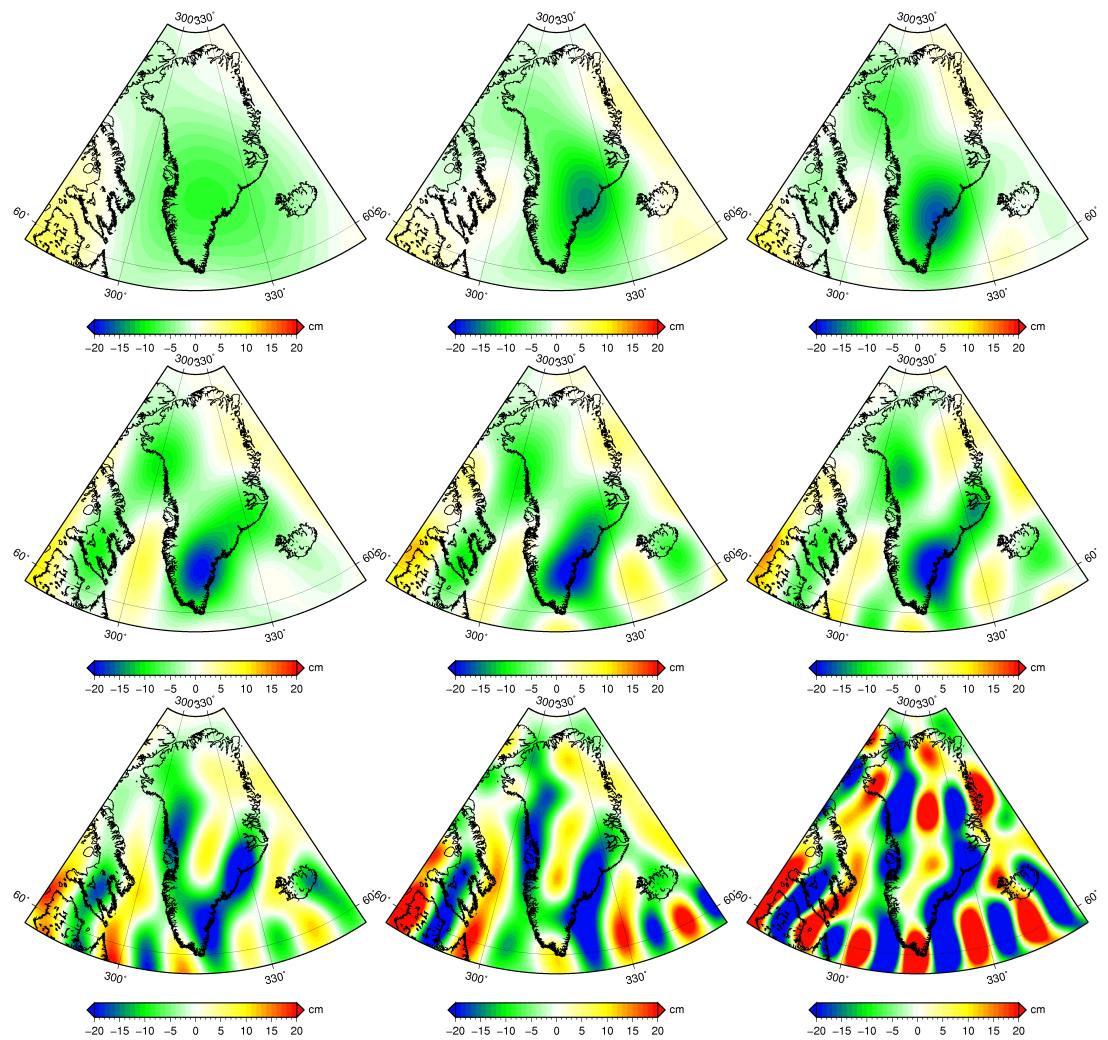
Figure 5.12: Yearly mass variation in terms of EWH computed from three-year difference (October 2006 - October 2003) from various grid levels, in EWH. From left to right: 20, 25, 30 (top row); 35, 40, 45 (centre row); 50, 55, 60 (bottom row)
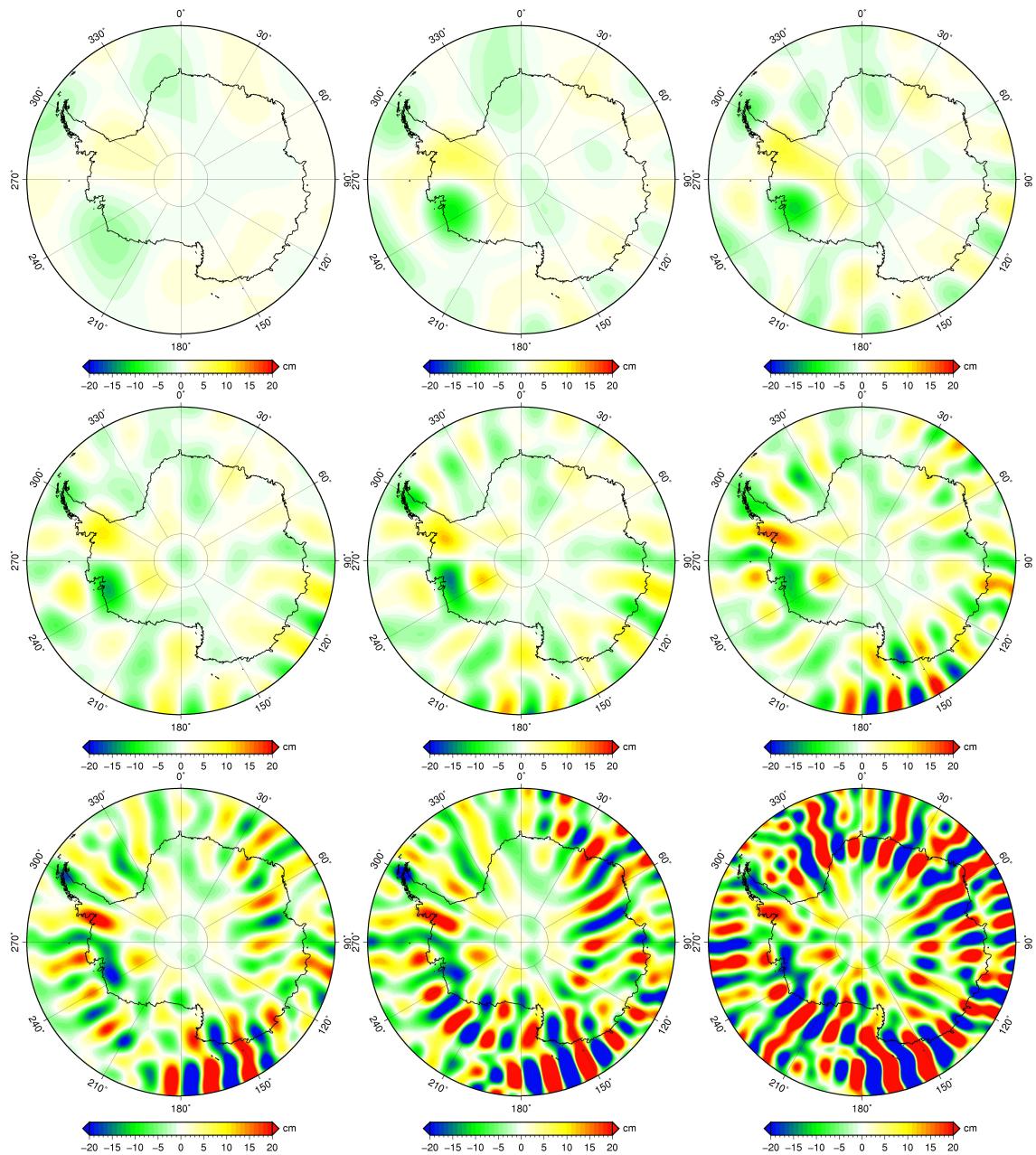
Figure 5.13: Yearly mass variation in terms of EWH computed from three-year difference (October 2006 - October 2003) from various grid levels, in EWH. From left to right: 20, 25, 30 (top row); 35, 40, 45 (centre row); 50, 55, 60 (bottom row)
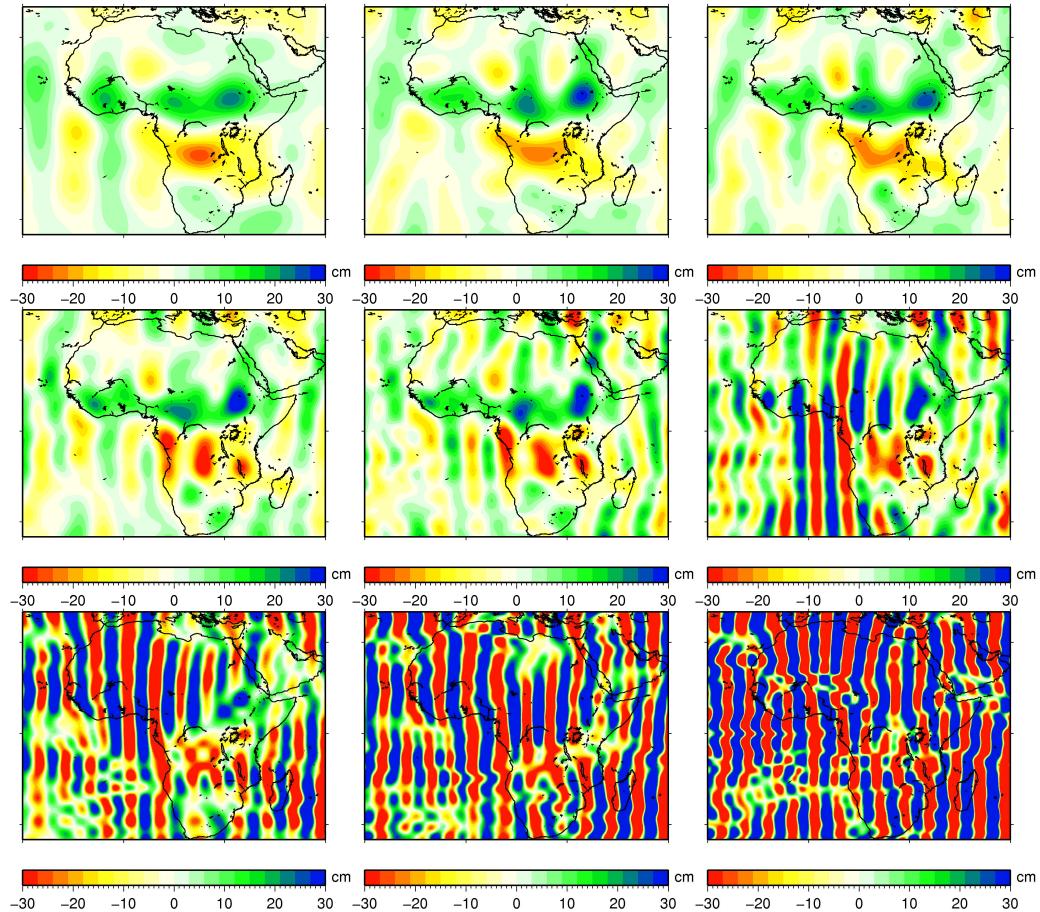
Figure 5.14: Unfiltered GRACE solutions for October 2006 and various grid levels, in EWH. From left to right: 20, 25, 30 (top row); 35, 40, 45 (centre row); 50, 55, 60 (bottom row)
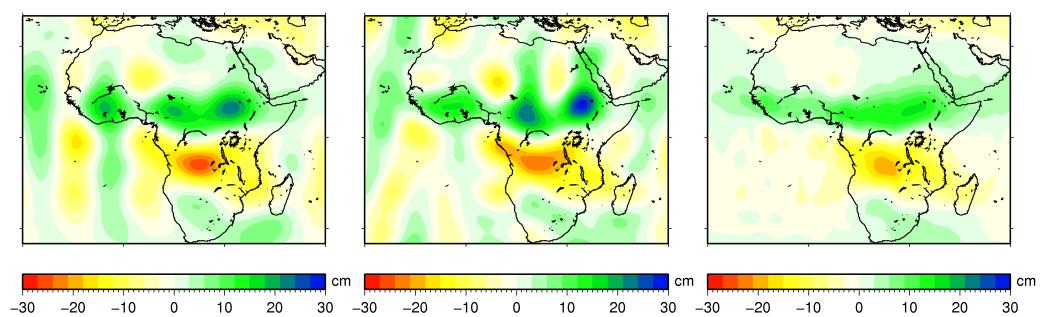
Figure 5.15: Unfiltered GRACE solutions for October 2006, in EWH. RBF on level 20 Reuter grid (left), RBF on level 25 Reuter grid (centre), DEOS global model, destriping and Gaussian 400 km smoothing applied (right)

Optimally-filtered solutions

The solutions in thre previous subsection were computed without optimal filtering and the findings are not applicable to optimally filtered solutions. Solutions without optimal filtering were shown because optimal filtering is a relatively new technique that is not yet part of the standard processing of GRACE solutions. The optimal filtering scheme eliminates high-frequency noise and retains the usable signal. As a result,

- much tighter grid spacings can be used and should be used to retrieve the complete signal;

- using a tighter grid spacing than required does not lead to noisy solutions.

Solutions for the whole time series were computed using three different global grids, with optimal depths as in table 5.2:

- Level 60 Reuter grid, 1,000 km depth, 4,582 basis functions;

- Level 90 Reuter grid, 700 km depth, 10,312 basis functions;

- Level 120 Reuter grid, 500 km depth, 18,334 basis functions.

Figure 5.16 shows the monthly solutions for April 2005. Figure 5.17 shows the resulting yearly trends, and the yearly signal amplitudes are shown in figure 5.18. All three grid spacings yield similar results. Level 90 and 120 show slightly larger signal in Antarctica in the April 2005 solution, and features like ice mass loss in Alaska and Greenland are defined somewhat more sharply than in the level 60 solution.

Since the grid levels chosen are much higher than those useful with unfiltered solutions, they can be considered sufficient to retrieve all signal at high latitudes. The optimal filter scheme suppresses noise that would otherwise results, especially at equatorial latitudes. Hence, no fine-tuning of the solution level to a specific area of interest is required. The level 90 solution is considered sufficient and offers numerical advantages over the level 120 solutions, and is used for all future optimally filtered results presented here.
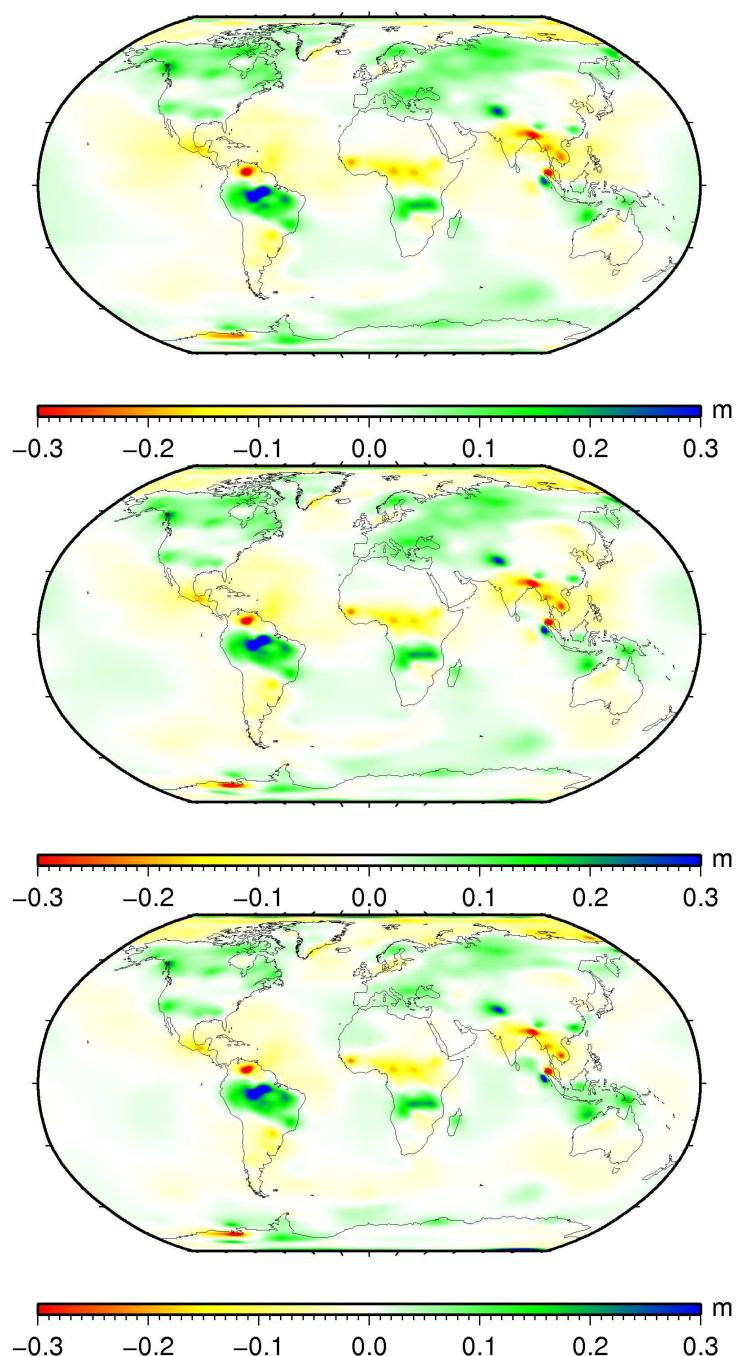
Figure 5.16: Optimally filtered monthly solutions (April 2005) for level 60 @ 1,000 km (top), 90 @ 700 km (centre), and 120 @ 500 km (bottom), in EWH.
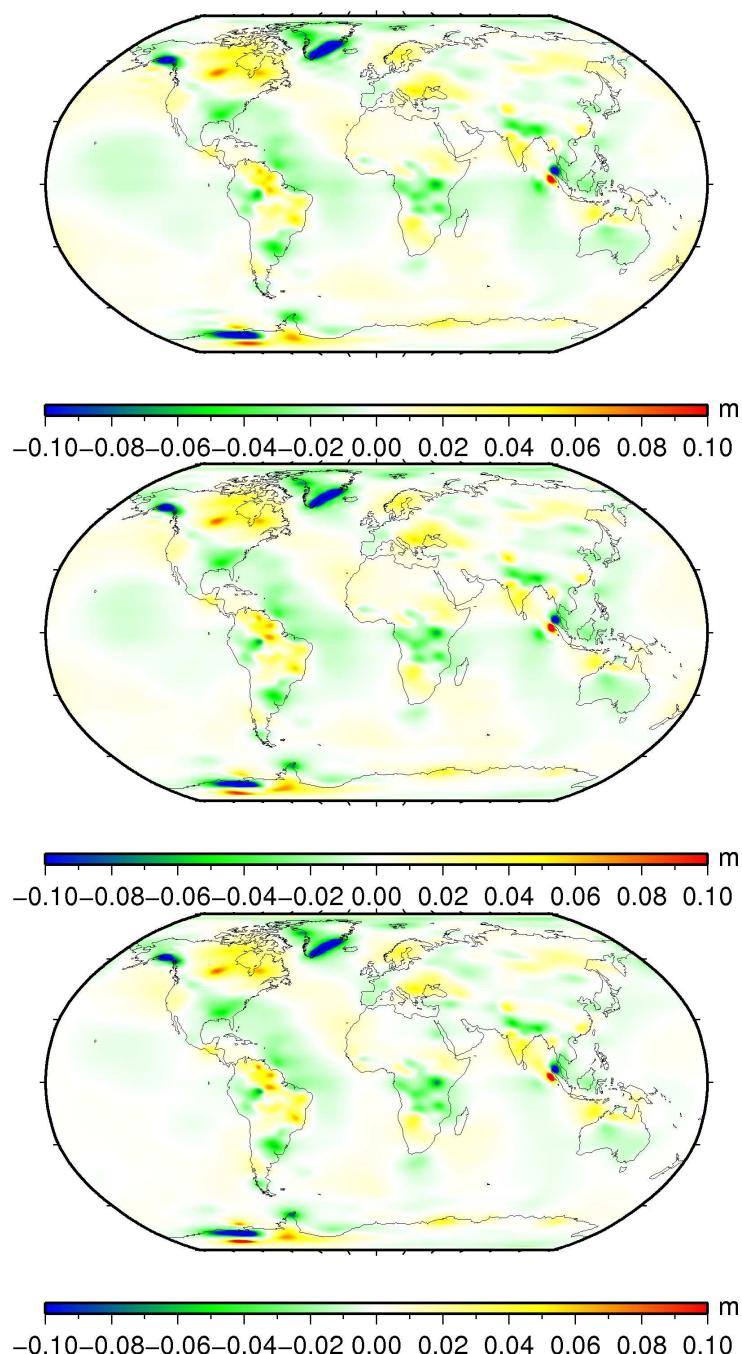
Figure 5.17: Yearly trends (2003-2006) for level 60 @ 1,000 km (top), 90 @ 700 km (centre), and 120 @ 500 km (bottom), in EWH.
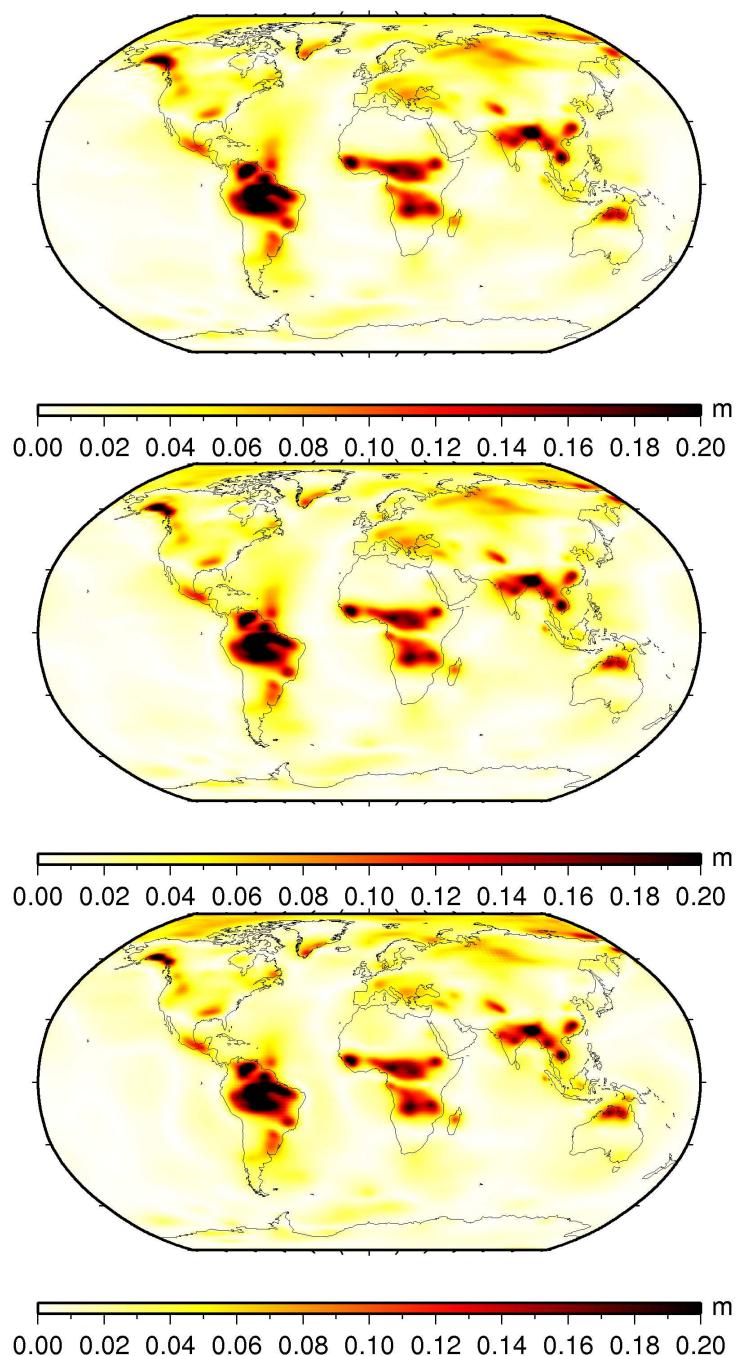
Figure 5.18: Yearly signal amplitudes (2003-2006) for level 60 @ 1,000 km (top), 90 @ 700 km (centre), and 120 @ 500 km (bottom), in EWH.

5.4.2 Data-adaptivity and local refinement

When using satellite data, we deal with a large amount of observations compared to a rather small number of basis functions. A typical month of GRACE data yields more than 500,000 rather uniformly distributed observations at a 5 second interval, while only a few thousand basis functions are used to recover the signal contained in the observations. Combined with the large bandwidth of the basis functions (see next section), there will be hundreds of observations within a basis function's influence radius. Adapting the basis function network to the data distribution will thus usually not be required. The smooth, long-wavelength characteristic of the signal compared to e.g. terrestrial data also imply that local refinement is not required.

5.4.3 Parametrised area

In regional gravity field modelling, we are only interested in a small region, not the whole Earth. It is thus not desirable to parametrise the whole Earth by placing basis functions on a global grid. Instead, we want to confine the parametrised area to our region of interest. Unfortunately, parametrising only the region of interest will lead to edge effects. It is thus necessary to extend the parametrised area.

To quantify the necessary extension, several test computations using RBFs on a level 35 at 1,600 km depth were made. Two unfiltered solutions were computed:

1. A global reference solution using global data coverage;

2. a regional solution parametrising only the Northern Hemisphere, once again using global data coverage.

Figure 5.19 shows the resulting differences in the monthly solution for April 2005. Aside from the expected differences over the Southern Hemisphere (which was not parametrised by the regional solution), there are also large errors over the Northern Hemisphere. Apparently, leakage (errors caused by unmodelled signal) affects the estimation. The large errors off the east coast of North America are caused by a strong signal in the Amazon river basin in the chosen month.

These leakage effects can be reduced by using only the observations over the parametrised area. Figure 5.20 shows the same difference plot (April 2005), but this time using only observations over the Northern Hemisphere for the regional solution. One can see that the differences become smaller at a larger distance from the edge of parametrisation (the equator). At about 30 degrees of Northern latitude, differences do not exceed two centimetres. It has to be noted that leakage still seems to be an issue, and is not confined to the border of the area, as can be seen by the stripes at high latitudes. This is caused by the fact that the space-localising RBFs used here have global support.

For the same scenario, the differences in terms of linear trend and yearly signal amplitudes are shown in figure 5.21. The yearly trend exhibits very small differences in regions further than 20 degrees from the border of parametrisation. The yearly signal
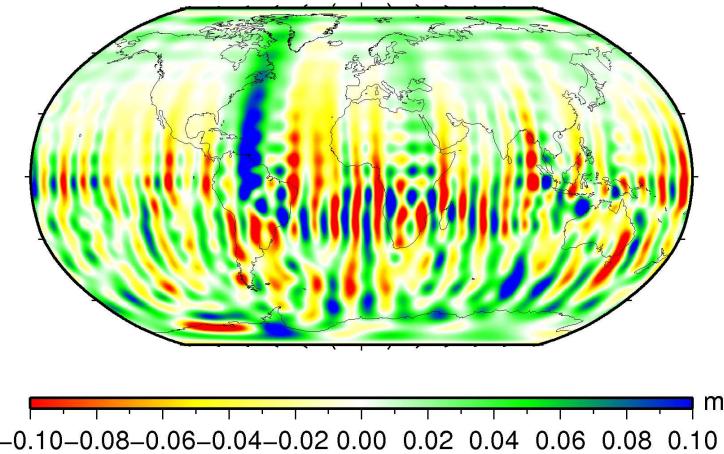
Figure 5.19: Differences between unfiltered solutions (April 2005) parametrising the whole globe and only the Northern Hemisphere, in EWH; global data coverage was used.
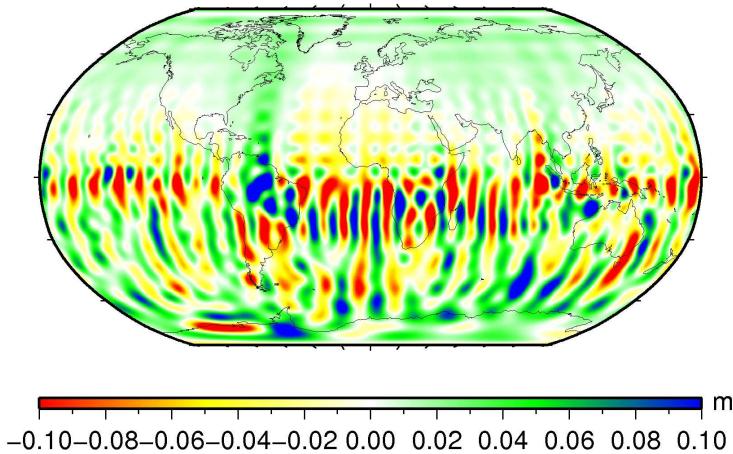

Figure 5.20: Differences between unfiltered monthly solutions (April 2005) parametrising the whole globe and only the Northern Hemisphere, in EWH; only observations over the Northern Hemisphere were used for the regional solution.

amplitudes suffer somewhat more from leakage out of the Amazon basin. Looking at figure 5.21, about 30 degrees of distance from the edge are required to reduce edge effects to below the centimetre-level.

The results for optimally filtered solutions are similar. Figure 5.22 shows the differences between a global solution and a solution covering only the Northern Hemisphere, in terms of a monthly solution and linear trends. The monthly solution exhibits maximum differences of approx. 2 cm beyond a 30 degree border, errors for the linear trend are below a centimetre.

It can be summarised that about 30 degrees of additional parametrised area are required around the desired region of interest. In the test cases, the edge of parametrisation was located next to an area with large signals (the Amazon river basin). This results in leakage. Only observations over the parametrised area should be used.
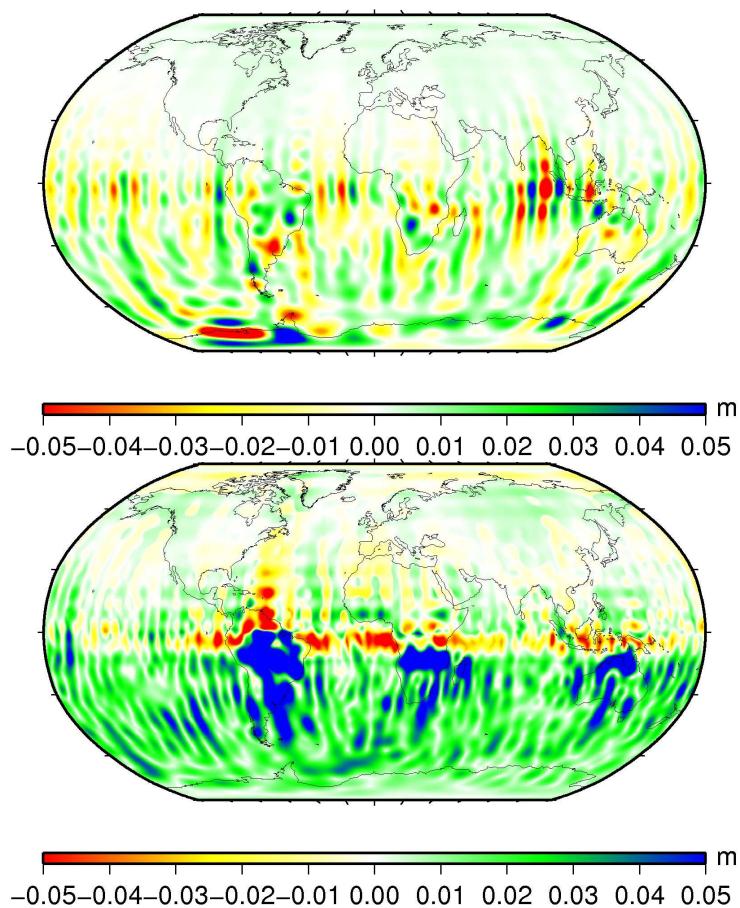
Figure 5.21: Differences between unfiltered solution parametrising the whole globe and only the Northern Hemisphere, in EWH: In terms of trend (top) and yearly signal amplitudes (bottom). Only observations over the Northern Hemisphere were used for the regional solution.
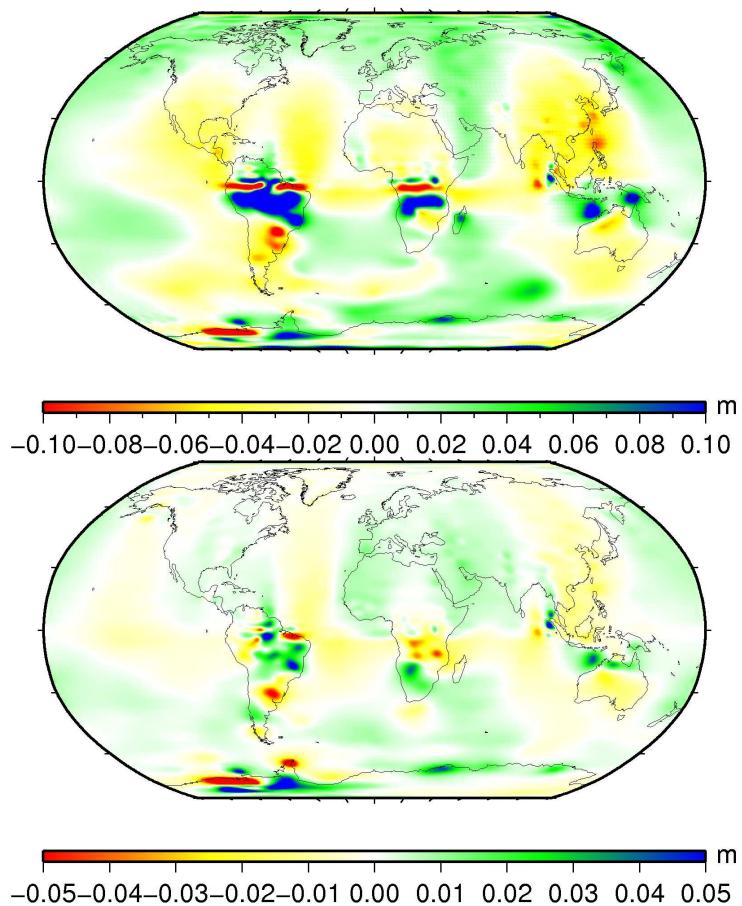
Figure 5.22: Differences between optimally filtered solutions parametrising the whole globe and only the Northern Hemisphere, in EWH; monthly solution for April 2005 (top) and yearly trend (2003-2006, bottom); only observations over the Northern Hemisphere were used for the regional solution.

5.5   Bandwidth selection

Choosing the optimal bandwidth in order to get the optimal solution is an integral part of the method presented here (see section 2.5.2). The optimal bandwidth is dependent on the spectral content of the data and the density of the RBFs, and is directly related to the depth of the RBFs. It is thus necessary to compute the optimal depth for each realisation of an RBF network.

Table 5.1 lists the optimal depths and corresponding bandwidths for several Reuter grid levels obtained while computing an unfiltered global GRACE monthly solution for February 2006. These numbers can be used for other months without significant impact on the quality of the solution. Optimally filtered solutions make use of basis functions placed somewhat deeper, and can be computed up to higher grid levels (table 5.2).

The effect of choosing a suboptimal depth has also been investigated using numerical experiments. Figure 5.23 shows the difference between two level 35 monthly solutions, at 1,600 km at 2,400 km depth. In most areas, the difference does not exceed a few cm of EWH. Differences are even smaller in terms of trends and yearly signal amplitudes. When in doubt, it is better to place the basis functions deep rather than shallow, as it yields a smoother solution.

level	depth	bandwidth
20	2500 km	920 km
30	1900 km	660 km
35	1600 km	550 km
40	1400 km	470 km
50	1100 km	360 km
60	900 km	290 km

Table 5.1: Optimal depth and resulting bandwidth for unfiltered monthly solutions depending on Reuter grid level.

level	depth	bandwidth
60	1,000 km	330 km
90	700 km	230 km
120	500 km	160 km

Table 5.2: Optimal depth and resulting bandwidth for optimally filtered monthly solutions depending on Reuter grid level.

This also holds for optimally filtered solutions. Figure 5.24 shows the differences between monthly solutions (April 2005) using an optimal depth of 700 km and a suboptimal depth of 600 km. Differences reach several centimetres in certain areas, but are on average below one centimetre. The red spots at high latitudes around 175 degrees of eastern longitude are a result of gaps in the Reuter grid used (see section 2.6.1).

Figure 5.25 shows the spectrum of the RBFs for depths of 2,500 km, 1,400 km, and 900 km. These are the optimal depths for Reuter grids of level 20, 40, and 60, respectively.
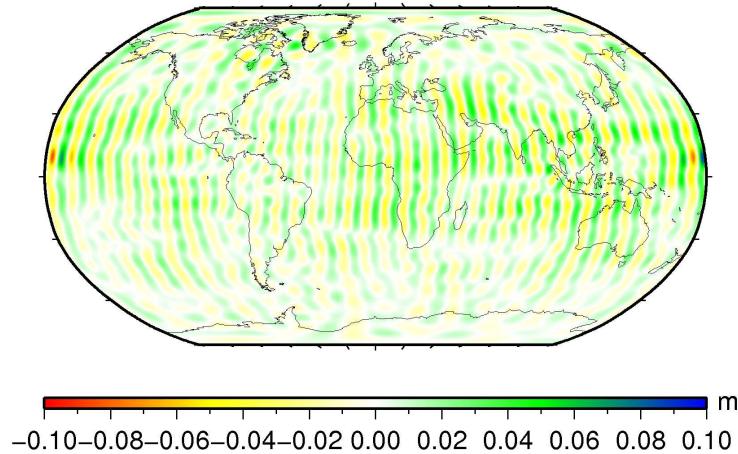
Figure 5.23: Difference between unfiltered level 35 @ 1,600 km and 35 @ 2,400 km solutions for April 2005 in EWH.
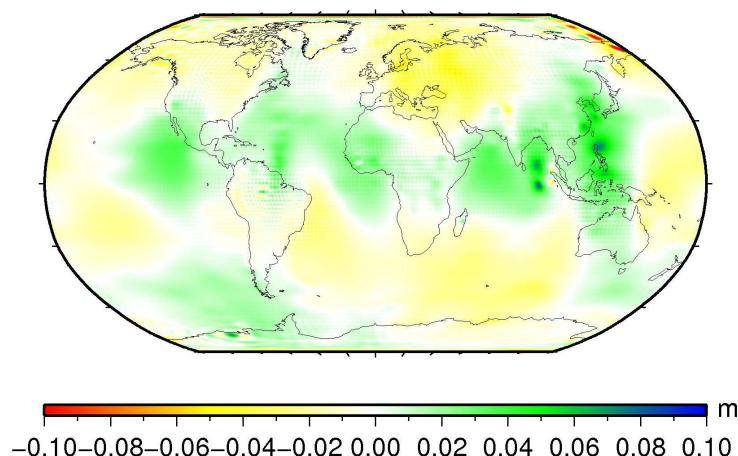

Figure 5.24: Difference between optimally filtered level 90 @ 600 km and 90 @ 700 km solutions for April 2005 in EWH.

