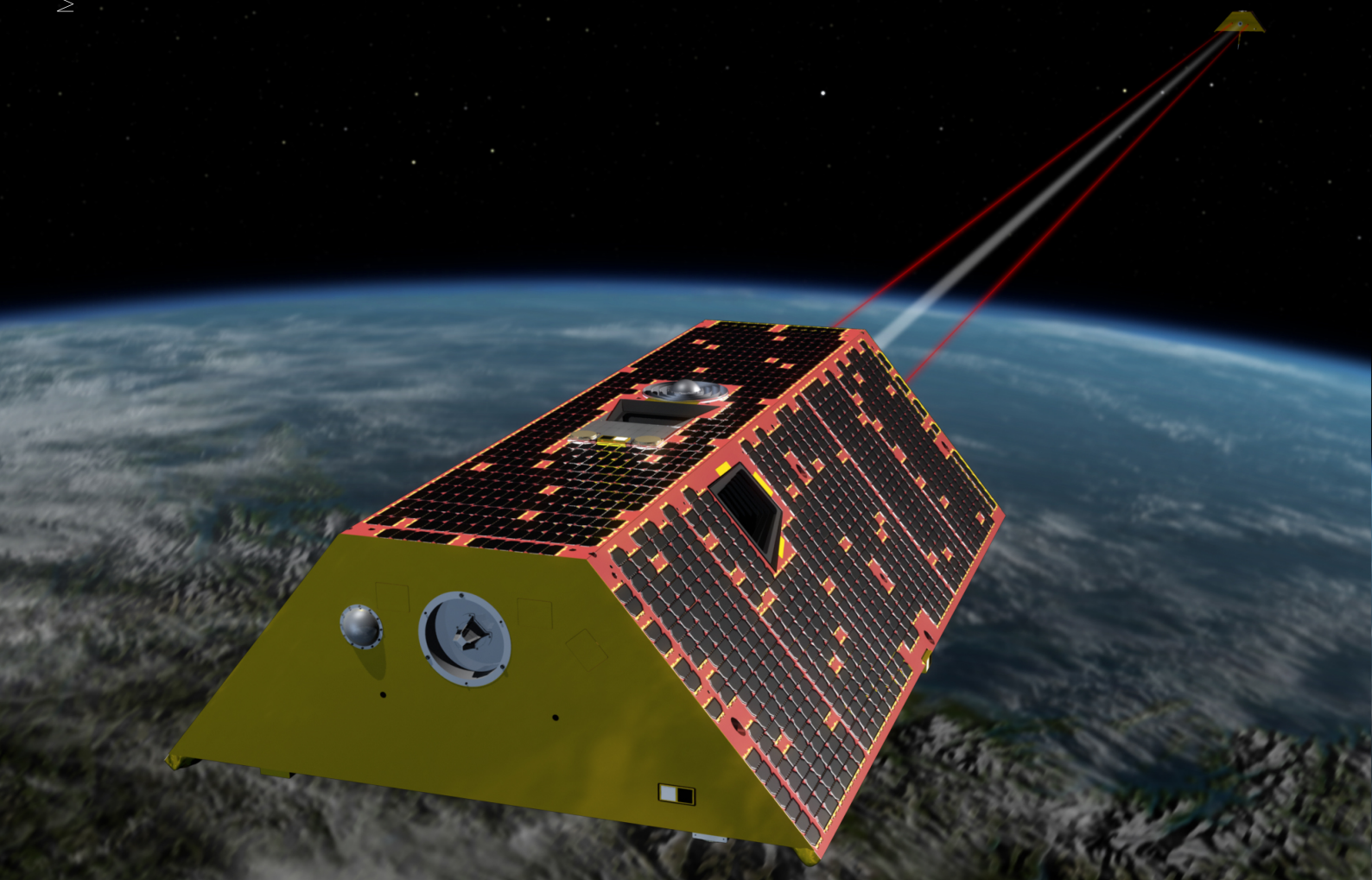
# Bridging the GRACE gap

Validation of satellite gravity observations via glacial isostatic adjustment

Sander W. L. Vermeulen

Master's Thesis



TU Delft

CSR

# Bridging the GRACE gap

## Validation of satellite gravity observations via glacial isostatic adjustment

by

Sander W. L. Vermeulen

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on 21 August 2019 at 15:00 CET.

| | | |
|---|---|---|
| Student number: | 4235819 | |
| Project duration: | December 1, 2018 – August 21, 2019 | |
| Thesis committee: | Dr. ir. W. van der Wal, | TU Delft |
| | Dr. ir. B. C. Root, | TU Delft |
| | Dr. P. G. Ditmar, | TU Delft |
| | Dr. ir. J. de Teixeira da Encarnação, | CSR, The University of Texas at Austin / TU Delft |

*This thesis is confidential and cannot be made public until 1 November 2019*



TU Delft Delft University of Technology

CSR Center for Space Research

# Executive summary

The GRACE mission (2002-2017) provided unprecedented insights into mass redistribution in the Earth system. Following a strong call for continuation of the mass observations, the GRACE-Follow On mission was launched in May 2018, leaving a coverage gap of ca. 1 year between GRACE and GRACE-FO. Geopotential solutions derived from data of the Swarm mission (2013-present) are candidate data to bridge this gap. This report presents a study comparing GRACE and Swarm observations of the gravity trend induced by glacial isostatic adjustment (GIA). Mantle viscosity studies suggest the Hudson Bay area GIA-induced gravity trend is linear on the multi-decadal time scale, which suggests we can extrapolate the GRACE-derived GIA observations into the Swarm time period. The goal of this study is **to use the mass change rate induced by glacial isostatic adjustment to examine the continuity between time series of mass change inferred from the GRACE and Swarm satellite missions**. Such a continuity would validate the sensitivity of the Swarm measurement system to mass rates, a key requirement for Swarm to successfully bridge the GRACE gap.

We identify three steps required to successfully extract an agreeing GIA-induced gravity trend estimate from the GRACE- and Swarm-derived geopotential solutions. First, all observations must be carefully weighted in the least squares regression. Especially the Swarm-derived Stokes coefficient time series show substantial heteroscedasticity, presumably due to the sensitivity of the Swarm measurement system to plasma gradients in the ionosphere [86, 122], and to a lesser extent due to the sub-optimal GPS-receiver settings aboard the Swarm satellites in the period before May 2015 [104]. We use the RMS of the post-fit residuals as uncertainty estimates [114], that we find after subtracting a 1-year moving mean and (semi-)annual components estimated from a 2-year moving window from the Stokes coefficient time series. Second, the gravity fields must be truncated to a maximum degree of 13. In degrees above 13, increased noise levels complicate inferring true geophysical signal from the Swarm-derived gravity fields. To avoid a discrepancy in spatial resolution, the GRACE fields are truncated accordingly. Third, all Stokes coefficients below degree 4 must be truncated. We suspect that long-wavelength leakage from the mass change rate induced by ice mass loss on the Antarctic continent affects the mass change rate that we observe in North America. As the Antarctic mass loss is non-stationary Shepherd et al. [89], the magnitude of the long-wavelength leakage likely also varies across the time periods of respectively the GRACE and Swarm mission.

In a novel experiment we analyze the spatial distribution of the variability of the GRACE-derived gravity trend and show that the Hudson Bay area positive trend anomaly coincides with an area of minimal trend variability. Nonetheless, we show that the difference between the highest and lowest trend in this area from any 5-year window of GRACE data equals ca. 0.7 $\mu$Gal a$^{-1}$. This is high w.r.t. the amplitude of the GIA-induced trend (ca. 2 $\mu$Gal a$^{-1}$). We argue the inter-annual variability is due to hydrology variations. A correction based on GLDAS-NOAH deteriorates the agreement, where we should make the side note that we have identified irregularities in the GLDAS source data. We assess the influence of undulations induced by ice mass loss in Greenland and Alaska and find negligible effects.

We find a latitude-weighted mean residual of 0.0629 $\mu$Gal a$^{-1}$ between the GRACE- and Swarm-derived gravity trends of the positive gravity trend anomaly around Hudson Bay, when we define the limit of this anomaly at >+1.0 $\mu$Gal a$^{-1}$ as computed from the GRACE data. In the same area, we find a latitude-weighted mean standard error of the GRACE- and Swarm-derived gravity trends of 0.0153 and 0.400 $\mu$Gal a$^{-1}$, respectively. The mean misfit equals 0.426 standard errors of the Swarm error Accordingly, we conclude that there is strong evidence of a secular gravity signal in the Hudson Bay area that shows continuity between its GRACE- and Swarm-derived observations.

Finally, we validate our conclusions via the GRACE-FO geopotential solutions. The results are difficult to interpret, most likely due to the limited amount of available GRACE-FO data. Nonetheless, we find that Swarm and GRACE-FO observe various large scale inter-annual mass change processes in a similar way. We conclude that the GRACE and Swarm measurement systems have a similar sensitivity to mass change trends of a spatial scale up to ca. 1500 km. This newfound continuity between GRACE and Swarm is a step towards a multi-decadal record of mass change observations.

# Preface

Many space flight students are drawn to the subject because of an innate curiosity for exploring the deep, unknown void that is space. I used to be no different. Yet, I slowly learned that the concreteness of studying the planet beneath my very own feet beats the appeal of anything far away in said unknown void. Endeavors like the quest of British colonel George Everest to map the tallest peaks of the Himalayan, the Indian surveys of George Airy and John Pratt, or the submarine voyage across the Earth's oceans of TU Delft alumnus Felix Vening Meinesz illustrate how the history of geodesy is rich with adventure.

The fact that modern-day geodetic research does not require sturdy hiking boots anymore does not make it any less exciting. In a time where society is in a heated debate about the impact of our presence on planet Earth, a debate often dominated by emotion instead of reason, thorough understanding of the complex interactions that govern the dynamics of the Earth system is more important than ever. Less than two decades ago a completely new Earth observable was introduced: global time-resolved gravity measurements. Contributions from authors all over the world have shown that these measurements can find a use in an ever-expanding array of research fields, many of which related to the issue of climate change. I firmly believe we are only at the beginning of a new and exciting era in understanding our home planet, a notion I suspect aforementioned pioneers felt too in their respective days. May this newly fueled interest in mother Earth be forwarded to any reader of this document.

First and foremost, I am very thankful for the guidance, motivation, and support of dr. ir. Bart Root. In one of my first lectures at the TU Delft Faculty of Aerospace Engineering, I was struck by Bart's passion for satellite geodesy, and I am glad to admit that part of this enthusiasm has transferred to me in the last year. The majority of the research of this study was executed during a stay at The Center for Space Research of the University of Texas at Austin (CSR). This opportunity was kindly offered to me by dr. Srinivas Bettadpur, the director of CSR. I am very grateful for all the help I have received from dr. João de Teixeira da Encarnação and dr. Mark Tamisiea, both with arranging my stay and with realizing my research ambitions after arriving in Austin. It has been an honor and a pleasure to work in a place that has such a rich history of accomplishments in the field of space research. I would like to thank dr. ir. Wouter van der Wal, who kindly shared an idea that became the foundation for this study, and later provided assistance that shaped the research outlined in this report.

The TU Delft STIR Fund and the Van der Maasfonds have kindly provided financial aid that allowed me to travel to Austin to stay at CSR. Besides the aforementioned benefits for my master's thesis project, this trip has left me with many new friends and memories that I will cherish for years to come.

Finally, I want to express my greatest gratitude to my loving parents, who have been there for me during every minute of the roaring years that were my college career. Their personal and financial support allowed me to flourish as a person and as a scholar, and most importantly, to do so while living the best days of my life.

I am excited about what the future will bring, and what better way to describe my sentiment than to let the song that served as a wake-up call for the astronauts aboard the Apollo 10 mission do so:

*"The best is yet to come"*

Ol' blue eyes

*Sander W. L. Vermeulen*
*Austin, Texas, July 2019*

# Contents

# List of figures

# List of tables

# List of acronyms

| Acronym | Description |
|---|---|
| AIUB | Astronomic Institute of the University of Bern |
| ASU | Astronomical Institute of the Czech Academy of Sciences |
| CSR | Center for Space Research of The University of Texas at Austin |
| CWS | Canopy Water Storage |
| DLR | *Deutsches Zentrum für Luft- und Raumfahrt*; German Aerospace Center |
| DORIS | *Détermination d'Orbite et Radiopositionnement Intégré par Satellite*; Doppler Orbitography and Radio positioning Integrated by Satellite |
| dRMS | Degree Root Mean Square |
| EGSIEM | European Gravity Service for Improved Emergency Management |
| ESA | European Space Agency |
| EWH | Equivalent Water Height |
| GES DISC | Goddard Earth Sciences Data and Information Services Center |
| GFZ | *Das Deutsche GeoForschungsZentrum*; German Research Center for Geo-sciences |
| GGM | GRACE Gravity Model |
| GHWRM | Global Hydrological and Water Resource Model |
| GIA | Glacial Isostatic Adjustment |
| GLDAS(-NOAH) | Global Land Data Assimilation System (postfix *-NOAH* refers to the Noah land surface model) |
| GNSS | Global Navigation Satellite System |
| GPS | Global Positioning System |
| GRACE | Gravity Recovery And Climate Experiment |
| GRACE-FO | *idem*-Follow On |
| GSFC | Goddard Space Flight Center |
| ICGEM | International Center for Global Earth Models |
| IERS | International Earth rotation and Reference Systems service |
| IfG | Institute of Geodesy of the Graz University of Technology |
| JPL | Jet Propulsion Laboratory |
| KBR | K-Band Ranging system |
| LSM | Land Surface Model |
| NASA | National Aeronautics and Space Administration |
| NB | *Nota Bene*; take special notice |
| OSU | The Ohio State University |
| PCA | Principal Component Analysis |
| PODAAC | Physical Oceanography Distributed Active Archive Center |
| POI | Point Of Interest |
| PREM | Preliminary Reference Earth Model |
| RMS | Root Mean Square |
| RMSE | Root Mean Square Error |
| RSL | Relative Sea Level |
| SE | Standard Error |
| SH | Spherical Harmonics |
| SLR | Satellite Laser Ranging |
| SNR | Signal-to-Noise Ratio |
| SST | Satellite-to-Satellite Tracking (prefixes *ll-*: low-low ~, *hl-*: high-low ~) |
| SWC | Soil Water Content |
| SWE | Snow Water Equivalent |
| TBD | To Be Determined |
| TWS | Total Water Storage |
| VCE | Variance Component Estimation |
| WGHM | WaterGAP Hydrology Model |
| WGMS | World Glacier Monitoring Service |

# List of abbreviations

| Abbreviation | Description |
|---|---|
| ca. | *circa*; approximately |
| coeff. | coefficient(s) |
| e.g. | *exempli gratia*; for example |
| et al. | *et alia*; and others |
| etc. | *et cetera*; and so forth |
| excl. | excluding |
| i.e. | *id est*; in other words |
| incl. | including |
| lat. | latitude |
| lon. | longitude |
| n/a | not available |
| p. | page |
| pp. | pages |
| resp. | respectively |
| vs. | *versus*; against |
| w.r.t. | with respect to |

# List of symbols

## Latin symbols

| Symbol | Description | Remarks | Unit (SI unit) |
|---|---|---|---|
| $C_{l,m}$ | Cosine Stokes coefficient | Degree $l$ and order $m$ | TBD |
| $E[\,]$ | Expected value operator | | [ ] |
| $\dot{\bar{g}}^{(GIA)}$ | Latitude-weighted mean gravity trend | See Equation 3.5 | TBD |
| $G$ | Standard gravitational constant | $= 6.67428 \cdot 10^{-11}$ | $m^3\ kg^{-1}\ s^{-2}$ |
| $GM$ | Earth standard gravitational parameter | $= 3.986004415 \cdot 10^{14}$ | $m^3\ s^{-2}$ |
| $i$ | Orbital inclination | | degree ($= \frac{\pi}{180}$) |
| $J$ | Cost function | | [ ] |
| $k_l$ | Love number | Degree $l$, see Table A.1 | [ ] |
| $l$ | Degree | | [ ] |
| $l_{fwhm}$ | Gaussian filter full width half maximum | Expressed as SH degree | [ ] |
| $m$ | Order | | [ ] |
| $M$ | Mask function | | [ ] |
| $MA_\delta$ | Moving average | Window length $\delta$ | TBD |
| $n$ | Sample size | | [ ] |
| $N(\mu,\sigma^2)$ | Normal distribution | Mean $\mu$ and variance $\sigma^2$ | |
| $N_{l,m}$ | Normalization factor | Degree $l$ and order $m$ | [ ] |
| $R$ | Earth mean radius | = 6378136.3 | m |
| $P$ | Covariance matrix | | |
| $P_{l,m}$ | Legendre function | Degree $l$ and order $m$ | [ ] |
| $P_l$ | Legendre polynomial | Degree $l$ | [ ] |
| $r_l$ | Degree correlation coefficient | Degree $l$ | [ ] |
| $r_g$ | Low-pass Gaussian filter radius | | m |
| $r_{hp,g}$ | High-pass Gaussian filter radius | | m |
| $S_{l,m}$ | Sine Stokes coefficient | Degree $l$ and order $m$ | TBD |
| $s_g(l)$ | Low-pass Gaussian filter scaling factor | Degree $l$ | [ ] |
| $s_t(l)$ | Low-pass truncation scaling factor | Degree $l$ | [ ] |
| $s_{hp,g}(l)$ | High-pass Gaussian filter scaling factor | $= 1 - s_g(l)$, degree $l$ | [ ] |
| $s_{hp,t}(l)$ | High-pass truncation scaling factor | $= 1 - s_t(l)$, degree $l$ | [ ] |
| $t$ | Time | | y ($\approx 31556952$ s) |
| $z$ | Arbitrary constant or function | Does not denote any particular parameter | TBD |
| $T$ | Period | | y ($\approx 31556952$ s) |
| $T(\alpha,\nu)$ | Student's T-statistic | Confidence level $\alpha$ and $\nu$ degrees of freedom | [ ] |
| $U$ | Amplitude | | TBD |
| $A$ | Design matrix | | [ ] |
| $\mathbf{x}$ | Parameter vector | | TBD |
| $\mathbf{y}$ | Observation vector | | TBD |

$G$, $GM$, and $R$ are sourced from the 2010 IERS Conventions [62], in line with the GRACE Science Data System (SDS) RL06 processing standards [123]. Units denoted as *TBD* are to be determined from the input data, mathematical context, etc.

## Greek symbols

| Symbol | Description | Remarks | Unit (SI unit) |
|---|---|---|---|
| $\varepsilon$ | Residuals vector | | TBD |
| $\Delta$ | Differential trend | See Equation 3.4 | TBD |
| $\Delta^{(GIA)}$ | Latitude-weighted RMS of differential trend | See Equation 3.7 | TBD |
| $\Theta$ | Sum of all periodic components | See Equation 2.33 | TBD |
| $\mu$ | Arithmetic mean | | |
| $\nu$ | Degrees of freedom | (= $n$ - number of estimated parameters) | [ ] |
| $\rho_E$ | Mean density of the Earth | = $\frac{GM}{G}\frac{3}{4\pi R^3}$ = 5514.32310829 | kg m$^{-3}$ |
| $\rho_{water}$ | Density of water | = 1000 | kg m$^{-3}$ |
| $\sigma$ | Standard deviation | | [ ] |
| $\sigma^2$ | Variance | | [ ] |
| $\tau$ | Student's T-distribution | | |
| $\Phi$ | F-statistic | | [ ] |
| $\chi^2$ | Chi-squared statistic | | [ ] |
| $\omega$ | Phase | | y ($\approx$ 31556952 s) |

$\rho_{water}$ is sourced from Wahr et al. [112]. Units denoted as *TBD* are to be determined from the input data, mathematical context, etc.

## Symbol annotations

| Annotation | Description | Remarks |
|---|---|---|
| $\hat{(\,)}$ | Least squares estimate | See Section 2.5 |
| $\bar{(\,)}$ | Normalized | |
| $\dot{(\,)}$ | First derivative w.r.t. time | |
| $(\,)^{(Comb)}$ | Derived from combined data | See Section 3.6.4 |
| $(\,)^{(GIA)}$ | Latitude-weighted mean or RMS in area where GRACE CSR RL06 ($l_{max}$ = 13) gravity trend >+1.0 $\mu$Gal a$^{-1}$ | See Equation 3.5 and Figure 3.8 |
| $(\,)^{(GRACE)}$ | Derived from GRACE data | |
| $(\,)^{(Swarm)}$ | Derived from Swarm data | |
| $(\,)^T$ | Transpose (matrix operation) | |

# 1 Introduction

*"Circling the Earth in my orbital spaceship, I marveled at the beauty of our planet.*
*People of the world, let us safeguard and enhance this beauty - not destroy it."*

Yuri Garagin, after returning to Earth as the first person in space on 12 April 1961 [45].

## 1.1. A new Earth observable

The NASA/DLR Gravity Recovery And Climate Experiment (GRACE) marked an important milestone in Earth observation. For the first time, mass movements in the Earth system could be studied directly by observing the induced changes in the Earth's gravity field. These time-resolved gravity observations meant a new and unique addition to the existing suite of Earth observables. Initially a pioneering geodesy experiment, the GRACE measurements are now regarded as a reliable mass transport product that has been used to study a wide range of geophysical processes including ocean mass changes [44] ice mass changes [89, 111], terrestrial water storage changes [85, 100], response of the solid Earth to the growth and decay of ice sheets (glacial isostatic adjustment; GIA) [98, 108, 115, 119], and a number of other processes. Studies on changes in fresh water storage basins vital to communities in the local area [69, 73], or studies that link GRACE observations to the complex interactions and transitions associated with today's changing climate [101] have potential for substantial impact. Tangible societal benefits can already be seen in the improvements of service applications such as the United States Drought Monitor that GRACE data helped realize [6], or the European Gravity Service for Improved Emergency Management (EGSIEM), which aims to use mass redistribution measurements to warn the public for future hydrological extremes [39].

Table 1.1 gives a rough estimate of the amplitude, spatial scale, and temporal scale of the gravitational variations induced by various geophysical processes. The table is shown to give the reader a general idea on what kind of gravitational variations to expect. Figure 1.1 shows an example analysis of the Earth gravity observations of GRACE. The left side of Figure 1.1 shows the gravitational acceleration trend. Note how processes like GIA (e.g., around Hudson Bay), and ice melt (e.g., Greenland) are observable in this signal component. Via Newton's law of universal gravitational the magnitude of the observed gravity trend can be correlated to the magnitude of the mass change rate, which allows for direct quantification of aforementioned geophysical processes. The right side of Figure 1.1 shows the amplitude of yearly variation in gravitational acceleration. Note how high-precipitation areas (e.g., the Amazon) can be differentiated from areas with low annual rainfall (e.g., Sahara Desert). Again, the GRACE data can be used to directly quantify how these mass transport processes develop through time.

To compute the results in Figure 1.1, a trend, annual, and semi-annual component are estimated from the CSR RL06 geopotential solutions spanning April 2002 to June 2017 via least-squares regression. This approach will be elaborated upon in Section 2.5. The geopotential solutions we use are published as a set of *Stokes coefficients*, which can be used to weigh a Fourier series of spherically-defined orthogonal basis functions called the *spherical harmonics* to find the spatial maps shown in Figure 1.1. This report assumes a working knowledge of the decomposition of a function defined on the sphere or a potential field into a set of spherical harmonics. For additional information about spherical harmonics we refer the reader to Heiskanen and Moritz [37].

| **Geophysical process** | **Amplitude** [m s$^{-2}$] | | **Spatial scale** [km] | | **Temporal scale** [y] | |
|---|---|---|---|---|---|---|
| | Min | Max | Min | Max | Min | Max |
| Mass transport in atmosphere and oceans | $3 \times 10^{-10}$ | $7 \times 10^{-9}$ | 20 | $1 \times 10^{4}$ | $1 \times 10^{-3}$ | 2 |
| Hydrological mass redistribution in or on the terrestrial surface | $3 \times 10^{-9}$ | $2 \times 10^{-7}$ | $\approx 0$ | $1 \times 10^{4}$ | $2 \times 10^{-2}$ | 10 |
| Time evolution of glaciers | $4 \times 10^{-9}$ | $4 \times 10^{-7}$ | $\approx 0$ | $1 \times 10^{3}$ | $8 \times 10^{-2}$ | 100 |
| Glacial isostatic adjustment | $4 \times 10^{-10}$ | $4 \times 10^{-8}$ | 500 | $1 \times 10^{4}$ | $1 \times 10^{4}$ | $1 \times 10^{5}$ |
| Mass transport due to earthquakes | $9 \times 10^{-9}$ | $4 \times 10^{-6}$ | $\approx 0$ | $2 \times 10^{3}$ | $\approx 0$ | 0.5 |

Table 1.1: Overview of typical values of amplitude (in terms of gravitational acceleration), spatial scale, and temporal scale of the gravitational variation induced by selected geophysical processes. Table contents adapted from de Teixeira da Encarnação [22] with the author's permission.

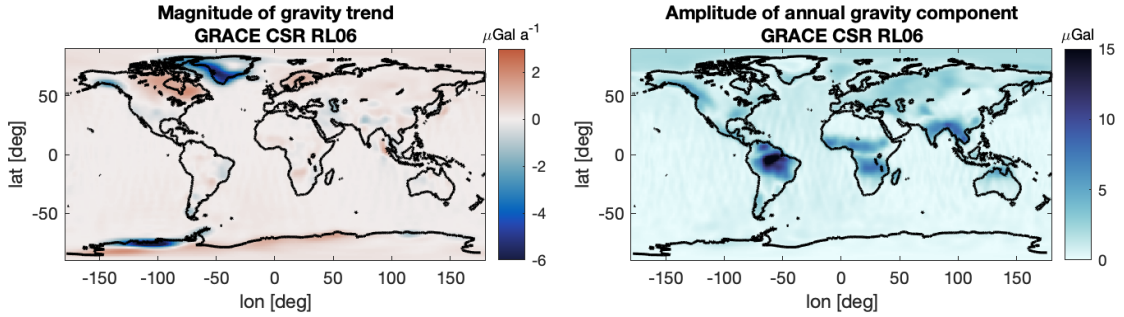



Figure 1.1: Example of time-variable gravity analysis via GRACE data. Left side: gravitational acceleration trend. Right side: amplitude of yearly variation in gravitational acceleration. Computed by fitting a trend, annual, and semi-annual component via least-squares regression to the CSR RL06 Stokes coefficients, which span April 2002 to June 2017. We apply a 300 km half-width Gaussian filter for smoothing (will be covered in Section 2.4.1.

## 1.2. Continuing GRACE

The GRACE mission was operational from 2002 to 2017. As the GRACE mission results were accepted by the scientific community, there was a strong call for continuation of the time series of mass transport observations [13, 54]. In fact, the 2017 NASA Decadal Survey[1] [54] placed a *mass change continuity mission* among the top five priorities for continued Earth observations. The GRACE Follow On (GRACE-FO) mission was launched in May 2018, which prioritized prolonging the measurements over revolutionizing the mission concept and therefore employs roughly the same mission configuration.

Uninterrupted continuation of the time series of mass transport observations that GRACE started would require bridging the data gap of ca. 1 year between GRACE and GRACE-FO (the *GRACE gap*). This is currently a subject of broad interest in the field of satellite gravimetry, as many users would benefit from longer time series. In the last two decades we have seen that the insights from satellite gravimetry extend far beyond our innate curiosity for planet Earth. Continued monitoring of variations in the hydrological cycle of the Earth system can aid humankind in predicting and managing the effects of large-scale weather phenomena, and in understanding the processes that constitute what has been called *the defining problem of our generation*: anthropogenic climate change [101]. As such, a continuous, multi-decadal record of Earth mass transport observations is a project with potential for considerable scientific, societal, and political impact.

Various methods have been proposed to bridge the GRACE gap. Geopotential solutions based on surface loading inversion seem ineffective due to the sparsely and heterogeneously distributed ground-based GPS stations [71, 126]. For similar reasons, in addition to the relatively high orbit of the used satellites, inversion of satellite laser ranging (SLR) position observations of Earth orbiters is limited to very coarse resolutions (degrees < 5, ca. 5000 km at the equator) [11, 91, 125]. The altitude of the orbiter affects the spatial resolution as the decay of gravitational potential with altitude is inversely related to the wavelength of a particular harmonic [37]. The limited resolution of SLR-derived results makes it hard to invert to reliable mass transport estimates [14].