The degree variances of the resulting solutions are shown in figure 5.26. As expected, solutions generated with a higher-level grid of basis functions located less deep contain higher frequencies.
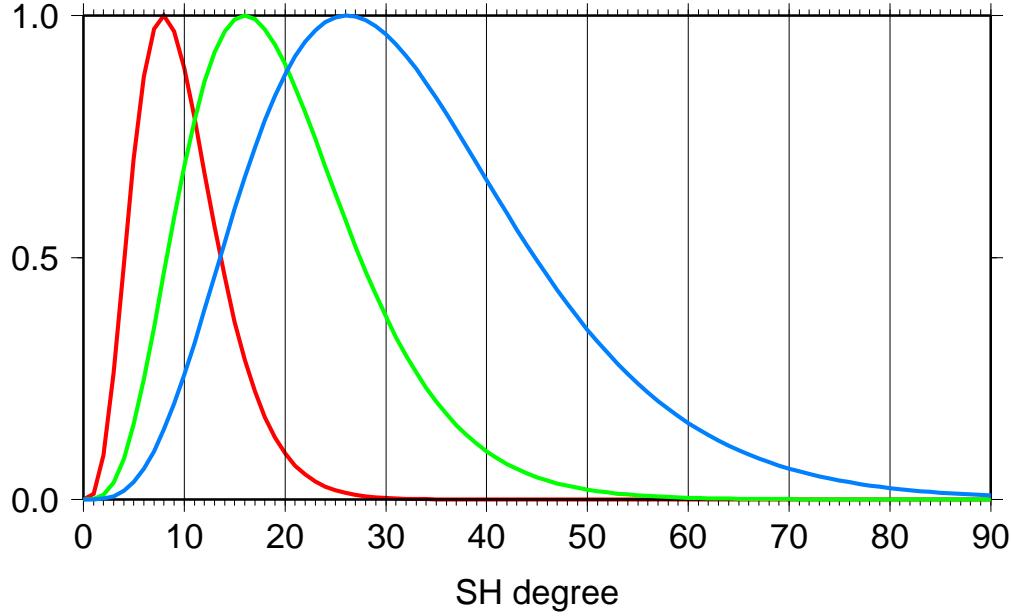
Figure 5.25: Spectrum of Poisson Wavelets of degree 3 and depth 2500 km (red), 1400 km (green), and 900 km (blue)
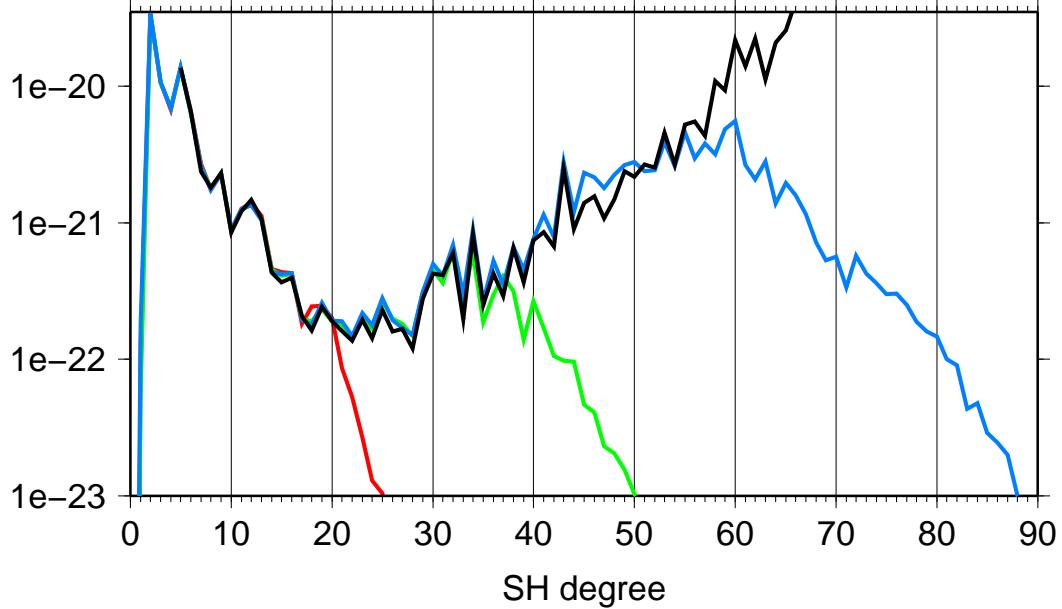

Figure 5.26: Degree variances of solutions with Reuter grid level 20 (red), 40 (green), 60 (blue), as well as an SH solution (black).

The cumulative degree variances for the three solutions, as well as those of a solution in terms of spherical harmonics up to degree 120, are shown in figure 5.27. It is interesting to see that the signal of a solution with a Reuter grid of level γ has a signal content equivalent to the spherical harmonic solution at degree γ.
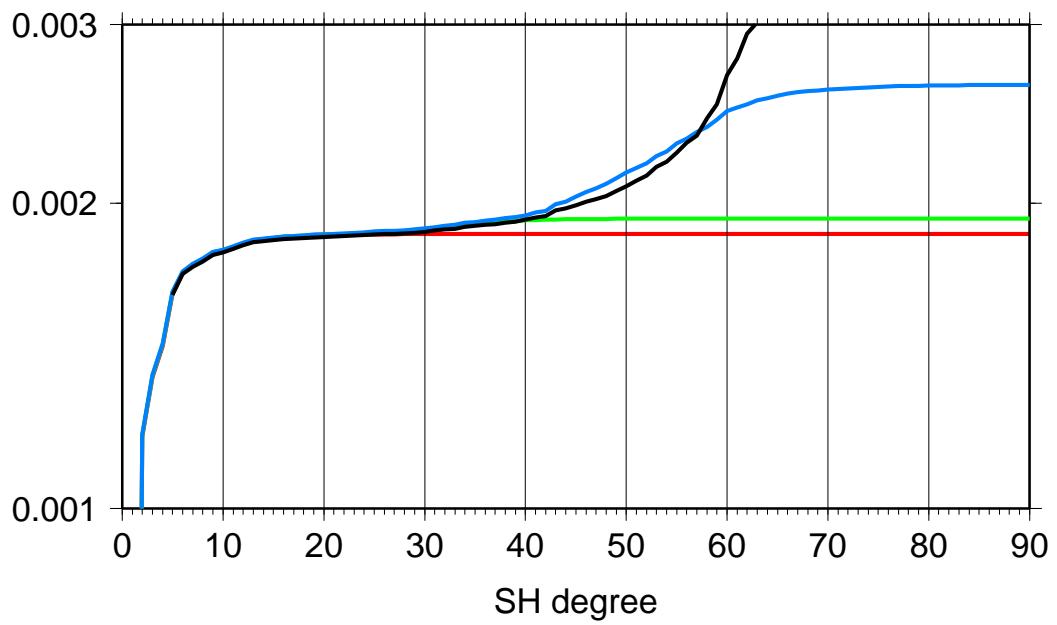
Figure 5.27: Cumulative degree signal [m] of solutions with Reuter grid level 20 (red), 40 (green), 60 (blue), as well as an SH solution (black).

5.6   Results

5.6.1   Comparison of unfiltered RBF and spherical harmonic solution

Spherical harmonics are the standard parametrisation used for global gravity field modelling. It is thus interesting to verify the methodology using RBFs by comparing it to spherical harmonic solutions. In this section, it will be investigated whether a global RBF representation is capable of producing a solution that is of equal quality as a spherical harmonic solution using the same data.

This test compared unfiltered monthly GRACE solutions computed using spherical harmonics and RBFs. The month chosen was April 2005, with 483,840 GRACE observables. A spherical harmonic solution was computed up to degree 120, resulting in 14,641 spherical harmonic coefficients. An RBF solution was computed with basis functions located at a depth of 1,600 km on a Reuter grid of level 35, using 1,542 basis functions.

Figure 5.28 shows the degree variances for both solutions. The solutions are almost identical up to degree 30, after which the degrees variance of the spherical harmonic solution increase. These coefficients beyond degree 30 are dominated by noise. The RBF solution contains no significant signal beyond degree 40 due to the chosen parametrisation. Note that no hard cut-off at degree 35 is noticeable - a result of the Poisson wavelet's spectral characteristics (see figure 2.8 and 5.25).
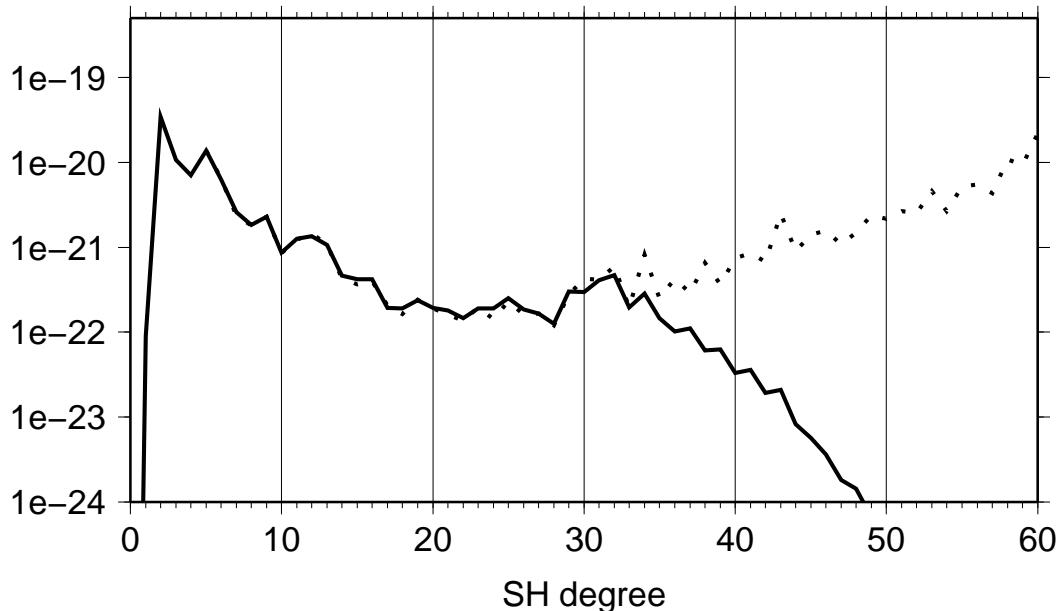

Figure 5.28: Dimensionless degree variances, April 2005 GRACE solution using RBFs (solid) and spherical harmonics (dotted).

The corresponding cumulative degree variances are shown in figure 5.29. The spherical harmonic solution contains a little less signal up to degree 35, where the two curves intersect. The signal contained in the SH solution beyond that is mostly noise. This is obvious when looking at the resulting EWH fields in figures 5.30 and 5.31.
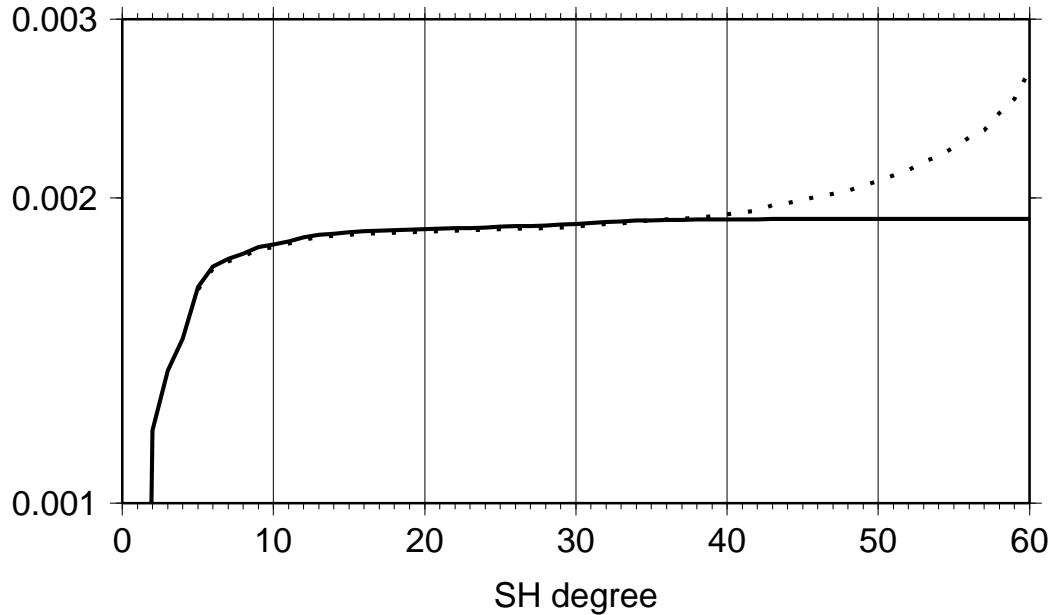
Figure 5.29: Cumulative degree variances [m], April 2005 GRACE solution using RBFs (solid) and spherical harmonics (dotted).

Figure 5.30 shows the EWHs obtained when using the potential coefficients up to degree 35, while figure 5.31 has been generated using the coefficients up to degree 50. It is obvious that going from degree 35 to 50 adds little or no useful information, while distorting the model with stripe patterns.
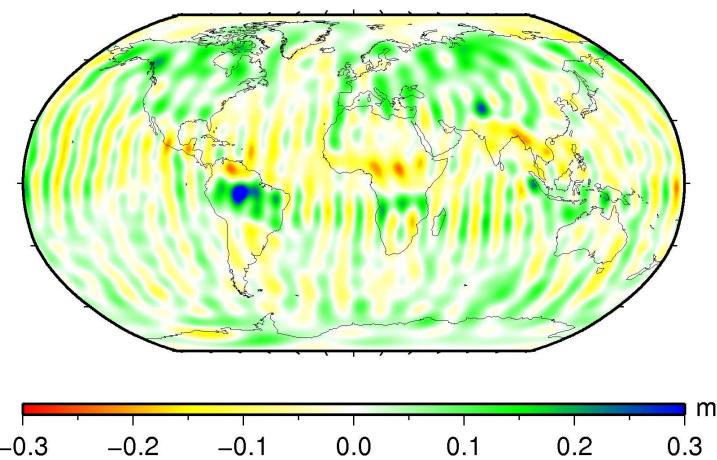

Figure 5.30: EWH resulting from spherical harmonic solution computed up to degree 120, only degrees up to 35 are shown.
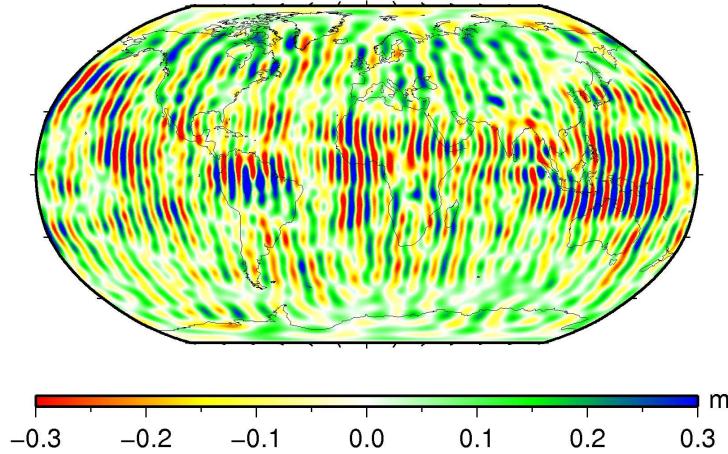
Figure 5.31: EWH resulting from spherical harmonic solution computed up to degree 120, only degrees up to 50 are shown.

The EWHs obtained from the RBF solution are shown in figure 5.32. When comparing it to figure 5.30, the SH solution up to degree 35, the results appear to be very similar.
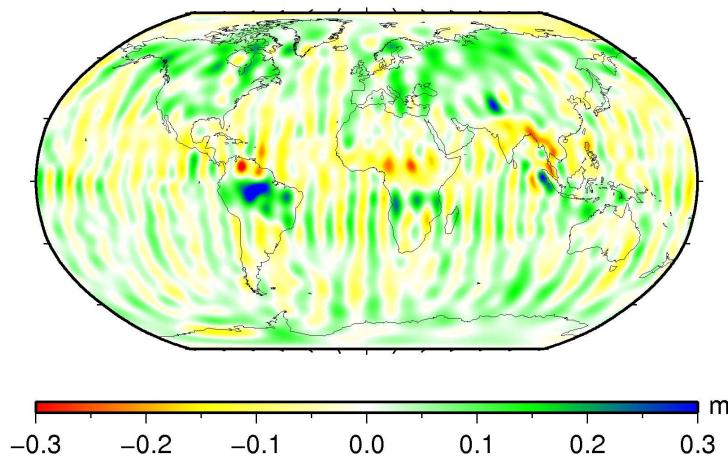

Figure 5.32: EWH resulting from level 35 @ 1,600 km RBF solution.

The differences between the spherical harmonic solution truncated at degree 35 and the RBF solution in terms of EWH are shown in figure 5.33. The most obvious difference are the stripe patterns, which are stronger in the SH solution over the Pacific than in the RBF solution. The RBF solution captures more signal of e.g. the Sumatra earthquake and hydrological features in South America and Africa. The rms of the difference is 3.3 cm in terms of EWH.

We can conclude that with monthly solutions, the use of RBFs delivers a solution comparable to a solutions with SHs, although the spatial plots showed differences in certain areas. These differences can however be attributed to larger errors in the SH solution. There is no evidence that an RBF solution is not as good as an SH solution.
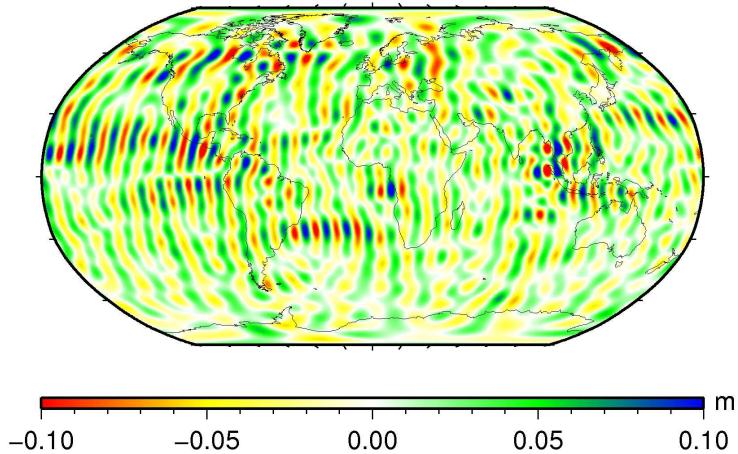
Figure 5.33: Difference between SH solution truncated at degree 35 and level 35 @ 1600 km RBF solution, in EWH.

5.6.2 Models used for comparison

The next two sections make use of results obtained from several GRACE time-series. These include the following RBF solutions:

- Unfiltered global RBF solutions computed on a level 35 Reuter grid at a depth of 1,600 km. **RBF 35** refers to the solution without frequency-dependent data weighting, and **RBF 35 FW** refers to the solution with frequency-dependent data weighting (the noise model, see section 5.2).

- Optimally filtered global RBF solutions computed on a level 90 Reuter grid at a depth of 700 km. These are referred to as **RBF 90** (without frequency-dependent data weighting) and **RBF 90 FW** (with frequency-dependent data weighting).

- For Greenland and Antarctica, regional RBF solutions are also shown. Both were computed using RBFs on a level 90 Reuter grid at a depth of 700 km and made use of optimal filtering. For Greenland, the solution was computed using only RBFs in the Northern Hemisphere. Similarly, for Antarctica the regional solution was computed using RBFs only in the Southern Hemisphere. In both cases only data over the parametrised area was used. No frequency-dependent data weighting was applied. These solutions are referred to as **RBF 90 REG**.

To assess the performance of the RBFs, the RBF solutions were compared to global SH solutions from three sources:

- Release 4 monthly solutions of the Centre for Space Research (CSR) (Bettadpur, 2007). These degree-60 solutions were destriped and smoothed with a 400 km Gaussian filter, and are referred to as **CSR DS400**.

- The monthly solution of DEOS (Liu, 2008). These degree-70 solutions were also destriped and smoothed with a 400 km Gaussian, and are referred to as **DEOS DS400**.

- The optimally-filtered monthly DMT-1 models (Klees et al., 2008b). They are referenced to as **DMT-1**.

The DEOS DS400 and DMT-1 spherical harmonic models were computed using the same data, processing strategy, and functional model as used for the RBF solutions. The DMT-1 models used the same optimal filter approach as the the optimally-filtered RBF solutions RBF 90, RBF 90 FW, and RBF 90 REG. Differences between these models and the RBF models are due to the different representation (and different observation weights in the case of optimally filtered solutions) only and are thus well-suited to showcasing the differences caused by the use RBFs instead of spherical harmonics. The CSR DS400 models are used as an external reference.

All comparisons were made using 46 monthly solutions, from **February 2003 to December 2006** with the exception of June 2003. Using several regions of interest, the results will be used to answer the following questions:

- Are RBFs capable of outperforming SHs for GRACE modelling, or at least able to produce solutions of comparable quality?

- Does frequency-dependent data weighting offer a benefit over the assumption of white noise?

- What is the benefit resulting from the use of optimal filtering?

Liu et al. (2009) have investigated the quality of the DMT-1 models. Errors of 2 to 3 cm are given for point-wise values and small basins of the order of 10^5 km², whereas the error for mean mass estimates for larger basins are somewhat smaller, 1.5-2 cm. These values will be used when judging the quality of the RBF GRACE solutions.

5.6.3   Recovery of ice mass loss in Greenland and Antarctica

Climate change is an important matter of public concern. One consequence of climate change with immediate repercussions is sea level change. Sea level change can be accurately measured with tide gauges and, after corrections, with satellite radar altimetry. Since they only provide information about the total sea level change without being able to discriminate between the contributors, other measurements are needed to quantify the factors contributing to sea level change. It is clear that the melting of polar ice due to global warming is one such contributing factor, but the exact magnitude of the contribution is not known. It is thus important to actually measure ice mass loss in order to quantify its effect on global sea level change.

Due to the large areas involved, their remoteness, and the harsh weather conditions, only satellite methods can provide estimates with a sufficient temporal and spatial resolution. Terrestrial GPS measurements are restricted to a limited number of points, and airborne laser altimetry campaigns are too expensive and time-consuming to execute them regularly and with the desired area coverage. GRACE allows continuous ice mass monitoring, and the results show good agreement with altimetry measurements (Cazenave et al., 2009).

Comparison to ICESat

While GRACE measures gravity and gravity change, ICESat, launched in January of 2003, is a laser altimetry satellite designed specifically for measuring ice surface heights. It provides data of ice surface altitudes and altitude change with almost global coverage (the orbit inclination is 86°) and a high spatial resolution, due to the laser footprint of only 50 metres at the Earth's surface. Unfortunately, problems with the satellite's lasers (one of the three lasers failed almost immediately, and premature aging of the others is a concern) limits it to three 35-day measurement campaign per year. The computation of ice mass loss from the altimetry measurements is also not straightforward. Assumptions about the snow density have to be made, and snow compaction may be interpreted as ice mass loss. Also, strong slopes, which are present in coastal areas, where rapid melting occurs, may lead to inaccurate distance and thus altitude measurements (Slobbe et al., 2008).

To provide corrections for these effects, and to fill the gaps between ICESat measurement campaigns, as well as to extend the measurements beyond ICESat's lifespan, the GRACE satellites can be used. They measure gravity and gravity change, which can be translated into mass change. The mass estimates derived from GRACE are not affected by assumptions about snow density or snow compaction (but these are required for discriminating between ice mass and firn variations), and can be computed continuously since 2003. The combination of GRACE and ICESat can be used to determine the magnitude of glacial-isostatic adjustment (GIA) effects (Riva et al., 2009). Unfortunately, the spatial resolution that can be achieved is significantly lower than what can be achieved with ICESat. GRACE and GRACE model computation also has its own error sources, such as leakage and striping.

Leakage

The computation of ice mass loss requires the integration over a specific area. Unfortunately, the RBFs work as smoothing filter. As a result, the signal is "smeared out" over a large area. This means that signal originating inside the area may leak out of the area, and will thus not be considered when doing the integration over the original target area. Signal from outside may also leak into the area of interest and be taken into account.

To assess the effect of leakage when RBFs are used, a simulation was done for Greenland. Figure 5.34 shows mass loss for several basins in Greenland derived from ICESat measurements. The total mass loss according to ICESat amounts to 145 Gigatons per year (Slobbe et al., 2009a). It is visible that most of the mass loss happens along the coast.

The leakage effect was quantified using the ICEsat data set as true signal. When computing unfiltered monthly GRACE solutions, Poisson wavelets at a depth of 1,100 km on a level 45 Reuter grid are used. Using this parametrisation to approximate the ICESat data set causes the same leakage effect that is also existent when using GRACE measurements, as the same smoothing of the signal occurs. The resulting ice mass loss over Greenland is computed as 84 Gigatons per year, only 58% of the original input signal. Baur et al. (2009) state that GRACE models retrieve only around 50% of the actual mass
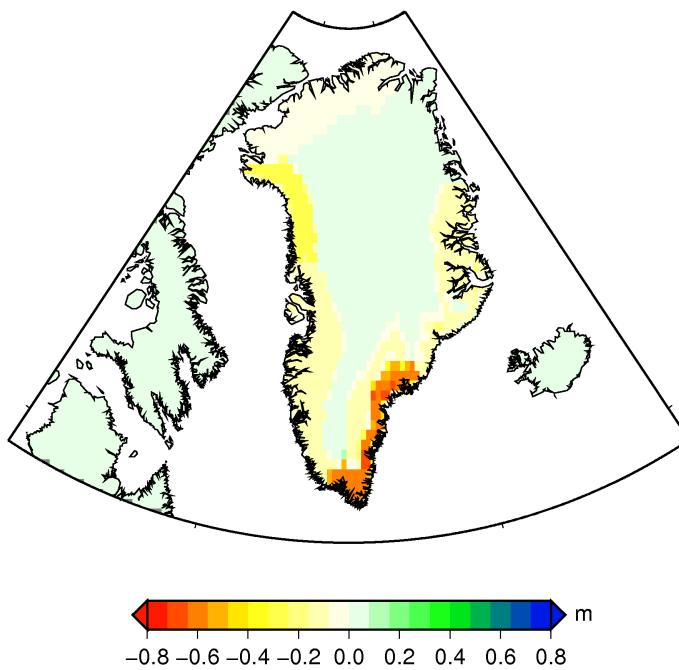
Figure 5.34: Mass loss per basin [m] for Greenland derived from ICESat, in EWH (Slobbe et al., 2009a).

loss signal. As figure 5.35 shows, the large signal in southeastern Greenland has partly leaked into the ocean.

One might be inclined to apply a scaling factor to compensate for the mass loss due to leakage (Velicogna and Wahr, 2005). It is however obvious that the leakage effect is governed by the spatial distribution of the signal. A strong signal in central Greenland will not leak out into the ocean. Applying a constant scaling factor will in this case overcorrect and lead to overestimation of the signal. Scaling factors should only be used in the case of homogeneous mass distributions inside the target area, and if there is no danger of signal outside of the area leaking into the area.

Instead, the integration area should be enlarged to also capture signal that has leaked out. The Poisson wavelets at 1,100 km depth have a correlation length of about 300 km. Increasing the integration area by 300 km yields a mass loss estimate of 147 Gigatons in the simulation. This is almost identical to the input signal of 145 Gigatons. It has to be noted that the use of this increased integration area is only possible since no other signals are present inside the 300 km area. Otherwise, these would falsely be taken into account.

With optimally filtered solutions, a different parametrisation is used: RBFs on a level 90 Reuter grid at a depth of 700 km. These are narrower than the RBFs at 1,100 km depth (figure 5.36), and the optimal filter leads to increased spatial resolution. As a consequence, the area extension used is 200 km degrees instead of the 300 km used for the unfiltered solutions.
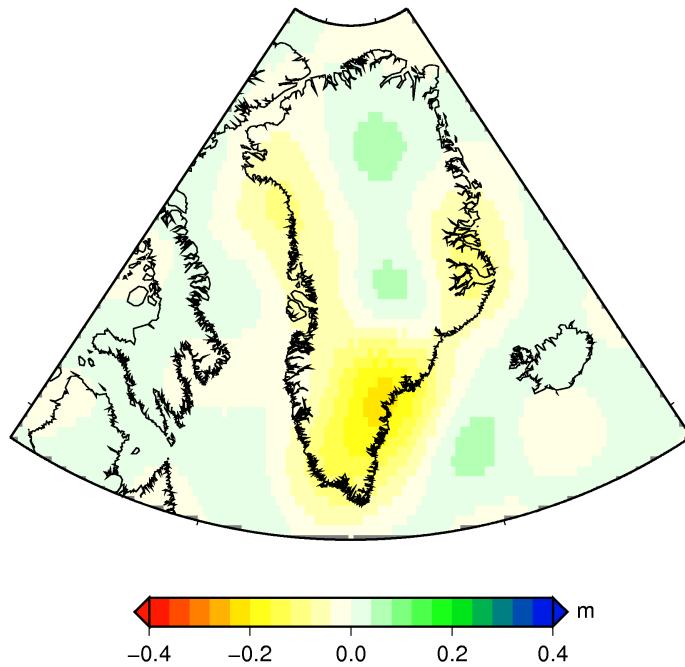
Figure 5.35: Signal recovered by RBFs [m] on level 45 Reuter grid @ 1100 km depth, in EWH.

Greenland

Figure 5.37 shows time series of mass estimates as obtained from various GRACE models. There is little difference between RBF solutions with and without frequency-dependent weighting. The unfiltered and filtered RBF solutions and the DMT-1 solution show similar behaviour. The DEOS DS400 solution (i.e. without optimal filtering) and the CSR DS400 solution differ significantly for certain months. This can be attributed to different processing strategies (Klees et al., 2008c).

The resulting trend and annual signal estimates are shown in table 5.3. With the exception of the DEOS DS400 and CSR DS400 models, all solutions yields a yearly trend between -139 and -150 Gt per year. Slobbe et al. (2008) estimated a trend of -147 Gt/yr from ICESat data over a slightly longer time span (February 2003 - April 2007).

Besides time-series and trends for all of Greenland, it is also interesting to look at spatial plots of the trend. Figure 5.38 shows these plots in terms of EWH for the four models computed without the use of optimal filtering. The two RBF solutions offer much better estimation and localisation of the trend in southeastern Greenland than the two spherical harmonic solutions. The RBF solutions are affected by noise artifacts, and the signal in northwest Greenland is shifted towards the ocean.

Likewise, plots for the optimally filtered solutions are shown in figure 5.39. All four solutions produce similar plots. The RBF solution with frequency-dependent weighting (top left) shows stronger signals in northeast and northwest Greenland than the others. The localisation of signal is better and almost no noise is visible over the oceans compared to the unfiltered solutions.

For comparison, figure 5.40 shows the yearly trend of altitude change as computed from
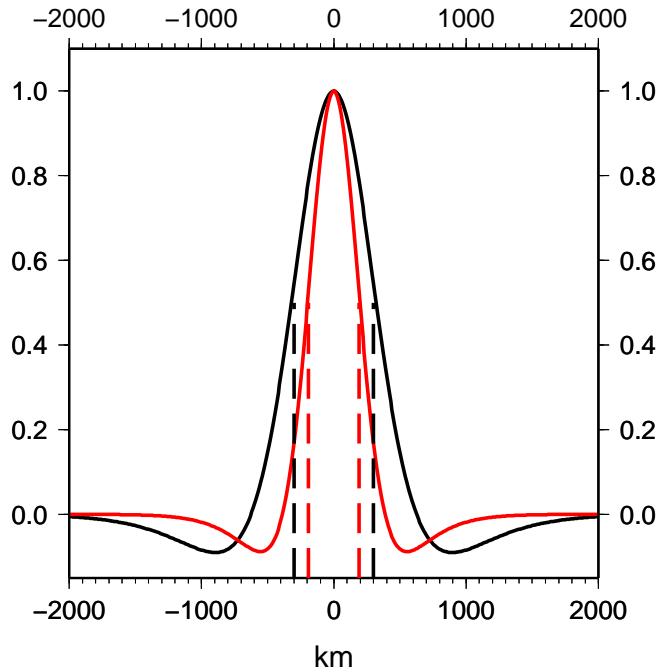
Figure 5.36: RBFs at a depth of 1,1000 km (black) and 700 km (red) with correlation lengths (dashed).

model	trend	annual amplitude
RBF 35	-150.2 Gt/yr	71.0 Gt
RBF 35 FW	-147.6 Gt/yr	73.7 Gt
DEOS DS400	-122.1 Gt/yr	46.4 Gt
CSR DS400	-171.2 Gt/yr	60.7 Gt
RBF 90	-146.8 Gt/yr	76.6 Gt
RBF 90 FW	-138.7 Gt/yr	69.6 Gt
RBF 90 REG	-140.5 Gt/yr	44.9 Gt
DMT-1	-139.0 Gt/yr	74.3 Gt

Table 5.3: Yearly trends and signal amplitudes for Greenland, computed from GRACE time series over the February 2003-December 2006 time span.
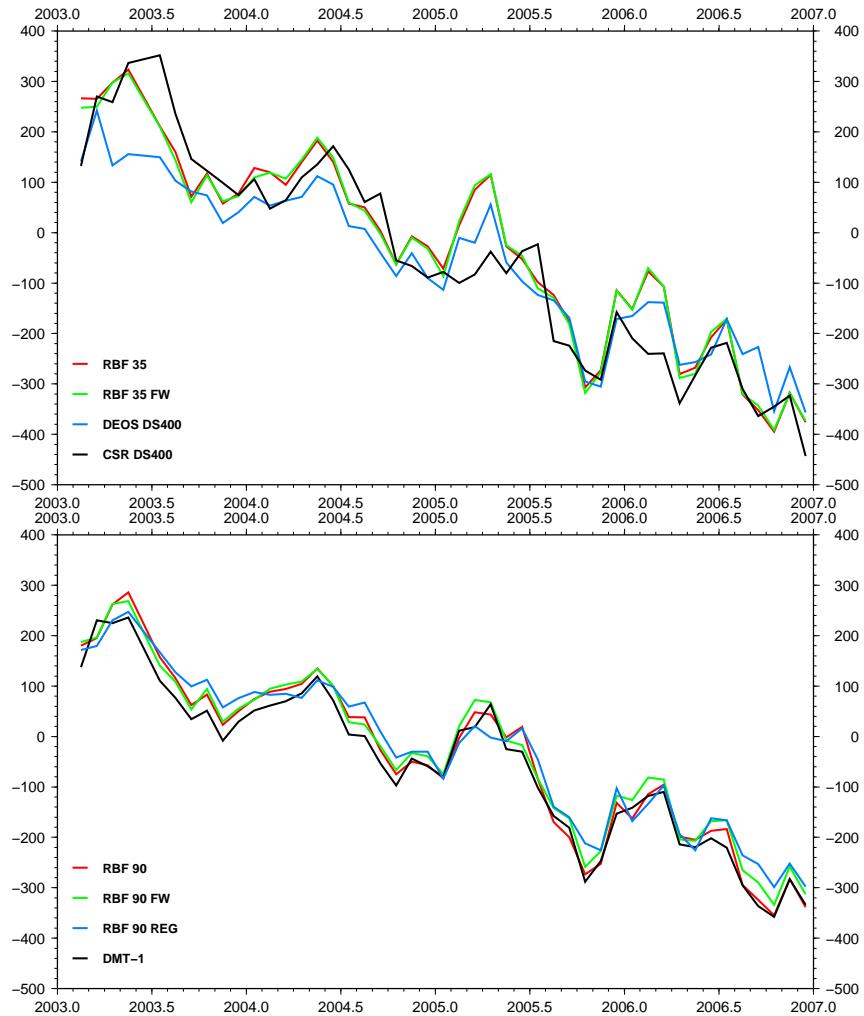
Figure 5.37: Time series [Gt] for Greenland for unfiltered solutions (top) and filtered solutions (bottom).
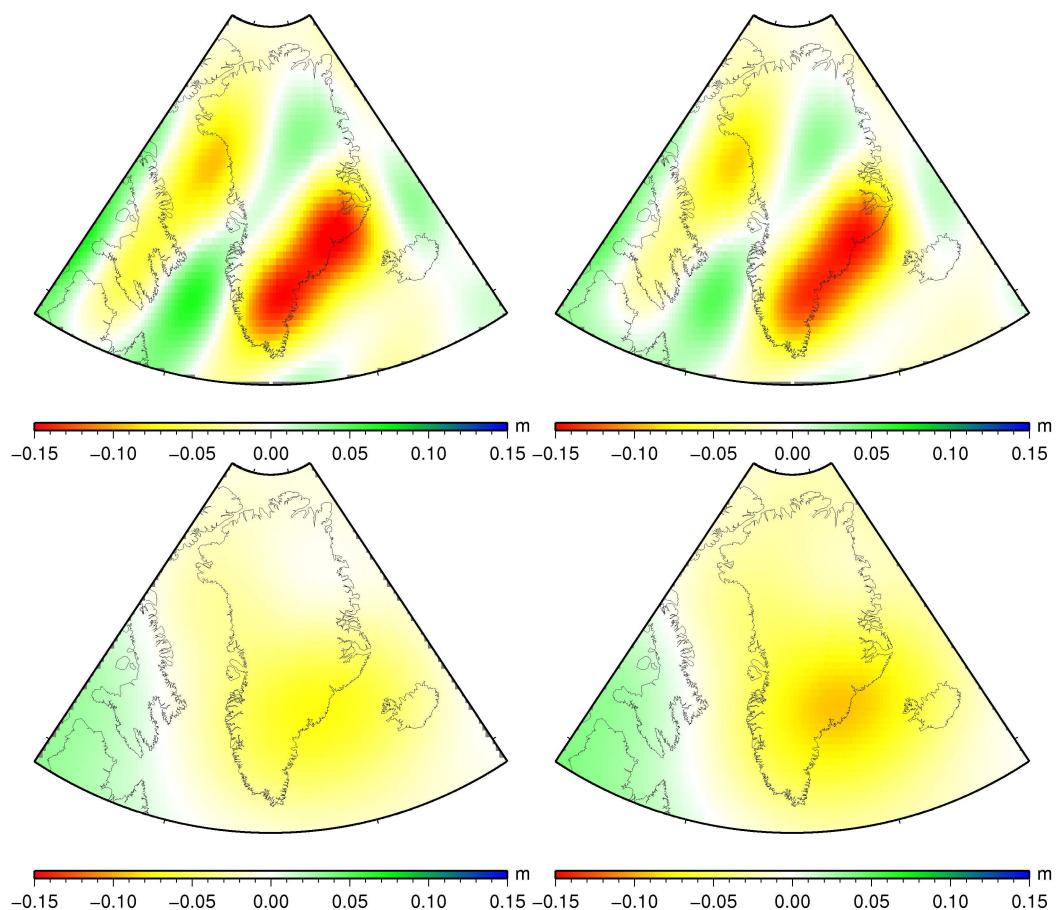
Figure 5.38: Yearly trends, unfiltered solutions: RBF 35 (top left), RBF 35 FW (top right), DEOS DS400 (bottom left), CSR DS400 (bottom right).
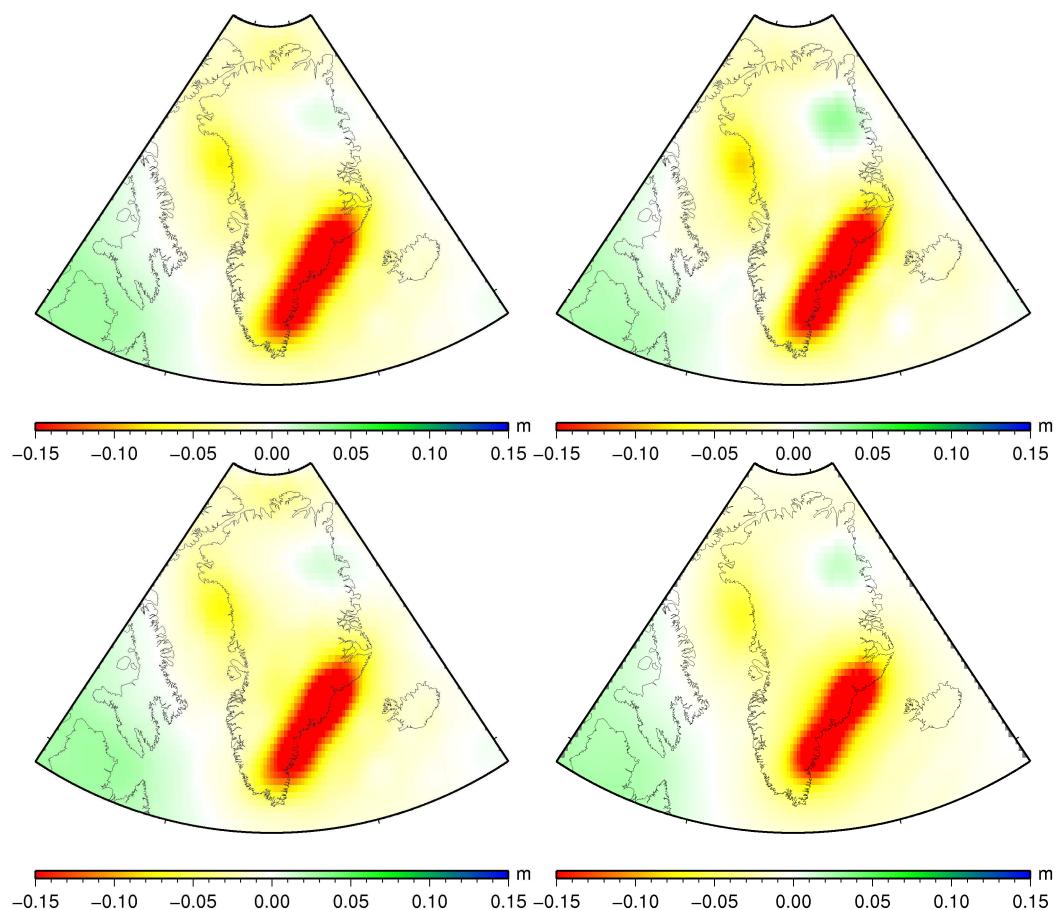
Figure 5.39: Yearly trends, filtered solutions: RBF 90 (top left), RBF 90 FW (top right), RBF 90 regional (bottom left), DMT-1 (bottom right).