Using satellite position observations of the Global Positioning System (GPS; also described as high-low satellite-to-satellite tracking [hl-SST]) yields more promising results. Here, the ESA Swarm satellites are particularly interesting for three rea-

[1] `http://science.nasa.gov/earth-science/decadal-surveys`

sons. The Swarm mission time frame (active since late 2013) covers the GRACE gap, the satellites are equipped with geodetic quality GPS-receivers, and the satellites fly at a low orbital altitude (ca. <500 km) [35]. Numerous authors have computed gravity fields from GPS position observations of the Swarm satellites at intervals short enough to assess the time dependency of Earth's gravity field on a monthly temporal scale [10, 24, 38, 124]. Combining the geopotential solutions with results derived from SLR further improves the geopotential solutions, especially during periods of increased ionospheric activity when the Swarm fields notably degrade [49, 91, 126].

At the moment, various groups routinely compute gravity fields from Swarm satellite data. We focus on solutions that are part of an ongoing project of de Teixeira da Encarnação et al. [23], where individual Swarm-derived gravity fields of groups at the University of Bern (AIUB), the Czech Academy of Sciences (ASU), Graz University of Technology (IfG), and The Ohio State University (OSU) are combined with weights derived from variance component estimation (i.e., combination on the *solutions level*). It is worth noting that these four individual solutions use different approaches to invert the GPS position observations into Earth gravity fields, namely the celestial mechanics approach [8], the decorrelated acceleration approach [9], the short-arcs approach [50], and the improved energy balance approach [88], respectively. It is believed that combining the four solutions best exploits the particular sensitivity of each of the four approaches, yielding the highest quality gravity fields [23]. These combined Swarm solutions do not contain SLR-derived information. Weigelt et al. [117] offer a data product that does combine hl-SST and SLR, but for reasons that will be given in Section 2.1.4 we mainly focus on the combined Swarm fields.

Like the Swarm-derived fields, the GRACE field are also based on inverting observations of the kinematic motion of the respective satellites into potential fields. The primary source of these observations are micrometer-level measurements of the inter-satellite distance [99]. For Swarm, these source data are centimeter-level GPS position observations [124]. Understandably, this accuracy discrepancy manifests itself as a substantial difference in quality of the respective geopotential solutions. Most notably, an increase in non-geophysical noise in the short-wavelength coefficients of the Swarm fields reduces the spatial resolution of the usable signal to ca. 1500 km (vs. ca. 300 km for GRACE). Nonetheless, Lück et al. [46] derived time series of ocean mass from Swarm-derived gravity fields, and Meyer et al. [51] investigated cryospheric mass change in this way.

Figure 1.2 shows the same time-dependent components of the Earth's gravity fields as Figure 1.1 (i.e., gravity trend and amplitude of annual variation), but now comparing these components across GRACE- and Swarm-derived gravity fields. We have applied a 1000 km Gaussian filter to decrease aforementioned high-degree noise [112] (this method will be elaborated upon in Section 2.4). Note how the amplitude of the annual component shows good agreement between Swarm and GRACE. The results are less congruent when comparing global maps of the gravity trend. For a great part, this is explained by the fact that the GRACE and Swarm missions cover mostly different time frames and that most geophysical processes observable in the gravity trend component (e.g., ice loss [52, 89], terrestrial water storage changes [83]) do not have a linear time dependence, but instead show substantial inter-annual variability.

Some geophysical processes induce a secular gravity signal that shows little inter-annual variability. An example of this is glacial isostatic adjustment, a term used to define the deformation of the Earth's crust in response to ice mass redistribution. The viscoelastic response of the solid Earth to changing loads, coupled to motion of mass underneath Earth's crust, induces both positive and negative gravity trends across the Earth [92]. For two reasons, we focus on the gravity trend around Hudson Bay (this area is also referred to as *Laurentia*). First, as the magnitude of the GIA-induced mass change rate is the largest here, this location gives us the best chance of successfully isolating the GIA-induced gravity trend from the relatively noisy Swarm fields. Second, studies using various GIA-related observations (e.g., relative sea level change, GPS-observations of surface deformation, [time-variable] gravity, etc.) to infer estimates of the (local) mantle viscosity yield viscosity results high enough to allow the exponential time-dependence of the mass change in the Hudson Bay area to be approximated as linear on a multi-decadal scale [42, 60, 98]. For more information on GIA studies we refer the reader to Whitehouse [119]. This linear time-dependency is a key assumption in our study, as it allows measurements of the GIA-induced gravity trend to be extrapolated on said multi-decadal scale. Figure 1.3 shows a close-up of the GRACE observations of the gravity trend in Laurentia. This figure was computed from the GRACE CSR RL06 Stokes coefficients spanning April 2002 to June 2017 by estimating a trend, annual, and semi-annual component via least squared regression after applying a 300 km half width Gaussian filter.

A requirement for Swarm to successfully bridge the GRACE gap is similar sensitivity of the measurement system to terrestrial mass change. This study focuses on sensitivity of the respective measurement systems to the mass change *rate*. We use the mass change rate induced by GIA to assess the continuity between time series of gravity observations inferred from GRACE and Swarm satellite data. The fact that the Laurentian GIA-induced gravity trend may be approximated as linear on a multi-decadal time scale suggests that a similar gravity trend should be retrievable from the GRACE and Swarm gravity observations, regardless of the fact that their observations cover mostly different time frames (2002-2017 vs. 2013-present). Identifying a continuity between the GRACE- and Swarm-derived gravity trend observations, and further establishing Swarm as the solution to the GRACE gap, would mean moving one step closer to a multi-decadal record of mass transport in the Earth system.

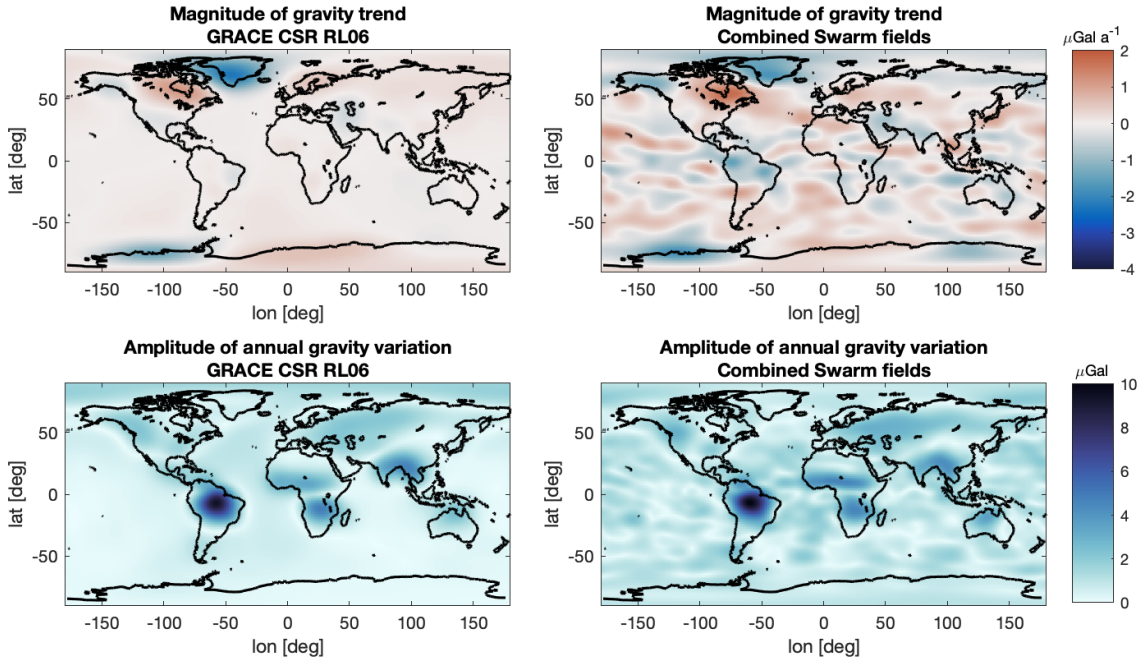



Figure 1.2: Comparison of the gravity trend (top panels) and the amplitude of the yearly gravity variation (bottom panels) between CSR GRACE RL06 (spanning April 2002 to June 2017, shown in left panels) and the combined Swarm fields (spanning December 2013 to March 2019, shown in right panels). Computed by fitting a trend, annual, and semi-annual component via least squares regression. We apply a 1000 km half-width Gaussian filter for smoothing. NB: these are different filter settings, along with different color scales, as compared to Figure 1.1.

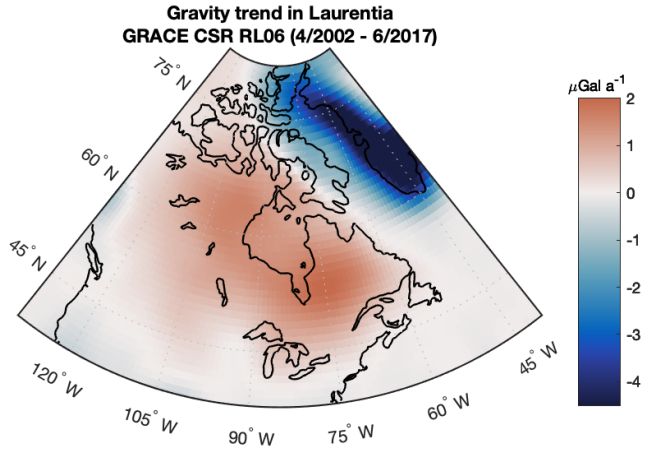



Figure 1.3: GRACE-derived gravity trend in Laurentia. Computed from CSR GRACE RL06 data by fitting a trend, annual, and semi-annual component via least squared regression. We apply a 300 km half-width Gaussian filter for smoothing. General scientific consensus is that the positive gravity trend anomaly visible around Hudson Bay is dominated by GIA.

## 1.3. Research question

We summarize our research objective in the following main research question:

*Does a continuity of the gravitational signal induced by glacial isostatic adjustment exist between time series of mass change inferred from GRACE and Swarm satellite data?*

Our hypothesis is confirmed if the difference between the GRACE-derived and the Swarm-derived observation of the GIA-induced trend is smaller than the uncertainty of the computed trends. This makes error and uncertainty analysis an important part of this research. The time span of the data that we use is April 2002 to June 2017 for GRACE, and December 2013 to March 2019 for Swarm.

We split the main research question into the following sub-questions:

**Q.1** *What is the mass change rate in Laurentia as observed by GRACE and Swarm?*
We compute mass change rate estimates via gravity fields solutions based on GRACE and Swarm satellite data. These results will contain the GIA-induced gravity trend, but also trend signals caused by other geophysical processes.

(a) What is the magnitude and spatial distribution of the gravity trend in the area around Hudson Bay?

(b) What truncating or Gaussian filtering parameter settings are required to remove the most noise from the gravity field solutions, while leaving the GIA-induced trend estimate untouched?

(c) What is the influence on the trend result of including the $S_2$, $K_1$, and $K_2$ tidal aliases as components in the fitted model?

(d) What is the uncertainty of the gravity trend estimates?

(e) What is the influence of weighting the observations when computing the trend?

(f) What is the influence of the length of the gravity field time series on the trend estimate?

**Q.2** *What is the influence from other geophysical signals on the observed mass change rate in the Laurentian area?*
The common procedure to isolate the GIA-induced component from gravity observations involves correcting these observations for other influences from other geophysical processes.

(a) What is the variability between gravity trends computed from different time windows?

(b) What is the influence from inter-annual hydrology variations, and how can this be corrected?

(c) What is the influence from signal leakage and undulations caused by mass change in Greenland and Alaska, and how can this be corrected?

(d) What is the influence from long-wavelength leakage caused by mass change in other areas of the Earth, and how can this be corrected?

**Q.3** *How do the GRACE- and Swarm-derived mass change rates in Laurentia compare?*
We use the results from the previous two sub-questions to evaluate the differences between the GRACE and Swarm observations of the GIA-induced gravity trend in the Hudson Bay area.

(a) What is the difference in magnitude and spatial distribution between the GRACE- and Swarm-derived gravity trends in the area around Hudson Bay?

(b) Is the difference between the GRACE and Swarm observations of the Hudson Bay area gravity trend statistically significant?

**Q.4** *Can data inferred from observations of the GRACE-FO satellite mission be used to validate the conclusions derived from the combined Swarm fields?*
In this final research question, we aim to validate our findings via geopotential solutions derived from the successor of GRACE: GRACE-FO.

(a) How to the mass observations after the GRACE period of Swarm and GRACE-FO compare?

(b) How does adding Swarm resp. GRACE-FO information to the GRACE time series affect the mass rate estimates?

## 1.4. A note on the structure of the thesis

The remainder of this report is structured as follows: Chapter 2 and Chapter 3 cover the methodology of this study, coarsely distinguishing between general gravity field processing methods and methods specific to GIA study, respectively. Note how this aligns with the difference between research sub-questions Q.1 and Q.2. Readers familiar with the application of satellite-derived gravity fields to geophysical study might want to skip Chapter 2, but do note our error estimation method in Section 2.6.1 as it diverts slightly from traditional practice. Similarly, readers experienced with the application of GRACE data to GIA study might want to skip the first sections of Chapter 3. Exceptions are our interpretation of

an existing method to quantify the total influence of all non-GIA processes on the observed gravity trend in Section 3.1, and Section 3.2.3, where we identify and correct inconsistencies in the commonly used GLDAS-NOAH hydrology model. We also suggest reading Section 3.3 while anticipating that the application of Swarm-derived gravity fields happens at a drastically lower maximum degree, compared to modern-day GRACE applications, which amplifies the unwanted effects described in that section. Furthermore, we think that Section 3.6, where we explain how we will use the recently released GRACE-FO geopotential solutions to validate our conclusions, are worthwhile for any reader as this is a first-of-a-kind experiment.

Note that when we show an application of a particular method in the methodology chapters, this is only to motivate our choice for this technique, to show how various techniques compare, or to verify correct implementation. We do not intend to perform any analysis related to our research question in Chapter 2 and Chapter 3.

Chapter 4 shows and discusses the results of our experiments. We would like to highlight the results of our *trend variability* experiment, covered in Section 4.5, as we believe it proves the feasibility of the central idea of this study: extrapolating the GRACE observations of the GIA-induced gravity trend in North America into the Swarm era (and vice versa). Section 4.9 combines all of the study's findings at that point and presents a quantitative assessment of the discrepancy between the GRACE- and Swarm-derived gravity trend observations, in light of the uncertainty of the trend estimates. As such, this section could be regarded as a quantitative summary of the study. Finally, we would like to point Section 4.10, where present the results of aforementioned GRACE-FO-based validation exercise.

To aid the readability of our figures and to align with literature, we will mostly use the non-SI unit of microGal ($\mu$Gal) to describe (gravitational) acceleration, where 1 $\mu$Gal = $1 \times 10^{-8}$ m s$^{-2}$. Traditionally in geodesy, *gravity* refers to the resultant force of the gravitational attraction in accordance with Newton's law of universal gravitation and the centrifugal acceleration due to the rotation of the particular body [34]. In this report we use an Earth-fixed reference frame (as per GRACE RL06 processing standards [123]), which means the centrifugal component reduces to 0. We will therefore use *gravity* and *gravitational acceleration* interchangeably in this text. Similarly, we will not cover nor discern between the various gravity differentials commonly used in geodesy (e.g., gravity anomaly, gravity disturbance), as we are almost exclusively studying gravity trends.

# 2

# Data and general processing methods

This first methodology chapter covers general processing, regression, and uncertainty computation methods applied to the geopotential solutions. At the end of this chapter we have covered all necessary methodology to answer research sub-question Q.1. Processing steps applied to Stokes coefficients are only shown for the cosine coefficient $C_{l,m}$, but are identical for the sine Stokes coefficient $S_{l,m}$.

Definitions of acronyms are given in the List of Acronyms on page xiii. Symbol definitions and units, and values used for constants can be found in the List of Symbols on page xv.

## 2.1. Gravity field solutions

The main data source for the experiments in this study are satellite-derived solutions of the Earth's gravity field. This section introduces the geopotential solutions that we use. Auxiliary data sets used for, e.g., corrections to metrics derived from the geopotential solutions or data sets used for verification or validation purposes will be introduced when applicable.

### 2.1.1. GRACE

Unless noted otherwise, we use the GRACE CSR RL06 geopotential solutions in Stokes coefficient format [102]. All gravity solutions that we denounce as *GRACE data* are derived using the K-band ranging (KBR) observations as primary data source. We use the solutions with a maximum spherical harmonics degree of 60. For more background information on GRACE data processing we refer the reader to Bettadpur [7]. All GRACE geopotential solutions in Stokes coefficient format that we use, including older releases we sporadically use for verification purposes, are publicly available via the NASA Physical Oceanography Distributed Active Archive Center (PODAAC)[1].

At certain points in this study we will use the GRACE CSR RL06m mascon solutions [80, 82] for illustrative or correctional purposes. These solutions are also computed with the KBR observations as primary data source but use a different inversion scheme that is based on estimating local mass concentrations on a geodetic grid (the *mascons*) directly from the GRACE KBR measurements. This method was first applied to GRACE by Luthcke et al. [47]. The CSR RL06m mascon solutions are publicly available via the CSR GRACE website[2].

### 2.1.2. GRACE-FO

We use the GRACE-FO CSR [103], JPL [53], and GFZ [19] RL06 geopotential solutions in Stokes coefficient format. Similar to the GRACE solutions we use, these solutions are derived using the K-band ranging (KBR) observations as primary data source. We use the solutions with a maximum spherical harmonics degree of 60. We refer readers interested in more background information about GRACE-FO data processing to Yuan [123]. The CSR, JPL, and GFZ solutions all cover June 2018 to April 2019, not including August and September 2018.

In an attempt optimally use all available data, we do not use the individual solutions but instead take the arithmetic

[1] `http://podaac-tools.jpl.nasa.gov/drive/`
[2] `http://www2.csr.utexas.edu/grace/`

mean of the CSR, JPL, and GFZ data products. As the epochs line up perfectly this does not require any interpolation. For GRACE, this approach was shown to be effective in the early days of the mission when data was scarce [114]. All GRACE-FO geopotential solutions that we use are also publicly available via the NASA PODAAC[3].

### 2.1.3. Swarm

We use geopotential solutions in Stokes coefficient format derived from GPS position observations of the ESA Swarm satellites. In a project managed by de Teixeira da Encarnação et al. [23] four independently computed time series of gravity fields are combined at the solutions level, as was mentioned in Section 1.2. The combined Swarm fields have a maximum spherical harmonics degree of 40. The combined Swarm solutions are publicly available via the ESA Swarm Mission website[4]. Readers interested in more background information about these fields or this project are also referred to aforementioned website. With the exception of Appendix B, we do not use the individual Swarm solutions anywhere in this study. As such, we will refer to the combined Swarm fields as simply the *Swarm* fields hereafter.

### 2.1.4. GeoQ

The final satellite-derived geopotential product that we use are the 2019 *GeoQ* geopotential solutions of Weigelt et al. [116, 117]. These solutions are computed from GPS position observations of 27 different Earth orbiters using the accelerations approach [9]. The combination is done on the normal equations level with weights derived from variance component estimation. In degrees ≤ 6, the authors combine the GPS-derived Stokes coefficients with counterparts derived from SLR observations of 9 satellites. The GeoQ fields are publicly available via the International Center for Global Earth Models (ICGEM) FTP server[5]. Weigelt et al. [117] showed that the observations of the positive trend anomaly around Hudson Bay derived from these solutions show very good agreement with GRACE observations after applying a 750 km half width Gaussian filter (virtually identical spatial distribution, discrepancy in trend amplitude <1%).

It is important to note that the publicly available GeoQ fields have had substantial post-processing applied to them, most importantly a temporal filtering process that includes a Hodrick-Prescott filter (a deseasoning filter commonly used in macroeconomic studies) and a Kalman-type filter. This means the epochs are no longer independent observations, which makes them fundamentally different from the GRACE- or Swarm-derived fields. In this study we are not arguing in favor or against such processing, but we do reason that comparing independent GRACE or Swarm observations to the GeoQ solutions would be a faulty comparison. Additionally, the GeoQ fields contain GPS position observations and accelerometer data of the GRACE satellites. Both these data will not be available during the GRACE gap, which makes it hard to extend any conclusions derived from the full GeoQ data set to its use as a GRACE gap filler. For these two reasons, we will not focus on the GeoQ fields in this study. Instead, we will use a version of the GeoQ solutions that did not have the temporal filter applied to them as an auxiliary data set to investigate how a longer time series of GPS-derived geopotential solutions compares to the combined Swarm fields[6]. The filtered GeoQ fields are not used anywhere in this study, and therefore we will refer to the unfiltered GeoQ fields as simply the *GeoQ* fields hereafter. For more background information about the GeoQ fields we direct the reader to Weigelt et al. [116, 117].

### 2.1.5. Pre-processing

In early GRACE studies it became apparent that the second-degree zonal coefficient, $C_{2,0}$, of the GRACE-derived gravity fields has a particularly high error [e.g. 43]. Cheng and Ries [17] executed an analysis of cross-track forces acting on GRACE, and the cross-track accelerometer data, and concluded that a systematic, temperature dependent error in the accelerometer data could be the cause of this problem with this particular coefficient. Most GRACE application studies replace $C_{2,0}$ with results derived from SLR position observations of other Earth orbiters. However, as the goal of this study is to infer the continuity between observations of specifically the GRACE and Swarm mission, we choose not to introduce any outside information into the geopotential solutions. We therefore set $C_{2,0}$ to 0 in all geopotential solutions that we use, regardless if they are based on GRACE, Swarm, or any other observation source.

The origin of our reference frame is the Earth's center of mass [123], which means the degree 1 coefficients all reduce to 0. Even though Chambers et al. [15] showed that the degree 1 coefficients have a non-negligible effect on the results of mass transport inversions, we choose to keep the 0 values as a correction would again require introducing external data sources [121].

[3] `http://podaac-tools.jpl.nasa.gov/drive/`
[4] `http://earth.esa.int/swarm/`
[5] `ftp://icgem.gfz-potsdam.de/`
[6] These data are not publicly available, but were kindly provided to us by dr. Matthias Weigelt at the Institute of Geodesy of the Leibniz University in Hannover, Germany.

## 2.2. Dimensionalizing and expanding Stokes coefficients

The satellite-derived Earth gravity fields that we use are distributed in the form of sets of non-dimensional Stokes coefficients $C_{l,m}$ and $S_{l,m}$. These coefficients can be dimensionalized into various gravity functionals. We choose to use gravitational acceleration, and not the geoid height, as gravitational acceleration provides a higher spatial resolution [55]. Equivalent water height (EWH), a unit where the gravitational change is converted into a fictitious water layer of a certain height, is often use in literature. We use comparison to results from literature as a verification method throughout this report, and when applicable we convert our results to EWH to match the literature.

Even though Ditmar [26] showed using an oblate reference spheroid yields more accurate mass estimates, we choose to use a spherical reference Earth. The goal of this study is to compare mass rate estimates inferred from different satellite data sets, not to find the most accurate estimate in the absolute sense. As it simplifies computations, we therefore opt to use a spherical reference Earth. We dimensionalize the Stokes coefficients via:

$$\text{For gravitational acceleration in m s}^{-2}\text{:} \quad C^{(g)}_{l,m} = \frac{GM}{R^2}(l+1)\cdot C_{l,m} \tag{2.1}$$

$$\text{For equivalent water height in m:} \quad C^{(EWH)}_{l,m} = \frac{R}{3}\frac{\rho_E}{\rho_{water}}\frac{2l+1}{l+k_l}\cdot C_{l,m} \tag{2.2}$$

where $GM$ is the standard gravitational parameter of the Earth, $R$ is the mean radius of the Earth, and $l$ and $m$ are respectively the degree and order of the particular Stokes coefficient. $\rho_E$ and $\rho_{water}$ are respectively the mean density of the Earth and the density of water in kg m$^{-3}$. Finally, $k_l$ is the Love number corresponding to degree $l$. The Love numbers define how the Earth deforms elastically under a changing surface load. Equation 2.2 is based on the work of Wahr et al. [112], who use Love numbers computed from the *PREM* Earth model of Dziewonski and Anderson [29]. We use the same numbers, which are listed in Table A.1 in Appendix A.1. To use the method of Wahr et al. [112] we make three assumptions:

- We assume the Earth to be spherically symmetrical, i.e., its material properties are only a function of the distance w.r.t. the center of the Earth.
- We assume the Earth's mantle and lithosphere to behave viscoelastically, as governed by Maxwell's viscoelastic constitutive law.
- We assume the displacements of the Earth's surface to be small, which means we may use a linear relationship between strain and displacement.

We expand the spherical harmonic coefficients into a spatial map of gravitational acceleration or EWH at the surface of our spherical reference Earth via [37]:

$$f(\theta,\lambda) = \sum_{l=0}^{\infty}\sum_{m=0}^{l} P_{l,m}(\sin\theta)\left(C_{l,m}\cos m\lambda + S_{l,m}\sin m\lambda\right) \tag{2.3}$$

where $f(\theta,\lambda)$ is a gravity functional of which the unit is defined by the unit of the Stokes coefficients. $P_{l,m}(x)$ denotes the *associated Legendre functions*, for which the reader is referred to Appendix A.2. The geopotential solutions that we use employ normalization conventions as defined in IERS 2010 [62]. These conventions are given in Appendix A.3. The main text of the report always refers to normalized Stokes coefficients or Legendre functions, therefore we leave out the common bar notation (e.g., $\bar{C}_{l,m}$, $\bar{P}_{l,m}$, etc.).

Most figures of a spatial map of (information derived from) a gravity field solution in this study are computed using one or more of the equations in this section. All other verification plots can therefore simultaneously verify our implementation of the relations in this section, and therefore we will not do so in this section separately.

## 2.3. Field-derived quantities

For a quantitative, concise comparison of (gravity) fields, spatial maps of expanded Stokes coefficients are sometimes an unpractical tool. For this reason, we will define a number of quantities that allow us to assess various gravity fields more purposefully.

### 2.3.1. Degree root-mean-square

Similar to how the root-mean-square (RMS) is a measure of the average power of an alternating current, this metric can be used to quantify the *power* of a particular degree of a spectral decomposition of a field. Here, it is important to know that the degree of a particular Stokes coefficient correlates with the spatial artefacts it describes in the spatial domain. In practice, this means two things. First, the maximum degree $l_{max}$ of a spectral decomposition governs its

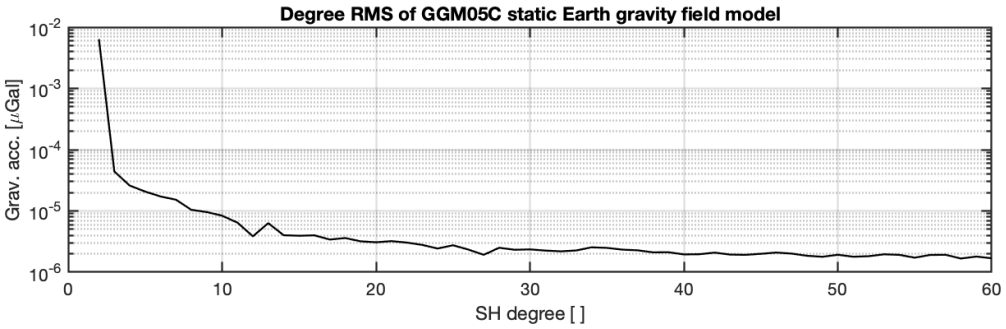



Figure 2.1: Degree RMS of the GGM05C static Earth gravity model of Ries et al. [70].

spatial resolution ($\approx \frac{20000\,\text{km}}{l_{max}}$), and second, the correlation between the degree RMS (dRMS) and the degree number tends to be negative for real geophysical signals. This makes sense, as gravitational potential is positively correlated to mass, which is in general positively correlated to size (of course, exceptions apply). We define the degree RMS of degree $l$ as the RMS of all orders in degree $l$, i.e.:

$$dRMS_l = \sqrt{\frac{\sum_{m=0}^{l}\left(C_{l,m}^2 + S_{l,m}^2\right)}{2\cdot(l+1)-1}} \quad \text{for} \quad l \geq 1 \tag{2.4}$$

As an example, Figure 2.1 shows the dRMS of the GGM05C static Earth gravity field model of Ries et al. [70]. Note how, generally speaking, the dRMS and degree number are negatively correlated. The static part of the gravity field dominates any gravity observations that we use (see also Table 1.1). To focus our analysis on the time-variable part of the epochs, we subtract a static gravity field model such as GGM05C before computing a particular field-derived quantity. We will show this for GRACE and Swarm in Section 2.4.