ICESat data. It is difficult to compute accurate mass loss from ICESat due to unknown snow densities. ICESat yields however a much higher spatial resolution than GRACE, and can be used to judge the spatial resolution of GRACE solutions. It is evident that most ice mass loss happens along the southeastern coast, while the mass loss along the northwestern coast that is detected by some of the GRACE solutions is also visible. ICESat shows a positive trend, snow accumulation, on the southern tip of Greenland. This positive signal can not be found in GRACE plots. Some positive signal is also visible in northeastern Greenland where it is located especially by the RBF 90 FW solution.
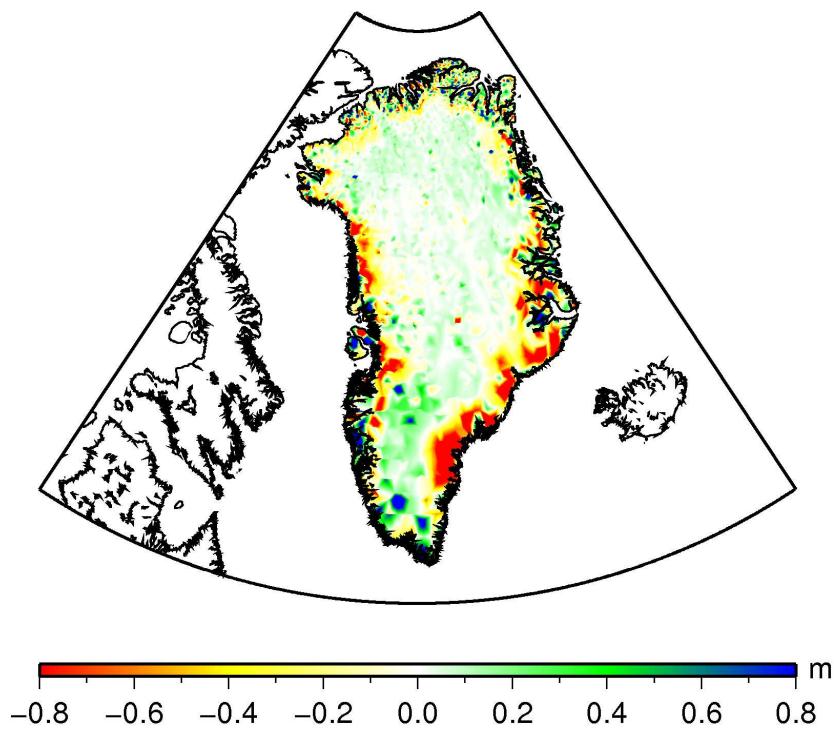
Figure 5.40: Yearly trend [m] of altitude change from ICESat.

Antarctica

A similar series of computations as has been made for Greenland has also been made for Antarctica. Figure 5.41 shows mass estimate time series for Antarctica. Several things are noticeable from this plot:

- For the unfiltered level 35 solutions, whether or not frequency-dependent weighting is used does not matter. The optimally-filtered level 90 solutions for the first time show significant differences between solutions with and without frequency-dependent weighting. The solutions with frequency-dependent data weighting (RBF 90 FW) show considerably larger maximum amplitudes.

- The differences between the various models are quite large, in 2006 even between the RBF 90 FW and DMT-1 solutions that produce very similar results in other regions (see Greenland and section 5.6.4). The differences might be caused by different stochastic models.

- The RBF regional solution (RBF 90 REG) shows significant differences with respect to the other RBF 90 solutions. The regional parametrisation causes a reduced sensitivity to long wavelengths.

- A seasonal signal, as can be seen with Greenland and other areas, is not clearly visible in most solutions. An exception are the CSR DS400 and RBF 90 regional solution, which show annual signals three time as large as other solutions, but on the other hand result in much smaller trend estimates.

The yearly trends and annual signal amplitudes resulting from the time series are listed in table 5.4. These values have not been corrected for glacial-isostatic adjustment (GIA), and thus show the total mass variation. Estimates vary between 16 Gt/yr and 74 Gt/yr. The optimally-filtered solutions are in the range of 44 to 58 Gt/yr, with the exception of the regional solution that yields the overall smallest mass gain of all optimally-filtered solutions (and thus the largest mass loss after correction for GIA).

GIA corrections for Antarctica differ depending on the model used and amount to -107 Gt/yr with the IJ05 model (Ivins and James, 2005) and -149 Gt/yr with ICE-5G (Peltier, 2004), but are required when comparing mass balance estimates from GRACE and ICE-Sat. Likely mass balance estimates from ICESat after GIA correction vary between -55 and -85 Gt/yr (Gunter et al., 2009). The RBF 90, RBF 90 FW, and DMT-1 solutions fall into this range after GIA correction, but the uncertainties involved make it impossible to use ICESat as validation for the GRACE mass balance estimate.

Once again, spatial plots of the estimated trends have been made. The plots for the unfiltered solutions are shown in figure 5.42. The two RBF 35 solutions show much more signal than the SH solutions, at the cost of more noise. The positive signals along the eastern coast lead to the largest positive trend estimates of all models presented here.

The plots for the optimally-filtered solutions are shown in figure 5.43. It is clearly visible that the spatial resolution is much higher and that noise is well suppressed compared to the unfiltered solutions. The three RBF solutions, especially the regional solution, suffer
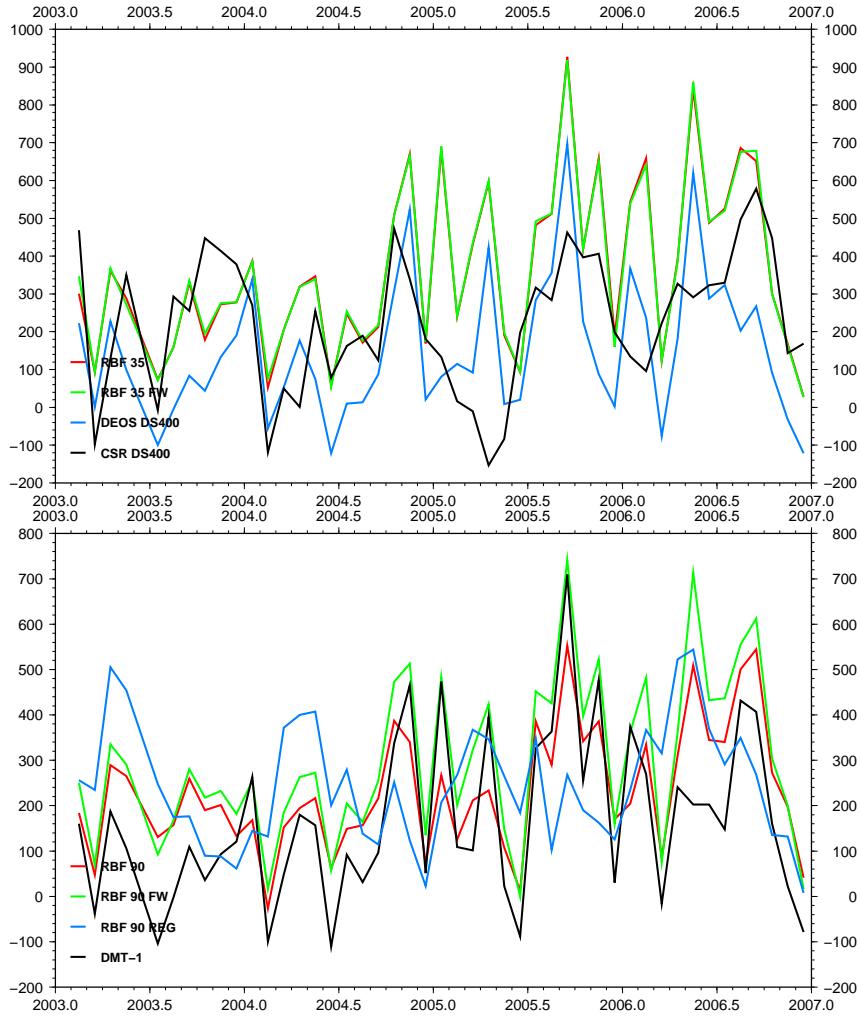
Figure 5.41: Time series [Gt] for Antarctica for unfiltered solutions (top) and filtered solutions (bottom).

model	trend	annual amplitude
RBF 35	74.4 Gt/yr	23.1 Gt
RBF 35 FW	72.0 Gt/yr	23.1 Gt
DEOS DS400	34.6 Gt/yr	17.8 Gt
CSR DS400	16.7 Gt/yr	145.0 Gt
RBF 90	43.8 Gt/yr	67.0 Gt
RBF 90 FW	58.0 Gt/yr	49.9 Gt
RBF 90 REG	19.6 Gt/yr	144.3 Gt
DMT-1	44.7 Gt/yr	52.0 Gt

Table 5.4: Yearly trends and signal amplitudes for Antarctica, computed from GRACE time series over the February 2003-December 2006 time span.
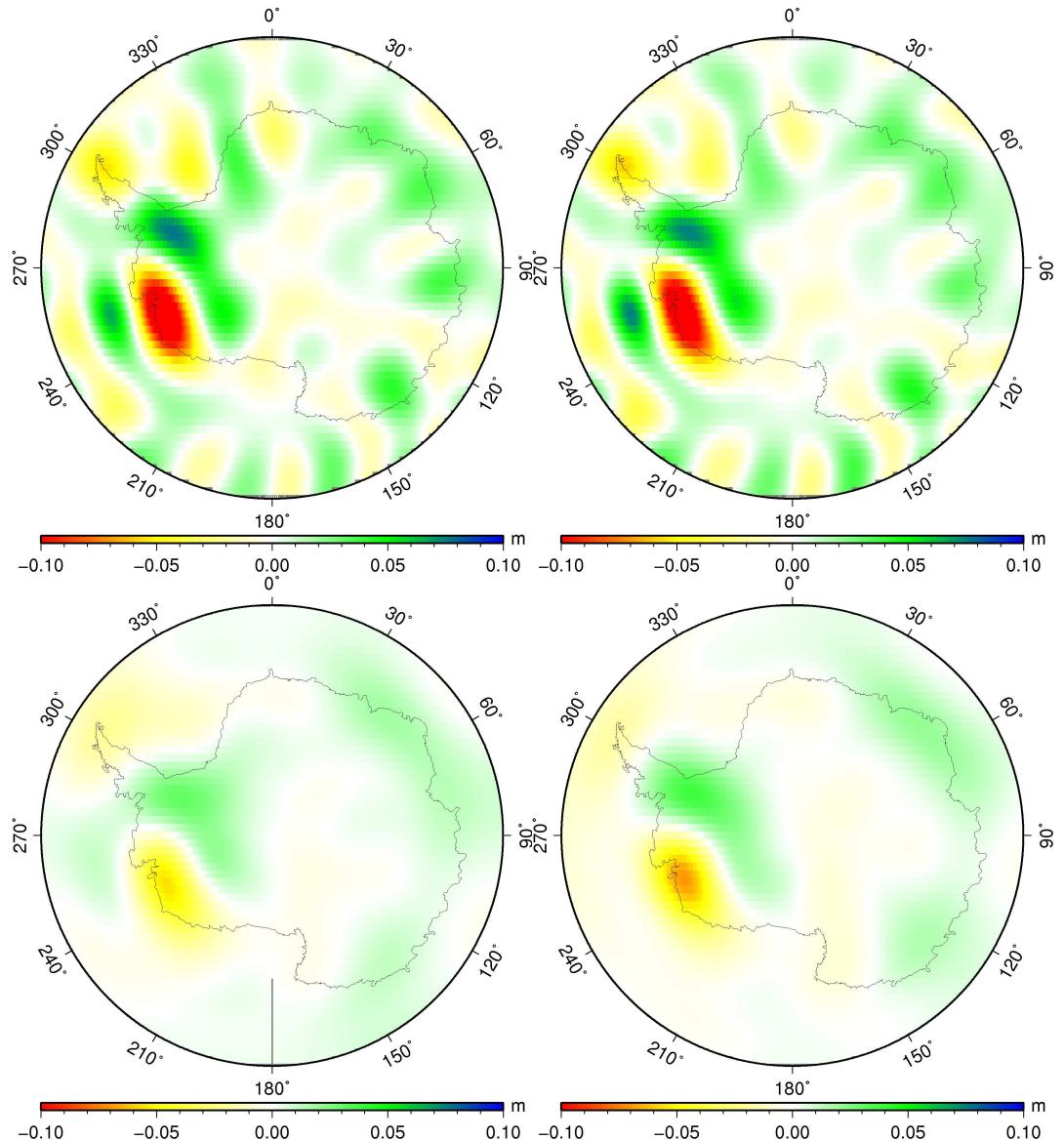
Figure 5.42: Yearly trends, unfiltered solutions: RBF 35 (top left), RBF 35 FW (top right), DEOS DS400 (bottom left), CSR DS400 (bottom right).

from some artifacts around the South Pole. These have been artificially suppressed in the computation of DMT-1 by assigning small signal variances to very high latitudes, gradually decreasing to a variance of 0 at the pole. Their exact cause is unknown; possible explanations are deficiencies in the stochastic model (too optimistic at high latitudes) and a small polar gap, since GRACE's orbit inclination is not exactly 90 degrees.

All four solutions clearly separate the positive signal over West Antarctica into two distinct spots and see mass loss at the tip of the Antarctica peninsula. This signal is stronger in the RBF solutions than in the DMT-1 model. All four solutions also detect a positive signal off the coast of west Antarctica. Only the RBF solutions detect a very slight mass loss at the coast at 0 degrees longitude (top of the plot).

More insight is gained by comparing the spatial plots to a plot of mass change derived from ICESat data, figure 5.44. The actual height change as measured by ICESat has been multiplied with the presumed densities at the points to yield a plot of mass change. Comparing this plot to figure 5.43, we see that most of the larger features seen by ICESat are also visible in the optimally-filtered GRACE solutions. Only the RBF solutions see the negative signal at 0 degrees longitude. As a result, the RBF solutions exhibit higher correlation to ICESat in space than the DMT-1 solution (see figure 5.5). This indicates that RBFs may provide better resolution at high latitudes than an SH solution.

GRACE model	Correlation coefficient
RBF 90	0.81
RBF 90 REG	0.79
DMT-1	0.75
CSR DS400	0.78

Table 5.5: Spatial correlation coefficients between various GRACE solutions and ICESat over Antarctica.
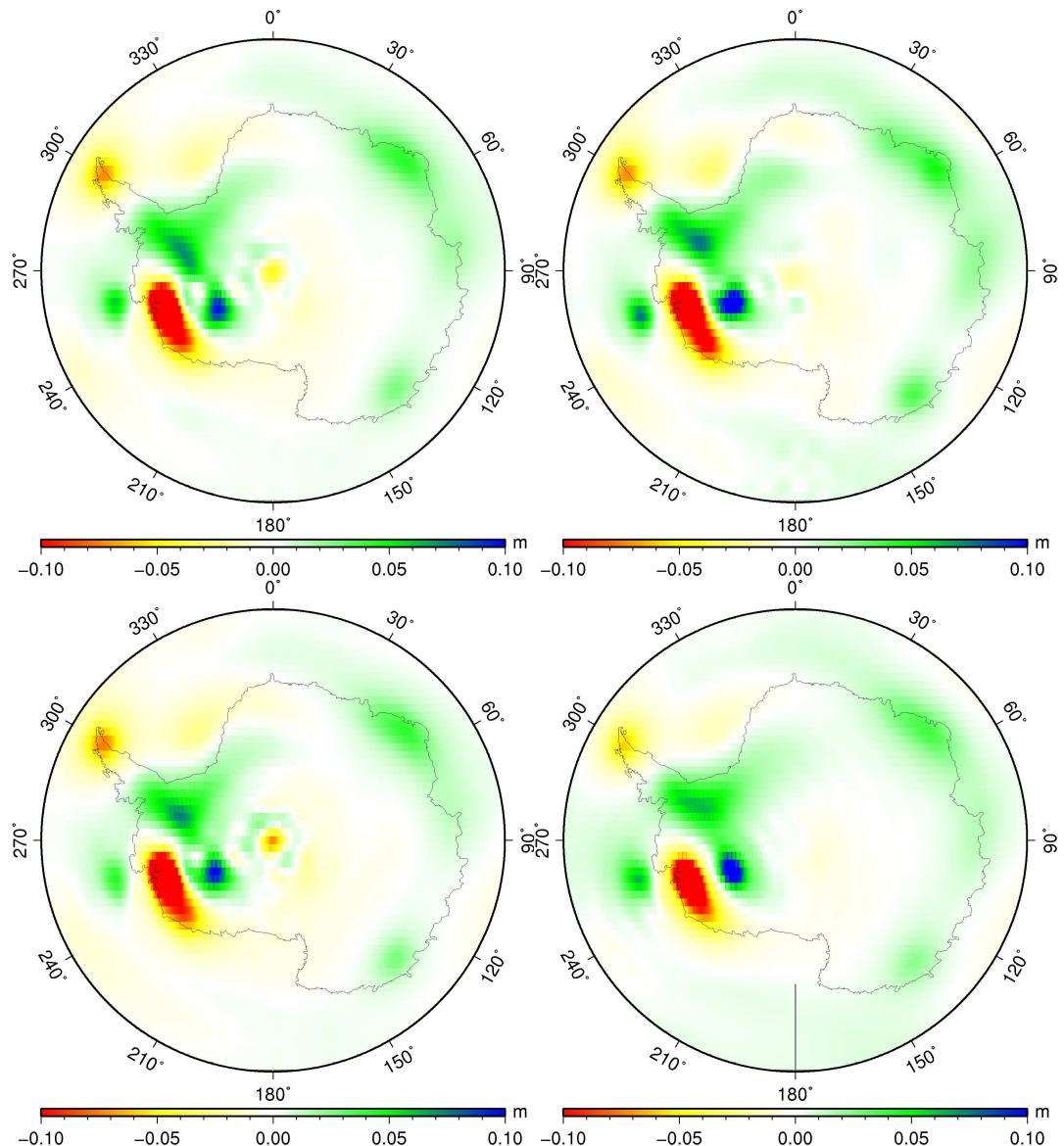
Figure 5.43: Yearly trends, filtered solutions: RBF 90 (top left), RBF 90 FW (top right), RBF 90 regional (bottom left), DMT-1 (bottom right).
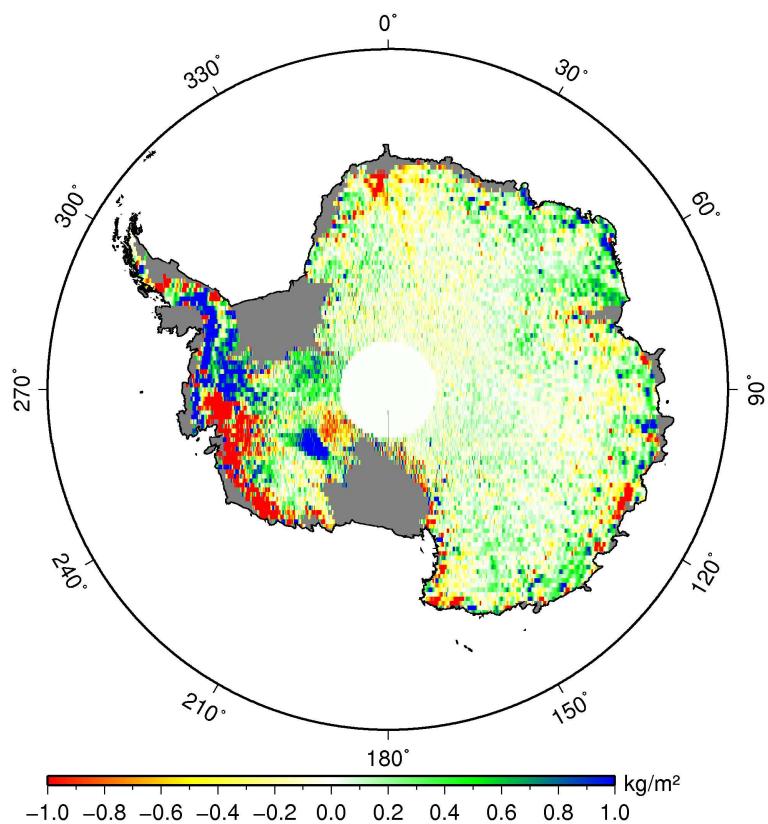
Figure 5.44: Yearly trend of mass variation computed from ICESat data (Gunter et al., 2009).

5.6.4 Recovery of terrestrial water storage variations

Besides ice mass loss, another mass transport process that can be monitored using GRACE are water storage variations at river basin scale. GRACE makes it possible to detect these worldwide and is not bound to terrestrial measurements. Results from GRACE can be used to understand hydrological cycles, to monitor changes in hydrology such as lake levels and irregular events (extreme rainfalls, draughts), and to validate and calibrate hydrological models.

Using the GRACE monthly models described in section 5.6.2, four-year time series (2003-2006) of mean monthly mass variations were computed for four river basins (fig. 5.45):

- The Mississippi river basin (area: 2,980,000 km²), an area that is well-covered by hydrology data;

- The Ob river basin (Siberia, 2,990,000 km²), where snowfall is the dominant source of signal, and rapid melting occurs in spring;

- The Rhine river basin, a comparatively small area in central Europe (185,000 km²);

- The Rio de la Plata river basin (3,100,000 km²), an area where the GRACE models may be affected from leakage due to the proximity of the Amazon river basin.

These time series are used for comparing the various GRACE models among each other and for comparison with the PCR-GLOBWB hydrological model. In addition to these four river basins, Lake Victoria in Africa is used for a comparison to satellite radar altimetry.

Comparison of RBF solutions to other GRACE models

Besides trend and annual amplitudes, this section also uses the difference rms with respect to a mean model as means of comparison:

$$\mathrm{rms}_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\delta h_{j,i} - \Delta\delta h_{mean,i}}, \qquad (5.29)$$

for each model j and n months. The mean model δh_{mean} is the average of all $u = 7$ models used,

$$\delta h_{mean,i} = \frac{1}{u}\sum_{j=1}^{u}\delta h_{j,i}. \qquad (5.30)$$

Figure 5.46 shows the time series of monthly water storage variations for the Mississippi river basin, and table 5.6 shows the related trends, amplitudes, and differences with respect to the mean model. All models show similar behaviour. Only the maximum
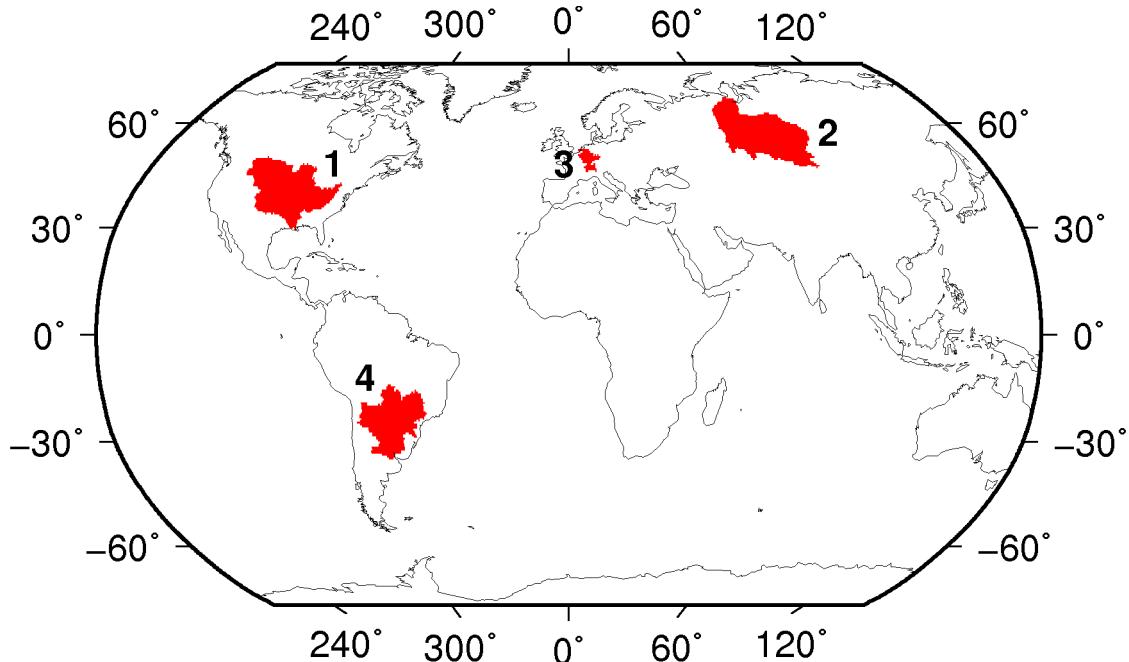
Figure 5.45: River basins used in the comparison: 1) Mississippi, 2) Ob, 3) Rhine, 4) Parana/Rio de la Plata.

amplitudes differ in some years, especially for the CSR DS400 solutions (spring 2004 and 2006). The two unfiltered SH solutions (DEOS DS400 and CSR DS400) estimate a linear trend that differs by one centimetre with respect to the other solutions, and differ the most from the mean model. RBF solutions with (RBF 35 FW, RBF 90 FW) and without frequency-dependent weighting (RBF 35, RBF 90) can be considered identical. The RBF 90 and DMT-1 solutions are also almost identical, which should not come as a surprise given that both use the same data, functional model, and optimal filtering approach.

Figure 5.47 shows the time series of monthly water storage variations for the Ob river basin, and table 5.7 shows the related trends, amplitudes, and differences with respect to the mean model. The DEOS DS400, DMT-1, and all RBF models show very similar behaviour. The CSR DS400 solutions result in slightly different maximum amplitudes, especially in the spring of 2006, with a much smaller maximum accumulation. This results in a larger rms with respect to the mean model. Solutions with and without frequency-dependent data weighting show no significant differences.

The time series of monthly water storage variations for the Rhine river basin is shown in figure 5.48, with table 5.8 showing the related trends, amplitudes, and differences with respect to the mean model. The three optimally-filtered solutions result once again in very similar time series. The DEOS DS400 and CSR DS400 solutions also show the same seasonal behaviour, but smaller signal amplitudes. Here, the unfiltered level 35 RBF solutions show somewhat erratic behaviour and much too large minimum amplitudes in the summers of 2003 and 2005. Obviously the chosen parametrisation results in a solution that is not smooth enough for this particular small area, as oscillations are not averaged over a large enough area. Interestingly, the amplitude estimates are compara-
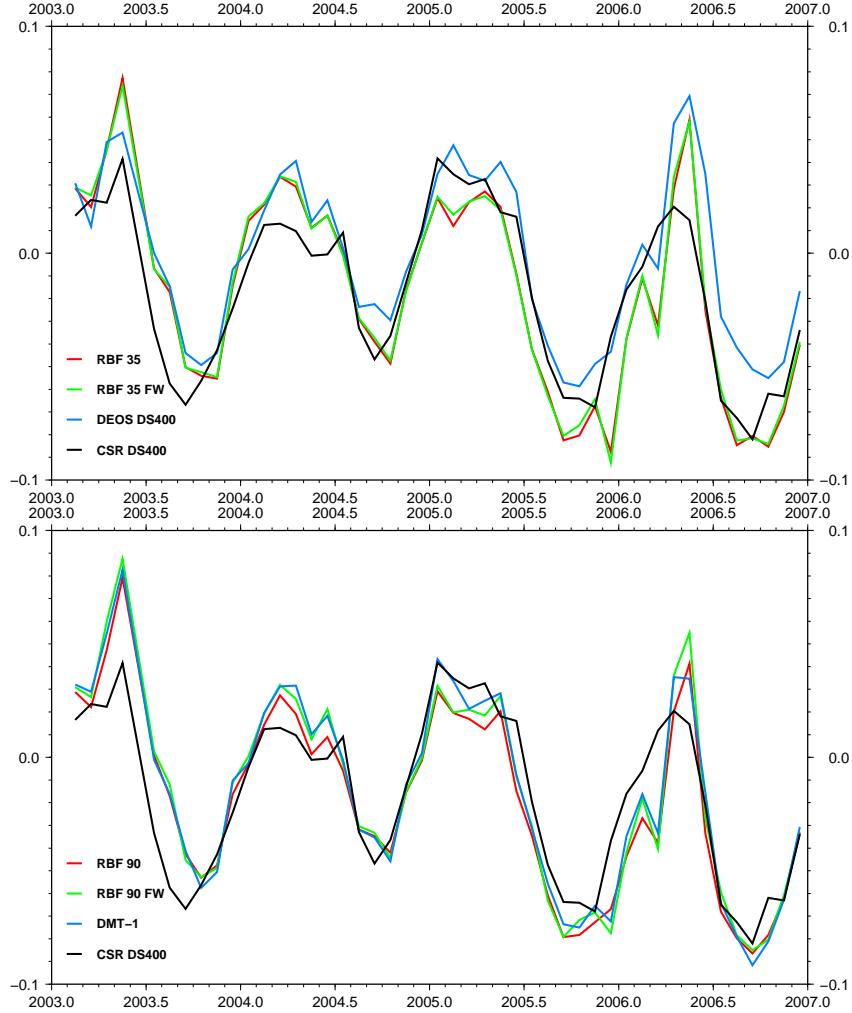
Figure 5.46: Time series in EWH [m] of monthly water storage variations for the Mississippi river basin; unfiltered solutions (top) and filtered solutions plus CSR DS400 (bottom).

ble to the three optimally-filtered solutions. Trends are positive instead of negative, but absolute differences in the trend estimate amount to only a few millimetres.

Finally, figure 5.49 shows the time series of monthly water storage variations for the Rio de la Plata river basin, and table 5.9 shows the related trends, amplitudes, and differences with respect to the mean model. The four RBF solutions, the DEOS DS400 models, and DMT-1 result in very similar time series, with identical solutions regardless whether or not frequency-dependent weighting is used. The CSR DS400 solutions behave quite differently and show a clear phase shift and even a reversed seasonal cycle, which leads to a large rms with respect to the mean model. Klees et al. (2008c) have attributed the different behaviour to different processing strategies.

A spatial plot of October 2005, a month exhibiting large differences between different GRACE solutions, is shown in figure 5.50. The CSR DS400 solution only barely captures the large positive signal modelled in the two RBF solutions. In addition to this, the signal in the Amazon river basin is smeared out by the Gaussian smoothing, resulting in leakage and a negative mass balance estimate for the CSR DS400 solution instead of the

model	trend	annual amplitude	rms to mean model
RBF 35	-1.4 cm/yr	4.5 cm	1.2 cm
RBF 35 FW	-1.4 cm/yr	4.5 cm	1.1 cm
DEOS DS400	-0.4 cm/yr	4.3 cm	1.9 cm
CSR DS400	-0.6 cm/yr	4.2 cm	2.6 cm
RBF 90	-1.5 cm/yr	3.9 cm	1.3 cm
RBF 90 FW	-1.5 cm/yr	4.3 cm	1.0 cm
DMT-1	-1.4 cm/yr	4.4 cm	0.6 cm

Table 5.6: Yearly trends, signal amplitudes, and rms to mean model for the Mississippi river basin in terms of EWH, computed from GRACE time series over the February 2003-December 2006 time span.

model	trend	annual amplitude	rms to mean model
RBF 35	0.5 cm/yr	4.9 cm	0.5 cm
RBF 35 FW	0.6 cm/yr	5.0 cm	0.7 cm
DEOS DS400	0.4 cm/yr	5.3 cm	1.4 cm
CSR DS400	-0.2 cm/yr	5.1 cm	1.7 cm
RBF 90	0.5 cm/yr	4.7 cm	0.5 cm
RBF 90 FW	0.6 cm/yr	5.0 cm	0.5 cm
DMT-1	0.6 cm/yr	5.1 cm	0.9 cm

Table 5.7: Yearly trends, signal amplitudes, and rms to mean model for Ob river basin in terms of EWH, computed from GRACE time series over the February 2003-December 2006 time span.

strong positive signal seen in the other solutions.

Several conclusions can be drawn from the comparisons of the various GRACE solutions:

- Whether or not frequency-dependent data weighting is used does not matter for basin averages. The plots showed almost identical time series, and amplitude estimates differ only by 0% to 10%.

- The optimally filtered RBF 90 / RBF 90 FW and DMT-1 solutions yield similar to identical basin averages. Amplitude estimates are within a few %, with the exception of the Rhine river basin. For this small basin a 15% lower amplitude estimate results from the DMT-1 solution.

- For larger basins, the unfiltered level 35 RBF solutions perform similar to other solutions. With small basins, as could be seen with the Rhine basin, noise has too much influence. In order to achieve a good mass estimate, a smaller grid level (larger grid spacing) and corresponding wider bandwidth need to be chosen.