### 2.3.2. Degree correlation coefficient

Comparing the dRMS of two potential solutions can infer some conclusions about their differences in signal spread across the various degrees, but not about how this difference is spread across the various orders in each degree. In Tapley et al. [99, supporting online material] we find a more effective way of comparing two fields by using the Pearson correlation coefficient of each particular degree (from hereon: degree correlation coefficient). This coefficient is defined as the fraction of the shared variance within a degree (i.e., the covariance) and the root product of the individual variances of two sets of Stokes coefficients. The degree correlation coefficient is a value between 1 and -1, where 1 means perfect correlation between the two fields, -1 means perfect anti-correlation, and 0 means no correlation. Tapley et al. [99] use this method for validation of early GRACE-derived fields w.r.t. a hydrology model and Bezděk et al. [10] use it to validate Swarm-derived fields w.r.t. GRACE [10]. We compute the degree correlation coefficient $r_l$ of degree $l$ between field $A$ and $B$ via:

$$r_l^{(A,B)} = \frac{1}{\sigma_l^{(A)}\sigma_l^{(B)}} \sum_{m=0}^{l}\left(C_{l,m}^{(A)}C_{l,m}^{(B)} + S_{l,m}^{(A)}S_{l,m}^{(B)}\right) \quad \text{for} \quad l \geq 1 \tag{2.5}$$

where $\sigma_l$ is the square root of the degree variance of degree $l$, that is defined as:

$$\sigma_l^2 = \sum_{m=0}^{l}\left(C_{l,m}^2 + S_{l,m}^2\right) \quad \text{for} \quad l \geq 1 \tag{2.6}$$

Martinec [48] gives a method to assign a confidence interval to the degree correlation values. This method assumes the observations uncertainties come from a normal distribution. We will test this assumption in light of the available GRACE and Swarm data in Section 2.6.1 and Section 4.2. The critical correlation coefficient for degree $l$, $r_l^{(crit)}$, is given by:

$$r_l^{(crit)} = \frac{T(\alpha,\nu)}{\nu + T(\alpha,\nu)^2} \tag{2.7}$$

where $T(\alpha,\nu)$ denotes the Student's T statistic for $\nu$ degrees of freedom and confidence level $\alpha$. We use $\alpha = 5\%$ and $\nu = 2l$. We can find $T(\alpha,\nu)$ via the inverse cumulative distribution function of a Student's T distribution. This function is given in Appendix A.6. Evaluating this function is done conveniently via MATLAB's `tinv` function. The critical correlation coefficient for $\alpha = 5\%$ can be interpreted (and computationally verified, which we do not show) as the correlation coefficient that has a 5% probability of occurring if the true correlation coefficient of the data at hand is 0.

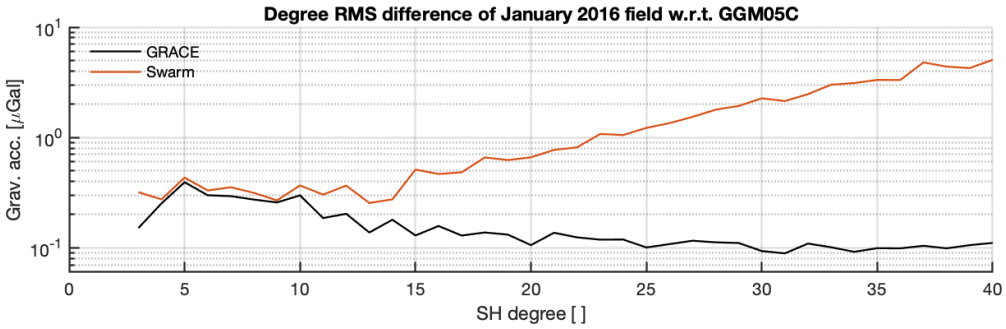



Figure 2.2: Degree RMS of the difference between the January 2016 fields of GRACE CSR RL06 (black line) and Swarm (red line) and GGM05C.

The degree correlation coefficient is a versatile tool, as it provides an opportunity to quantitatively assess not only the difference between individual gravity fields [24], but also the difference between quantities derived from those fields. For example, aforementioned GRACE validation study of Tapley et al. [100] assesses the annual periodic component w.r.t. a hydrology model, and Sasgen et al. [79] compare gravity trends from a GIA model to gravity trends derived from GRACE data.

## 2.4. Noise reduction

Section 1.2 explained how the source data of the GRACE and Swarm gravity fields are multiple orders of magnitude apart in terms of accuracy, and how this translates in an increase in high-degree noise in the Swarm fields w.r.t. the GRACE fields. To illustrate this, we compute the degree RMS difference of the GRACE and Swarm fields w.r.t. a static background gravity model. We use GGM05C [70] again. Note that this is a different background model than is used in the computation of the Swarm gravity fields. For this analysis the influence of using two different background models is negligible.

The results for an arbitrarily chosen month are in Figure 2.2. The difference between GRACE resp. Swarm and the static models is caused by either the time-varying gravity signals or noise. We postulate that the negative correlation w.r.t. the SH degree also holds for time-variable gravity signals, which suggest GRACE is observing (mostly) real signal. We see that for degrees up to 10, Swarm follows GRACE reasonably well. However, where the GRACE degree RMS difference continues to be negatively correlated to the SH degree after degree ca. 13, we see a positive correlation between the SH degree and the dRMS of the Swarm-derived gravity field. This suggests noise in the Swarm-derived fields is dominant in the degrees past that point.

### 2.4.1. Gaussian filtering

A very widely used method to reduce the influence from high-degree noise is Gaussian filtering, introduced to GRACE data by Wahr et al. [112]. The gravity observations are weighted according to a Gaussian function, which can be applied in either the spatial or the spectral domain. From the convolution theorem we know that the Fourier transform of a convolution of two signals equals the point-wise product of their Fourier transforms, and vice versa. Having said that, applying the Gaussian filter in the spectral domain is in general computationally more efficient, and therefore our preferred method. We multiply each Stokes coefficient observation with a Gaussian scale factor $s_g$, which is a function of the coefficient's degree $l$:

$$s_g(l) = \exp\left[-4\ln(2)\left(\frac{l}{2l_{fwhm}}\right)^2\right] \tag{2.8}$$

where $l_{fwhm}$ is a constant defining the Gaussian function's full width at half maximum, expressed as a Stokes coefficient degree. It is more common to express this parameter in units of length, i.e., as the radius $r_g$ of the Gaussian function in the spatial domain. We define:

$$r_g = \pi\frac{R}{2l_{fwhm}} \tag{2.9}$$

where $R$ is the radius of the Earth. As Gaussian filtering is a very common technique, it is part of various other spatial plots that we show later in this report to compare to literature for verification of other methods. This means the correctness of our implementation of Gaussian filtering can be derived from those plots.

Figure 2.3 shows the effect of Gaussian filtering on the degree RMS difference w.r.t. GGM05C of the January 2016 Swarm gravity field. Note how the reduction in degree RMS scales exponentially while the filter radius scales linearly. Also, note

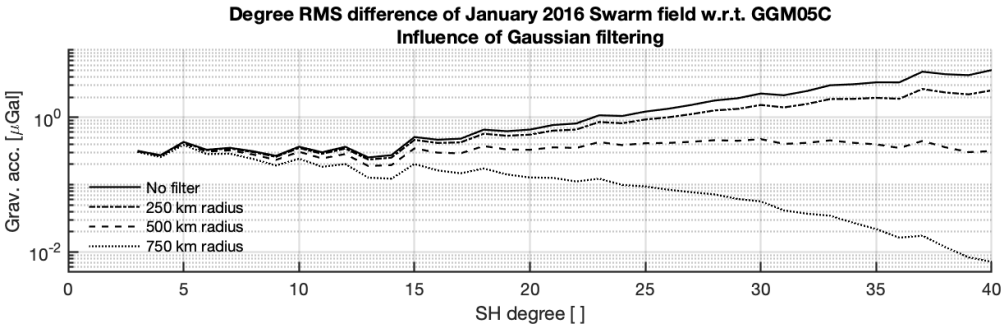



Figure 2.3: Effect of Gaussian filtering on the degree RMS difference w.r.t. GGM05C of the Swarm fields. Filter radii were chosen arbitrarily, only to give the reader an impression of the general effect of this filter type.

how the higher degrees are affected the most, but the lower degrees, which we postulate are mostly real signal, are also suppressed to some extent. This *signal attenuation* is a key issue when applying the Gaussian filter to a geopotential solution in a geophysical study. The configurations shown here are not necessarily the configurations that we will use in our research and are only shown here to give the reader a general idea of how Gaussian filtering affects the spectral representation of a (gravity) field.

Over the years, a number of intricate filters have been designed for the GRACE data. Many of these filters are non-isotropic, i.e., they are a function of more than just the degree number, to account for the correlation between the GRACE errors and using that correlation to optimize the filtering process [41, 96]. This correlation of the errors is apparent from the striping pattern that is seen when expanding GRACE gravity fields, which is caused by the anisotropic sensitivity of the GRACE measurement system [101]. As this sensitivity is different for the GPS-derived Swarm gravity fields, it is unlikely that the filters that are designed specifically for the GRACE correlated errors will yield good results when applied to Swarm data. For that reason, we will not look into those filters here. Note that this does not mean that anisotropic filtering in general cannot be applied to GPS-derived gravity fields. We will elaborate on this in Chapter 4 when we inspect the Swarm errors.

### 2.4.2. Truncation

Instead of applying a Gaussian filter, we can also choose to discard all coefficients past a particular degree. This is called truncation. When a (gravity) field is truncated, the degree RMS plots of Figure 2.2 and 2.4 would simply stop after a certain maximum degree. As this is trivial, we will not show these graphs. We can define the truncation scale factor $s_t$, which is again a function of the coefficient's degree $l$, as:

$$s_t(l) = \begin{cases} 1, & \text{if } l \leq l_{trunc} \\ 0, & \text{if } l > l_{trunc} \end{cases} \tag{2.10}$$

where $l_{trunc}$ is a constant equal to the degree of the shortest wavelength harmonic to not be truncated.

Truncation has the advantage that it leaves the often less noisy lower degree coefficients untouched. On the contrary, truncation of a gravity field does reduce its spatial resolution, as shorter wavelength signals can only be represented by higher degrees. This means small-scale artefacts are no longer discernible from the expanded Stokes coefficients. Decreasing the number of harmonics in a Fourier series also decreases the series capabilities of representing strong signal gradients and increases unwanted undulations in the spatial expansion. We will cover this in Section 3.3.

## 2.5. Signal decomposition via least squares regression

To isolate the GIA signal from the satellite-derived time-varying gravity fields, we start with extracting the secular component (the gravity trend) from the time series of gravity observations. This is commonly done via least squares regression [100]. Principal component analysis (PCA; also referred to as *empirical orthogonal functions*) [65] is sometimes used as a complimentary technique [106]. We decide to include only least squares regression for two reasons. First, PCA extracts the most dominant components in the data - not necessarily the linear one. Second, most authors agree that there are no substantial differences in the quality of the results of either approach when it comes to GIA study [107].

The least squares regression can be done either in the spectral domain, where a model is fitted to time series of the Stokes coefficients, or in the spatial domain, where a model is fitted to time series of the gravitational signal at each particular grid point. The least squares estimates of the model parameters can in turn be converted back and forward too, which

means there is no difference in results from either approach. It should be noted that working in the spectral domain tends to be computationally heavier, because the number of grid points increases faster than the number of Stokes coefficients for an increasing spatial resolution.

### 2.5.1. Linear least squares

Least squares regression is based on minimizing the sum of squares of the residuals after subtracting a model, $f(t)$, from the observations $\mathbf{y}$. To do so, we estimate weights for the various components of the model, which consists of a polynomial of degree $(n-1)$ and $m$ periodic components. We define:

$$f(t) = \sum_{i=1}^{n} \left[ A_i \cdot t^{(i-1)} \right] + \sum_{j=1}^{m} \left[ B_j \sin\left(\frac{2\pi}{T_j} t\right) + C_j \cos\left(\frac{2\pi}{T_j} t\right) \right] + \varepsilon \tag{2.11}$$

where $t$ is the time, $A_i$, $B_j$, and $C_j$ are the weights of polynomial component $i$ and periodic component $j$, and $T_j$ is the period of periodic component $j$. $\varepsilon$ are the residuals, which contain noise and unmodeled effects. We start by sequencing the weights of Equation 2.11 to form the parameter vector $\mathbf{x}$, so that:

$$\mathbf{x} = \begin{bmatrix} A_1 & A_2 & \dots & A_n & B_1 & C_1 & B_2 & C_2 & \dots & B_m & C_m \end{bmatrix}^T \tag{2.12}$$

Our observation vector, $\mathbf{y}$, with $k$ observations is defined as follows:

$$\mathbf{y} = \begin{bmatrix} y(t_1) & y(t_2) & \dots & y(t_k) \end{bmatrix}^T \tag{2.13}$$

Finally, we construct our design matrix $A$, which is based on our model Equation 2.11 and has time variable $t_k$ that corresponds to each observation $y_k$, as follows:

$$A = \begin{bmatrix} (t_1)^{1-1} & (t_1)^{2-1} & \dots & (t_1)^{n-1} & \sin\left(\frac{2\pi}{T_{j=1}} t_1\right) & \cos\left(\frac{2\pi}{T_{j=1}} t_1\right) & \dots & \sin\left(\frac{2\pi}{T_{j=m}} t_1\right) & \cos\left(\frac{2\pi}{T_{j=m}} t_1\right) \\ (t_2)^{1-1} & (t_2)^{2-1} & \dots & (t_2)^{n-1} & \sin\left(\frac{2\pi}{T_{j=1}} t_2\right) & \cos\left(\frac{2\pi}{T_{j=1}} t_2\right) & \dots & \sin\left(\frac{2\pi}{T_{j=m}} t_2\right) & \cos\left(\frac{2\pi}{T_{j=m}} t_2\right) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ (t_k)^{1-1} & (t_k)^{2-1} & \dots & (t_k)^{n-1} & \sin\left(\frac{2\pi}{T_{j=1}} t_k\right) & \cos\left(\frac{2\pi}{T_{j=1}} t_k\right) & \dots & \sin\left(\frac{2\pi}{T_{j=m}} t_k\right) & \cos\left(\frac{2\pi}{T_{j=m}} t_k\right) \end{bmatrix} \tag{2.14}$$

We can now state our least squares problem as:

$$\mathbf{x} = A\mathbf{y} + \varepsilon \tag{2.15}$$

where $\varepsilon$ is the residual vector. We seek to minimize our cost function $J$, which we rewrite using Equation 2.15 as:

$$J = \varepsilon^T \varepsilon \tag{2.16}$$

$$J = \mathbf{y}^T (\mathbf{y} - A\mathbf{x}) - \mathbf{x}^T (\mathbf{y} - A\mathbf{x}) \tag{2.17}$$

We minimize $J$ by seeking an $\hat{\mathbf{x}}$ that minimizes the second term on the right side of Equation 2.17, since the first term on the right side cannot be minimized when $A\hat{\mathbf{x}}$ approximates $\mathbf{y}$. We exclude the trivial solution $\hat{\mathbf{x}} = 0$ and define the least squares estimate of our parameter vector as:

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{y} \tag{2.18}$$

Finally, we define the post fit residuals

$$\hat{\varepsilon} = \mathbf{y} - A\hat{\mathbf{x}} \tag{2.19}$$

The residuals constitute all signal from the observations that is not captured by the model we defined in 2.11. Generally, this is either due to noise in the observations, or due to an inadequate model. For that reason, the residuals can be used in assessing the observation uncertainty, as well as the model fit. We will use the residuals to compute the standard deviation of the satellite-derived gravity observations in Section 2.6.1.