- With the exception of the Rio de la Plata river basin, all RBF models and DMT-1 showed generally similar behaviour to the CSR DS400 solutions, but often result in slightly larger signal amplitudes. For the Rio de la Plata river basin, the CSR
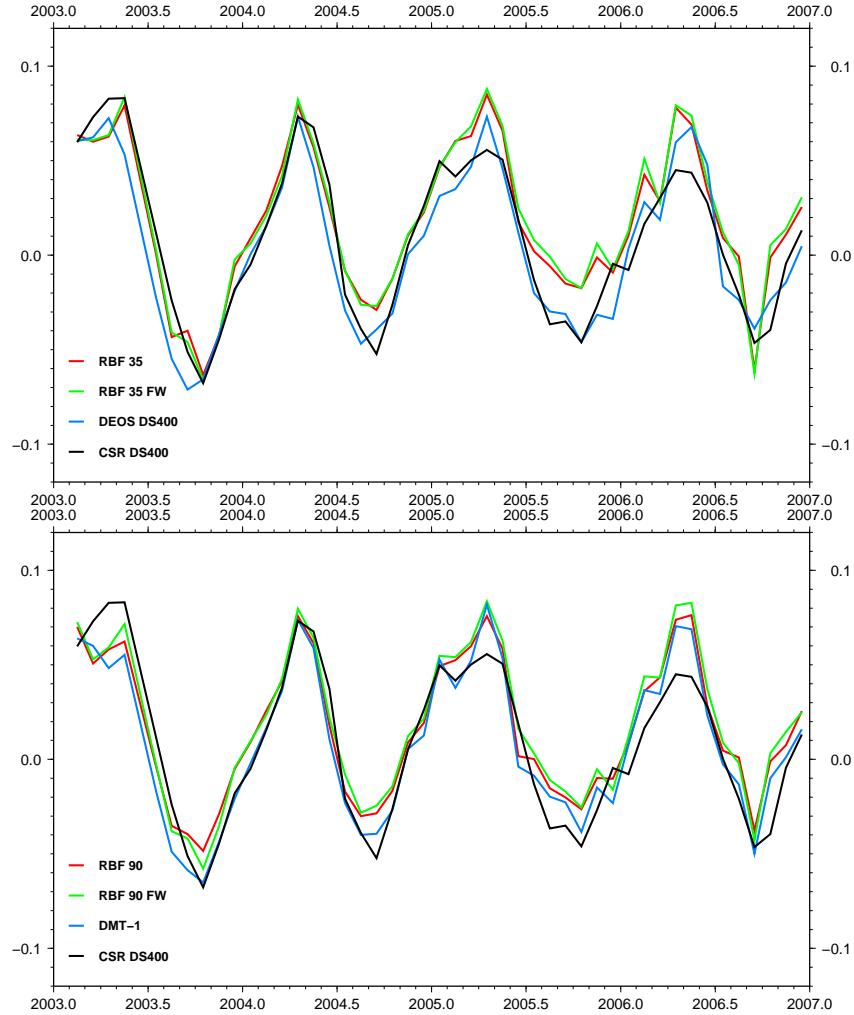
Figure 5.47: Time series in EWH [m] of monthly water storage variations for the the Ob river basin; unfiltered solutions (top) and filtered solutions plus CSR DS400 (bottom).

DS400 solution is clearly affected by leakage from the Amazon river basin. This leads to phase shifts with regard to both the filtered and unfiltered RBF solutions, and DMT-1.
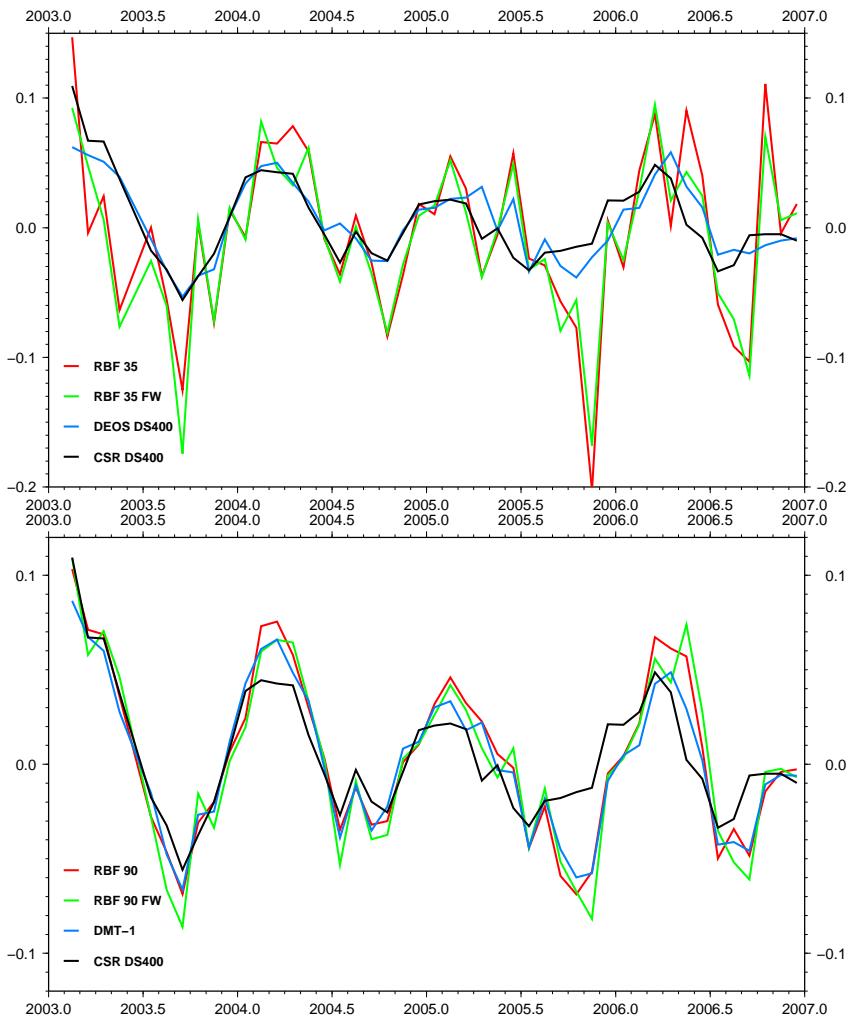
Figure 5.48: Time series in EWH [m] of monthly water storage variations for the Rhine river basin; unfiltered solutions (top) and filtered solutions (bottom).

model	trend	annual amplitude	rms to mean model
RBF 35	0.3 cm/yr	5.3 cm	3.5 cm
RBF 35 FW	0.5 cm/yr	5.3 cm	2.8 cm
DEOS DS400	0.2 cm/yr	3.7 cm	2.4 cm
CSR DS400	-0.5 cm/yr	3.6 cm	5.2 cm
RBF 90	-0.4 cm/yr	5.3 cm	1.8 cm
RBF 90 FW	-0.3 cm/yr	5.3 cm	1.5 cm
DMT-1	-0.6 cm/yr	4.5 cm	1.6 cm

Table 5.8: Yearly trends, signal amplitudes, and rms to mean model for the Rhine river basin in terms of EWH, computed from GRACE time series over the February 2003-December 2006 time span.
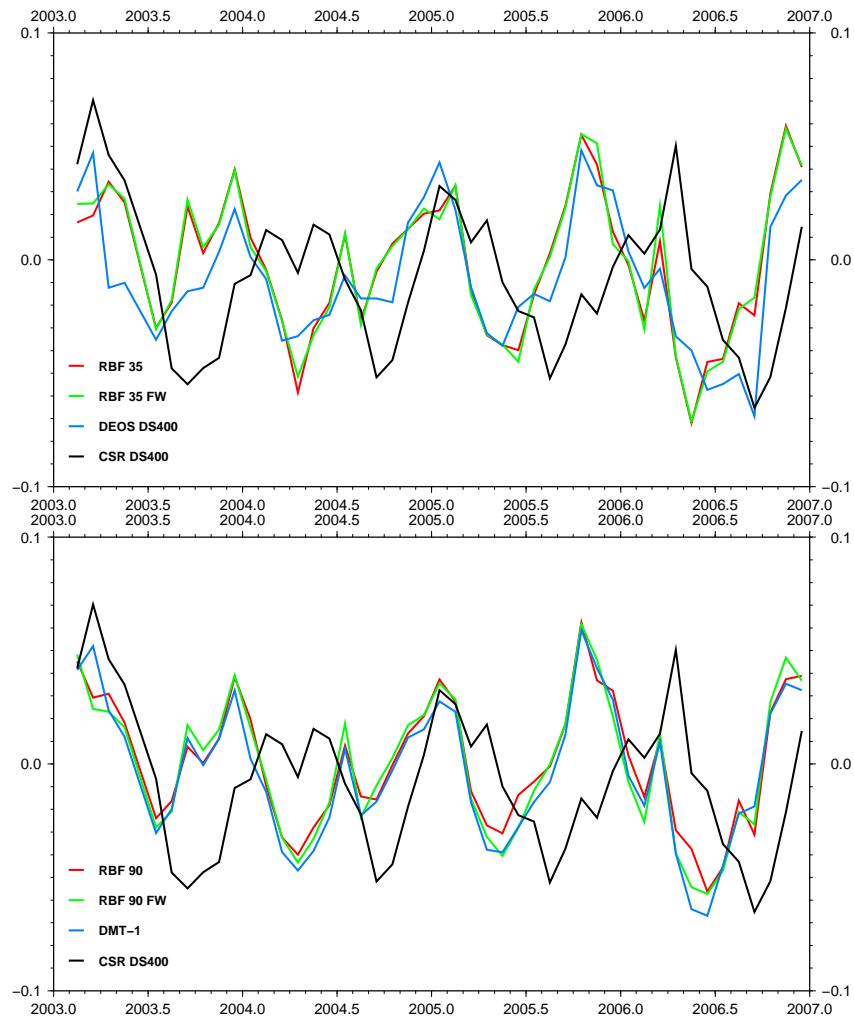
Figure 5.49: Time series [m] in EWH [m] of monthly water storage variations for the Rio de la Plata river basin; unfiltered solutions (top) and filtered solutions plus CSR DS400 (bottom).

model	trend	annual amplitude	rms to mean model
RBF 35	-0.5 cm/yr	3.1 cm	1.2 cm
RBF 35 FW	-0.5 cm/yr	3.2 cm	1.3 cm
DEOS DS400	-0.4 cm/yr	3.1 cm	1.1 cm
CSR DS400	-0.3 cm/yr	3.6 cm	4.6 cm
RBF 90	-0.5 cm/yr	2.8 cm	1.0 cm
RBF 90 FW	-0.6 cm/yr	3.2 cm	1.1 cm
DMT-1	-0.6 cm/yr	3.1 cm	0.9 cm

Table 5.9: Yearly trends, signal amplitudes, and rms to mean model for the Rio de la Plata river basin in terms of EWH, computed from GRACE time series over the February 2003-December 2006 time span.
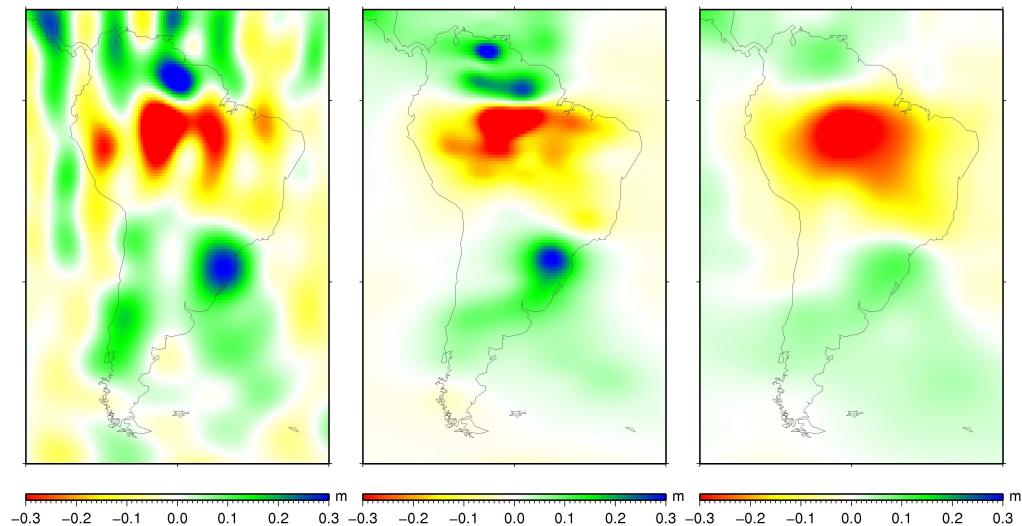
Figure 5.50: Monthly solutions for October 2005 over South America. RBF 35 (unfiltered, left), RBF 90 (optimally-filtered, centre), and CSR DS400 (destriped and 400km Gaussian smoothing, right).
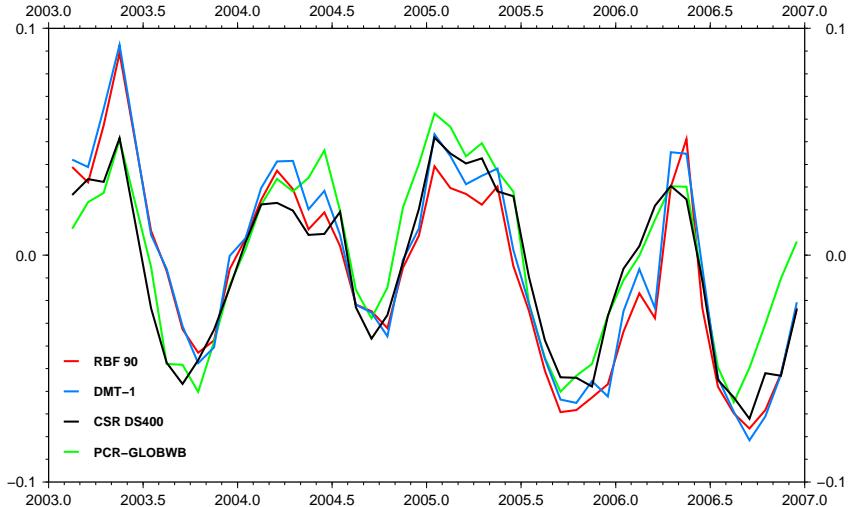
Figure 5.51: Time series in EWH [m] of monthly water storage variations for the Mississippi river basin, computed from GRACE and PCR-GLOBWB.

Comparison of GRACE solutions to the PCR-GLOBWB hydrological model

In addition to a comparison of different GRACE models, validation was attempted by comparison to a hydrological model. The PCR-GLOBWB model (van Beek, 2007) was used. PCR-GLOBWB is assumed to be of higher quality than e.g. the GLDAS (Rodell et al., 2004)global hydrology model. It was compared to the optimally filtered RBF level 90 solutions, the DMT-1, and the CSR DS400 (destriped, 400 km Gaussian smoothing) models, using the same four river basins as before.

model	trend	annual amplitude	rms
RBF 90	-1.5 cm/yr	3.9 cm	3.5 cm
DMT-1	-1.4 cm/yr	4.4 cm	2.5 cm
CSR DS400	-0.6 cm/yr	4.2 cm	1.3 cm
PCR-GLOBWB	-0.3 cm/yr	4.0 cm	

Table 5.10: Yearly trends, signal amplitudes, and rms error to PCR-GLOBWB for the Mississippi river basin in terms of EWH, computed from GRACE and PCR-GLOBWB time series over the time span February 2003-December 2006.

Figure 5.51 shows the time series of monthly water storage variations for the Mississippi river basin, the associated trend and amplitude estimates are listed in table 5.10. All models show similar behaviour. The maximum amplitudes differ somewhat, especially in the spring 2003 and autumn 2006, which also results in different trend estimates. These differences can be attributed to model deficiencies and poor rainfall data. The CSR DS400 solutions exhibits the best fit to the hydrological model.

The time series of monthly water storage variations for the Ob river basin are shown in figure 5.52, trend and amplitude estimates are shown in table 5.11. While the phases and general behaviour are similar, differences in maximum amplitudes exist. Spring 2003 and spring 2006 show more signal in GRACE than in PCR-GLOBWB. The RBF 90 and
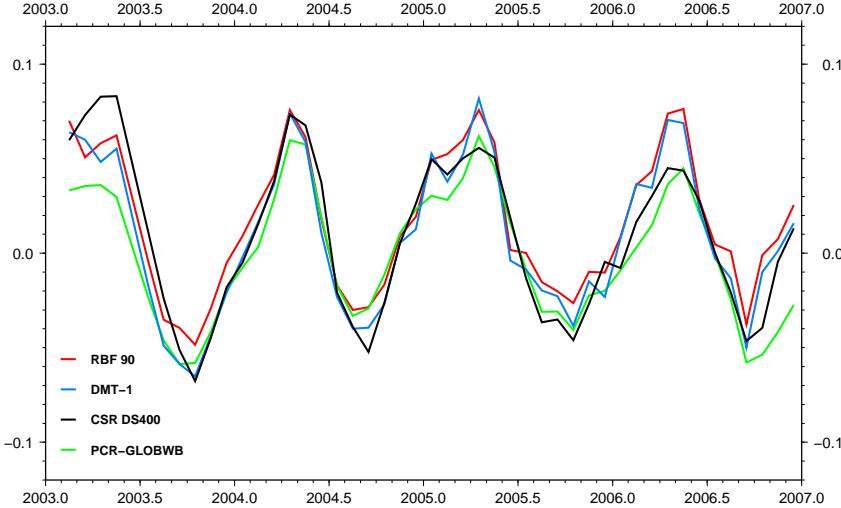
Figure 5.52: Time series in EWH [m] of monthly water storage variations for the Ob river basin, computed from GRACE and PCR-GLOBWB.

model	trend	annual amplitude	rms
RBF 90	0.5 cm/yr	4.7 cm	2.2 cm
DMT-1	0.6 cm/yr	5.1 cm	1.8 cm
CSR DS400	0.4 cm/yr	5.3 cm	3.3 cm
PCR-GLOBWB	0.0 cm/yr	4.3 cm	

Table 5.11: Yearly trends, signal amplitudes, and rms error to PCR-GLOBWB for the Ob river basin in terms of EWH, computed from GRACE and PCR-GLOBWB time series over the February 2003-December 2006 time span.

DMT-1 models exhibit a significantly better fit to PCR-GLOBWB than the CSR DS400 model.

model	trend	annual amplitude	rms
RBF 90	-0.4 cm/yr	5.3 cm	1.8 cm
DMT-1	-0.6 cm/yr	4.5 cm	1.6 cm
CSR DS400	-0.5 cm/yr	3.6 cm	5.7 cm
PCR-GLOBWB	-0.4 cm/yr	4.8 cm	

Table 5.12: Yearly trends, signal amplitudes, and rms error to PCR-GLOBWB for the Rhine river basin in terms of EWH, computed from GRACE and PCR-GLOBWB time series over the February 2003-December 2006 time span.

Figure 5.53 shows the time series of monthly water storage variations for the Rhine river basin, table 5.12 lists the trends and signal amplitudes. In the winter of 2004/2005, PCR-GLOBWB yields a significantly larger maximum amplitude than the GRACE models, with the RBF solution exhibiting the smallest difference. In the summer of 2005, the CSR DS400 solution deviates from the other models. This results in a smaller amplitude estimate for the CSR DS400 solution compared to the other models, and a much larger rms error with respect to PCR-GLOBWB.
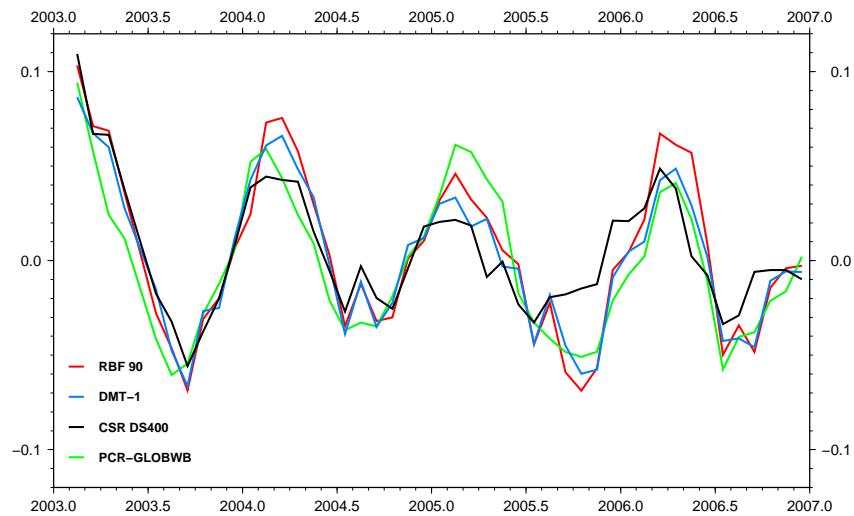
Figure 5.53: Time series in EWH [m] of monthly water storage variations for the Rhine river basin, computed from GRACE and PCR-GLOBWB.
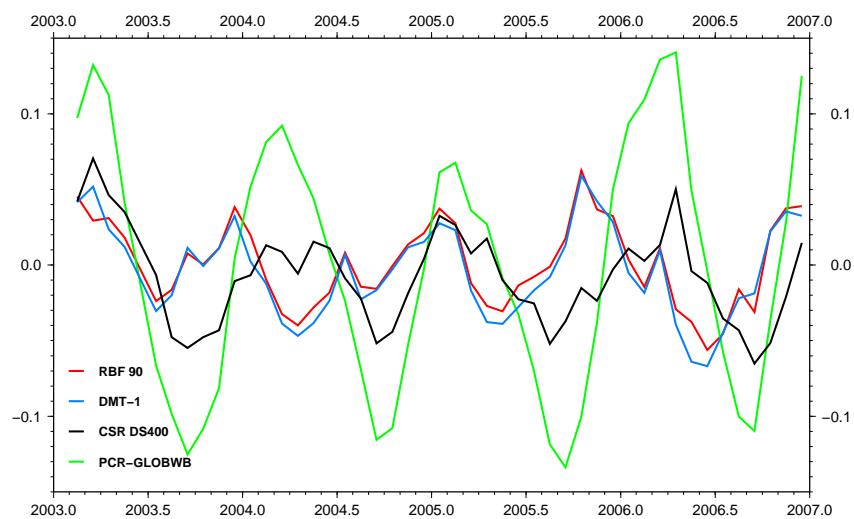

Figure 5.54: Time series in EWH [m] of monthly water storage variations for Rio de la Plata river basin, computed from GRACE and PCR-GLOBWB.

model	trend	annual amplitude	rms
RBF 90	-0.5 cm/yr	2.8 cm	8.4 cm
DMT-1	-0.6 cm/yr	3.1 cm	8.6 cm
CSR DS400	-0.4 cm/yr	3.1 cm	6.9 cm
PCR-GLOBWB	1.0 cm/yr	10.9 cm	

Table 5.13: Yearly trends, signal amplitudes, and rms error to PCR-GLOBWB for the Rio de la Plata river basin in terms of EWH, computed from GRACE and PCR-GLOBWB time series over the February 2003-December 2006 time span.

Finally, figure 5.54 shows the time series of monthly water storage variations for the Rio de la Plata river basin, the related trends and amplitudes are listed in table 5.13. It has previously been established that for this basin, the solutions computed at DEOS behave differently from those computed at CSR. In fact, it can be seen that the CSR DS400 solutions show a behaviour that more similar to the hydrological model (although with much smaller amplitudes), while the RBF and DMT-1 models yield a phase shift. One example is October 2005, where the RBF and DMT-1 solutions show a large positive signal, while the hydrological model produces a negative amplitude. The CSR DS400 solution is somewhere in between.

It has been seen before in figure 5.50 that the CSR DS400 solutions differ significantly for certain months as it is strongly affected by noise and has a lower spatial resolution. It is more difficult to assess why the hydrological model yields such different results with regard to the RBF 90 and DMT-1 solutions. It has been suggested by Klees et al. (2008c) that GRACE solutions may be affected by leakage from the Amazon river basin. Figure 5.50 shows that this is not the case for the RBF 90 solution. Instead, the time series of the PCR-GLOBWB model follows the general behaviour of the time series of the Amazon river basin with a slight phase shift (figure 5.55). This yields the question whether the hydrological model is accurate, since GRACE solutions show opposite cycles in the Rio de la Plata and Amazon river basins. The trustworthiness of the hydrological model for this particular area is further questioned by the large signal amplitudes computed from PCR-GLOBWB, as Berbery and Barros (2002) state that mean variations over de la Plata river basin are small due to different precipitation regimes within the basin. This is supported by the RBF and DMT-1 GRACE models.

Overall, it can be summarised that, with the exception of the Rio de la Plata river basin, de RBF 90 and DMT-1 solutions show good agreement with the PCR-GLOBWB hydrological model. The results obtained for the Rio de la Plata river basin question the reliability of the hydrological model for this particular area.
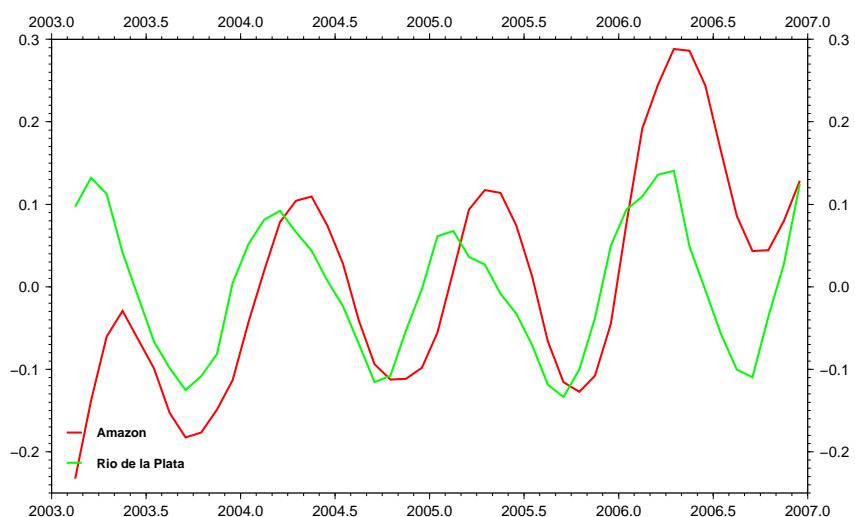
Figure 5.55: Time series in EWH [m] of monthly water storage variations for Amazon and Rio de la Plata river basins, computed from the PCR-GLOBWB hydrological model.

model	trend	annual amplitude
RBF 35	-5.9 cm/yr	1.2 cm
RBF 35 FW	-6.0 cm/yr	1.0 cm
DEOS DS400	-0.4 cm/yr	4.3 cm
CSR DS400	-3.6 cm/yr	1.4 cm
RBF 90	-5.5 cm/yr	0.9 cm
RBF 90 FW	-6.6 cm/yr	0.8 cm
DMT-1	-4.4 cm/yr	0.8 cm
altimetry	-35.1 cm/yr	9.4 cm

Table 5.14: Yearly trends and signal amplitudes computed from GRACE time series and radar altimetry over the February 2003-December 2006 time span; in terms of EWH (GRACE) and lake surface level change (altimetry).

Comparison of GRACE solutions to radar altimetry - Lake Victoria

Satellite radar altimetry makes it possible to accurately measure surface height changes of water bodies (Berry et al., 2005). These height changes can be used for comparison with EWH estimates obtained from GRACE. A suitable water body for such a comparison is Lake Victoria (Awange, 2006):

1. The size of Lake Victoria is on the edge of the spatial resolution we get from GRACE. Mass change is large enough to be detected by GRACE.

2. It exhibits a trend that can be recovered by GRACE and that can be used to estimate scaling factors.

No leakage correction (see section 5.6.3) has been applied. Leakage correction could be applied by expanding the integration area as has been done for Greenland and Antarctica, but this might result in signal from outside the area of interest being taken into account.

Figure 5.56 shows time series of mass variations for Lake Victoria obtained from the same GRACE models used previously and described in section 5.6.2. Behaviour is similar to what we have seen before: Little to no difference between solutions with and without frequency dependent weighting and good agreement between the optimally filtered RBF and DMT-1 solutions. Like with the Rhine river basin, the integration area is too small to average out noise in the RBF 35 solution, leading to a noisy time series.

Table 5.14 lists linear yearly trends and signal amplitudes as obtained from the GRACE solutions and radar altimetry. All trends resulting from GRACE solutions are significantly smaller than the trend observed by radar altimetry. The strongest trend of a GRACE solution is retrieved from the RBF 90 FW solution. Even though the RBF 90 FW time series plot looks very similar compared to the RBF 90 solution (no frequency-dependent data weighting), the trend estimate is different - mostly due to the lower value for November 2006 obtained from the RBF 90 FW solution.

Figure 5.57 shows the time series as measured by radar altimetry and the scaled GRACE time series. The scaling factor was computed from the ratio of GRACE to altimetry
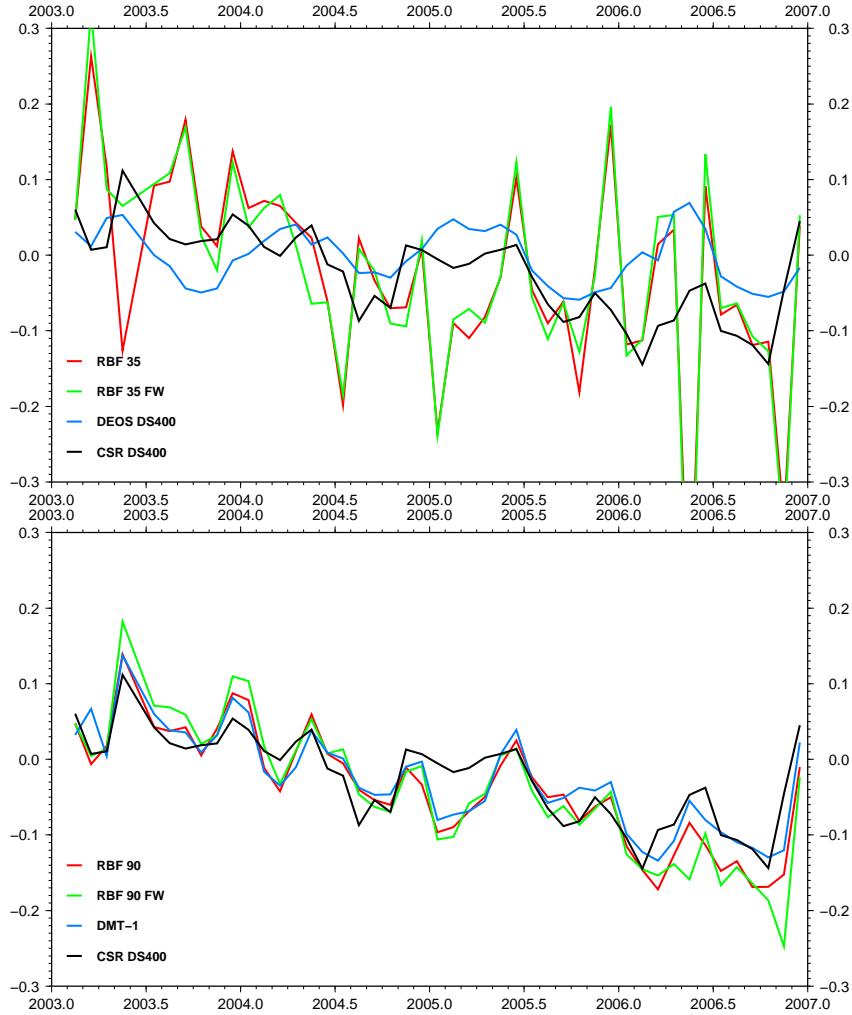
Figure 5.56: Time series in EWH [m] of monthly water storage variations for Lake Victoria; unfiltered solutions (top) and filtered solutions (bottom).

trend, $s_i = \frac{t_i}{t_{alti}}$. The GRACE models clearly echo the seasonal behaviour that is shown in the altimetry time series. Differences between the three GRACE time series shown are small.

5.7 Summary and conclusions

This chapter displayed the use of RBFs for modelling gravity from data collected by the GRACE satellite mission. For GRACE modelling, the functional model and optimal filtering approach developed at DEOS was used. Network design and bandwidth selection as it relates to GRACE modelling was covered:

- For GRACE gravity field modelling, the RBFs need to be placed on an equidistant grid. Reuter grids have this property; moreover, they allows for sufficiently fine adjustments in the number of RBFs used by changing the level. Additionally,
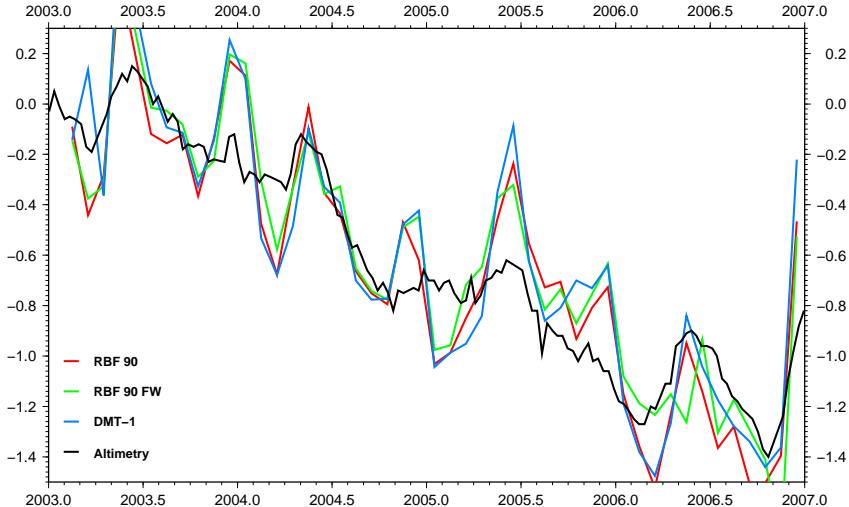
Figure 5.57: Altimetry time series and scaled GRACE time series [m].

the Reuter grid level chosen is approximately equal to the maximum spherical harmonic degree of the solution.

- Unfiltered GRACE solutions should make use of Reuter grid levels 25 to 35; level 45 may be attempted for high-latitude areas such as Greenland. Note that the choice of the Reuter grid level implies implicitly a filtering by parametrisation.

- Optimally filtered solutions can use much higher grid levels. Noise due to over-parametrisation is not an issue, as this is suppressed by the filter. For the optimal filter of Klees et al. (2008b), level 90 has proven to provide very good solutions.

- Data and parametrisation area should be larger than the target area. Empirically it was found than an area extension of 30° around the target area needs be used to suppress edge effects sufficiently well; only observations above this parametrised area should be used.

- The RBFs need to be placed deep. Typical depths are 1,600 km for a level 35 Reuter grid and 700 km for a level 90 Reuter grid when working with order-3 Poisson wavelets. Empirical testing has shown that depths that differ 10% (more in the case of unfiltered solutions) from the optimal depth have little adverse effect on the solution.

The first test presented here compared an unfiltered RBF solution to an unfiltered spherical harmonic solution. This comparison showed that the two solution were approximately equal up to a spherical harmonic degree equal to the chosen Reuter grid level. The results focused on the comparison of various GRACE solutions for two main applications: The detection and quantification of ice mass loss in Greenland and Antarctica, and the monitoring of hydrological signals.

- The optimally-filtered solutions clearly outperform unfiltered solutions, especially in terms of spatial resolution. The unfiltered models in most cases yield comparable time series behaviour, often with smaller maximum amplitudes. The level 35

RBF solution is too noisy for mean mass estimates over smaller basins, as noise is not averaged out over a larger area.

- In the case of unfiltered solutions, RBF representations yield much better spatial localisation of signals and higher amplitudes than spherical harmonic representations.

- With the exception of Antarctica, whether or not frequency-dependent data weighting was used did not result in significant differences in the monthly mean mass estimates.

- Mass loss estimates and trends for Greenland are similar for the RBF 90 and DMT-1 solutions. This also holds for signal amplitudes and spatial patterns.

- For Antarctica, there are significant differences between the SH and RBF solutions. The RBF 90 solutions seem to yield a slightly better spatial resolution and show a mass loss at 0 degree longitude that is also evident from ICESat measurements. As a result, the RBF 90 solutions feature higher correlation coefficients with regard to ICESat than DMT-1. This indicates that RBFs may offer advantages for GRACE modelling at high latitudes.

- For larger river basins, which are located at low- and mid-altitudes, the RBF 90 solutions and DMT-1 yield very similar results. This indicates that RBF representations produce solutions of the same quality as SH representations at equatorial latitudes.

- Small differences exist between the RBF 90 and DMT-1 solutions for the relatively small (185,000 km², 15 times smaller than the large basins used) Rhine river basin (15% difference in annual amplitude) and Lake Victoria (area: 68,800 km²); the RBF 90 FW yielded the largest trend estimate (50% larger than DMT-1) and was thus closest to the trend observed by altimetry. This indicates that RBF representations may yield better mass storage estimates for very small areas on the edge of GRACE's spatial resolution.

- With the exception of the Rio de la Plata river basin, the RBF solutions (as well as DMT-1) showed good agreement with the PCR-GLOBWB hydrological model (amplitude differences $< 10\%$). The discrepancies for the Rio de la Plata basin question the reliability of the hydrological model for this area.

Overall, it can be concluded that the use of RBFs for GRACE gravity field modelling results in solutions of at least equal quality compared to spherical harmonic solutions. It is difficult to assess whether RBFs can produce better solutions than spherical harmonics. There are some indications, such as the detection of mass loss in Antarctica at 0 degree latitude that is not visible in DMT-1, and the stronger trend observed for Lake Victoria. Note that the optimal filter applied to RBF solutions used a a noise covariance matrix $\mathbf{C}$ with a variance factor of $\sigma^2 = (18\,\mu m)^2$, whereas the DMT-1 solutions used a variance factor of $\sigma^2 = (40\,\mu m)^2$ (see section 5.3.3). It is thus not obvious whether the improvements in certain areas are due to the regional parametrisation or simply due to a higher observation weight in the estimation using RBFs.

For global optimally-filtered solutions, the numerical complexity slightly lower for computing RBF solutions (level 90: 10,000 unknowns) than for computing SH solutions ($l_{max} = 120$: 14,600 unknowns). RBFs offer an advantage when regional solutions are computed. Hence, RBFs are more flexible: They can be used for both global and regional solutions, whereas spherical harmonics are limited to global solutions.

Chapter 6

Local gravity field modelling from terrestrial data

Radial basis functions are a natural candidate for the computation of gravity fields from terrestrial gravity data. Unlike satellite data, the data coverage is confined to a small region. A regional parametrisation is thus a logical choice. Terrestrial data are often tightly spaced and make it possible to compute a high-resolution gravity field, corresponding to a very high degree of spherical harmonics. Computing such a field with spherical harmonics would require a much larger numerical effort than is necessary with RBFs.

Terrestrial data are almost always of a heterogeneous nature, both in terms of spatial distribution and quality. Data density will be less in remote and inaccessible areas such as mountains, with much higher data density and possibly oversampling in flat areas. This makes it necessary to fine-tune the RBF network, and indeed the algorithms of data-adaptivity (section 2.6.2) and local refinement (section 2.6.3) were motivated by the use of terrestrial data.