The amplitudes of the cosine and sine parts (i.e., $B_j$ and $C_j$) are used to compute the amplitude $U_j$ and phase $\omega_j$ of

periodic component $j$ via:

$$U_j = \sqrt{B_j^2 + C_j^2} \tag{2.20}$$

$$\omega_j = 0 \quad \text{if } B_j = 0 \wedge C_j = 0 \tag{2.21}$$

$$\omega_j = 0 \quad \text{if } B_j = 0 \wedge C_j > 0 \tag{2.22}$$

$$\omega_j = \frac{T_j}{2} \quad \text{if } B_j = 0 \wedge C_j < 0 \tag{2.23}$$

$$\omega_j = \frac{T_j}{2\pi} \cdot \arcsin\frac{B_j}{U_j} \quad \text{if } B_j \neq 0 \wedge C_j > 0 \tag{2.24}$$

$$\omega_j = \frac{T_j}{2\pi}\left(\pi - \arcsin\frac{B_j}{U_j}\right) \quad \text{if } B_j \neq 0 \wedge C_j < 0 \tag{2.25}$$

### 2.5.2. Weighted least squares

In some collections of observations, the uncertainty of the observations changes with time. This is known as observation heteroscedasticity (from Greek *hetero*; different, and *skedasis*; dispersion). If left unaccounted for, heteroscedasticity can deteriorate the parameter estimates of the least squares regression. By weighing the observations w.r.t. their uncertainty, this risk can be (partially) mitigated [118]. We define the covariance matrix of the observations $P_{yy}$ as follows:

$$P_{yy} = \begin{bmatrix} \sigma_{y1}^2 & & & \\ & \sigma_{y2}^2 & & \\ & & \ddots & \\ & & & \sigma_{yk}^2 \end{bmatrix} \tag{2.26}$$

The diagonal of this matrix contains the variances (i.e., squared standard deviations) of observations $y_1$ to $y_k$. By definition, the off-diagonal entries of a covariance matrix contain the covariance values of the observations. However, as covariance parameters are much harder to determine and Wahr et al. [114] suggest their impact on the propagated uncertainty of the expanded Stokes coefficients is negligible, we have decided to neglect the observation covariances in this study. We cover our method to populate the diagonal of $P_{yy}$ in Section 2.6.1. We define our weighted cost function as:

$$J = \varepsilon^T P_{yy}^{-1} \varepsilon \tag{2.27}$$

Similarly to Equation 2.17, we substitute Equation 2.15 and minimize $J$ by finding $\hat{\mathbf{x}}$ (excluding the trivial solution $\hat{\mathbf{x}} = 0$) via:

$$\hat{\mathbf{x}} = (A^T P_{yy}^{-1} A)^{-1} A^T P_{yy}^{-1} \mathbf{y} \tag{2.28}$$

Finally, we compute the covariance matrix of the parameter vector $P_{\hat{x}\hat{x}}$, which contains the variances of the elements in the parameter vector, via error propagation:

$$P_{\hat{x}\hat{x}} = (A^T P_{yy}^{-1} A)^{-1} \tag{2.29}$$

### 2.5.3. Comparing regressions via the F-test

Equation 2.11 showed the model that we will fit to the time series of Stokes coefficients. The second summation of the right side of the equation allows us to include periodic components in this model. Commonly, authors include an annual and a semi-annual component ($T$ = 1 y, 0.5 y, respectively), which makes sense in light of the periodicity of most climate cycles. Ray et al. [68] suggest that the $S_2$, $K_2$, and $K_1$ tides alias into the GRACE solutions. Steffen et al. [94] warned that not including such a periodic alias in the regression model could affect the trend estimate when the data set at hands spans a non-integer number of cycles of the periodic alias. Other authors trying to isolate the GIA-induced gravity trend from GRACE data assume this influence is small and neglect it in the regression model [79].

Even though [87] showed the amplitude of the aliases differs per Stokes coefficient, we aim to use the same regression model for all coefficients for the sake of computational simplicity. Before we define our definitive regression model, we will assess the reasonableness of aforementioned assumption by testing the Stokes coefficient time series for statistically significant periodic components with a period corresponding to aforementioned tidal aliases. We apply this test per Stokes coefficient time series and will include the particular periodic components into our regression model for all coefficients if this leads to a significant improvement in fit in more than 50% of the coefficients in both GRACE and Swarm. We take period values that are also used to de-alias the CSR RL06m mascon solutions, which equal 161, 1400, and 2800 days for respectively the $S_2$, $K_2$, and $K_1$ aliases[7]. In the Swarm case, comparing these periods to the length of the data (5.25 y = 1916 d) warns us to be critical of the results for especially the $K_1$ alias.

[7] Dr. Himanshu Save, personal communication.

We test for significant periodic components via an F-test. This test compares a restricted model $A$ to a full model $B$. Model $A$ must be nested in model $B$, i.e., model $B$ contains all terms of model $A$ and at least one term more. This test requires the post-fit residuals to be approximately Gaussian-distributed. We will examine the reasonableness of this assumption for the GRACE and the Swarm data in Section 4.2. We define the hypotheses:

- $H_{F0}$: the coefficients in $B$ that are not in $A$ are all not statistically significant.
- $H_{Fa}$: at least one coefficient in $B$ that is not in $A$ has a significant, non-zero value.

In literature $H_{Fa}$ is often interpreted as model B having a *statistically improved fit* w.r.t. model $A$. We compute the F-statistic $\Phi_{A,B}$ by dividing the variance reduction after going from model $A$ to model $B$ by the variance of model $A$, where all variances are standardized w.r.t. the degrees of freedom in the particular model. We define [57]:

$$\Phi_{A,B} = \frac{\left(\frac{\chi^2_A - \chi^2_B}{\nu_A - \nu_B}\right)}{\frac{\chi^2_B}{\nu_B}} \tag{2.30}$$

where $\chi^2$ is the post-fit unit weight residual sums of squares (also referred to as *chi-squared statistic*) [20], $\nu$ is degrees of freedom each model, defined as the number of observations minus the number of estimated parameters. We define $\chi^2$ as:

$$\chi^2 = \sum_{i=1}^{i=k} \left(\frac{y_i - \hat{y}_i}{\sigma_i}\right)^2 \tag{2.31}$$

where $y$, $\hat{y}$, and $\sigma$ are the observed value, the modeled value, and the standard deviation of the observed value, respectively. If $H_{F0}$ is true, F-test statistic $\Phi_{A,B}$ will follow a F-distribution with $\nu_A$ degrees of freedom in the numerator and $\nu_B$ degrees of freedom in the denominator [57]. This means we can calculate the probability $p_\Phi$ of observing a $\Phi$ within the interval $[0, \Phi_{A,B}]$ when $H_{F0}$ is true via the cumulative distribution function of the F-distribution. Going back to $A$ and $B$, this translates into the probability of observing a reduction in residuals (standardized to $\nu_2 - \nu_1$) equal to what we compute from comparing $A$ and $B$, or smaller. However, we are more interested in the complementary event: observing a certain or *greater* $\chi^2$ reduction if $H_{F0}$ is true, of which the probability is $1 - p_\Phi$. This would imply that the $\chi^2$ reduction is due to random chance, i.e., that the additional components in model $B$ are zero or not significant.

We predetermine our significance level at 5%, which means we reject $H_{F0}$ when $p_\Phi > 0.95$ and deem the improvement of the fit of the particular model $B$ w.r.t. model $A$ statistically significant. Note that we never *prove* that a component is significant. We compute the probability of observing the current statistic under the null hypothesis, and if this probability is below our confidence level we accept the alternative hypothesis (i.e., *reductio ad absurdum*) [25]. Our choice of significance level therefore means we have a 5% chance of deeming a model's particular increase in reduction of residuals significant, even though it happened due to chance (i.e., we reject $H_{F0}$ even though it is true - a *Type I error*).

In Appendix B we present a case study of an experiment where we apply inferential statistical analysis in an attempt to extract geophysical signal from the noisy high-degree Stokes coefficients. Instead of including all coefficients in Equation 2.3 when we expand a particular model parameter's least squares estimate, we include only the coefficients for which that estimate is statistically significant in that particular Stokes coefficient time series [20, 78, 79]. Readers with an interest in statistics might want to visit this part of the report.

## 2.6. Observation uncertainties

Section 2.5.2 explained how we weigh the observations w.r.t. they uncertainty in the regression process. In Equation 2.29 we saw how to propagate the observation uncertainties into the trend error. The following section covers the method we use to first find those observation errors, and then assess their probability distribution. The errors we estimate are the sum of measurement errors, processing errors, and errors in the geophysical models used to de-alias the satellite measurements prior to constructing gravity fields [113].

### 2.6.1. Magnitude and spatial distribution

Many GRACE data products include the standard deviation per Stokes coefficient per epoch. These error estimates are referred to as the calibrated errors, and are equal to the diagonal components of the covariance matrix, scaled by the GRACE processing centers to match certain characteristics in the gravity fields [114]. Applying any filtering to the gravity

fields yields these uncertainty values invalid. Additionally, some authors suggest that the calibrated errors underestimate the true error [e.g. 105]. Taking all this into account, we decide to recompute the uncertainty of the GRACE observations.

Calibrated errors for the combined Swarm fields are not provided by de Teixeira da Encarnação et al. [23]. Computing these from the individual Swarm gravity fields would be a cumbersome task as the combination is done using weights derived via variance component estimation (VCE). Firstly, this would require insight into the VCE methodology, and secondly, we would have to take into account that the errors of the individual fields are not independent. For these reasons, we also compute the Swarm observation uncertainties from the observations.

Wahr et al. [113] first suggested using the RMS of the post-fit residuals, the RMSE, as the error estimate of a particular Stokes coefficient. The RMSE is defined from Equation 2.19 as:

$$RMSE_i = \sqrt{\frac{\sum_{i=0}^{i=k} \left(\mathbf{y}_i - A\hat{\mathbf{x}}_i\right)^2}{k}} = \sqrt{\frac{\hat{\varepsilon}^T \hat{\varepsilon}}{k}} \tag{2.32}$$

where all parameters are as defined before.

To retain the dependence of the error on the Stokes coefficient order, Wahr et al. [114] opt to take the calibrated errors and scale them per degree with the degree RMS of the residuals. While we agree that this likely yields more realistic results, we decline to use this method as we do not have calibrated errors for our Swarm fields. We will therefore approximate the Stokes coefficient error as simply the RMSE of that particular Stokes coefficient time series.

The method of Wahr et al. [113] is based on the assumption that a certain model captures all signal from the observed mass transport processes, and that the residuals are thus all noise. Initially, Wahr et al. [113] used only a constant and an annual component. Over time, other authors extended this method to account for gravity trends [111] or variation of the amplitude of the yearly cycle [106]. We further expand the method of Wahr et al. [114] by subtracting a moving mean component of 1 year, taking into account the variable spacing of the GRACE epochs, to account for long-term variability of the Stokes coefficient. We choose a 1-year window because this would leave the annual- and semi-annual components intact. We base this approach on Velicogna and Wahr [109], who use the RMS of the difference between the observations and *"temporally smoothed version of those coefficients"* (no more information is given by the authors). After subtracting the moving mean, we subtract an annual and semi-annual cycle computed from a 2-year moving window, as in Van der Wal et al. [106]. Finally, we compute the RMS of the remaining residuals (i.e., the RMSE). The RMSE results need to be scaled by a constant to account for the RMS reduction due to subtracting the fitted model. We determine this constant by computing the RMS reduction after fitting our model to an arbitrarily selected high number ($n > 2000$) of randomly sequenced time series of gravity observations of a length equal to that of the GRACE or Swarm data products [106, 114]. We summarize this entire method via:

$$RMSE_i = \sqrt{\frac{1}{n}\sum_{t=0}^{k} \left[\mathbf{y}_i - MA_{1\,yr}(\mathbf{y}_i) - \Theta_i\right]^2} \cdot s_{RMS} \tag{2.33}$$

where $y_i$ is the vector containing $k$ observations of Stokes coefficient $i$ spanning time $\mathbf{t}^{(y)} = [1, 2, 3, ..., k]$, next, $MA_{1\,yr}(y_i)$ denotes the moving average of $y_i$ computed from a 1-year window, $s_{RMS}$ is the scaling factor to account for RMS reduction, and $\Theta_i$ denotes the vector containing the sum of all modeled periodic components of $\mathbf{y}_i$, as estimated from a 2-year moving window. We define $\Theta_i$ as:

$$\Theta_i = \sum_{j=1}^{m} \sum_{\gamma=0}^{\gamma_k} \left[\hat{\mathbf{x}}_{U_j}^{(\gamma)} \sin \frac{2\pi}{T_j}\left(\mathbf{t}_\gamma - \hat{\mathbf{x}}_{\omega_j}^{(\gamma)}\right)\right] \tag{2.34}$$

where $\sum_{j=1}^{m}$ denotes the summation of $m$ periodic components with respective periods $T_1$ to $T_m$ (see Equation 2.11), $\gamma_k$ is the number of consecutive 2-year windows that fits in the time span of the observations (rounded up to the nearest integer), and $\mathbf{t}_\gamma$ is a subset of $\mathbf{t}^{(y)}$, the time vector of the observations. We define:

$$\gamma_k = \left\lceil \frac{t_k^{(y)} - t_0^{(y)}}{[2\text{ y}]} \right\rceil \tag{2.35}$$

$$\mathbf{t}_\gamma = \left\{ t_i \subseteq \mathbf{t}^{(y)} \mid \left(t_0^{(y)} + \gamma \cdot [2\text{ y}]\right) \leq t_i < \left(t_0^{(y)} + (\gamma+1) \cdot [2\text{ y}]\right) \right\} \quad \text{for } i = 1, 2, 3, ..., k \tag{2.36}$$

In the case of unevenly spaced epochs we do not interpolate but rather use the closest available epoch, taking care not to include epochs twice. Finally, $\hat{\mathbf{x}}_{U_j}^{(\gamma)}$ and $\hat{\mathbf{x}}_{\omega_j}^{(\gamma)}$ are the least squares estimates of the amplitude and phase of periodic component $j$ (see Equation 2.20 to 2.25), computed from only the epochs in temporal subset $\mathbf{t}_\gamma$.