This chapter establishes how the RBF approach can be used for the fast computation of accurate local gravity fields. This includes the proper choice of RBF centres and bandwidths. The main use of local gravity fields computed from terrestrial data is to realise a height reference surface. The methodology presented here has already shown its usefulness for this task in computing a height reference surface for Germany (Klees et al., 2007) and the Netherlands (Klees et al., 2008a).

The data are usually obtained during gravimetry campaigns along levelling lines. These measurement campaigns may have been executed decades apart, using various instruments and different post-processing procedures. This results in data that vary significantly in quality, unlike satellite data that are assembled within a short time and usually are of consistent quality. This variation in data quality needs to be taken into account by proper weighting, desirably in a semi-automatic way such as variance component estimation (section 3.5).

The chapter starts with the derivation of the functional model for the observation types that are encountered when dealing with terrestrial data. Sections 6.2 and 6.3 cover the aspects of network design and bandwidth choice as they apply to terrestial data. The

approach is illustrated using two real data sets. One covering an area in the northeast of the USA (section 6.4.1), the other modelling a larger area in eastern Canada (section 6.4.2).

6.1 Functional model

The disturbing potential T is represented in terms of RBFs according to

$$T(\mathbf{x}) = \frac{GM}{R_E} \sum_{i=1}^{n} \alpha_i \Psi_i \left(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i \right), \tag{6.1}$$

where n is the number of basis functions, α_i is the coefficient of basis function i, $\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|}$ and $\hat{\mathbf{y}} = \frac{\mathbf{y}}{|\mathbf{y}|}$ are points on the unit sphere, and

$$\Psi_i (\mathbf{x}, \mathbf{y}_i) = \sum_{l=0}^{\infty} \psi_l \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} P_l \left(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i \right), \tag{6.2}$$

with ψ_l being the Legendre coefficients of the basis function, R the radius of the Bjerhammar sphere, and P_l the Legendre polynomial of degree l.

Throughout the computations presented here, the Poisson wavelets of order 3 have been used as basis functions. Their Legendre coefficients are

$$\psi_l = \frac{2l+1}{4\pi} l^n |\mathbf{y}|^l. \tag{6.3}$$

For all computations, analytical expressions of the basis functions have been employed. For the Poisson wavelets used here, the analytical expressions have been derived as (Klees et al., 2008a)

$$\Psi(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi R^2} \left(2\chi_{n+1}(\mathbf{x}, \mathbf{y}) + \chi_n(\mathbf{x}, \mathbf{y}) \right) \tag{6.4}$$

with

$$\chi_n(\mathbf{x}, \mathbf{y}) = 2 |\mathbf{y}|^{n+1} b_{n+1} + \sum_{i=1}^{n} \beta_{n,i} |\mathbf{y}|^i b_i. \tag{6.5}$$

6.1.1 Functional model for gravity disturbances

In spherical approximation, gravity disturbances δg are related to the disturbing potential T by

$$\delta g(\mathbf{x}) = -\frac{\partial T(\mathbf{x})}{\partial |\mathbf{x}|}. \tag{6.6}$$

When we apply the operator $-\frac{\partial}{\partial |\mathbf{x}|}$ to eq. (6.2), we obtain

$$-\frac{\partial}{\partial |\mathbf{x}|} \Psi_i(\mathbf{x}, \mathbf{y}_i) = \sum_{l=0}^{\infty} \psi_l \frac{l+1}{|\mathbf{x}|} \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} P_l\left(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i \right), \tag{6.7}$$

independent of the choice of the kernel and its Legendre coefficients ψ_l.

For the analytical expressions of the Poisson wavelets, we need the radial derivates of b_0 and b_1, eq. (5.10):

$$\frac{\partial b_0}{\partial |\mathbf{x}|} = -\frac{|\mathbf{x}| - |\mathbf{y}| \cos \vartheta}{|\mathbf{x} - \mathbf{y}|^3}, \tag{6.8}$$

$$\frac{\partial b_1}{\partial |\mathbf{x}|} = \frac{\cos \vartheta}{|\mathbf{x} - \mathbf{y}|^3} - \frac{3}{|\mathbf{x} - \mathbf{y}|^5} \left(|\mathbf{x}| \cos \vartheta - |\mathbf{y}| \right) \left(|\mathbf{x}| - |\mathbf{y}| \cos \vartheta \right). \tag{6.9}$$

The basis functions are then computed according to eq. (6.4) with

$$-\frac{\partial}{\partial |\mathbf{x}|} \chi_n(\mathbf{x}, \mathbf{y}) = -2 |\mathbf{y}|^{n+1} \frac{\partial b_{n+1}}{\partial |\mathbf{x}|} - \sum_{i=1}^{n} \beta_{n+1} |\mathbf{y}|^i \frac{\partial b_i}{\partial |\mathbf{x}|}. \tag{6.10}$$

6.1.2 Functional model for gravity anomalies

In spherical approximation, gravity anomalies Δg are related to the disturbing potential T by

$$\Delta g(\mathbf{x}) = -\frac{\partial T(\mathbf{x})}{\partial |\mathbf{x}|} - \frac{2}{|\mathbf{x}|} T(\mathbf{x}). \tag{6.11}$$

For the spectral representation, it follows that

$$-\frac{\partial}{\partial |\mathbf{x}|} \Psi_i(\mathbf{x}, \mathbf{y}_i) - \frac{2}{|\mathbf{x}|} \Psi_i(\mathbf{x}, \mathbf{y}_i) = \sum_{l=0}^{\infty} \psi_l \frac{(l-1)}{|\mathbf{x}|} \left(\frac{R}{|\mathbf{x}|} \right)^{l+1} P_l\left(\hat{\mathbf{x}}^T \hat{\mathbf{y}}_i \right). \tag{6.12}$$

The analytical expressions are once again computed using eq. (6.4), with

$$-\frac{\partial}{\partial |\mathbf{x}|} \chi_n(\mathbf{x}, \mathbf{y}_i) - \frac{2}{|\mathbf{x}|} \chi_n(\mathbf{x}, \mathbf{y}_i) = -2 |\mathbf{y}|^{n+1} \frac{\partial b_{n+1}}{|\partial \mathbf{x}|} + \frac{2}{|\mathbf{x}|} b_{n+1} - \sum_{i=1}^{n} \beta_{n+1} |\mathbf{y}|^i \left(\frac{\partial b_i}{\partial |\mathbf{x}|} + \frac{2}{|\mathbf{x}|} b_i \right). \tag{6.13}$$

6.1.3   Functional model for height anomalies

Height anomalies are related to the disturbing potential T by

$$\zeta(\mathbf{x}) = \frac{T(\mathbf{x})}{\gamma(\mathbf{x}')}, \tag{6.14}$$

where γ is the normal gravity and $\mathbf{x}'$ is the point on the telluroid associated with the surface point $\mathbf{x}$ by telluroid mapping.

6.2   RBF network design

6.2.1   Grid choice

When processing terrestrial data, usually a dense distribution of observations over a relatively small area has to be dealt with. This makes it possible to compute a high-resolution gravity field. Since we only deal with small areas, a global grid is not required. Overparametrisation due to increasingly dense basis function placement at higher latitudes will usually also not be a problem, so often an equiangular grid (see section 2.6) will be an adequate choice. At higher latitudes, an equidistant grid like the Reuter grid will offer advantages. In this chapter, both grid types have been used.

The grid spacing should be chosen in such a way that all signal in the smoother areas can be modelled. It is not advisable to use a grid spacing aimed at recovering all signal in areas with high signal variation, as such a grid will lead to overparametrisation in areas with smoother signal characteristics.

As a rule of thumb, choosing a grid with an amount of basis functions equal to one third the number of observations is a good starting point. In case of serious oversampling, such a grid will be too dense. An optimal grid should be constructed by trial-and-error, by experimenting with various grid spacings and comparing the solutions obtained. When in doubt, it is advisable to use the coarser grid. This reduces the numerical complexity and the danger of over-parametrisation, and unmodeled signal can be considered by local refinement.

6.2.2   Data-adaptivity and local refinement

Point-wise terrestrial gravity measurements are usually inhomogeneously distributed. Figure 6.1 shows gravity anomalies in the central United States of America obtained from the PACES database (`http://paces.geo.utep.edu/gdrp/`). Some areas have very sparse data coverage, with very dense coverage over other areas, and along levelling lines.

Basis functions should only be used when there are observations within their bandwidth, as a basis function without observations within its bandwidth will lead to numerical instability and require regularisation. To achieve this, the basis function network should be
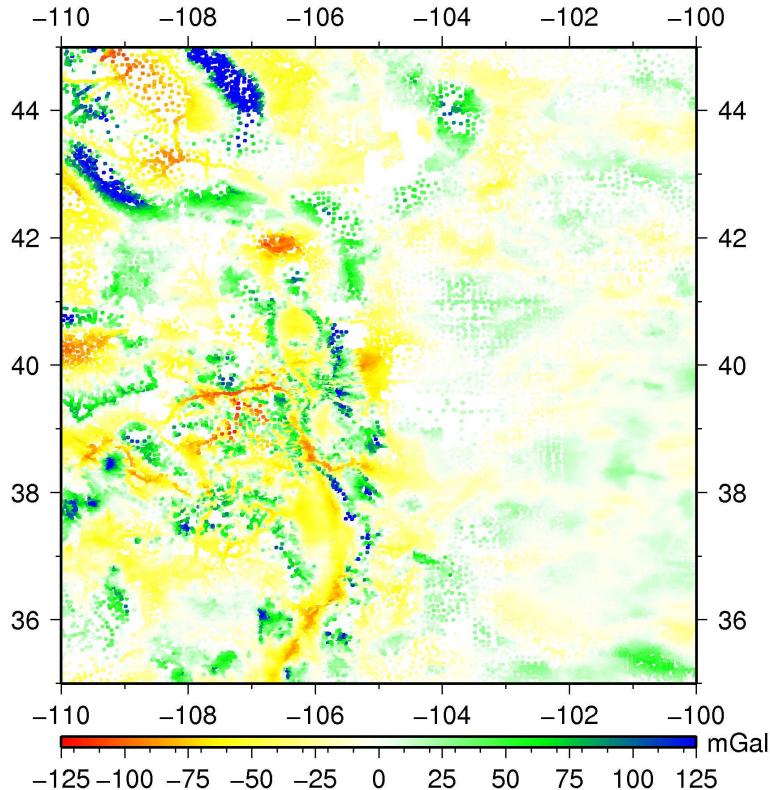
Figure 6.1: Gravity anomalies in the central USA

adapted to the data distribution. This can be achieved with the data-adaptivity algorithm discussed in section 2.6.2.

For test regions with heterogeneous signal content (i.e. strong gradients in some areas and smooth features in other area, as in figure 6.1), the RBF grid should be chosen properly for areas of smooth signal variation, not the areas with high signal variation. Choosing such a grid will lead to underparametrisation in the latter areas. This can be addressed by using the local refinement algorithm described in section 2.6.3.

6.2.3 Parametrised area

With satellite observations, the parametrised area had to be larger than the target area, in order to avoid edge effects. Terrestrial data, on the contrary, contain mostly high-frequency signal after removal of a high-quality long-wavelength model. As a result, edge effects due to truncation of the area are much smaller. When an accurate global reference model, such as EGM2008 (Pavlis et al., 2008) or EIGEN-GL04C (Foerste et al., 2008) is subtracted, edge effects can be limited to an area of at most one degree.

The parametrised area needs to be equal to the data area. If observations covering an area larger than the area of interest are available, these can be used to reduce edge effects. Using a data area that is one degree larger than the area of interest is sufficient to completely suppress edge effects.

6.3   Bandwidth selection

Dealing with terrestrial measurements is very different from dealing with satellite observations. Unlike with satellite measurements, we do not have a similar data distribution and spectral signal content for every region of interest. Each case has its own data set with specific characteristics in terms of spatial distribution and spectral content. Because of this, no recommendation for basis function distribution and associated depth / bandwidth can be given. For each case, the optimal bandwidth should be estimated, using the algorithm described in section 2.5.2.

6.4 Results

6.4.1 Northeastern USA

Area of interest

A regional gravity field was computed for an area of 10 x 10 degrees in the northeastern United States and parts of Canada (fig. 6.2). The area includes a small part of ocean, Lake Erie, Lake Ontario, Lake Huron, and the Appalachian Mountains.
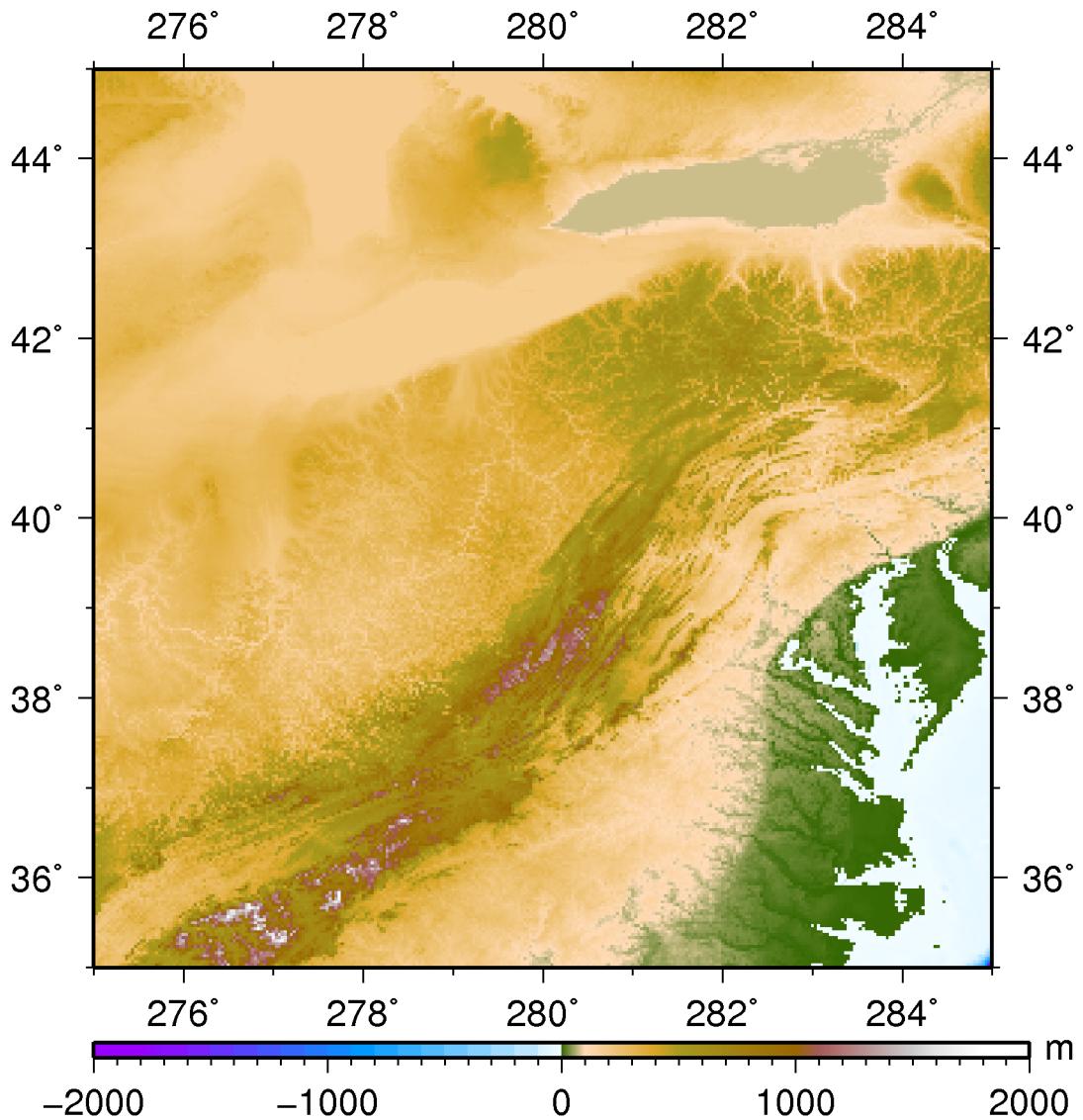

Figure 6.2: Topography of the test area in the northeastern USA.

Input data

As input data, surface gravity anomalies from three sources were used:

- 129,466 pointwise values in the USA from the PACES database (`http://paces.geo.utep.edu/gdrp/`),

- 12,002 pointwise values in Canada and over the Great Lakes from the Geoscience Data Repository (`http://gdrdap.agg.nrcan.gc.ca/geodap/index_e.html`),

- 3,838 gridded values over the Atlantic ocean from the KMS02 data set (`ftp://ftp.spacecenter.dk/pub/GRAVITY/KMS02/`).

Terrain corrections supplied with the data sets were applied, and EIGEN-GL04C up to degree 150 was subtracted as reference model. The resulting residual gravity anomalies, as used for the computation, are shown in figure 6.3. Statistics are listed in table 6.1.

In addition to the gravity anomalies, 1,069 GPS-derived height anomalies were used (`http://www.ngs.noaa.gov/GEOID/GPSonBM03/`). The results were compared against the official American geoid, GEOID03 (Roman et al., 2004), which is available from `http://www.ngs.noaa.gov/GEOID/GEOID03/`. Figure 6.4 shows the geoid heights from GEOID03 after the EIGEN-GL04C contribution up to degree 150 has been removed.

data set	minimum	maximum	mean	standard deviation
USA	-63.3 mGal	123.1 mGal	-1.5 mGal	20.6 mGal
Canada	-33.0 mGal	28.5 mGal	-0.5 mGal	9.2 mGal
KMS02	-56.5 mGal	71.5 mGal	-5.4 mGal	21.2 mGal

Table 6.1: Statistics of residual gravity anomaly data sets for USA test area after the subtraction of EIGEN-GL04C up to degree 150.

Grid choice and basis function placement

A variety of grids were used, from a grid spacing of 30' (approximately 55 km) up to 6' (11 km). The optimal depths / bandwidths for all grids were estimated using the method described in section 2.5.2. The resulting depths are shown in table 6.2.

Data-adaptive network design was used in all cases. Results were computed both with and without local refinement. For grid spacings below 24', a threshold of 5 mGal was used for the local refinement. For 24' and 30' grid spacing, the threshold had to be increased to 10 mGal, and the minimum basis function spacing had to be increased to avoid overparametrisation. The deep placement and wide influence radius of the basis functions would otherwise lead to a singular normal equation system and require very strong regularisation with adverse effects on the solution.

Results

All computations were performed using regularisation. The regularisation parameter was estimated using variance component estimation with stochastic trace estimation (MC-VCE, see section 3.5). The MC-VCE was also employed to estimate optimal weights

grid spacing	approximate point distance	optimal depth
30'	55 km	220 km
24'	44 km	170 km
18'	33 km	100 km
12'	22 km	70 km
6'	11 km	30km

Table 6.2: Optimal depths for various grid spacings

for the three observation groups. The gravimetric field and the GPS/levelling height anomalies were combined by means of a corrector surface. The corrector surface was determined by approximating the differences between the height anomalies computed from a gravity-only solution and the GPS/levelling-derived height anomalies with a tension spline interpolant, with a tension factor of 0.25. This was done with the *surface* command of GMT (Wessel and Smith, 1991). Figure 6.5 shows both the gravimetric solution and the corrector surface resulting from the solution with 12' grid spacing and local refinement.

The results of the computations are listed in table 6.3. Using a larger number of basis functions reduces the data rms. Figure 6.6 shows the errors for the 12' grid with and without local refinement, with full statistics for these computations shown in table 6.4. The application of local refinement clearly reduces the gravity anomaly errors, as more signal gets modelled.

grid	LR	# of RBFs	USA	Canada	rms KMS02	GEOID03	8°x8°
30'	no	441	9.6 mGal	4.7 mGal	11.0 mGal	14.9 cm	12.2 cm
30'	yes	705	7.9 mGal	4.2 mGal	7.8 mGal	19.8 cm	9.0 cm
24'	no	676	8.4 mGal	4.1 mGal	9.0 mGal	12.4 cm	9.3 cm
24'	yes	3592	5.0 mGal	3.4 mGal	4.4 mGal	9.5 cm	7.1 cm
18'	no	1156	7.1 mGal	3.6 mGal	7.2 mGal	9.7 cm	7.9 cm
18'	yes	7529	4.1 mGal	2.5 mGal	2.8 mGal	10.0 cm	7.2 cm
12'	no	2599	6.1 mGal	3.0 mGal	5.3 mGal	8.7 cm	7.5 cm
12'	yes	8569	4.0 mGal	2.9 mGal	2.9 mGal	8.3 cm	7.2 cm
6'	no	9710	4.6 mGal	2.0 mGal	3.0 mGal	9.0 cm	7.6 cm
6'	yes	12073	3.9 mGal	2.0 mGal	1.9 mGal	9.6 cm	8.0 cm

Table 6.3: Results for various grid spacings, with and without local refinement (LR). Number of basis functions used, rms error for the three data sets, and geoid height rms error for the full area, and a sub-area of 8°x8° that is not affected by edge effects.

The basis function network for a grid spacing of 12' with local refinement is shown in figure 6.7. Local refinement basis functions are placed in areas that had large residuals previous to local refinement.

Results are different concerning the geoid height error - all grid spacings from 24' to 12' yielded a geoid height error of about 7 cm with local refinement, for an area of 8°x8° that does not suffer from edge effects. Taking edge effects into account, the 12' solution

data set	minimum	maximum	mean	standard deviation
USA, no LR	-51.1 mGal	102.4 mGal	-0.5 mGal	6.1 mGal
Canada, no LR	-18.5 mGal	19.5 mGal	-0.1 mGal	3.0 mGal
KMS02, no LR	-39.6 mGal	24.0 mGal	-0.7 mGal	5.3 mGal
USA, with LR	-46.8 mGal	73.5 mGal	-0.4 mGal	4.0 mGal
Canada, with LR	-18.7 mGal	20.3 mGal	-0.1 mGal	2.9 mGal
KMS02, with LR	-42.3 mGal	17.1 mGal	-0.3 mGal	2.9 mGal

Table 6.4: Statistics of gravity anomaly errors for 12' grid spacing solution, with and without local refinement (LR).

yields the best quality. The errors resulting from the 24', 12', and 6' solutions are shown in figure 6.8. The overall picture is the same. The solutions with tighter grid spacing capture more signal, which reduces errors in some areas, but increases errors in other areas. This especially affects the 6' solution. Apparently, not all signal that is contained in the data is actually represented in GEOID03, or noise is modelled. Furthermore, specially computed terrain corrections that are not publicly available have been used in the computation of GEOID03. It is thus difficult to assess the quality of the solutions by comparing them to GEOID03. What is obvious is that the 6' solution suffers from slightly larger edge effects (especially in the lower left corner), and that the grid spacing that can be used to get a good solution is quite variable, if local refinement is applied. This illustrates that the local refinement algorithm serves its desired purpose.
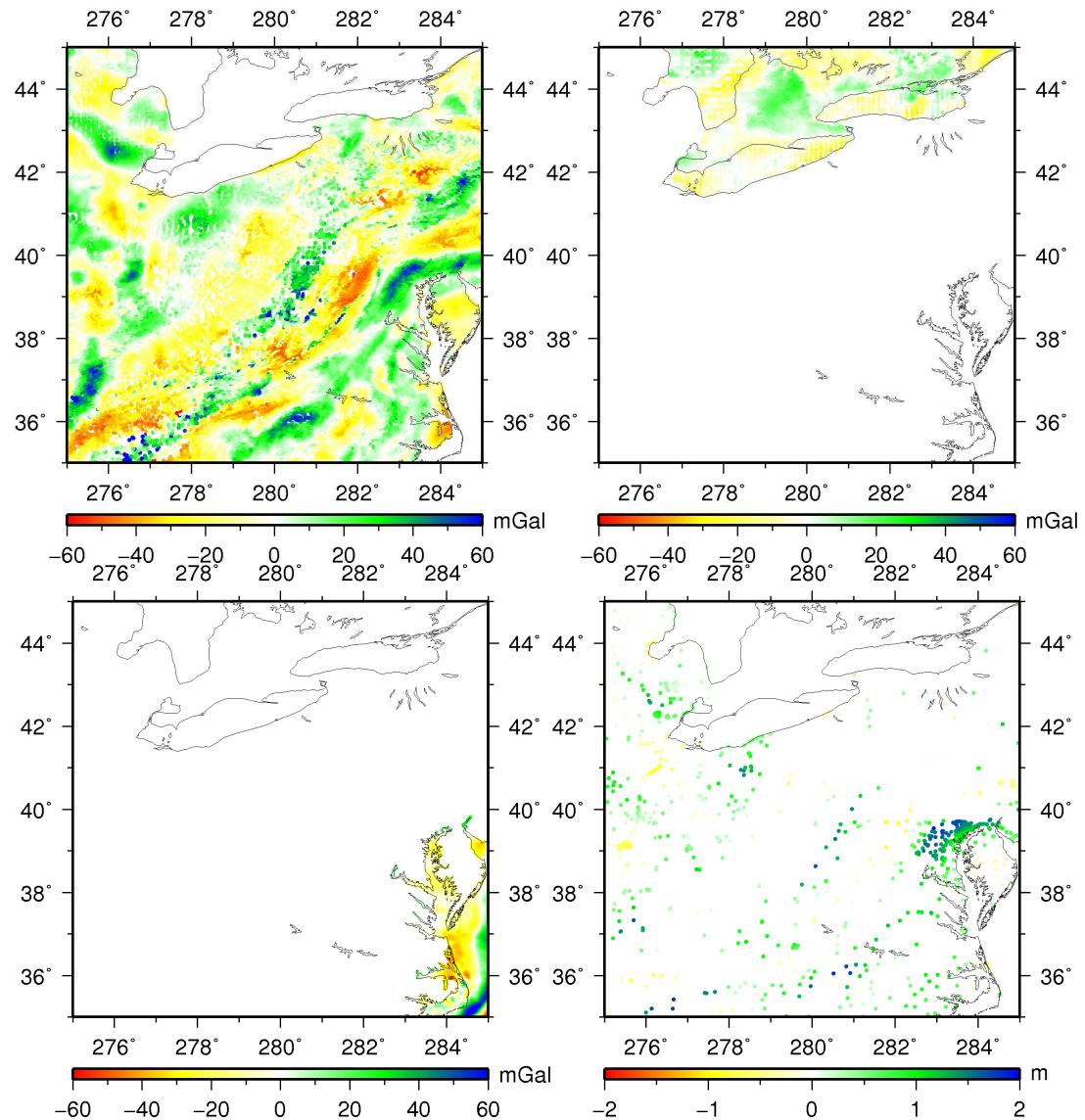
Figure 6.3: Residual gravity anomalies for the USA (top left), Canada (top right), KMS02 (bottom left), and residual GPS height anomalies (bottom right), after the subtraction of EIGEN-GL04C up to degree 150.
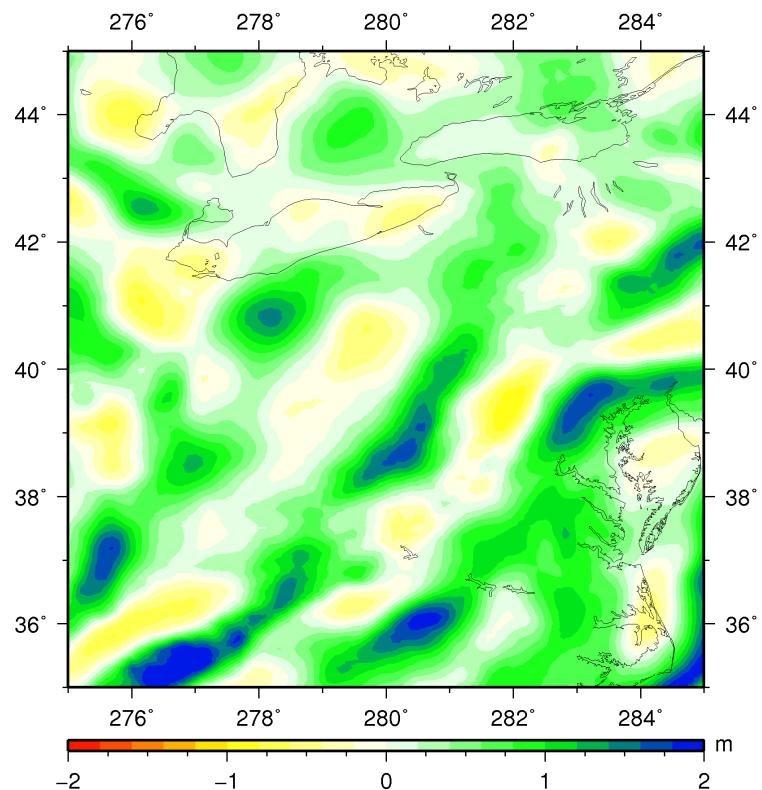
Figure 6.4: Residual GEOID03 geoid model after the subtraction of EIGEN-GL04C up to degree 150.
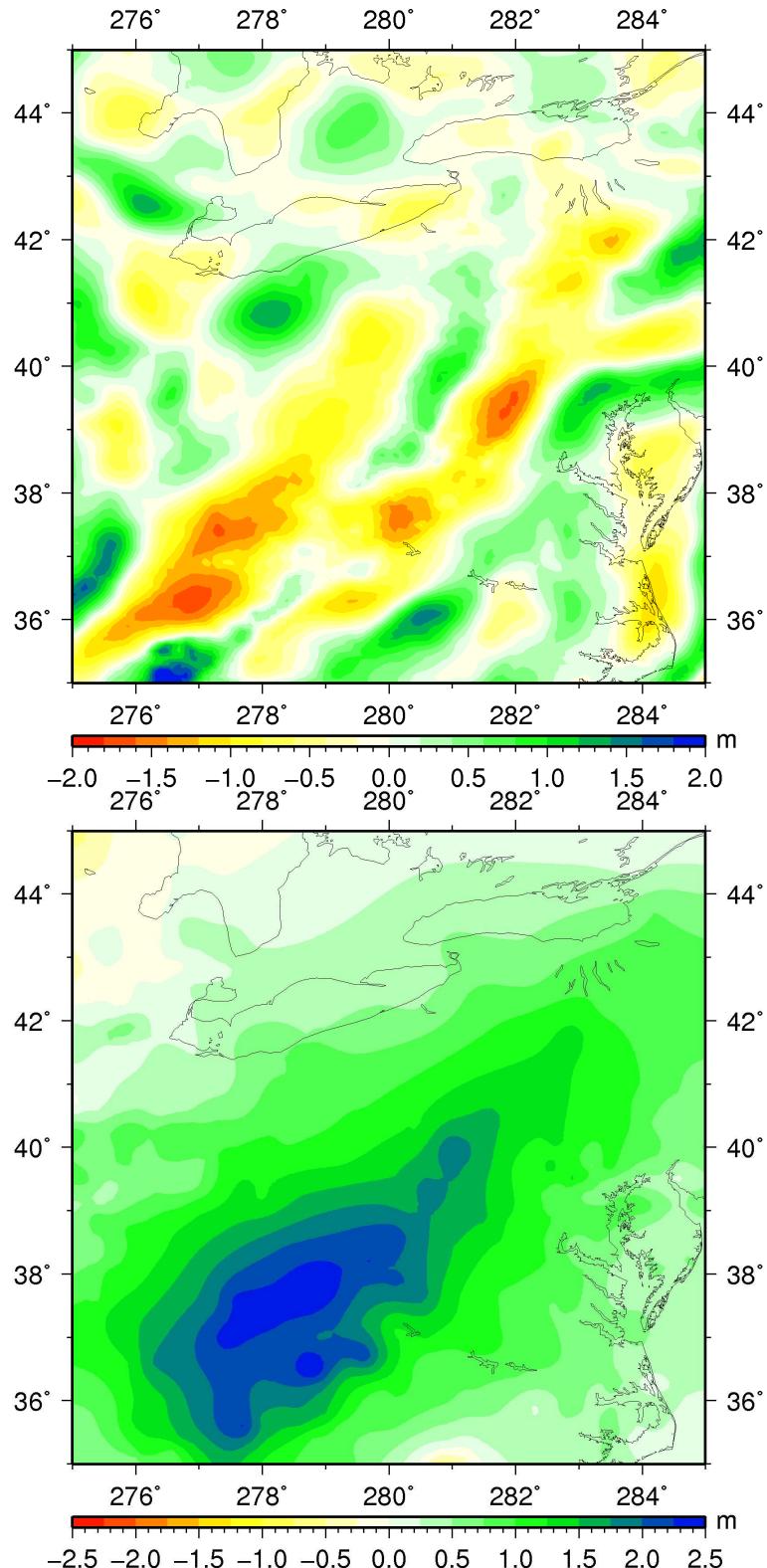
Figure 6.5: Gravimetric solution (top) and corrector surface (bottom), with 12' grid spacing and with local refinement.
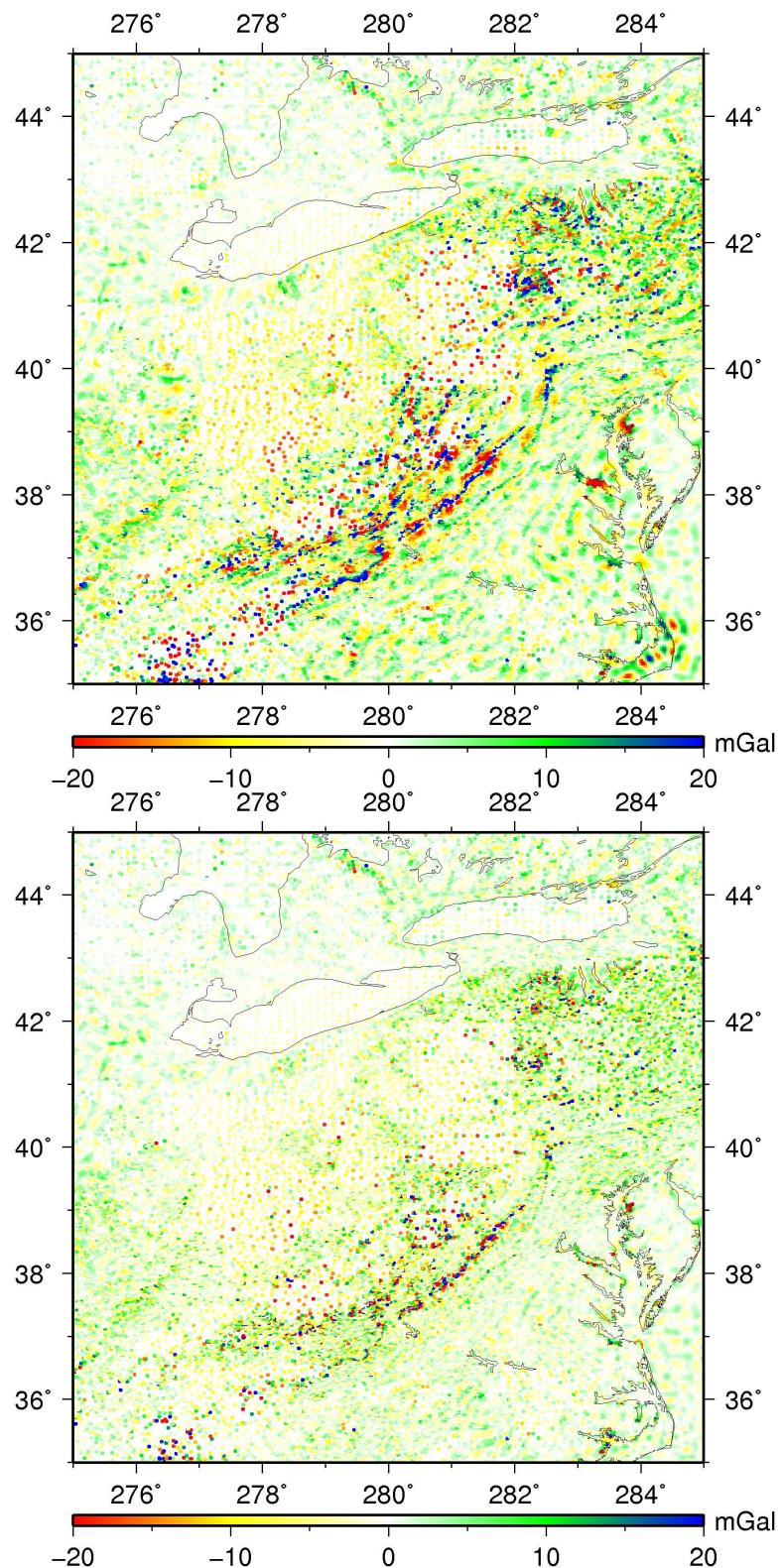
Figure 6.6: Gravity anomaly errors with 12' grid spacing, without (top) and with local refinement (bottom)
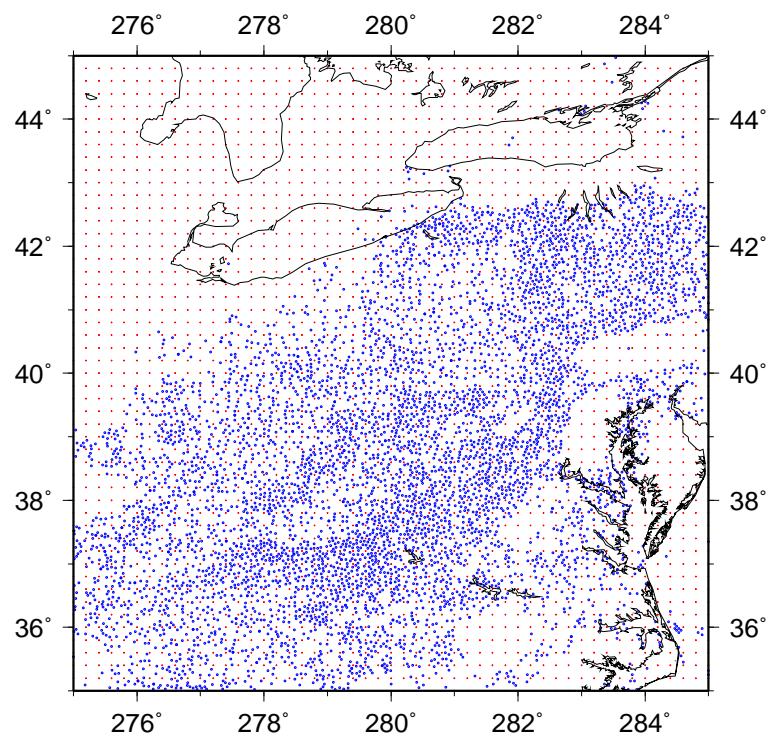
Figure 6.7: Basis function placement, 2,599 RBFs on 12' grid (red) and 5,970 local refinement RBFs (blue).
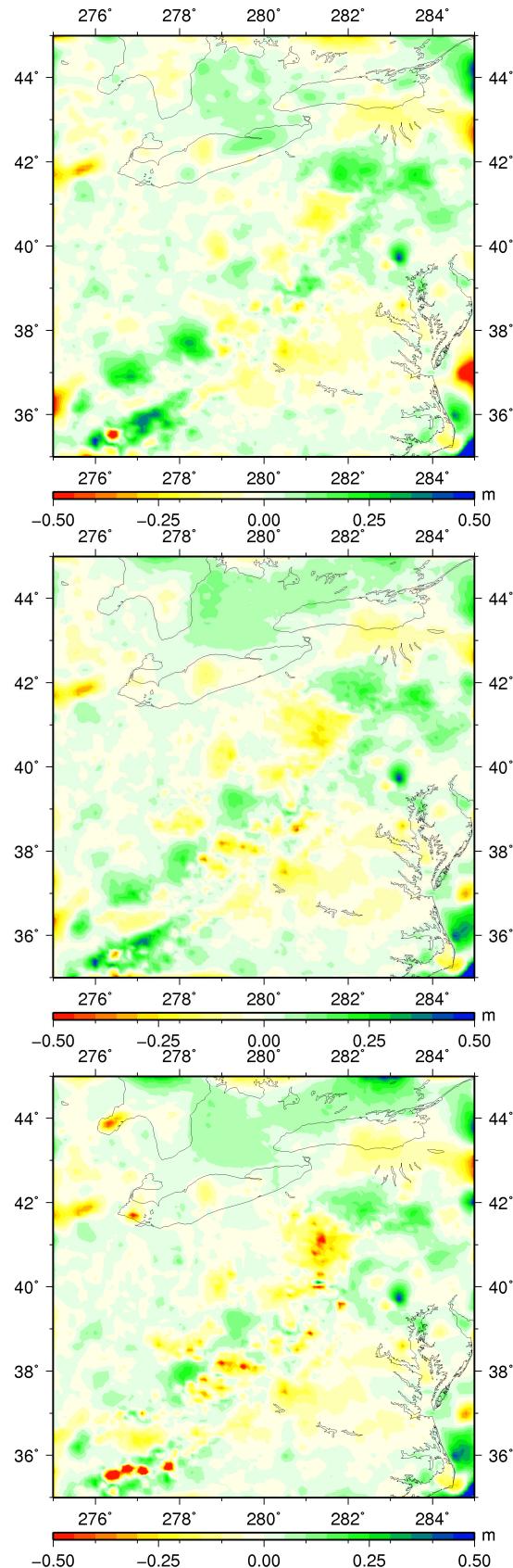
Figure 6.8: Differences between GEOID03 and RBF solutions (including corrector surface) for a grid spacing of 24' (top), 12' (centre), and 6' (bottom), after the application of local refinement.