Figure 2.4 compares our method to a method subtracting a bias, trend, and an annual component. We use the $S_{5,1}$ coefficient from the CSR GRACE RL06 fields as an example, as it has a long-term variability that we postulate is real signal.

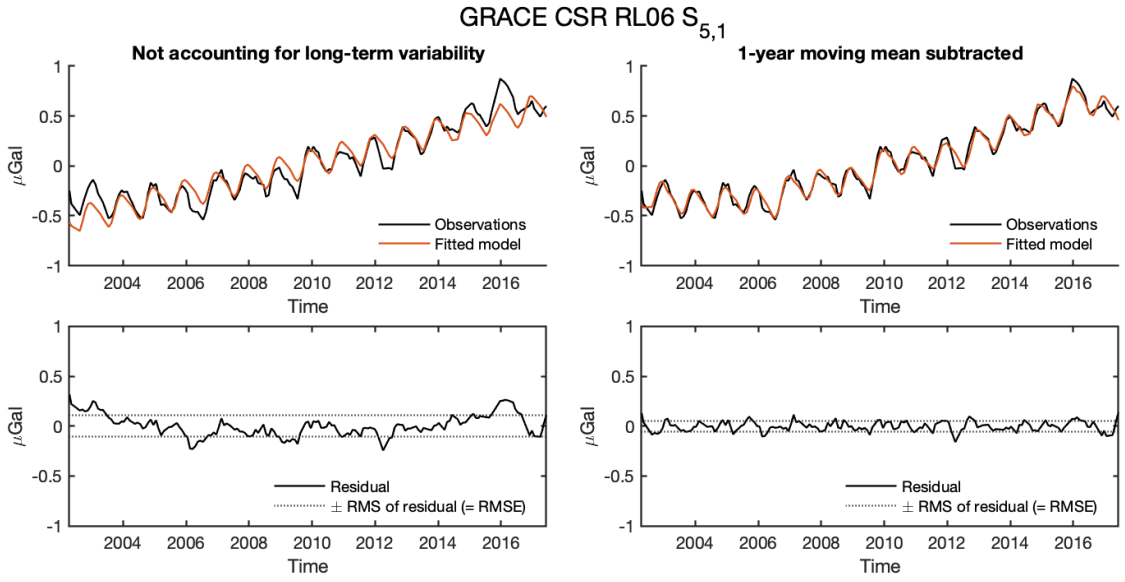



Figure 2.4: Left: time series of GRACE CSR RL06 observations of an exemplary Stokes coefficient ($S_{5,1}$), fitted model consists of a bias, trend and annual cycle, post-fit residuals, and RMSE estimate. Left: same but using a model consisting of a 1-year moving mean and an annual cycle from a 2-year moving window. Note how the subtracted model on the left does not capture all signal, and how this propagates into a higher RMSE result.

Note how the model on the left clearly captures less signal than the model on the right. This translates into an RMSE of 0.109 $\mu$Gal (left) versus 0.0546 $\mu$Gal (right).

The window length of 1 year for our moving average and (semi-)annual periodic component means that $S_2$ tidal alias (period 161 d) that exists in the GRACE solutions may affect our error estimate [94]. In Section 4.1 we will investigate how many GRACE and Swarm Stokes coefficient time series contain a significant $S_2$ (and $K_2$, $K_1$) alias, and decide if a correction to the error model is required.

We can expand the RMSE of the Stokes coefficients into the uncertainty of the gravity or mass estimates (see Section 2.2) via [114]:

$$\sigma(\theta,\lambda) = \sqrt{\sum_{l=0}^{\infty}\sum_{m=0}^{l} P^2_{l,m}(\sin\theta)\left(\delta C^2_{l,m}\cos^2 m\lambda + \delta S^2_{l,m}\sin^2 m\lambda\right)} \tag{2.37}$$

where $\delta C_{l,m}$ and $\delta S_{l,m}$ are the RMSE of the particular Stokes coefficients. To take into account the time-dependency of the Stokes coefficient errors, we follow Wahr et al. [114]. We compute a map of the gravity errors by inserting the residuals of a particular epoch in Equation 2.37 and take the global, latitude-weighted mean of that map. As we are using a spherical reference Earth, this latitude-weighing is done by simply multiplying the result with the cosine of the latitude of the center of the grid point. Next, we take our error estimate that we computed via the RMSE of the whole time series, and scale it w.r.t the monthly mean error. We verify our computation routines by recomputing Feng [32, Figure 9b], who uses the method of Wahr et al. [114], only with newer data (CSR RL05). As our method is largely based on the same idea (using the residuals), but uses a model that aims to subtract more geophysical signal before computing the RMS. We thus expect that the spatial error distribution is the same, but the overall magnitude is lower. A visual comparison with the results of Feng [32] should verify this.

The left panel of Figure 2.5 shows the spatial distribution of the computed errors of the November 2003 epoch of the CSR GRACE RL05 fields, computed via the method of Wahr et al. [114]. We invite the reader to compare this to Feng [32, Figure 9b], as the virtually identical results verify our correct implementation of the error computation method of Wahr et al. [114]. The right panel of Figure 2.5 shows the errors for the same GRACE epoch, as computed via our method. Note how the spatial distribution is the same, while the overall amplitude is slightly reduced. The latitude-weighted RMS of the left map is 5.053 cm EWH, and for the right map 4.877 cm EWH. This slight reduction in amplitude is exactly what we expected, and thus we conclude that we have also successfully verified the implementation of our own observation uncertainty estimation method.

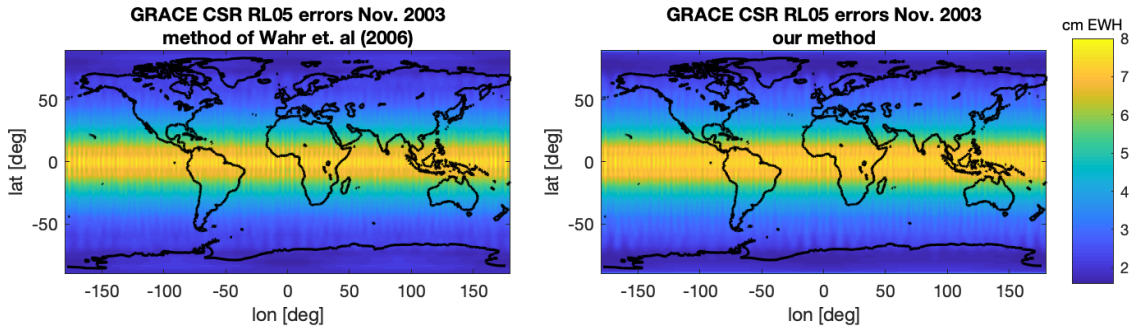



Figure 2.5: Verification of our error estimation method. Left shows the method of Wahr et al. [114], as shown in Figure 9b in Feng [32]. Shown here that our error estimation method yields the same spatial distribution, but a slightly lower amplitude. The latitude-weighted RMS equals 5.053 cm EWH (left) vs. 4.877 cm EWH (right).

### 2.6.2. Probability distribution

To use our error estimates to assign confidence levels to the mass change results computed from the Stokes coefficients, we need to investigate their probability distribution. Also, normally distributed residuals are a prerequisite for statistical tests such as the F-test in Section 2.5.3. Note that in the latter case we are referring to the post-fit residuals of the least squares regression (Equation 2.19), not the RMSE mentioned in the last section. This section describes the computation of the probability distribution of the RMSE but the same can be applied to the post-fit residuals.

We use a histogram of the normalized residuals to assess their distribution. To compute this histogram, we first use the method from the last section to compute the post-fit residuals of each Stokes coefficient. For each time step, we convert the Stokes coefficient residuals into a map of gravity variations via Equation 2.3. We normalize the gravity variations by dividing each map by a map of the expected error, that we find by expanding the RMSE (Equation 2.33) of each coefficient time series in Equation 2.37. Finally, we plot a histogram of the latitude-weighted, normalized residuals and compute the best-fitting normal distribution $N(\mu_{\bar{\varepsilon}}, \sigma^2_{\bar{\varepsilon}})$ from the mean $\mu_{\bar{\varepsilon}}$ and variance $\sigma^2_{\bar{\varepsilon}}$ of the normalized residuals [25]:

$$\mu_{\bar{\varepsilon}} = E[\bar{\varepsilon}] = \frac{1}{n}\sum_{i=0}^{n} \bar{\varepsilon}_i \tag{2.38}$$

$$\sigma^2_{\bar{\varepsilon}} = E[(\bar{\varepsilon} - \mu_{\bar{\varepsilon}})^2] = \frac{1}{n}\sum_{i=0}^{n} (\bar{\varepsilon}_i - \mu_{\bar{\varepsilon}})^2 \tag{2.39}$$

$$N(\mu_{\bar{\varepsilon}}, \sigma^2_{\bar{\varepsilon}}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\bar{\varepsilon} - \mu_{\bar{\varepsilon}}}{\sigma_{\bar{\varepsilon}}}\right)^2\right] \quad \text{for} \quad -\infty < \bar{\varepsilon} < \infty \tag{2.40}$$

where $E$ denotes the expected value operator.

To verify our computation routine we recompute the results of Wahr et al. [114, Figure 4], who apply aforementioned method to the first 22 GRACE months, smoothed with a 750 km radius Gaussian filter, to show their residuals follow a normal distribution. Aiming to use as much of the same data as possible, we use the first 22 months of the GRACE CSR RL01 Stokes coefficients. First, we scale the coefficients to EWH via Equation 2.2. We exclude the degree 1 and $C_{2,0}$ coefficients and apply a 750 km radius Gaussian filter. Figure 2.6 shows the histogram of the normalized residuals. Note how it follows the normal distribution well. This is in good accordance with the results of Wahr et al. [114, Figure 4]. We attribute the small discrepancies between our results and the results in Wahr et al. [114] to the fact that the authors use the formal errors scaled to match the RMSE, whereas we use the RMSE itself (see Section 2.6.1).

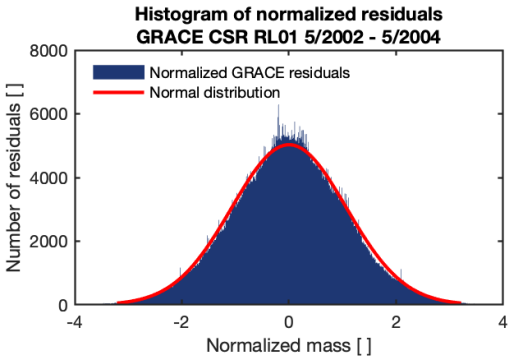



Figure 2.6: Histogram of normalized mass residuals of the first 22 epochs of the CSR RL01 GRACE Stokes coefficients and fitted normal distribution. Shown here to verify our residual distribution test routine by recomputing Wahr et al. [114, Figure 4]

# 3

# Isolating and assessing the GIA-induced gravity trend

One of the key limitations of the GRACE measurement principle, and of using gravimetry for geophysical study in general, is that there is no vertical resolution. The geopotential observed at a given location is the result of all mass in the vertical column encompassing that point. In our case, this means that the gravity trends that we compute will contain the GIA-induced gravity signal, but also various other signals layered on top of it. Common practice to isolate the GIA-induced trend component includes reducing the observed gravity trends with modeled gravity trends of the other signals. However, our research is slightly different from other GIA studies in the sense that we are trying to isolate a congruent GIA-induced gravity trend observation from two independent data sources, and not necessarily strive for the most accurate or *cleanest* GIA measurement. This means the time dependency of any of the diluting signals is more important than its absolute magnitude. Constant errors factor out once we subtract the GRACE and Swarm observations. This increases the potential impact of, for example, surface hydrology variations, but decreases the impact of approximations such as the spherical Earth mentioned in Section 2.1.5.

In the following chapter, Section 3.1 introduces a novel method to assess the variability of the gravity trend due to all non-linear influences combined. Section 3.2 presents our correction for inter-annual surface hydrology variations, Section 3.3 does this for interference from Greenland and Alaska mass change, and Section 3.4 does this for other long-wavelength leakage. At this point, we have covered all necessary methodology to answer research sub-question Q.2. Next, Section 3.5 shows the methods we use to quantitatively assess the discrepancy between the GRACE- and Swarm-derived observations of the GIA-induced gravity trend and answer research sub-question Q.3. Finally, Section 3.6 explains how we will validate our results using the recently released GRACE-FO geopotential solutions to answer research sub-question Q.4.

Definitions of acronyms are given in the List of Acronyms on page xiii. Symbol definitions and units, and values used for constants can be found in the List of Symbols on page xv.