6.4.2 Canada

Area of interest

A regional gravity field was computed for an area of 40 by 20 degrees in Canada, between 110 to 70 degrees of western longitude, and 50 to 70 degrees of northern latitude (figure 6.9). The dominant geographic feature in this area is the Hudson Bay.
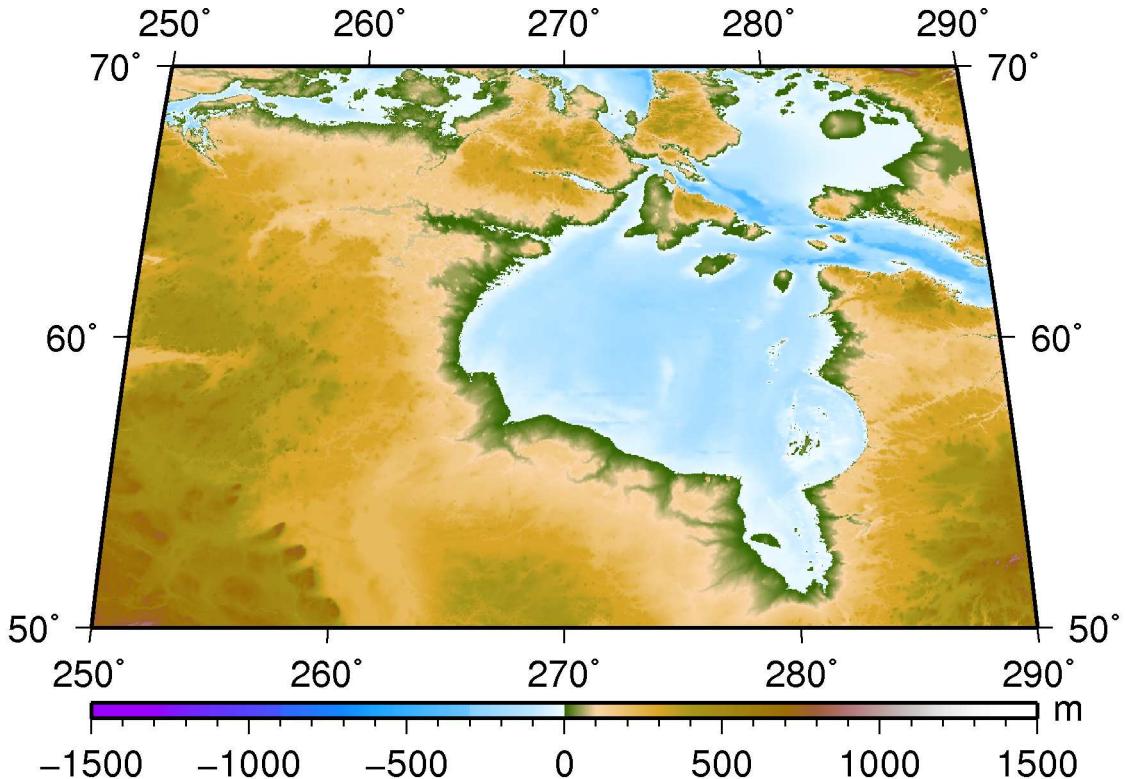
Figure 6.9: Topography of the test area in Canada.

Input data

As input data, surface gravity anomalies from two sources were used:

- 112,628 pointwise values in Canada from the Geoscience Data Repository (`http://gdrdap.agg.nrcan.gc.ca/geodap/index_e.html`), and

- 199,924 gridded values over the Hudson bay and other water bodies from the Sandwell/Smith data set (Sandwell and Smith , 1997).

EIGEN-GL04C (Foerste et al., 2008) up to degree 150 was subtracted as reference model. The resulting gravity anomalies, as used for the computation, are shown in figure 6.10. Statistics for the data sets are listed in table 6.5.

The results were compared against the latest realisation of the official Canadian gravimetric geoid, CGG05 (Véronneau and Huang, 2005). Figure 6.11 shows the geoid

data set	minimum	maximum	mean	standard deviation
Canada	-59.9 mGal	85.6 mGal	-2.0 mGal	11.1 mGal
Sandwell/Smith	-47.9 mGal	47.0 mGal	-1.7 mGal	11.2 mGal

Table 6.5: Statistics of reduced gravity anomaly data sets for Canada test area.

heights from CGG05 after the EIGEN-GL04C contribution up to degree 150 has been removed. To make a comparison between the quasi-geoid computed by the approach described here and the CGG05 geoid, corrections between quasi-geoid and geoid were computed according to Heiskanen and Moritz (1967):

$$(\zeta - N)_{[m]} \doteq \frac{-\Delta g_{B[gal]}}{\gamma_{[\frac{m}{s^2}]}} \cdot H_{[km]},$$ (6.15)

where Δg_B is the Bouguer anomaly, H is the height of the point, and γ is the Normal gravity. The Bouguer anomalies were also retrieved from the Geoscience Data Repository. All comparisons were then made on the basis of height anomalies.

Grid choice and basis function placement

Due to the large area and high latitudes involved, an equiangular grid is not well suited to this case. Using this grid type would lead to overparametrisation at the higher latitudes. Instead, Reuter grids of level 360, 540, and 720 were used. The optimal depths / bandwidths for all grids were estimated using the method described in section 2.5.2, using the rms fit to CGG05 as the optimality criterion. The resulting depths are shown in table 6.6.

Data-adaptive network design was used in all cases. Results were computed both with and without local refinement. A threshold of 10 mGal was used for the local refinement, as using a 5 mGal threshold yielded a lower geoid accuracy. By setting the threshold for the Sandwell / Smith data set to 100 mGal, no local refinement was applied to this data set. This was done in order to avoid overfitting to this very dense gridded data set.

Reuter grid level	approximate point distance	optimal depth
360	55 km	85 km
540	37 km	60 km
720	28 km	45 km

Table 6.6: Optimal depths for the used grid levels.

Results

The computations without local refinement were performed without regularisation. With local refinement turned on, regularisation was applied, with the regularisation parameter being estimated by VCE. VCE was not used for estimating weights for the two observation groups, as it proved unreliable due to the large number of observations in the
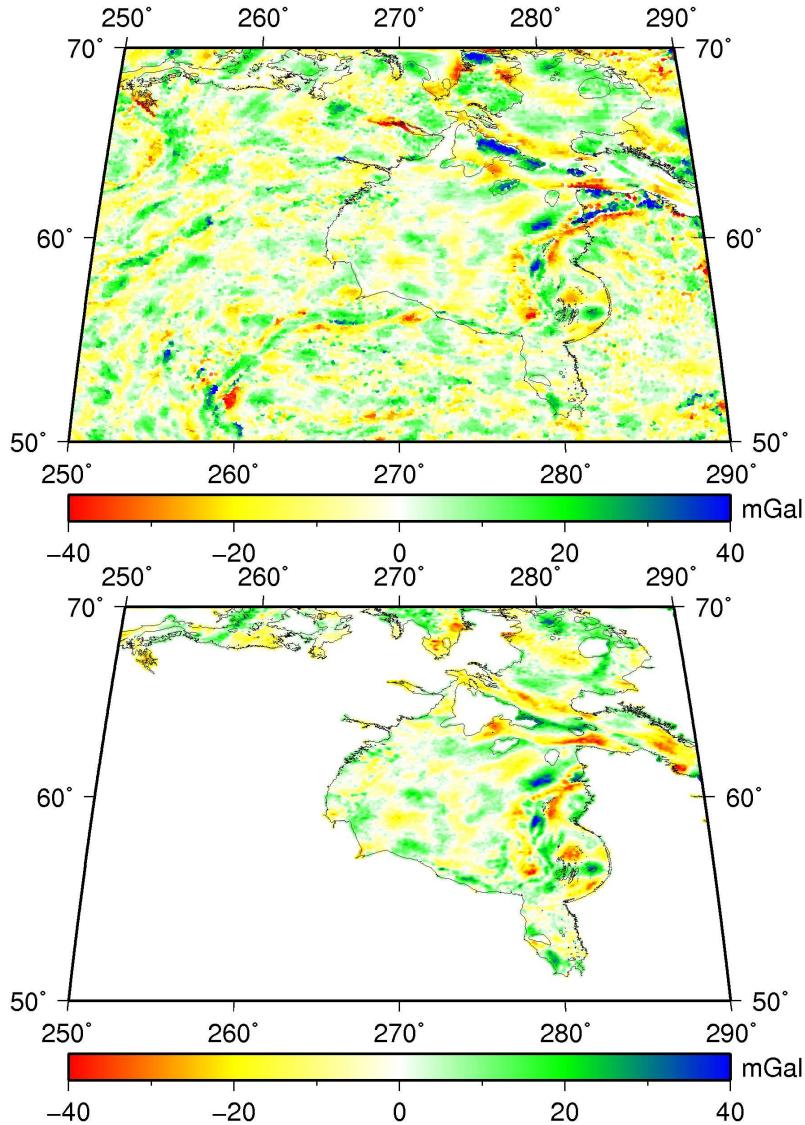
Figure 6.10: EIGEN-GL04C (up to degree 150) reduced gravity anomalies: Canadian data (top) and Sandwell/Smith data (bottom).

Sandwell / Smith data set. Instead, a fixed weight of 5 for the Canadian terrestrial data and unit weight for the Sandwell / Smith data was used, as this yielded better results.

Table 6.7 lists the results of the computations using the three different grid levels, with and without local refinement. The resulting error (rms) for both data sets, as well as the height anomaly error relative to CGG05 are given.

Without the application of local refinement, using a higher grid level, and thus more basis functions, results in a smaller gravity anomaly error. The geoid height error is also smaller, but the percentage of the reduction is smaller. With local refinement turned on, all three grid levels yield similar gravity anomaly errors.

Figure 6.12 shows the gravity anomaly residuals with and without local refinement, for Reuter grid level 540. Full error statistics are listed in table 6.8. It can be seen that the errors are reduced significantly by using local refinement. Figure 6.13 shows the associated distribution of grid basis functions (red) and local refinement basis functions (blue).
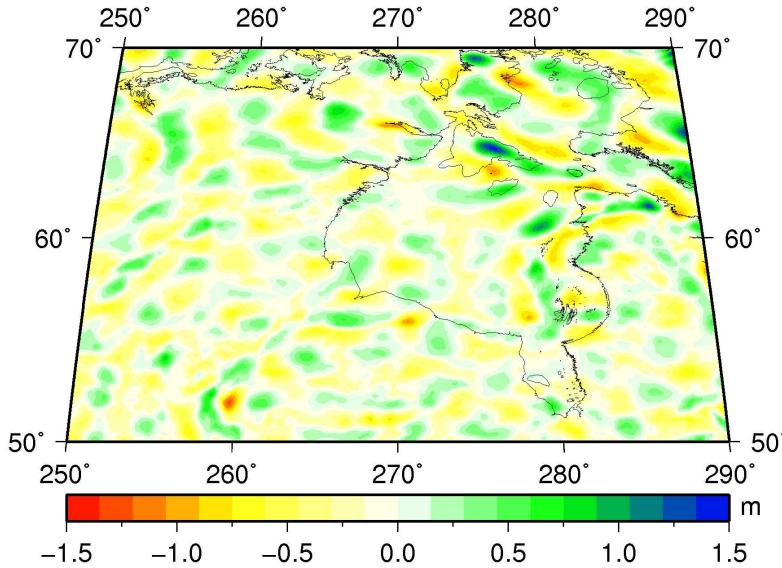
Figure 6.11: EIGEN-GL04C (up to degree 150) reduced CGG05 gravity field.

grid	LR	# of RBFs	rms Canada	rms Sa/Sm	rms CGG05
360	no	1636	6.8 mGal	6.6 mGal	14.0 cm
360	yes	6703	3.8 mGal	5.4 mGal	9.9 cm
540	no	3641	5.4 mGal	5.7 mGal	10.2 cm
540	yes	6885	3.7 mGal	5.2 mGal	9.0 cm
720	no	6447	4.7 mGal	5.4 mGal	9.5 cm
720	yes	8931	3.5 mGal	5.1 mGal	9.0 cm

Table 6.7: Results for various grid levels, with and without local refinement (LR). Number of basis functions used, rms error for the Canadian and the Sandwell/Smith data set, and height anomaly rms error.

Local refinement basis functions were only placed in areas of large gravity anomaly residuals.

The benefit of local refinement concerning the height anomaly error depends on the initial grid level. The level 360 grid is clearly not sufficient to model the signal without local refinement. With local refinement, the height anomaly error drops by approx. 30%. For the level 540 grid, local refinement reduces the height anomaly error by 10%. Figure 6.14 shows the height anomaly errors before and after local refinement. Differences are evident in areas of previously large gravity anomaly errors. Most significant are the reduced errors on the north shore of Southampton Island, and the reduced edge effects in the northeast corner of the computation area.

For the level 720 grid, local refinement does not significantly improve the solution.

With local refinement turned on, the level 540 and 720 solutions deliver similar geoid height errors. The level 360 solution's error is about one centimetre larger. Looking at figure 6.15, we see that the level 360 solution produces the smoothest error plot, albeit with slightly larger errors in the north-east corner than the level 540 solution. The level 720 solution produces a few artifacts. These might be a result of overparametrisation,
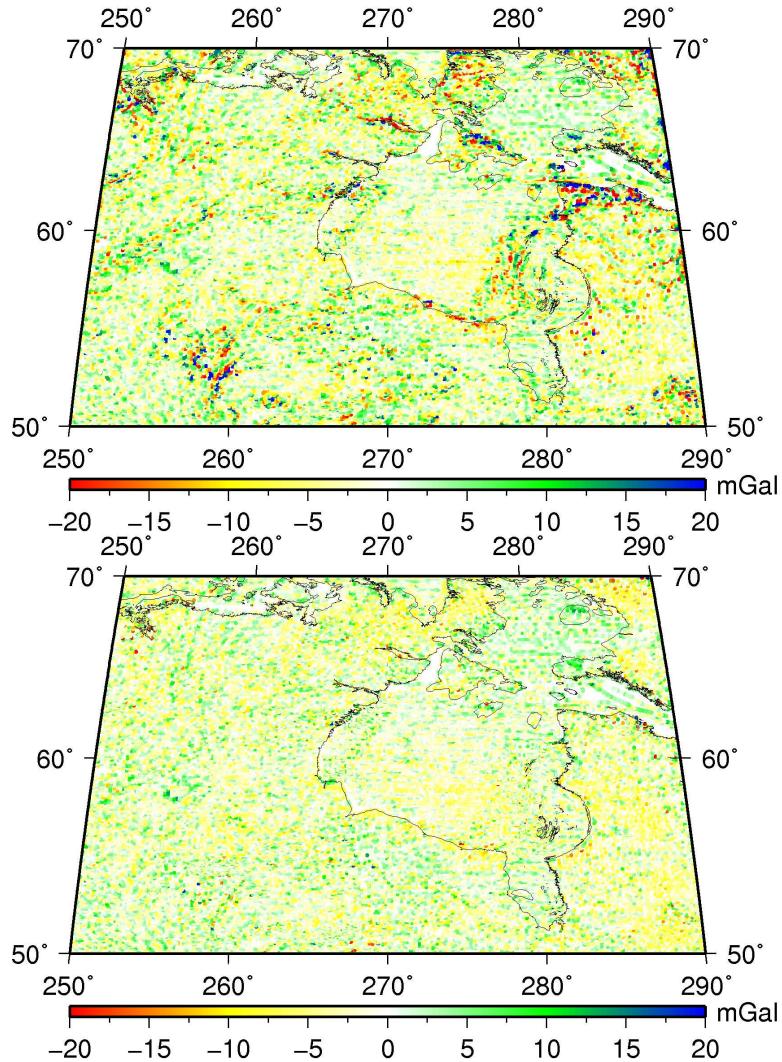
Figure 6.12: Gravity anomaly errors with level 540 grid, without (top) and with local refinement (bottom).

or the solution captures more signal than contained in CGG05. Overall, all three local-refinement solutions can be considered very similar.

Conclusions

From the results obtained with this test data set, it can be concluded that the approximation quality for gravity anomalies and height anomalies are mostly independent of the initial grid choice, if the basis function depth (and thus bandwidth) is correctly determined, local refinement is turned on, and overparametrisation is avoided. The resulting number of basis functions, and thus the numerical complexity, is also approximately the same.

Height anomaly rms errors resulted to about 9 cm in all cases with local refinement, and could not be eliminated using more basis functions. Two explanations for these errors can be given:

data set	minimum	maximum	mean	standard deviation
Canada, no LR	-80.8 mGal	56.4 mGal	-0.8 mGal	5.4 mGal
Sandwell/Smith, no LR	-39.8 mGal	35.7 mGal	-0.2 mGal	5.7 mGal
Canada, with LR	-36.2 mGal	25.5 mGal	-0.7 mGal	3.7 mGal
Sandwell/Smith, with LR	-51.2 mGal	71.5 mGal	-0.0 mGal	5.2 mGal

Table 6.8: Statistics of gravity anomaly errors for level 540 grid solution, with and without local refinement (LR).
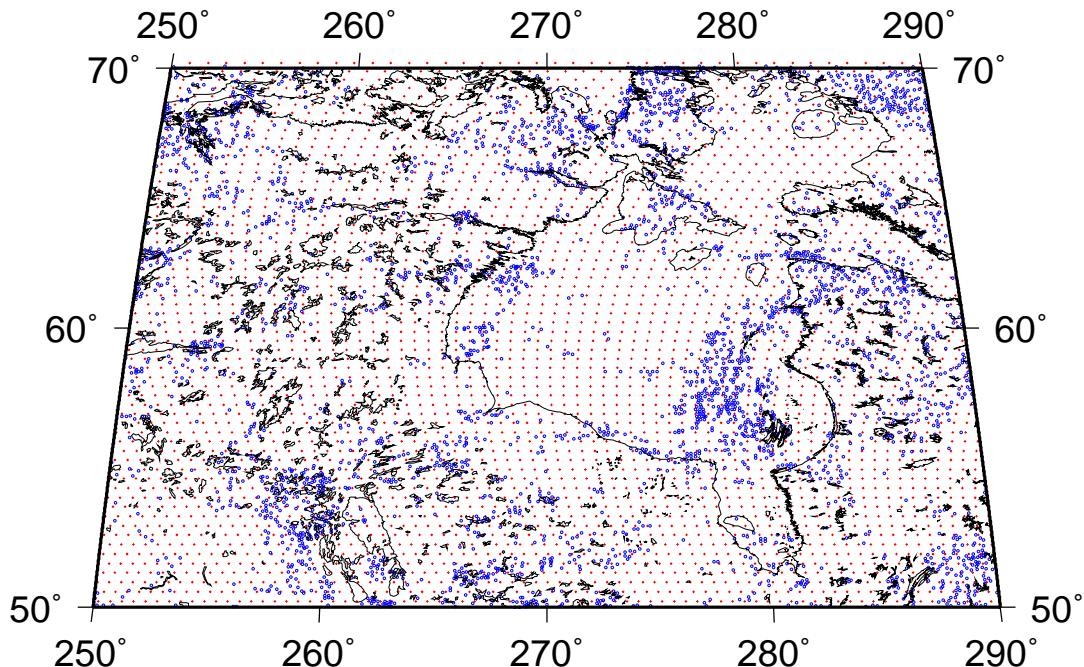

Figure 6.13: Basis function placement, 3,641 level 540 grid functions (red) and 3,244 local refinement functions (blue).

- CGG05 combined gravity anomalies from various sources with a global gravity model (GGM02C up to degree 200), while for the computations presented here, only gravity anomalies from two sources where used.

- For the computation of CGG05, specifically computed terrain corrections were used. These are not identical to the publicly available terrain corrections from the Canadian Geoscience Data Repository. Even though the area used here features rather smooth topography, this may explain some of the high-frequency errors.

- CGG05 is also effected by data and processing errors and thus not error-free.

Finally, slight edge effects can be noticed, especially in the northeast corner of the computation area. Using a 38° x 18° sub-area for computing the rms reduces the rms by about 1.5 cm.
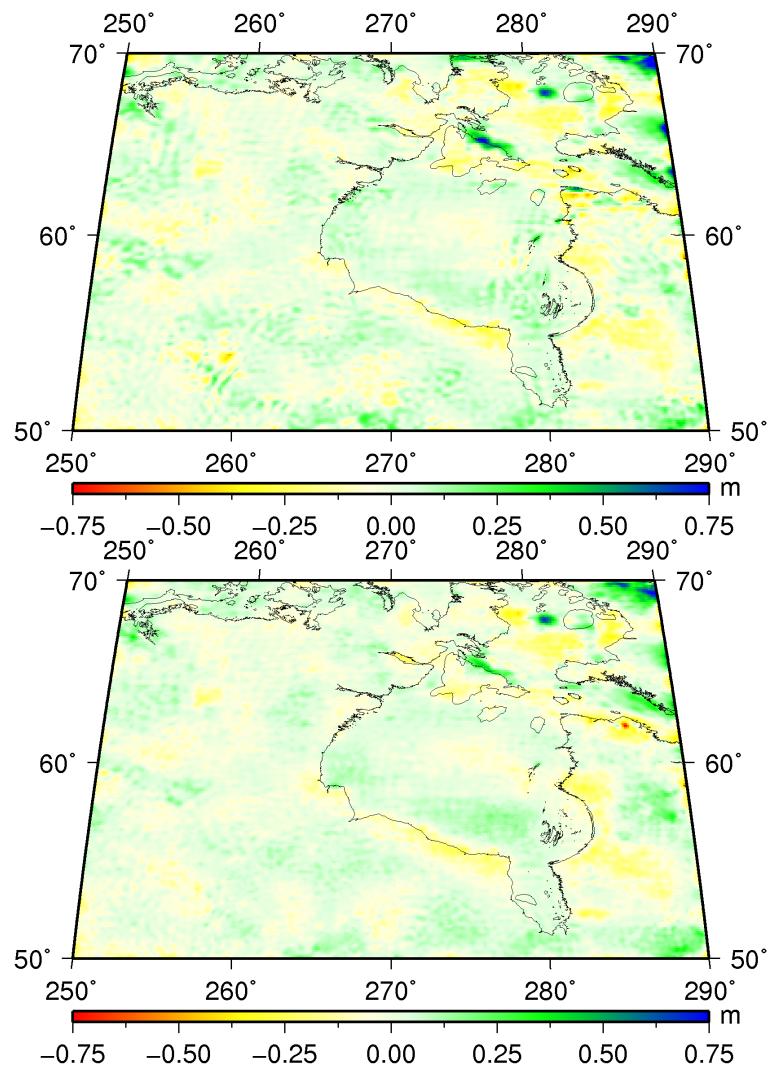
Figure 6.14: Geoid height errors with level 540 grid, without (top) and with local refinement (bottom).
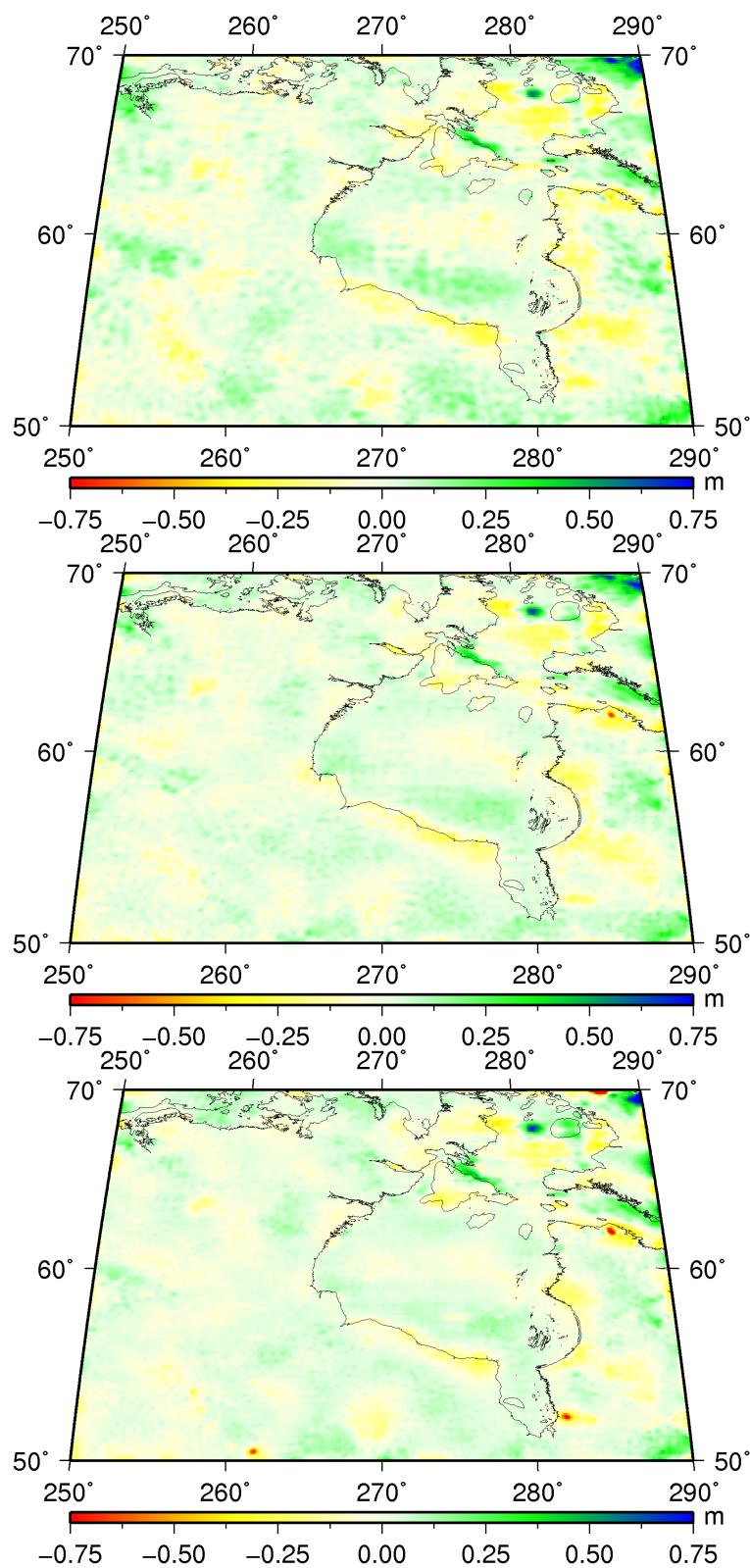
Figure 6.15: Height anomaly errors for a level 360 grid (top), level 540 grid (centre), and level 720 grid (bottom), after application of local refinement.

6.5 Summary and conclusions

In this chapter, the computation of regional gravity field models from terrestrial gravity data has been shown using two test areas, the northeastern USA and eastern Canada. It can be seen that the processing of large amounts of data, more than 300,000 observations over an area of 40° x 20° in the case of Canada, can easily be handled by the software developed in the course of the research presented here.

For the US data set, variance component estimation (VCE) has been used for determining the proper weights for observations from various sources (see table 6.3), as well as for the determination of the regularisation parameter. A numerically efficient implementation using Monte Carlo techniques makes it possible to use VCE even with the large data sets encountered here.

The technique of local refinement has been implemented and evaluated with both data sets. Together with the data-adaptive network design (DAND), the use of local refinement automatically leads to a good approximation regardless of the chosen coarse grid resolution.

Combination with GPS/levelling-derived height anomalies was achieved using a simple corrector surface, computed with the Generic Mapping Tools (Wessel and Smith, 1991).

With both test cases it was not possible to exactly reproduce the reference solutions GEOID03 and CGG05. Reasons for this are:

- For the computations shown here, publicly available gravity data was used. Other data were used in the computation of GEOID03 and CGG05.

- No accurate terrain corrections could be used. Even though both test areas featured mostly flat terrain, the approximation quality could be improved by using accurate high-resolution terrain corrections that lead to smoother input data.

- The two official fields were computed with different methodologies, which can cause additional differences, and they have their own errors.

Chapter 7

Combined modelling of satellite and terrestrial data

Satellite and terrestrial data both have inherent limitations. Satellite data are available globally at high quality, but can yield only low resolution gravity fields, unable to resolve structures of sizes smaller than approx. 150 km in the case of GRACE. Terrestrial data allow higher resolution, but may be of heterogeneous quality and spatial distribution.

To overcome the limitations of either data source, they need to be combined to estimate a high-quality gravity field. The traditional approach for combined solutions uses Stoke's Integral (Stokes, 1849; Molodensky et al., 1962; Heiskanen and Moritz, 1967). An overview of the related computational issues is given by Featherstone (2003).

This chapter investigates strategies for the combination of these data sets to produce a joint solution using the previously established RBF methodology. Section 7.1 describes possible combination strategies. Network design and bandwidth determination are covered in section 7.2. The combination strategies are investigated using both a global and a regional setting in section 7.3. Both settings make use of simulated data. This makes it possible to compare the results to a true field for error analysis.

7.1 Combination strategies

The aim is the computation of a combined solution $\mathbf{x}_C$ using the data sets $\mathbf{y}_S$ (satellite data) and $\mathbf{y}_T$ (terrestrial data) with their respective covariance matrices $\mathbf{C}_S$ and $\mathbf{C}_T$.

7.1.1 Remove-restore approach

In the remove-restore approach, the contribution of a satellite-only solution $\mathbf{x}_S$ is subtracted from the terrestrial data to remove low frequencies before the estimation:

$$\mathbf{y}_{T,res} = \mathbf{y}_T - \mathbf{A}_S\hat{\mathbf{x}}_S. \tag{7.1}$$

155

This is done in order to reduce edge effects. The stochasticity of the satellite solution $\hat{\mathbf{x}}_s$ and thus the satellite data $\mathbf{y}_s$ is not taken into account. After the computation of the solution

$$\hat{\mathbf{x}}_{T,res} = \left(\mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T\right)^{-1} \mathbf{A}^T \mathbf{C}_T^{-1} \mathbf{y}_{T,res} \tag{7.2}$$

the satellite contribution is added back, resulting in a combined solution

$$\hat{\mathbf{x}}_C = \hat{\mathbf{x}}_{T,rem} + \hat{\mathbf{x}}_s \tag{7.3}$$

with the covariance matrix

$$\mathbf{C}_C = \left(\mathbf{A}_T \mathbf{C}_T^{-1} \mathbf{A}_T^T\right)^{-1}. \tag{7.4}$$

The covariance matrix of the SH coefficients of the satellite solution is not used in the standard remove-restore approach. This is sub-optimal from a statistical point of view. An inadequate satellite model that is unable to remove a sufficient amount of low-frequency signal will also lead to edge effects.

An RBF solution can attempt to suppress long-wavelength errors that remain after subtraction of an accurate reference model (i.e. long-wavelength errors in the terrestrial data). This requires insensitivity of the RBFs to the long wavelengths by the choice of the parametrisation. This insensitivity can be achieved by 1) placing basis function shallow rather than deep and 2) starting the summation in eq. (2.3) at a degree l_{min} other than 0, preferably starting with degree $l_{min} = l_{max,ref} + 1$, where $l_{max,ref}$ is the degree up to which the reference model has been subtracted. In both cases, errors will not be totally suppressed, but may affect the quality of the solution.

7.1.2   High-pass filtering

Low-frequency errors in the terrestrial data can be removed by high-pass filtering the data (Hipkin and Hunegnaw, 2005). The actual combination is achieved by low-pass filtering both satellite and terrestrial solutions and subtracting the low-pass filtered part of the terrestrial solution:

$$\hat{\mathbf{x}}_C = \hat{\mathbf{x}}_T + f\left(\hat{\mathbf{x}}_S\right) - f\left(\hat{\mathbf{x}}_T\right). \tag{7.5}$$

If the filter is linear, eq. (7.5) is equivalent to

$$\hat{\mathbf{x}}_C = \hat{\mathbf{x}}_T + f\left(\hat{\mathbf{x}}_S - \hat{\mathbf{x}}_T\right). \tag{7.6}$$

While such a solution is easy to compute, it faces two problems:

1. The filter type and parameters need to be carefully chosen.

2. The stochasticity of both data sets is not taken into account, the satellite data is considered error-free.

7.1.3 Direct combination

The direct combination estimates a combined solution using both data sets as input data for a joint estimation:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_S \\ \mathbf{y}_T \end{bmatrix}, \quad \mathbf{C}_y = \begin{bmatrix} \mathbf{C}_S & 0 \\ 0 & \mathbf{C}_T \end{bmatrix}. \tag{7.7}$$

With the design matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_S \\ \mathbf{A}_T \end{bmatrix} \tag{7.8}$$

we obtain the solution

$$\begin{aligned} \hat{\mathbf{x}}_C &= \left(\mathbf{A}^T \mathbf{C}_y^{-1} \mathbf{A} \right)^{-1} \mathbf{A}^T \mathbf{C}_y^{-1} \mathbf{y} \\ &= \left(\mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{A}_s + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T \right)^{-1} \left(\mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{y}_s + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{y}_T \right) \end{aligned} \tag{7.9}$$

with the covariance matrix

$$\mathbf{C}_C = \left(\mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{A}_s + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T \right)^{-1}. \tag{7.10}$$

This solution takes the stochasticity of both data sets into account. Such a solution may be difficult and time-intensive to compute due to the size of $\mathbf{C}_s$. The effect of assuming $\mathbf{C}_s = diag\,(\mathbf{C}_s)$ will be investigated. Note that one parametrisation that can approximate both data sets is required.

7.1.4 Combination with satellite-only solution

Another approach is to combine a satellite-only solution $\hat{\mathbf{x}}_S$ with terrestrial data $\mathbf{y}_T$, which may be numerically easier to handle. The observations are

$$\mathbf{y} = \begin{bmatrix} \hat{\mathbf{x}}_S \\ \mathbf{y}_T \end{bmatrix}, \quad \mathbf{C}_y = \begin{bmatrix} \mathbf{C}_{\mathbf{x}_s} & 0 \\ 0 & \mathbf{C}_T \end{bmatrix}. \tag{7.11}$$

The satellite-only solutions must have been computed using the same basis functions that are used for the combined solution (sequential adjustment), which will usually require regularisation. This yields the design matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} \\ \mathbf{A}_T \end{bmatrix} \tag{7.12}$$

and the solution

$$\hat{\mathbf{x}}_C = \left(\mathbf{A}^T \mathbf{C}_y^{-1} \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{C}_y^{-1} \mathbf{y}$$

$$= \left(\mathbf{C}_{\mathbf{x}_s}^{-1} + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T\right)^{-1} \left(\mathbf{C}_{\mathbf{x}_s}^{-1} \hat{\mathbf{x}}_s + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{y}_T\right). \tag{7.13}$$

Since

$$\mathbf{C}_{\mathbf{x}_s}^{-1} \hat{\mathbf{x}}_s = \left(\mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{A}_s\right) \left(\mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{A}_s\right)^{-1} \mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{y}_s = \mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{y}_s \tag{7.14}$$

we see that eq. (7.13) is equivalent to eq. (7.9) and both solutions are identical. This also holds for the covariance matrix

$$\mathbf{C}_C = \left(\mathbf{C}_{\mathbf{x}_s}^{-1} + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T\right)^{-1} = \left(\mathbf{A}_S^T \mathbf{C}_s^{-1} \mathbf{A}_s + \mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T\right)^{-1}. \tag{7.15}$$

The effect of assuming $\mathbf{C}_{x_s} = diag\left(\mathbf{C}_{x_s}\right)$ will be investigated.

7.2   RBF network design and bandwidth selection

An RBF network for a combined solution must be able to model both satellite and terrestrial data. The two logical network choices are:

1. A single-layer RBF network with RBFs that are capable of modelling both satellite and terrestrial data.

2. Two layers of RBFs, with each capable of modelling satellite data and terrestrial data, respectively.

At first, it was attempted to compute a combined global solution that used GRACE data globally and terrestrial data regionally. No reference model was subtracted. Such a solution requires at least two layers of basis functions:

1. A global layer of deep basis functions that model the GRACE data on a coarse grid.

2. A regional layer of shallow basis functions that model the terrestrial data on a fine grid.

Two configurations of layers were attempted. Configuration 1 (figure 7.1) uses 1) a global grid of deep basis functions and a 2) regional grid of shallow basis functions, whose size is equal to the size of the terrestrial data area. Computations showed that numerical instabilities produced by the overlapping basis functions made it very difficult to find a working network configuration.

Configuration 2 (figure 7.2) uses the same regional grid. The global grid does not place basis functions in the area covered by the regional grid in order to avoid the previously
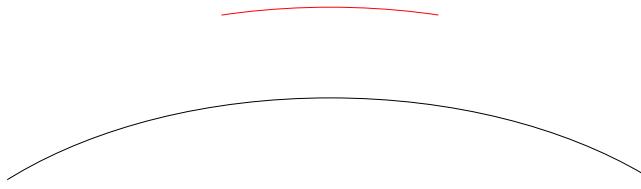
Figure 7.1: Two RBF layers, global (black) and regional (red), configuration 1.
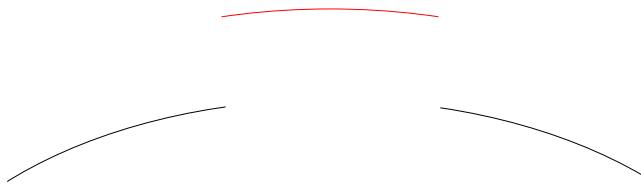
Figure 7.2: Two RBF layers, global (black) and regional (red), configuration 2.

experienced numerical instabilities. This choice of RBF centres generates oscillations along the border of the two layers, which degrades the quality of the the gravity field solution significantly.

All attempts at generating a network consisting of two layers that could produce a good solution failed. This led to the following conclusions:

1. Combined solutions will have to use only one layer of basis functions; the grid spacing and depth of the basis functions will be governed by the (high-resolution) terrestrial data.

2. The basis function network will have to be constrained to the size of the area covered by terrestrial data. Attempting to use this high-resolution grid globally would result in overparametrisation of the satellite data and a too large numerical complexity.

3. As a result of 2., since a regional solution will be computed, a sufficiently accurate reference model will have to be removed from the input data to suppress edge effects.

Since we have to deal with terrestrial data, the network design follows the guidelines described in section 6.2. Bandwidths are estimated as outlined in section 6.3.

7.3 Results

7.3.1 Global test

The first of two tests performed for testing various combination strategies uses a global setting. A global test setup makes it possible to use spherical harmonic analysis for

spectral analysis of input data and errors, and avoids edge effects.

Data

This test uses simulated GRACE and terrestrial data sets. This allows comparison with a true model for error analysis. In this case, EGM08 up to degree 170 (Pavlis et al., 2008) was chosen as true model. The true model was used for generating both satellite and terrestrial data.

The simulated GRACE data consisted of the EGM08 potential coefficients up to degree 170. Coloured noise was added to the coefficients. The noise realisation $\mathbf{e}_S$ was taken from the covariance matrix $\mathbf{C}_{GRACE}$ of ITG-Grace03s (Mayer-Gürr, 2007):

$$\mathbf{e}_S = \mathbf{Gz}, \tag{7.16}$$

where $\mathbf{C}_{GRACE} = \mathbf{GG}^T$ is the Cholesky decomposition (e.g. Golub and van Loan, 1996), and $\mathbf{z}$ is a vector of random numbers of the standard normal distribution ($\sigma^2 = 1$).

$$\mathbf{y}_S = \text{EGM08}_0^{170} + \mathbf{e}_S. \tag{7.17}$$

For the terrestrial data set, gravity disturbances (δg) were synthesised from EGM08 up to degree 170 (with the GRS80 normal field subtracted to obtain the disturbance quantities) on a global equiangular 1°-spaced grid. White noise with a standard deviation of 2 mGal was added as noise realisation $\mathbf{e}_T$:

$$\mathbf{y}_T = \delta g \left(\text{EGM08}_0^{170} \right) + \mathbf{e}_T. \tag{7.18}$$

Figure 7.3 shows the resulting noise spectra for both data sets. The simulated GRACE data set is less noisy below approx. degree 130, while the gravity disturbances are better beyond that. A combined solution should thus be able to yield smaller errors than solutions using only one of the two data sets.
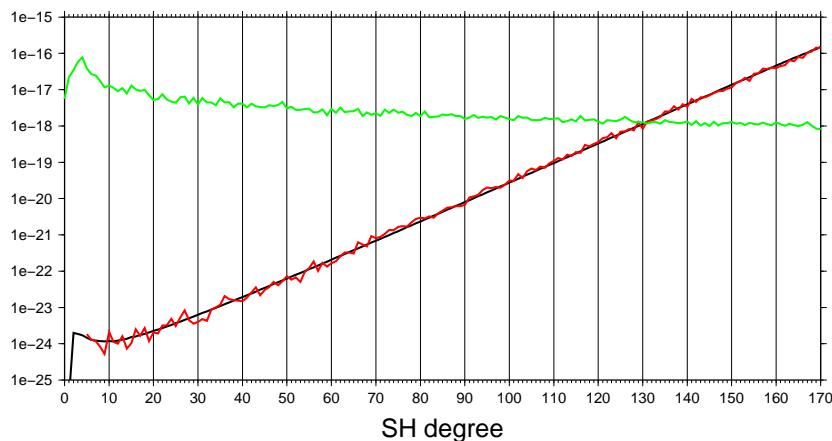

Figure 7.3: Degree variances of EGM08 errors (black), actual noise realisation for simulated GRACE coefficients (red), and simulated global grid of gravity disturbances (green).

Reference solutions

All computations were performed using a level 175 Reuter grid with Poisson wavelets of order 3 placed at a depth of 600 km. This choice of grid level and depth yielded a good approximation. Three reference solutions were computed:

1. A solution using noise-free gravity disturbances as input. This solution verifies whether the choice of the RBF parametrisation is capable of modelling the input data.

2. A solution using only the gravity disturbances with the white noise realisation as input. This solution shows the effect of the noise in the gravity disturbances on the estimated gravity field. The same errors would be present in a gravity disturbances-only solution.

3. A solution using only the GRACE potential coefficients with coloured noise realisation as input. This solution shows the quality that can be obtained for a GRACE-only solution.

The potential errors resulting from a reference solution using a global grid of noise-free gravity disturbances is shown in figure 7.4. The standard deviation is 0.5 mGal, the rms error with respect to EGM08 geoid heights is 3 mm. The differences are due to small numerical errors, as data created from spherical harmonics cannot be represented exactly by RBFs.
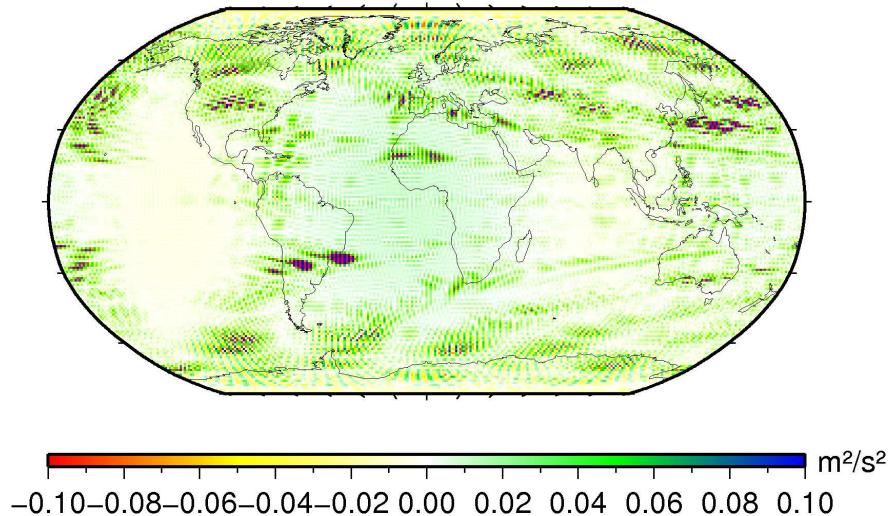

Figure 7.4: Gravity potential error resulting from noise-free gravity disturbances.

A solution using the noisy gravity disturbances results in an approximation error of 1.3 mGal and a rms geoid height error of 15.8 cm. The potential errors are shown in figure 7.5.

A GRACE-only solution using the noisy potential coefficients results in a rms geoid height error of 24 cm. Regularisation was necessary in order to achieve a stable solution. Figure 7.6 shows that the resulting errors are of a high-frequency nature.
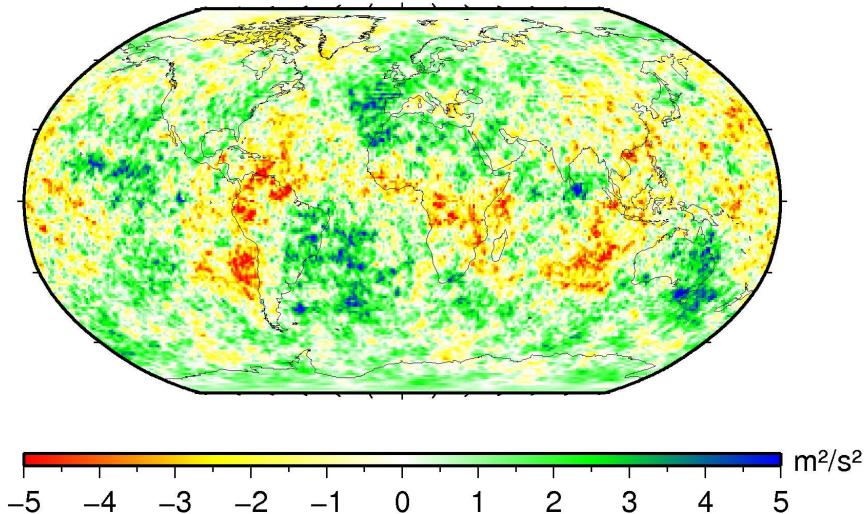
Figure 7.5: Gravity potential error resulting from gravity disturbances with 2 mGal white noise.
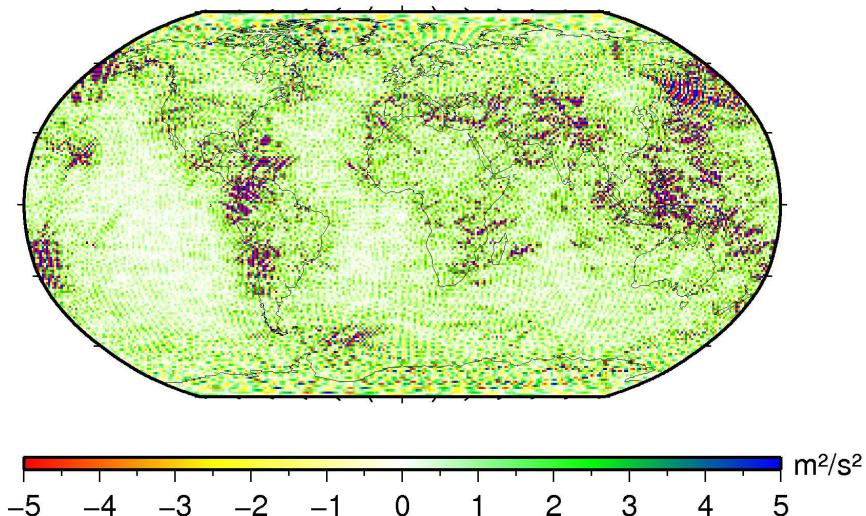
Figure 7.6: Gravity potential error of the GRACE-only solution.

A combined solution should improve with respect to the gravity disturbance-only and satellite-only solutions.

Combined solution I

A combined solution was computed using both the simulated GRACE observations (potential coefficients with colored noise realisation) and gravity disturbances as observations in a joint estimation according to eq. (7.9). Two scenarios were considered:

1. Using only a diagonal covariance matrix for the GRACE observations.

2. Using the full GRACE covariance matrix.

With both scenarios, the aforementioned parametrisation was used, with Poisson wavelets of order 3 placed at a depth of 600 km on a level 175 Reuter grid. As shown, this parametrisation is able to produce a good solution in the noiseless test case. The resulting geoid height errors are listed in table 7.1.

scenario	rms geoid height error
δg only	15.8 cm
GRACE only	24.0 cm
δg + GRACE, diagonal covariance matrix	5.9 cm
δg + GRACE, full covariance matrix	5.4 cm

Table 7.1: Resulting geoid errors for different scenarios.

Using a combination of gravity disturbances and GRACE-derived potential coefficients significantly improves the quality of the solution. Geoid height errors drop from almost 16 cm in the δg-only-case to below 6 cm. Using the full covariance matrix results in a 10% improvement with respect to the use of a diagonal covariance matrix.

The spectral distribution of the errors is shown in figure 7.7, and the cumulative errors are shown in figure 7.8. It is obvious that the addition of GRACE observations significantly reduces the errors below degree 120. The slight increase in error beyond degree 165 is caused by regularisation. A spatial plot of the errors of the combined solution is shown in figure 7.9.
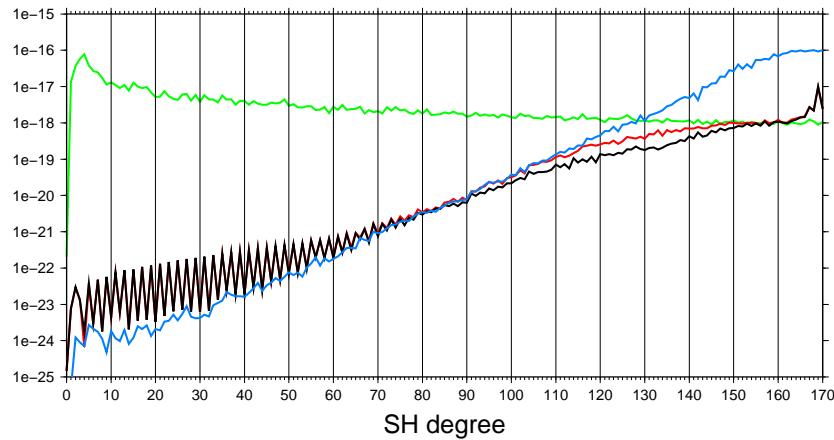
Figure 7.7: Dimensionless error degree variances of terrestrial data solution (green), satellite data solution (blue), combined solution with diagonal GRACE covariance matrix (red), and combined solution with full GRACE covariance matrix (black).

Combined solution II

A second series of computations was performed to investigate the effects of combining a satellite-only RBF solution with terrestrial data, instead of using the actual satellite data as input data. The following input data sets are used:

1. The same terrestrial data set, gravity disturbances with 2 mGal white noise.
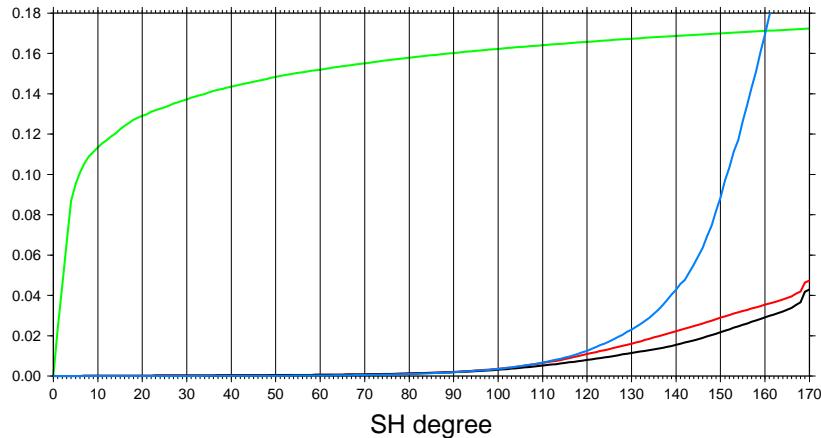
Figure 7.8: Cumulative geoid height degree errors [m] of terrestrial data solution (green), satellite data solution (blue), combined solution with diagonal GRACE covariance matrix (red), and combined solution with full GRACE covariance matrix (black).
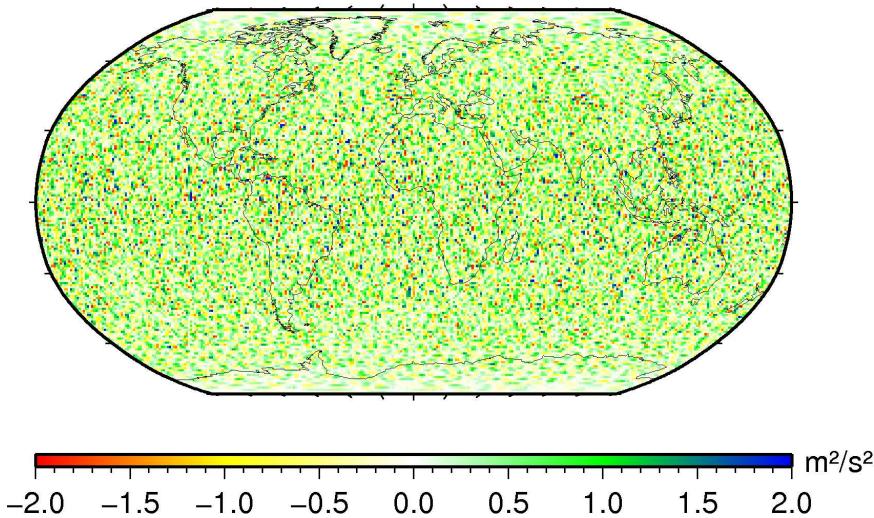
Figure 7.9: Gravity potential error resulting from gravity anomalies with 2 mGal white noise combined with GRACE potential coefficients, full GRACE covariance matrix.

2. An RBF solution computed from the GRACE potential coefficients that were previously used as observations.

The two data sets were combined according to eq. (7.13). Four scenarios were tested using these input data:

1. Combination of a) terrestrial data and b) RBFs estimated from satellite data with diagonal covariance matrix for GRACE; the diagonal covariance matrix of the RBFs is used.

2. Combination of a) terrestrial data and b) RBFs estimated from satellite data with diagonal covariance matrix for GRACE; the full covariance matrix of the RBFs is used.

3. Combination of a) terrestrial data and b) RBFs estimated from satellite data with full covariance matrix for GRACE; the diagonal covariance matrix of the RBFs is used.

4. Combination of a) terrestrial data and b) RBFs estimated from satellite data with full covariance matrix for GRACE; the full covariance matrix of the RBFs is used.

Scenarios 2 and 4 follow eq. (7.13) and eq. (7.14) and thus yield solutions identical to those previously computed by direct combination of satellite and terrestrial observations.

Using only a diagonal covariance matrix for the RBFs (scenarios 1 and 3) resulted in solutions of poor quality. Figure 7.10 shows the error degree variances of scenario 1 (green), scenario 3 (red), and scenario 4 (black). Cumulative errors for these solutions are shown in figure 7.11.
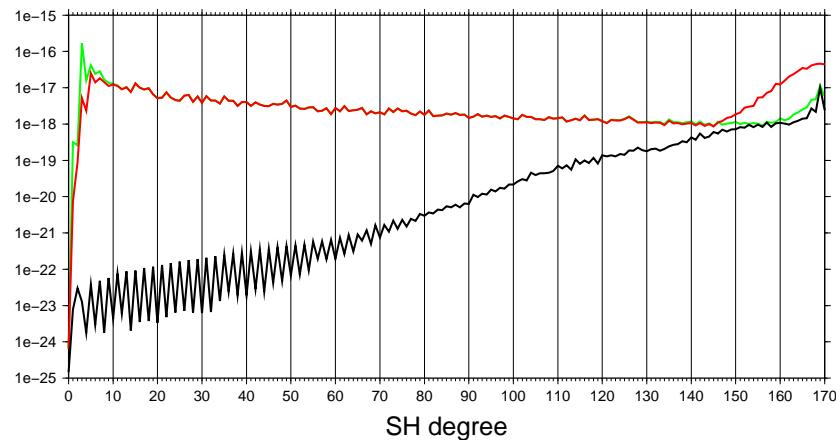
Figure 7.10: Dimensionless error degree variances of solution with diagonal covariance matrices for RBFs and GRACE (scenario 1, green), diagonal covariance matrix for RBFs and full covariance matrix for GRACE (scenario 3, red), and full covariance matrices for RBFs and GRACE (scenario 4, black).

Using a diagonal covariance matrix for the input RBFs computed from GRACE results in a poor solution because the dependence of the quality of the RBF satellite solution on the frequency is not taken into account. In order to achieve a good solution, it is necessary to use the full covariance matrix of the input RBFs.
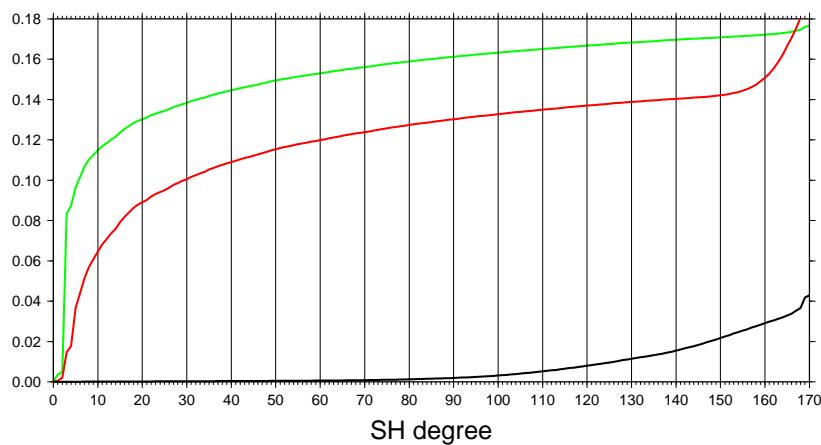
Figure 7.11: Cumulative geoid height degree errors [m] of solution with diagonal covariance matrices for RBFs and GRACE (scenario 1, green), diagonal covariance matrix for RBFs and full covariance matrix for GRACE (scenario 3, red), and full covariance matrices for RBFs and GRACE (scenario 4, black).

7.3.2 Regional test

While the previous test showed that the combination of GRACE and terrestrial data does indeed lead to an improved solution, it was not an entirely realistic test. The terrestrial data were only contaminated by white noise, which may not be a realistic assumption. In reality, we know little about the actual noise in terrestrial data sets. A good noise model is however necessary for achieving good results when combining different data sets, i.e. terrestrial and satellite data in this case.

This regional test has three goals:

1. Use a more realistic noise realisation for the terrestrial data set, and quantify the effects resulting from this noise realisation in a combined solution.

2. Show other effects of a combination in a regional setting, especially edge effects.

3. Prove that a single-layer RBF parametrisation is capable of modelling both satellite data and terrestrial data of significantly higher resolution.

Data

For this second test, EGM08 is once again used as a true model. In order to obtain a more realistic noise model for the gravity anomalies, the difference between the Canadian gravity data used in section 6.4.2 and gravity anomalies synthesised from EGM08 up to degree 720 was computed. The differences are shown in figure 7.12, the rms difference is 5.3 mGal.
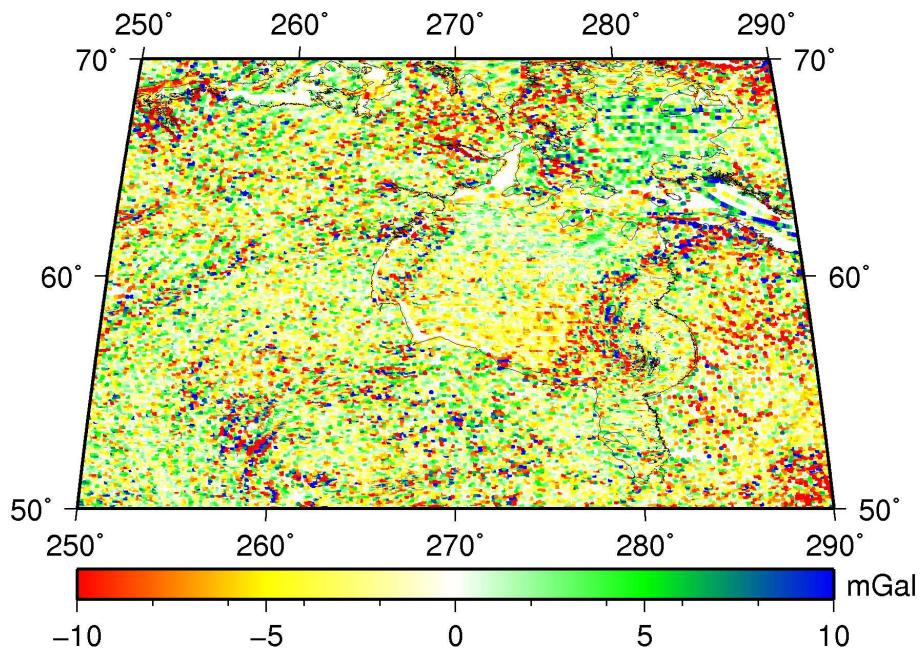

Figure 7.12: Pointwise differences between Canadian data and EGM08 gravity anomalies up to degree 720.

The differences contain both high- and low-frequency noise. Since we are interested in the effect that the addition of GRACE data has, the noise has to be contained to a part of the spectrum where GRACE can contribute to the solution. To achieve this, the difference between the Canadian and the EGM08 gravity anomalies has been smoothed with a Gaussian filter with 200km correlation length (figure 7.13). It is assumed that EGM08 is an accurate gravity field model especially at low frequencies. This low-frequency error must be caused by errors in the Canadian gravity anomalies, and is thus an accurate noise model. A similar aprpoach has been used by Huang et al. (2008) to quantify long-wavelength errors in North American gravity data.
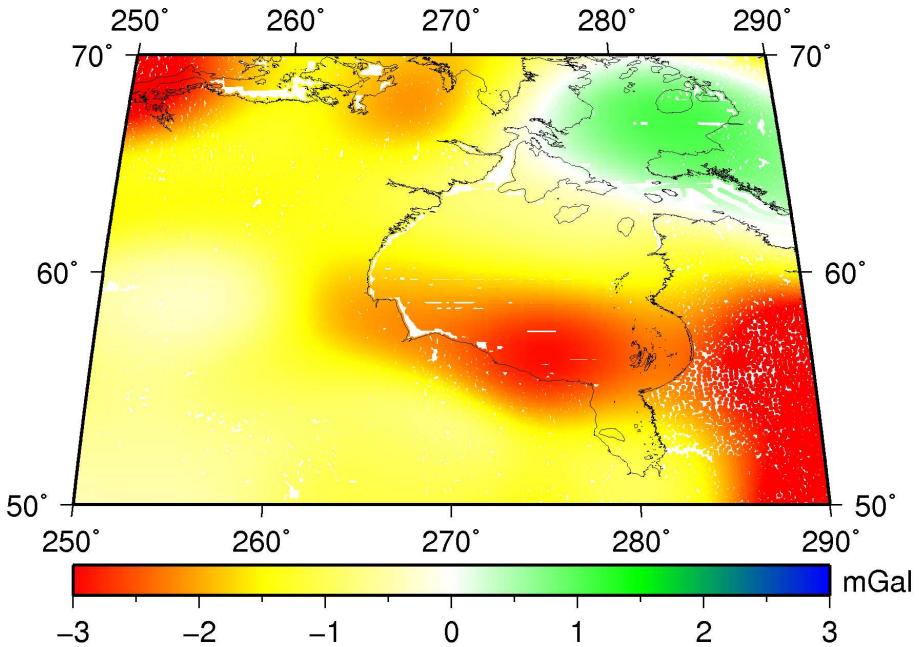
Figure 7.13: Pointwise differences between Canadian data and EGM08 gravity anomalies up to degree 720, smoothed with a Gaussian filter with 200km correlation length.

This difference is added as noise to synthetic gravity anomalies computed from EGM08 up to degree 720. The resulting data set, with EGM08 up to degree 150 removed to avoid edge effects, represents terrestrial data in the test. These data contains no high-frequency errors since we are only interested in the effect of the combination with GRACE, which does not contribute to the high frequencies.

$$\mathbf{y}_T = \Delta g \left(\text{EGM08}_{151}^{720} \right) + f_{200km} \left(\Delta g_{Canada} - \Delta g \left(\text{EGM08}_0^{720} \right) \right). \tag{7.19}$$

GRACE data are expressed using a set of EGM08 potential coefficients up to degree 170 corrupted by a noise model. The noise model is described in section 7.3.1. EGM08 up to degree 150 is used as background model. Hence, the observations are given as

$$\mathbf{y}_S = \text{EGM08}_{151}^{170} + \mathbf{e}_S. \tag{7.20}$$

The noise spectrum of the terrestrial data set was computed by approximating the noise realisation with RBFs and translating the solution into spherical harmonics according to

eq. (2.26). Figure 7.14 shows the spectra of the noise realisation for both data sets. Both lines intersect at degree 120, meaning that the GRACE data should improve the solution up to degree 120. It is interesting to note that the spectrum of the noise realisation for the terrestrial data is approximately flat. This means that the use of white noise in the previous test was a valid one, and that the assumption of white noise for the stochastic model in this test should yield good solutions. Other real terrestrial data sets may however behave differently.
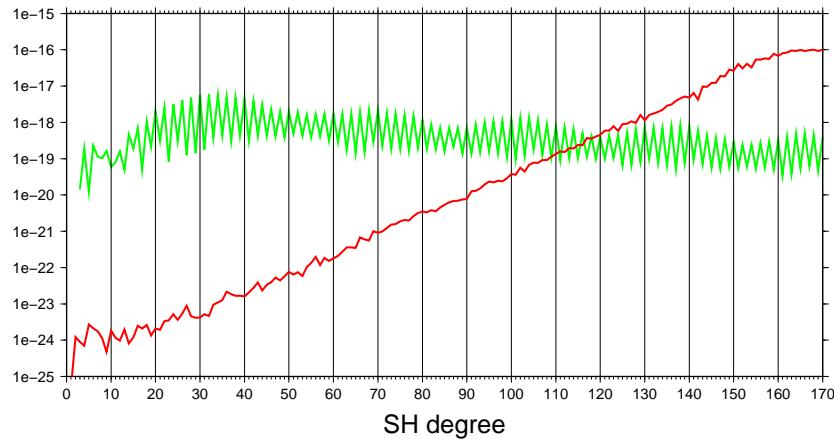
Figure 7.14: Dimensionless degree variances of noise realisations for simulated satellite data (red) and simulated terrestrial gravity anomalies (green).

Results will be compared to EGM08 from degree 151 to degree 720.

Reference solutions

All computations were performed using a level 900 Reuter grid with Poisson wavelets of order 3 placed at a depth of 120 km. This parametrisation yielded a good approximation. Three reference solutions were computed:

1. A solution using noise-free gravity anomalies as input data. This solution verifies whether the chosen RBF parametrisation is capable of modelling the input data.

2. A solution using only the gravity anomalies with the white noise realisation as input data. This solution shows the effect of the noise in the gravity anomalies on the estimated gravity field.

3. A solution using only the GRACE potential coefficients with coloured noise realisation as input. This solution shows the quality that can be obtained for a GRACE-only gravity field.

The potential errors resulting from a reference solution using the noise-free gravity anomalies is shown in figure 7.15. The RBFs were placed at a depth of 120 km. The approximation error is 0.06 mGal, the rms with respect to EGM08 geoid heights is 2.9
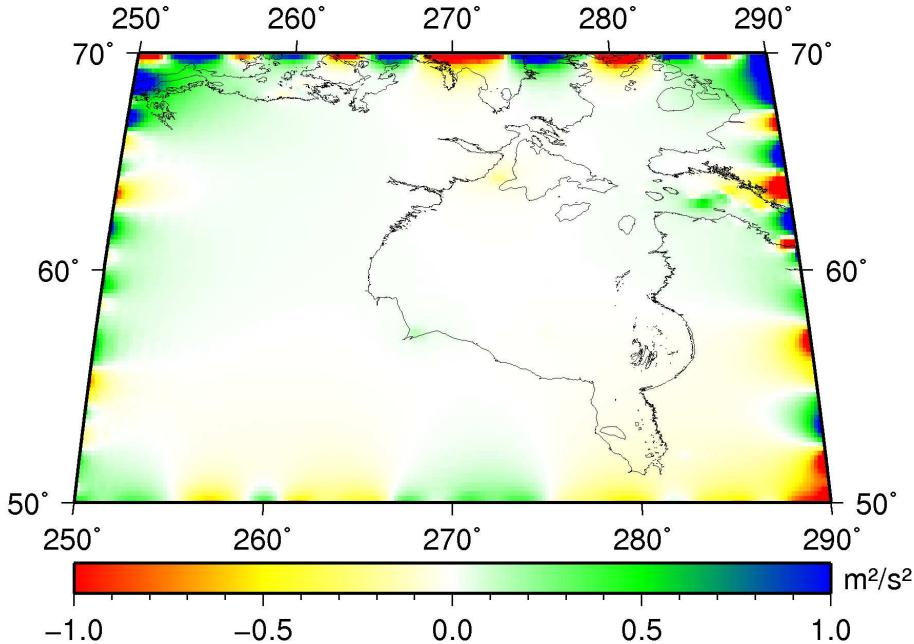
Figure 7.15: Gravity potential error using noise-free gravity anomalies.

cm. The errors are mostly due to edge effects (confined to within $1°$ of the edge), with small errors visible in areas with data gaps.

The error rms of a solution computed from the gravity anomalies contaminated with the chosen noise depends on the parametrisation chosen. Using the same level 900 @ 120 km depth solution, the rms geoid height errors amount to 50 cm. The errors can be reduced by using RBFs with decreased sensitivity to long wavelengths. The rms error is reduced to 17.6 cm by choosing 80 km as depth, and further to 5.7 cm by starting the summation in eq. (6.7) at degree $l_{min} = 151$. The potential errors for all three solutions are shown in figure 7.16.

A GRACE-only solution using the potential coefficients with colored noise added results in a rms geoid height error of 28.3 cm, using RBFs at 600 km depth on a level 180 grid. Regularisation was necessary in order to achieve a stable solution. The error plot (figure 7.6) shows that the resulting errors are of a high-frequency nature.

Combined solutions should improve with respect to these solutions.

Combined solution

A combined solution was computed by using both the simulated GRACE observations (potential coefficients with colored noise realisation) and gravity anomalies as observations in a joint estimation according to eq. (7.9). Two scenarios were considered:

1. Using only a diagonal covariance matrix for the GRACE observations.

2. Using the full GRACE covariance matrix.

With both scenarios, the Poisson wavelets were placed at a depth of 120 km on a level 900 Reuter grid. This parametrisation was able to produce a good solution in the noise-less test case. The resulting geoid height errors are listed in table 7.2. Error plots for the two combined solutions are shown in figure 7.18.

scenario	geoid height error rms
Δg only, level 900 @ depth 120 km	50.0 cm
Δg only, level 900 @ depth 80 km	17.6 cm
Δg only, level 900 @ depth 80 km $l_{min} = 151$	5.7 cm
GRACE only	28.3 cm
Δg + GRACE, diagonal covariance matrix	5.8 cm
Δg + GRACE, full covariance matrix	6.0 cm

Table 7.2: Resulting geoid errors for different scenarios.

Errors for the combined solution can be reduced by a few mm by choosing 80 km instead of 120 km as depth. While both solutions offer a significant improvement over the Δg-only solutions with $l_{min} = 0$, they are of the same quality as the Δg-only solution at 80 km depth with $l_{min} = 151$. Suppressing the long-wavelength errors by parametrisation has in this case the same effect as adding more accurate observations.

Using the full GRACE covariance matrix does lead to slightly larger errors than using the diagonal covariance matrix. This is most probably caused by an inaccurate stochastic model for the terrestrial data, since the assumption of white noise is only approximately correct.

No combination with an RBF solution computed only from GRACE data has been performed as no new insights compared to section 7.3.1 would result from such a test: A good solution can only be obtained by using the full covariance matrix for the RBF coefficients, which lead to the identical solution as using the SH coefficient.
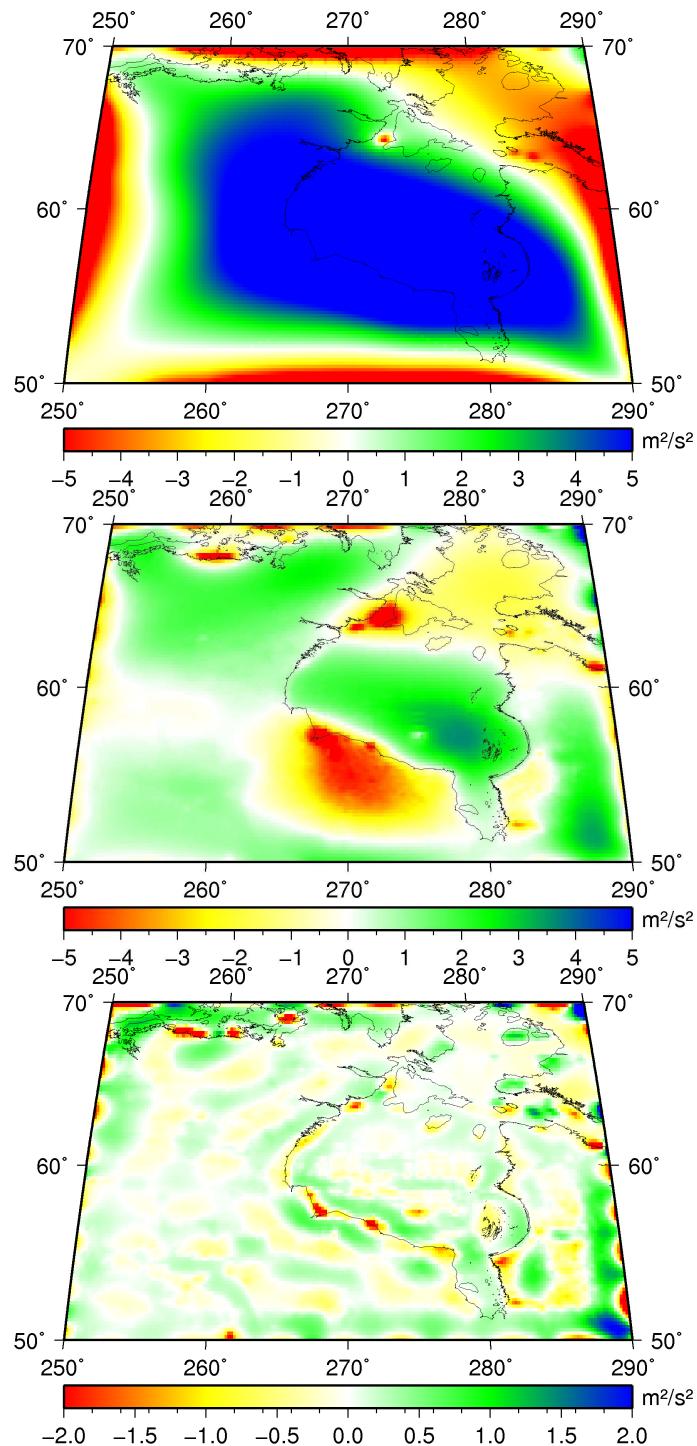
Figure 7.16: Gravity potential error of gravity anomaly-only solutions with noise realisation: 900 @ 120 km solution (top), 900 @ 80 km solution (centre), 900 @ 80 km solution with $l_{min} = 151$. Note the different scaling of the colorbars.
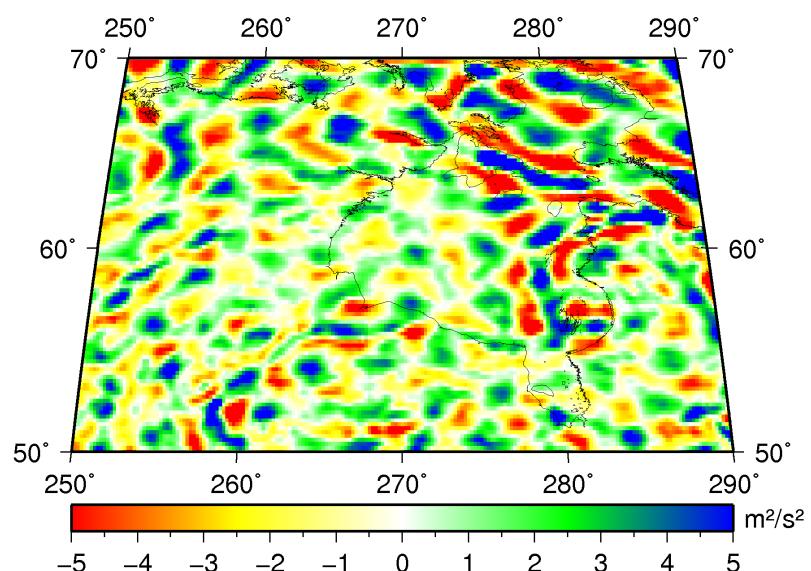
Figure 7.17: Gravity potential error of the GRACE-only solution.
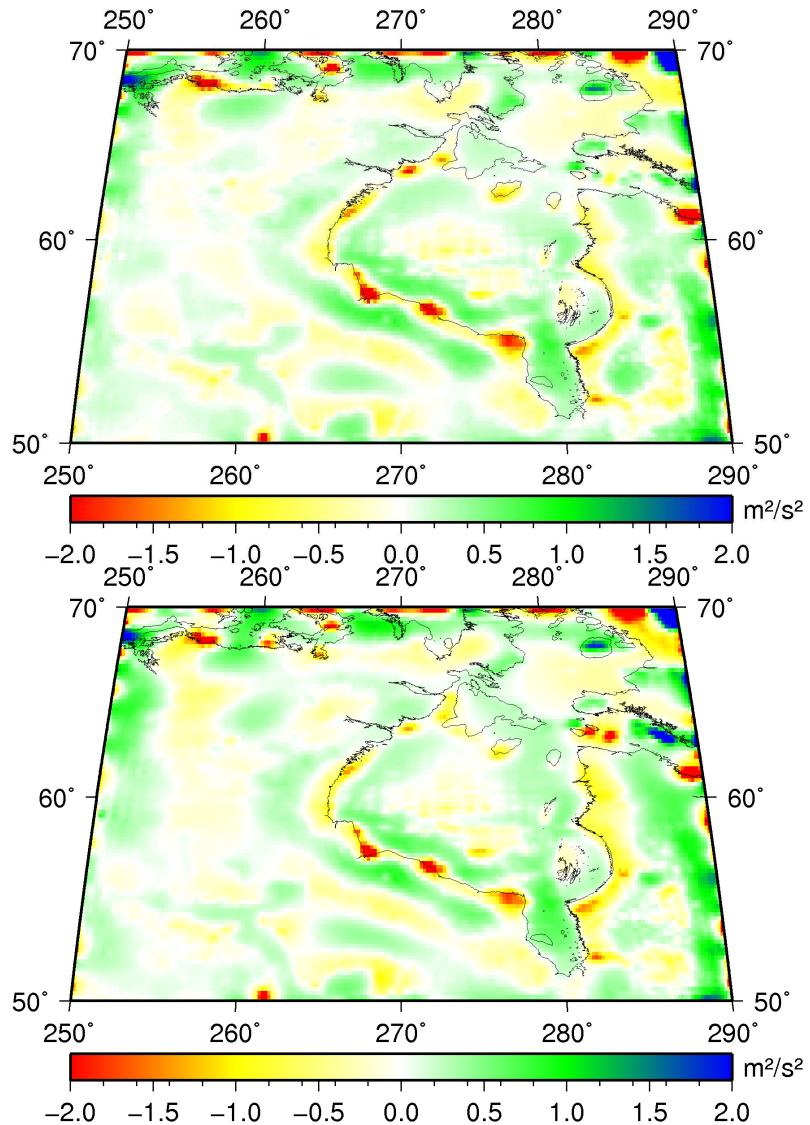
Figure 7.18: Gravity potential error of the combined solutions with diagonal (top) and full (bottom) covariance matrix for GRACE data.

7.4 Summary and conclusions

This chapter investigated methods of combining satellite and terrestrial data into one solution. It has been established that a remove-restore approach only helps in suppressing edge effects, but not in removing long-wavelength error in the terrestrial data by combination with more accurate satellite data.

A better combination was attempted by computing a combined solution using both satellite and terrestrial data as input in a joint estimation. This was performed using two test cases: A global test case with a white noise realisation contaminating the terrestrial data; and a regional test case with simulated low-frequency noise contaminating the terrestrial data. In both cases, realistic noise computed from the ITG-Grace03s covariance matrix (Mayer-Gürr, 2007) was applied to the satellite data. By working in a simulated environment, comparisons to a true field could be made.

Both test cases showed that a combined solution computed in this manner results in significantly improved quality compared to a remove-restore combination. Whether or not the full covariance matrix was used for the GRACE data had only a small impact on the quality of the solution. Both cases also proved that a single layer of RBFs is capable of modelling the entire signal of both satellite and terrestrial data.

The global test case made use of white noise for the terrestrial data. The regional test case tried to use a more realistic low-frequency noise realisation, which also exhibited white noise characteristics. As a result, the assumption of white noise in the stochastic model of the terrestrial data was sufficient in both cases, since white noise was the stochastic model used for the terrestrial data in the estimation. The presence of coloured noise could lead to a significantly worse quality of the combined solution unless a more accurate stochastic model for the terrestrial data is used. Such a model could be applied by making use of a frequency-dependent weighting approach as described in section 5.2. The difficulty would lie in the estimation of the noise model.

For the global test case, the use of an RBF solution estimated only from satellite data as input for a joint estimation was investigated. A good quality of the solution could only be achieved when the full covariance matrix for the input RBFs was used. There are indications that the correlations between RBF coefficients α_i are much stronger than between SH coefficients c_{lm}. Hence, a data combination on the level of a RBF solution needs to take the full covariance matrix into account. This does not result in a reduced numerical complexity compared to the use of SH potential coefficients, which can be used with a diagonal covariance matrix.

The regional test case showed that a solution which suppresses long-wavelength errors in the terrestrial data by the choice of the parametrisation is of comparable quality as a joint solution making use of satellite data. The suppression was achieved by placing the basis functions shallow and starting the summation at a minimum degree l_{min} that corresponded to the truncation degree of the subtracted reference model. Other data sets may however lead to different results. When more energy is contained in the low-frequency errors than in the test case, leakage may occur and negatively affect the solution. This has previously been experienced (in the other direction, with high-frequency signal leaking into the estimation) during the computation of multi-scale solutions (section 2.7).

Chapter 8

Summary, conclusions and recommendations

8.1   Summary and conclusions

The research presented here aimed at developing a methodology for regional gravity field modelling. At the start of the research, a choice was made to use RBFs as a representation. Poisson wavelets of order 3 were chosen as kernel function, after which two main issues were identified that have to be optimised for each data set: RBF network design and bandwidth.

Network and bandwidth selection

Testing with various grid types has proven that Reuter grids are well suited for RBF placement. An RBF solution obtained from a level γ Reuter grid is comparable to a SH solution of degree γ. If sufficient information about the data density and spectral content is present, a choice of the Reuter grid level can be made with little experimentation. Two semi-automatic algorithms were developed to counter over- and underparametrisation: Data adaptation and local refinement. Data adaptation removes RBFs that do not have observations within their influence radius from the estimation. Local refinement places additional basis functions in areas of large residuals after an initial estimation, and yields a good approximation regardless of the initial grid choice.

The optimal bandwidth is estimated using an optimality criterion. This can either be the data fit or the fit to a reference model, if available. Smooth solutions are obtained when the basis functions are placed deep, but at depths that still allow for stable solutions for the chosen network.

Parameter estimation

RBF coefficients are estimated by least-squares using a Cholesky solver. To handle the large amounts of data that are encountered, the analysis software has been numerically

optimised and parallelised for hybrid high performance computer architectures. This results in a program that can handle very large data sets in a reasonable amount, i.e. minutes, of computing time. This is especially important since multiple program runs may be required to find the optimal RBF network and bandwidth.

Modelling of satellite data

The RBF approach was used for gravity field modelling from data collected by the GRACE satellite mission. It made use of the functional model and data previously developed at DEOS for use with spherical harmonics. This included the implementation of the optimal filter scheme of Klees et al. (2008b) for the use with RBFs. The approach was tested using two real-world applications: The recovery of ice mass loss in Greenland and Antarctica, and the monitoring of water storage variations. For unfiltered solutions, it could be shown that RBF solutions outperform spherical harmonics, especially in terms of spatial resolution, and that no additional smoothing is required. Optimally filtered solutions are of much higher quality than unfiltered ones, and reduce the differences between spherical harmonic and RBF solutions. Except for some minor indications (a negative signal in Antarctica that is both seen by ICESat and the RBF solutions, but not by DMT-1; better trend estimate for Lake Victoria), no significant benefit due to the use of RBFs instead of spherical harmonics could be found, unless a reduced numerical complexity by computing regional solutions is taken into account.

Modelling of terrestrial data

Besides satellite data, the RBFs were also employed for the modelling of terrestrial data. Computations were performed using two data set, covering the northeastern USA and eastern Canada. With both data sets it was proven that RBFs can be used for computing high-resolution gravity fields from large amounts of data. The use of the data adaptivity and local refinement algorithms yielded good solutions regardless of the grid spacing chosen for RBF placement. The qualitative comparison to the two official geoids GEOID03 and CGG05 was not entirely satisfactory, which is not due to the RBF approach, but can be attributed to different data and processing strategies: rms errors of several centimetres remained.

Combined modelling of satellite and terrestrial data

Finally, approaches for the combination of satellite and terrestrial data where investigated in both a global and a regional simulated setting. It could be shown that a joint estimation using satellite and terrestrial data and the corresponding stochastic models yielded a significantly better quality of the solution than the remove-restore approach, since at long wavelengths satellite data is of higher quality than terrestrial data. The combined solutions were computed using one layer of RBFs, which as test computations showed was the only working parametrisation.

It can be summarised that RBFs can be considered as an established parametrisation for both regional and global gravity field modelling. The methodology developed with

this research can be employed for all types of gravity data sets, from local terrestrial measurements to global satellite observations. A numerically efficient implementation that can handle large data sets exists.

8.2 Recommendations for further research

During the course of the research presented here, additional research questions were formulated that could not be dealt with in a timely manner and are suggested for further research.

General issues

The choice of the bandwidth of the RBFs is important for the quality of the RBF solution. The bandwidth estimation using a post-fit criterion may not always be satisfactory. It should be investigated if it is possible to relate the optimal bandwidth to the data's auto-covariance function; the bandwidth estimate could be used to derive the necessary grid spacing.

Modelling of satellite data and optimal filtering

Here, GRACE modelling made use of the assumption of white noise, or a frequency-dependent weighting scheme (see section 5.2) using a simple noise model. GRACE modelling could benefit from an improved stochastic model that is constructed from post-fit residuals (Siemes and Liu, 2009). Basin averages will most likely not be significantly affected.

The determination of the observation accuracy σ is a critical factor in the optimal filtering scheme (see section 5.3.3), since $\hat{\mathbf{x}} = (\frac{1}{\sigma^2}\mathbf{N} + \mathbf{D}^{-1})^{-1}\mathbf{b}$. Here, both an external calibration and the post-adjustment noise level were used. It is suggested to look into the possibility of using variance component estimation for this purpose; the noise level should be individually estimated and applied for each month.

Siemes et al. (2009) have developed an improved trend and signal amplitude estimation algorithm. The full covariance matrix of each monthly solution is used in the estimation of the trend and amplitude parameters in eq. (5.20), and unfiltered SH coefficients are used as input. The resulting trend is then optimally filtered. This algorithm yields improved spatial resolution and larger signal amplitudes. This approach needs to be extended to RBF parametrisations.

Currently, the signal covariance matrix that is needed for the optimal filter is computed in the spatial domain (see section 5.3). Each monthly solution is used to synthesise a spatial representation in EWH; the variance of each grid point is computed, and this spatial signal covariance matrix is then transformed into the RBF basis by RBF analysis. Using above improved trend and signal amplitude estimation algorithm, it is possible to compute the signal covariance matrix directly from the RBF coefficients. This approach

would eliminate the synthesis/analysis steps and would not neglect correlations between the coefficients.

A longer GRACE time series would make it possible to compute not only one signal covariance matrix for the whole time span, but individual matrices e.g. for each month. This might improve the quality of the solution by allowing for more signal to be modelled in months of strong signal variation.

It has proven difficult to compare and validate GRACE models due to a lack of accurate reference data and models. This makes it difficult to compare e.g. SH and RBF solutions and prove the benefits (or drawbacks) of RBF modelling. More and higher quality external data sets that can be used for validation are required to definitely assess whether RBFs offer benefits over spherical harmonics in regard to gravity field modelling from GRACE.

Modelling of terrestrial data

Accurate terrain corrections might be required to smooth the input data for terrestrial gravity field modelling. This would lead to an improved approximation quality in regions of rugged topography (mountains). One method for computation of terrain corrections that can be used is the RTM (residual terrain modelling) approach by Forsberg and Tscherning (1981). Using this approach might also require the reduction of satellite data prior to computing a combined solution.

Combined solutions

So far, computing combined solutions using both satellite and terrestrial data with the RBF approach has only been investigated using simulated data sets. The findings should be verified using a suitable setting with real data sets. A method of computing an accurate stochastic model of the terrestrial data set may be required.

Other applications

Gravity field modelling with RBFs can now be considered an established and proven methodology. It is suggested to use them for other applications than those covered here. All computations presented here made use of Poisson wavelets of order 3 as kernel choice. These might not be a satisfactory choice for applications other than those presented here, in which case alternatives need to be investigated. It is then recommended to try harmonic splines Eicker (2008) as representation.

RBFs can be used for the computation of static gravity fields from several years of GRACE data. They are a natural candidate for modelling GOCE data, either for full modelling or for regional refinement of lower-resolution spherical harmonic models as done by Eicker (2008).

Currently, all RBF models computed from GRACE data are static models, since the signal is assumed to be constant over the time span of the input data, i.e. one month.

At DEOS, an approach has been implemented that takes time variability into account (Siemes and Liu, 2009). While this approach makes use of spherical harmonics, it could easily be applied to RBFs.

At DEOS, RBFs are already employed for marine geoid modelling, using a combination of radar altimetry and gravimetry data (Slobbe et al., 2009b). The goal is the joint estimation of the geoid and the mean dynamic topography.

While the research presented here has focused exclusively on modelling of satellite and terrestrial gravity data, a third method of gravity data collection is increasingly used: Airborne gravimetry. Using aircraft as measurement platforms allows for the gathering of high-resolution data over larger areas than terrestrial methods in much shorter amounts of time. Data gathering is not limited to areas accessible by vehicle.

Much of the challenge in airborne gravity field modelling lies in the pre-processing of the data. Alberts (2009) has developed a new methodology that combines the pre-processing steps and gravity field estimation, and performs at least equal compared to traditional methods. The chosen gravity field representation, using fundamental solutions of Laplace's equation in Cartesian coordinates, is however somewhat limited to rectangular areas and homogeneous data distributions. The periodicity of the representation also leads to edge effects. RBFs, employing the data-adaptive and local refinement algorithms developed here, suffer no such limitation. It is recommended to combine the approach of Alberts (2009) with RBFs as gravity field representation for airborne gravity field modelling.

Bibliography

Alberts BA (2009). Regional gravity field modeling using airborne gravimetry data. PhD Thesis, Delft University of Technology.

Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Sorensen D (1999). LAPACK Users' Guide. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Antunes C, Rail P, Catalão J (2003). Point mass method applied to the regional gravimetric determination of the geoid. Studia Geophysica et Geodaetica 47, 495-509.

Awange JL (2006). Lake Victoria: Ecology, resources, environment. Springer.

Barthelmes F (1986). Untersuchungen zur Approximation des äusseren Gravitationsfeldes der Erde durch Punktmassen mit optimierten Positionen. Veröffentlichungen des Zentralinstituts für Physik der Erde, Nr. 02, Potsdam.

Baur O, Kuhn M, Featherstone WE (2009). GRACE-derived ice-mass variations over Greenland by accounting for leakage effects. Journal of Geophysical Research 114, B06407.

Van Beek LPH (2007). PCR-GLOBWB model description. Integration of GFS Data with PCR-GLOBWB using FEWS, report by WL-Delft Hydraulics, Delft.

Berbery EH, Barros VR (2002). The hydrologic cycle of the La Plata Basin in South America. Journal of Hydrometeorology 3, 630-645.

Berry PAM, Garlick JD, Freeman JA, Mathers EL (2005). Global inland water monitoring from multi-mission altimetry. Geophysical Research Letters 32, L16401.

Bettadpur S (2007). GRACE 327-742, UTCSR level-2 processing standards document for level-2 product release 0004. Centre for Space Research, University of Texas at Austin.

Biancale R, Lemoine JM, Balmino G, Bruinsma S, Perosanz F, Marty JC, Loyer S, Bourgogne S, Gégout P (2007). 5 years of gravity variations from GRACE and LAGEOS data at 10-day intervals over the period from July 29th 2002 to June 22nd 2007. Centre Nationale D'Études Spatiales, Paris.

Blackford, L. S., Choi, J., Cleary, A., D'Azevedo, E., Demmel, J., Dhillon, I., Dongarra, J., Hammarling, S., Henry, G., Petitet, A., Stanley, K., Walker, D., Whaley, R. C.

(1997). ScaLAPACK Users' Guide. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Cazenave A, Dominh K, Guinehut S, Berthier E, Llovel W, Ramillien G, Ablain M, Larnicol G (2009). Sea level budget over 2003–2008: A reevaluation from GRACE space gravimetry, satellite altimetry and Argo. Global and Planetary Change 65, 83-88.

Chambodut A, Panet I, Manda M, Diament M, Holschneider M (2005). Wavelet frames: an alternative to spherical harmonic representation of potential fields. Geophysical Journal International 163, 875-899.

Chen JL, Wilson CR, Tapley BD, Grand S (2007). GRACE detects coseismic and post-seismic deformation from the Sumatra-Andaman earthquake. Geophysical Research Letters 34, L13302.

Ditmar P, Klees R (2002). A method to compute the Earth's gravity field from SGG/SST data to be acquired by the GOCE satellite. Delft University Press, Delft.

Ditmar P, Klees R, Kostenko F (2003). Fast and accurate computation of spherical harmonic coefficients from satellite gravity gradiometry data. Journal of Geodesy 76, 690-705.

Ditmar P, Kusche J, Klees R (2003). Computation of spherical harmonic coefficients from gravity gradiometry data to be acquired by the GOCE satellite: regularization issues. Journal of Geodesy 77, 465-477.

Ditmar P, van Eck van der Sluis AA (2004). A technique for modeling the Earth's gravity field on the basis of satellite accelerations. Journal of Geodesy 78, 12-33.

Ditmar P, Kuznetsov V, van Eck van der Sluijs AA, Schrama EJO, Klees R (2006). 'DEOS_CHAMP-01C_70': a model of the Earth's gravity field computed from accelerations of the CHAMP satellite. Journal of Geodesy 79, 586-601.

Ditmar P, Klees R, Liu X (2007). Frequency-dependent data weighting in global gravity field modeling from satellite data contaminated by non-stationary noise. Journal of Geodesy 81, 81-96.

Eicker A, Mayer-Gürr T, Ilk KH (2006). An integrated global/regional gravity field determination approach based on GOCE Observations. In: Observation of the Earth System from Space, Springer, 225-237.

Eicker A (2008). Gravity field refinement by radial basis functions from in-situ satellite data. PhD Thesis, Universität Bonn.

Featherstone WE (2003). Software for computing five existing types of deterministically modified integrations kernel for gravimetric geoid computation. Computers & Geosciences 29, 183-193.

Fengler MJ, Freeden W, Michel V (2003). The Kaiserslautern multiscale geopotential model SWITCH-03 from orbit pertubations of the satellite CHAMP and its comparison to the models EGM96, UCPH2002_02_0.5, EIGEN-1s and EIGEN2. Geophysical Journal International 157, 499-514.

Flechtner F (2007). GRACE 327-743, GFZ level-2 processing standards document for level-2 product release 0004. GeoForschungsZentrum Potsdam, Germany.

Fengler MJ, Freeden W, Kohlhaas A, Michel V, Peters T (2007). Wavelet modelling of regional and temporal variations of the Earth's gravitational potential observed by GRACE. Journal of Geodesy 81, 5-15.

Foerste C, Schmidt R, Stubenvoll R, Flechtner F, Meyer U, König R, Neumayer H, Biancale R, Lemoine JM, Bruinsma S, Loyer S, Barthelmes F, Esselborn S (2008). The GeoForschungsZentrum Potsdam/Groupe de Recherche de Geodesie Spatiale satellite-only and combined gravity field models: EIGEN-GL04S1 and EIGEN-GL04C. Journal of Geodesy 82, 331-346.

Foerste C, Flechtner F, Schmidt R, Stubenvoll R, Rothacher M, Kusche J, Neumayer KH, Biancale R, Lemoine JM, Barthelmes F, Bruinsma J, Koenig R, Meyer U. (2008), EIGEN-GL05C - A new global combined high-resolution GRACE-based gravity field model of the GFZ-GRGS cooperation. General Assembly European Geosciences Union, Vienna.

Forsberg R, Kenyon S (1994). Evaluation and downward continuation of airborne gravity data - the Greenland example. Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS94), 531-538, Banff.

Forsberg R, Tscherning CC (1981). The use of height data in gravity field approximation by collocation. Journal of Geophysical Research 86, 7843-7854.

Freeden W, Gervens T, Schreiner M (1998). Constructive approximation on the sphere (with applications to geomathematics). Oxford Science Publications, Clarendon.

Freeden W (1999). Multiscale modelling of spaceborne geodata. Teubner Stuttgart Leipzig.

Gerlach C, Földvary L, Svehla D, Gruber T, Wermuth M, Sneeuw N, Frommknecht B, Oberndorfer H, Peters T, Rothacher M, Rummel R, Steigenberger P (2003). A CHAMP-only gravity field model from kinematic orbits using the energy integral. Geophysical Research Letters 30, 2037.

Golub GH, Heath M, Wahba G (1979). Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21, 215-223.

Golub GH, van Loan CF (1996). Matrix computations. Johns Hopkins University Press, Baltimore, MD.

Goto K, van de Geijn R (2008). High-Performance implementation of the Level-3 BLAS. ACM Transactions on Mathematical Software 35, article 4.

Gunter B, Urban T, Riva R, Helsen M, Harpold R, Poole S, Nagel P, Schutz B, Tapley B (2009). A comparison of coincident GRACE and ICESat data over Antarctica. Journal of Geodesy 83, 1051-1060.

Han SC, Jekeli C, Shum C (2003). Efficient gracity field recovery using in-situ disturbing potential observables from CHAMP. Geophysical Research Letters 29, L1789.

Heikkinen M (1981). Solving the shape of the Earth by using digital density models. Finnish Geodetic Institute, Report 81:2, Helsinki.

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

Hestenes MR, Stiefel E (1952). Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards 49(6), 409–436.

Hipkin R, Hunegnaw A (2005). Mean dynamic topography by an iterative combination method. GOCINA Workshop "Improving Modelling of Ocean Transport and Climate Prediction in the North Atlantic Region Using GOCE Gravimetry", Luxembourg.

Holschneider M, Chambodut A, Mandea M (2003). From global to regional analysis of the magnetic field on the sphere using wavelet frames. Physics of the Earth and Planetary Interiors 135, 107-124.

Huang J, Véronneau M, Mainville A (2008). Assessment of systematic errors in the surface gravity anomalies over North America using the GRACE gravity model. Geophysical Journal International 175, 46-54.

Hutchinson MF (1990). A stochastic estimator of the trace of the influence matrix for Laplacian smoothing splines. Communications in Statistics - Simulation and Computation 19, 433-450.

Ilk KH (1993). Regularization for high resolution gravity field recovery by future satellite techniques. In: Inverse problems: principles and applications in geophysics, technology and medicine. Akademic Verlag Berlin, 189-241.

Ivins ER, James TS (2005). Antarctic glacial isostatic adjustment: a new assessment. Antarctic Science 17, 541-553.

Klees R, Wittwer T (2005). A data-adaptive design of a spherical basis function network for gravity field modelling. Proceedings Dynamic Planet 2005, "Monitoring and Understanding a Dynamic Planet with Geodetic and Oceanographic Tools", Cairns.

Klees R, Marchenko AN, Alberts B, Wittwer T (2005). Comparison of various methods for the inversion of airborne gravity data. Proceedings Dynamic Planet 2005, "Monitoring and Understanding a Dynamic Planet with Geodetic and Oceanographic Tools", Cairns.

Klees R, Wittwer T (2007). Local gravity field modelling with multi-pole wavelets. Dynamic Planet 2005, "Monitoring and Understanding a Dynamic Planet with Geodetic and Oceanographic Tools", 303-308. Springer.

Klees R, Prutkin I, Tenzer R, Wittwer T (2007). Development of a technique for combining parameters of the Earth's gravity field for quasi-geoid determination on the territory of the Federal Republic of Germany and Europe: Final report, Delft, 2007, 62 pages.

Klees R, Tenzer R, Prutkin I, Wittwer T (2008). A data-driven approach to local gravity field modelling using spherical radial basis functions. Journal of Geodesy 82, 457-471.

Klees R, Revtova EA, Gunter BC, Ditmar P, Oudman E, Winsemius HC, Savenije HHG (2008). The design of an optimal filter for monthly GRACE gravity models. Geophysical Journal International 175, 417-432.

Klees R, Liu X, Wittwer T, Gunter BC, Revtova EA, Tenzer R, Ditmar P, Winsemius HC, Savenije HHG (2008). A comparison of global and regional GRACE models for land hydrology. Surveys in Geophysics 29, 335-359.

Koch KR, Kusche J (2002). Regularization of geopotential determination from satellite data by variance components. Journal of Geodesy 76, 259-268.

Krarup T (1969). A contribution to the mathematical foundation of physical geodesy. Report 44, Danish Geodetic Institute, Copenhagen.

Kusche K, Klees R (2002). Regularization of gravity field estimation from satellite gravity gradients. Journal of Geodesy 76, 359-368.

Kusche J (2003). A Monte-Carlo technique for weight estimation in satellite geodesy. Journal of Geodesy 76, 641-652.

Kusche J (2007). Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. Journal of Geodesy 81, 733-749.

Lawson CL, Hanson RJ, Kincaid D, Krogh FT (1979). Basic Linear Algebra Subprograms for FORTRAN usage. ACM Transactions on Mathematical Software 5, 308-323.

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, Olson TR (1998). The development of the joint NASA GSFC and NIMA geopotential model EGM96. Technical paper 1998-206861, NASA Goddard Space Flight Centre.

Liu X (2008). Global gravity field recovery from satellite-to-satellite tracking data with the acceleration approach. PhD thesis, Delft University of Technology.

Liu X, Ditmar P, Siemes C, Slobbe DC, Revtova E, Klees R, Riva R, Zhao Q (2009). DEOS mass transport model (DMT-1) based on GRACE satellite data: methodology, validation, and application to estimating the ice mass balance in Greenland. Geophysical Journal International, submitted.

Marchenko AN (1998). Parametrization of the Earth's gravity field: point and line singularities. Lviv Astronomical and Geodetical Society, Lviv.

Marchenko AN, Barthelmes F, Meyer U, Schwintzer P (2001). Regional geoid determination: An application to airborne gravimetry data in the Skagerrak. Scientific Technical Report no. 01/07, GeoForschungsZentrum Potsdam.

Mayer-Gürr T (2007). ITG-Grace03s: The latest GRACE gravity field solution computed in Bonn. GRACE science team meeting, Potsdam.

Molodensky MS, Eremeev VF, Yurkina MI (1962). Methods for study of the external gravitational field and figure of the earth. Translated from the 1960 original, The Israeli Programme for the Translation of Scientific Publications, Jerusalem.

Moritz H (1980). Advanced physical geodesy. Wichmann, Karlsruhe.

Narcowich FJ, Ward, JD (1996). Nonstationary wavelets on the m-sphere for scattered data. Applied and Computational Harmonic Analysis 3, 324-336.

OpenMP Fortran application program interface, Version 2.0, 2002. http://www.openmp.org.

Pavlis NK, Holmes SA, Kenyon SC, Factor JK (2008). An Earth gravitational model to degree 2160: EGM2008. EGU General Assembly, Vienna.

Peltier W (2004). Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE. Annual Review of Earth and Planetary Sciences 32, 111-149.

Petrov L, Boy JP (2004). Study of the atmospheric pressure loading signal in VLBI observations, Journal of Geophysical Research 109, B03405.

Picard RR, Cook RD (1984). Cross-validation of regression models. Journal of the American Statistical Association 79, 575-583.

Reigber C, Balmino G, Schwintzer P, Biancale R, Bode A, Lemoine JM, König R, Loyer S, Neumayer KH, Marty JC, Barthelmes F, Perosanz F, Zhu SY (2002). A high-quality global gravity field model from CHAMP GPS tracking data and accelerometry (EIGEN-1S). Geophysical Research Letters 29, 1692.

Reigber C, Schmidt R, Flechtner F, König R, Meyer U, Neumayer KH, Schwintzer P, Zhu SY (2005). An Earth gravity field model complete to degree and order 150 from GRACE: EIGEN-GRACE02S. Journal of Geodynamics 39, 1-10.

Reilly JP, Herbrechtsmeier EH (1978). A systematic approach to modelling the geopotential with point mass anomalies. Journal of Geophysical Research 83, 841-844.

Reuter R (1982). Über Integralformeln der Einheitssphäre und harmonische Splinefunktionen. Veröffentlichungen des Geodätischen Instituts, RWTH Aachen, 33.

Riva REM, Gunter BC, Urban TJ, Vermeersen LLA, Lindenbergh RC, Helsen MM, Bamber JL, van de Wal RSW, van den Broeke, MR, Schutz BE (2009). Glacial isostatic adjustment over Antarctica from ICESat and GRACE satellite data. Earth and Planetary Science Letters, submitted.

Roman DR, Wang YM, Henning W, Hamilton J (2004). Assessment of the new national geoid height model, GEOID03. Proceedings of the American Congress on Surveying and Mapping 2004 meeting.

Rodell M, Houser PR, Jambor U, Gottschalck J, Mitchell K, Meng CJ, Arsenault K, Cosgrove B, Radakovich J, Bosilovich M, Entin JK, Walker JP, Lohmann D, Toll D (2004). The Global Land Data Assimilation System. Bulletin American Meteorological Society 85, 381-394.

Rummel R, Balmino G, Johannessen J, Visser P, Woodworth P (2002). Dedicated gravity field missions - principles and aims. Journal of Geodynamics 33, 3-20.

Saff EB, Kuijlaars ABJ (1997). Distributing many points on a sphere. The Mathematical Intelligencer 19, 5-11.

Sandwell DT, Smith WHF (1997). Marine gravity anomaly from Geosat and ERS 1 satellite altimetry. Journal of Geophysical Research 102, 10039-10054.

Schmidt M, Fengler M, Mayer-Gürr T, Eicker A, Kusche J, Sanchez L, Han S (2007). Regional gravity field modelling in terms of spherical base functions. Journal of Geodesy 81, 17-38.

Siemes C, Liu X, Ditmar P, Revtova E, Slobbe C, Klees R, Zhao Q (2009). Assessing mass change trends in GRACE models. Geophysical Research Abstracts 11, EGU2009-1748.

Siemes C, Liu X (2009). Probing the spatial and temporal resolution of the time-variable gravity field based on GRACE data. Submitted to Journal of Geodesy.

Slobbe DC, Lindenbergh RC, Ditmar P (2008). Estimation of volume change rates of Greenland's ice sheet from ICESat data using overlapping footprints. Remote Sensing of Environment 112, 4204-4213.

Slobbe DC, Ditmar P, Lindenbergh RC (2009). Estimating the rates of mass change, ice volume change and snow volume change in Greenland from ICESat and GRACE data. Geophysical Journal International 176, 95-106.

Slobbe DC, Klees R, Verlaan M, Gerritsen H (2009). Towards a marine geoid for the North Sea, consistent with the dynamics constraints. Workshop North Sea Baltic Sea, Delft.

Stokes GG (1849). On the variation of gravity on the surface of the Earth. Transactions of the Cambridge Philosophical Society 8, 672-695.

Stolk W (2009). An evaluation of the use of radial basis functions for mass change estimates at high latitudes. M.Sc. thesis, Delft University of Technology.

Swenson S, Wahr J (2006). Post-processing removal of correlated errors in GRACE data. Geophysical Research Letters 33, L08402.

Tapley B, Bettadpur S, Ries J, Thompson P, Watkins M (2004). GRACE measurements of mass variability in the Earth system. Science 305, 503-505.

Tapley B, Ries J, Bettadpur S, Chambers D, Cheng M, Condi F, Gunter B, Kang Z, Nagel P, Pastor R, Pekker T, Poole S, Wang F (2005). GGM02 – An improved Earth gravity field model from GRACE. Journal of Geodesy 79, 467-478.

Tenzer R, Klees R (2008). The choice of the spherical radial basis functions in local gravity field modeling. Studia Geophysica et Geodaetica 52, 287-304.

Velicogna I, Wahr J (2005). Greenland mass balance from GRACE. Geophysical Research Letters 32, L18505.

Vermeer M (1984). Geoid studies on Finland and the Baltic. Finnish Geodetic Institute, Report 84:3, Helsinki, Finland.

Vermeer M (1989). FGI studies on satellite gravity gradiometry. 1. Experiments with a 5-degree buried masses grid representation. Finnish Geodetic Institute, Report 89:3, Helsinki.

Vermeer M (1990). FGI studies on satellite gravity gradiometry. 2. Geopotential recovery at 0.5-degree resolution from global satellite gradiometry data sets. Finnish Geodetic Institute, Report 90:1, Helsinki.

Véronneau M, Huang J (2005). The Canadian Gravimetric Geoid Model 2005 (CGG2005). Geodetic Survey Division, Natural Resources Canada, Ottawa, Ontario.

Wahr J, Molenaar M, Bryan F (1998). Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. Journal of Geophysical Research 108 (B12), 30,205-20,229.

Watkins M, Yuan D (2007). GRACE 327-744, JPL level-2 processing standards document for level-2 product release 04. NASA Jet Propulsion Laboratory, Pasadena, CA, USA.

Weightman JA (1965). Gravity, geodesy and artifical satellites. A unified analytical approach. Proceedings Symposium on the use of artificial satellites for Geodesy, Athens.

Wendland H (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. Advances in Computational Mathematics 4, 389-396.

Wendland H (2005). Scattered data approximation. Cambridge monographs on applied and computational mathematics, Cambridge University Press.

Wessel P, Smith WHF (1991). Free software helps map and display data. Eos, Transactions, American Geophysical Union 72, 441, 445-446.

Wittwer T (2006). An introduction to parallel programming. VSSD Uitgeverij, Delft, ISBN 90-71301-78-8.

Curriculum Vitae

Tobias Wittwer was born on July 6th 1978 in Ostfildern/Ruit, a few kilometers south of Stuttgart, in southwest Germany. He attended the Hermann Hesse Gymnasium in Calw from 1989 to 1998. After his military service he started studying Geodesy at the Stuttgart University in 1999. He graduated in 2004 on "High Performance Computing in use for gravity field modelling with CHAMP, GRACE, and GOCE".

In October 2004 he started his PhD research at DEOS on regional gravity field modelling. He also served as webmaster and in the High Performance Computing committee of the Physical and Space Geodesy group. On October 1st 2009 he started working as specialist within the geodesy advisory group of ARCADIS Nederland BV.

Tobias is married to Baukje. They have five children: Two horses, two Guinea pigs, and a Golden Retriever. He is a member of the Academic Motorsports Group Stuttgart and the Dutch Gaming Association Ducosim. His other hobbies include photography and playing the guitar.