## 3.1. Quantifying trend variability

Van der Wal et al. [106] compared maximum gravity rates in Laurentia computed from 29 temporal subsets of a 3-year length from the GRACE CSR RL04 coefficients. The authors found that the trend maxima increase from ca. 1.2 $\mu$Gal a$^{-1}$ (using data from October 2002 to August 2005) to 1.9 $\mu$Gal a$^{-1}$ (using data from April 2003 to March 2006). This showed that the observed gravity trend in the area was not constant through time, as one would expect from the nature of the GIA process, which suggests mass change signals of other geophysical processes are layered on top of the GIA-induced trend. In this study, we hypothesize that this variability of the gravity trend is a function of the spatial resolution of the data at hand and the length of the data from which the trend is computed (i.e., the inter-annual variations roughly cancel out over time).

To investigate this, we will repeat and expand the experiment of Van der Wal et al. [106]. First, we will repeat the experiment using the latitude-weighted mean of the grid points bounded by a >+1.0 $\mu$Gal a$^{-1}$ gravity trend in GRACE CSR RL06 ($l_{max}$ = 13). This region will be elaborated upon and shown in Section 3.5. We argue that using the mean trend is less susceptible to outliers than the maximum gravity trend, although we do acknowledge that this could provide an underestimation of the trend variability as the amplitudes of very local inter-annual hydrology variations are likely reduced by the averaging process. Also, we will experiment with different window lengths to test the hypothesis that inter-annual

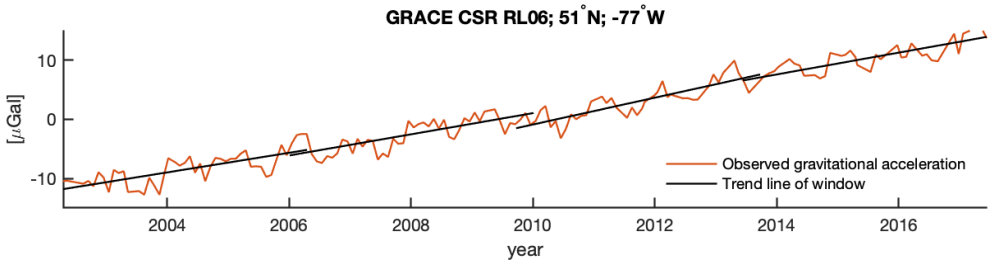



Figure 3.1: The concept of the trend variability experiment explained. Red line indicates a time series of the Earth's gravity signal (GRACE CSR RL06, 300 km Gaussian filter), centered around 0, at an arbitrary location (51°N; -77°W). Black lines indicate the trend in the gravity signal in a particular time window.

hydrology signals cancel out over increased lengths of time. Next, we expand the experiment to the spatial domain. Instead of computing the Laurentian mean gravity rate of each individual temporal subset, we will create spatial maps of the difference between the minimum and maximum observed gravity trend in any of the subsets. We call this value the *trend variability*. To account for outliers, we will discard the single lowest and single highest trend estimate. We will use 30 temporal subsets, evenly spaced across the entire span of the data set. We will again experiment with various window lengths to validate our findings of the first experiment regarding the canceling out of inter-annual signals diluting the GIA signal for time series of increased length.

Figure 3.1 explains the nature of these experiments by looking at a time series of an arbitrarily chosen grid point in the GRACE CSR RL06 data [7]. Only in this example we use 4 windows and take a window length of 4 years, to increase the readability of the figure. Note how the slope of the line starting in late 2009 is slightly steeper than the slope of the other lines. This suggests that there are inter-annual components superposed on the underlying secular component, which makes isolating that secular component non-trivial. For the first trend variability experiment, we will plot the slope of each of the black lines vs. their mid-point in the time domain. The second experiment will find the black line with the highest slope and subtract the black line with the lowest slope (after aforementioned outlier correction) for each individual grid point. This difference will then be plotted on a spatial map. Naturally, this means the outcome of this experiment will be highly dependent on the spatial resolution of the data we use, which is important to take into account once interpreting the results.

This experiment should help towards gaining a number of insights:

1. We will gather information about spatial distribution of the variability of the satellite-derived trend. This can help identify locations where the variability is low, i.e., where the trend is (relatively) constant through time. These locations would make ideal candidates for comparing data sets that cover different time frames (i.e., GRACE and Swarm). As was mentioned before, this conclusion will be restricted to the spatial resolution of the data used in this experiment.

2. We can compare the trend variability of the various satellite missions. A difference in trend variability, while using the same time span, window length, and data pre-processing steps, suggests that there is an influence on the trend that is specific to a measurement principle or data set (e.g., noise levels).

3. We will be able to assess if computing the trend from a longer time span results in less variability of the data, i.e., if the inter-annual influences on the gravity trend cancel out over time. It should be noted that comparing results from this experiment across window lengths should be done carefully, as the *samples* that we take are not independent of each other. Note how the trend lines in Figure 3.1 partially overlap. This will happen if the number of windows multiplied with the window length is more than the length of the entire data set. While we agree that the standard deviation of the trends of the particular windows would be a very interesting result, this sample dependence means that any statistics computed via generic methods are invalid. While we are aware that statistical methods for dependent data samples exist, we decide not to pursue these results for conciseness sake and leave this objective for future work.

The limited length of our time windows means we cannot include tidal aliases into our fitted curve, as their periods are far longer than our time window length. This means the tidal aliases, and thus their influence on the trend results we get from the GRACE data, could lead to an overestimation of the trend variability. For this reason, we will first compute the statistical test described in Section 2.5.3, before proceeding with the experiment described here.

## 3.2. Inter-annual hydrology variations

In the last section we learned how Van der Wal et al. [106] showed that the gravity trend in Laurentia is not constant through time. In their publication, the authors tied the trend increase to an increase in land water storage in the period 2003 to 2006 that they had identified in three different hydrology models. That authors concluded that surface hydrology variations have a non-negligible influence on the gravity trend observed by GRACE. Earlier studies using GRACE data to observe the GIA-induced gravity trend already identified that the GIA gravity signal was obscured by inter-annual hydrology variations (and had corrected for this, which will be covered later), but also claimed that the hydrology influence was minor [98] or less than 10% of the GIA-induced signal [55]. This latter statement does need the qualification that inter-annual hydrology variation in the time period and location used by Tamisiea et al. [98] was linear, and thus very hard to discern from a GIA-induced gravity trend. This illustrates how deriving a trend from a comparatively short (studies mentioned in this paragraph used 44 mo to 59 mo) can lead to spurious results. This warns us to be careful when interpreting the results of the Swarm-derived trends, as the Swarm time series spans 64 mo.

The most common way to correct for these influences is by subtracting gravity trends derived from a surface hydrology model from the satellite-derived trends [119]. Currently there is much debate in the scientific community about the usefulness of model-computed long-term total water storage (TWS) trends for correcting satellite gravity data. Both inter-comparison of hydrology models and comparison with GRACE measurements show substantial differences in results. Scanlon et al. [83] identify four main factors impacting the discrepancy between GRACE and the hydrology models: initial conditions and spin-up (number of years the model is run - this can be insufficient for the model to equilibrate with climate forcing), inaccuracies relating to model structure and water storage compartments, precipitation uncertainty, and model calibration.

The limited agreement between GRACE and continental hydrology models found by Scanlon et al. [83] is in line with the conclusions of authors who applied the hydrology models to isolate the GIA-induced trend from GRACE gravity observations, and found that the hydrology models were the Achilles heel of their study [60, 77, 92, 106]. Given the limited hydrological constraints and the large differences in methodology Roy and Peltier [77] observe between two prominent surface hydrology models (GLDAS and WGHM), the authors suggest that the current hydrological models are simply not robust enough to correct GRACE gravity measurements in order to isolate the GIA signal at this time.

Global hydrology models can be grouped into global Land Surface Models (LSM) or Global Hydrological and Water Resource Models (GHWRM). LSMs are integrated into general circulation models and aim to accurately estimate the fluxes between land surface and climate [83]. This makes long-term trends derived from LSMs susceptible for errors, as a small discrepancy in the estimated flux leads to a large inaccuracy once this flux error accumulates over time [30]. In theory, GHWRMs are better suited for this application as they use in-situ data to estimate basin sizes. An example of a GHWRM is the WaterGAP hydrology model (WGHM) [27], which has been used to isolate the GIA signal before [79, 93]. Some authors use an LSM and a GHWRM (often GLDAS and WGHM) and compare results [76]. Unfortunately for us, WGHM does not cover the Swarm mission time frame.

### 3.2.1. GLDAS-NOAH

In fact, the time frame of the Swarm data that we use (ending in March 2019) constrains our model choices to the point where options are very scarce. Additionally, we set out to use a model that has been used to isolate the GIA signal by previous authors, as it is not the goal of this study to assess the applicability of hydrology models not used in for GIA study before. Taking both these factors into account, we find that only the GLDAS-NOAH v2.1 [74] provides sufficient coverage to correct the Swarm derived trend and has been tried and tested in a GIA-context [60, 98, 106, 108]. GLDAS is an LSM, which warns us to be careful when interpreting inter-annual results. Including other hydrology models, either for a plain comparison of the results [77], or to estimate the error in the hydrology model used to correct the satellite data [106] would have been an informative addition to this study. Unfortunately, the time frame of the Swarm data makes this not possible. We intend to (partially) fill this void with the results from our trend variability experiment described in the last section.

To compute the hydrology-induced gravity trend, we first compute the vertical sum of various hydrology components in the GLDAS model, the total water storage TWS. We find the TWS of the column at longitude $\lambda$ and latitude $\theta$, at time $t$ via:

$$\mathrm{TWS}(\lambda,\theta,t) = \mathrm{SWC}(\lambda,\theta,t) + \mathrm{SWE}(\lambda,\theta,t) + \mathrm{CWS}(\lambda,\theta,t) \tag{3.1}$$

where SWC is the soil water content, SWE the snow water equivalent, and CWS is the canopy water storage. GLDAS provides these components in kg m$^{-2}$, on a 0.25° grid.

We compute monthly mean TWS grids from the daily GLDAS data files. In the case of GRACE, we exclude days from GLDAS that are also excluded from the GRACE fields. These days are listed in the GSM geopotential coefficient text files distributed by the GRACE processing centers. Next, we trim the GLDAS data set to the same end and start dates as the

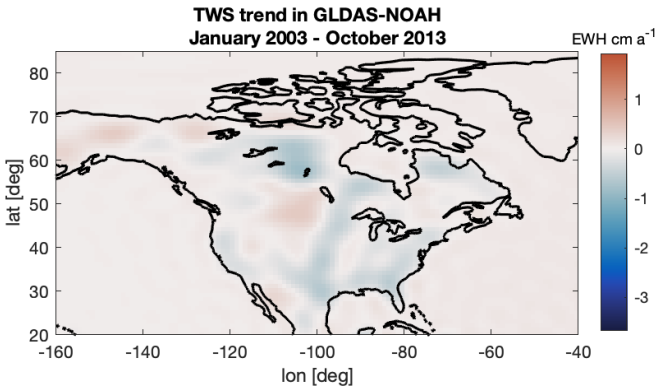



Figure 3.2: GLDAS TWS trend in the period January 2003 to October 2013. Shown here to verify our GLDAS processing routine by recomputing Roy [76, Figure 4.26].

satellite data set we are trying to correct. Finally, we apply the same processing (i.e., truncation, Gaussian filtering) as we do to our satellite-derived gravity fields to achieve equal comparison. We do this in the spectral domain and decompose the GLDAS grids into a degree 719 (NB: GLDAS-NOAH v2.1 is published as a series of 0.25°-grids) spectral representation. Using a water density of 1000 kg $m^{-2}$ and Equation 2.1 and 2.2 we convert mass change into gravity change. The methodology behind spectral decomposition is briefly covered in Section A.4.

We verify our implementation of the GLDAS-NOAH results by comparing to literature. We invite the reader to compare Figure 3.2 to Roy [76, Figure 4.26]. Here, we see the trend in GLDAS in the period January 2003 to October 2013, after truncating the model to degree 60 and applying a Gaussian filter with a 300 km radius. We see good agreement with the results of Roy [76] and conclude that our implementation of the GLDAS TWS trend computation is correct.

### 3.2.2. GLDAS pre-processing

To get more accurate long-term TWS trend from GLDAS, we apply a number of pre-processing steps. We start by following the advice on the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) FAQ[1] regarding long term trends derived from GLDAS. There is no ice sheet model that melts/displays snow over glaciers in GLDAS, which can lead to non-geophysical trends in SWE. To mitigate this risk, we mask all grid points that have a non-zero $SWE$ value for 90% of the epochs during our period of interest. This percentage was an arbitrary choice to give the correction some margin of error. Figure 3.3 shows the area that we mask due to permanent snow cover in white. Simply setting the masked points to zero would likely lead to strong gradients in our trend map, which could cause unwanted undulation effects (more in Section 3.3). For this reason, we set the grid points that are masked by the SWE mask equal to the arithmetic mean of the eight surrounding grid points.

Mostly due to ice mass loss, Greenland, Alaska, and some areas in western Canada show strong trends in TWS. This is visible in both the GRACE- and the Swarm-derived gravity trends in Figure 1.2 on page 4. These areas are mostly masked by our SWE-based mask, which means our GLDAS correction will not correct the strong trends shown in Figure 1.2. Even though they are not directly in our area of interest, we need to correct for these mass changes due the nature of the spherical harmonic decomposition. We will cover the reason for this and the correction that we propose in Section 3.3. As these are separate corrections, we must set the corresponding areas in GLDAS to zero to not end up with a double correction. GLDAS does not include Greenland (as defined by its coastlines), but as we also correct for Alaska separately, we must exclude the corresponding area there. We compute a mask from the GRACE CSR RL06m mascon solutions, as these are relatively free of long-wavelength leakage and have been corrected for GIA [80]. We define the limits of our mask as a < 0 $\mu$Gal $a^{-1}$ gravity trend as computed from the 4/2002-6/2017 epochs of RL06m, along with latitude and longitude bounds [45°N, 70°N] and [-160°W, -118°W], respectively. Figure 3.3 shows the areas in Greenland and Alaska that we mask in also white.

### 3.2.3. Irregularities in GLDAS-NOAH

During the first iterations of this study we noted that the GLDAS model, even after applying aforementioned pre-processing steps, yielded high-amplitude trends in a small number of very localized areas. Upon closer inspection, we found out this was caused by likely non-geophysical SWC jumps around the second half of 2016. We observe the same results in NASA's

[1] http://disc.gsfc.nasa.gov/information/faqs

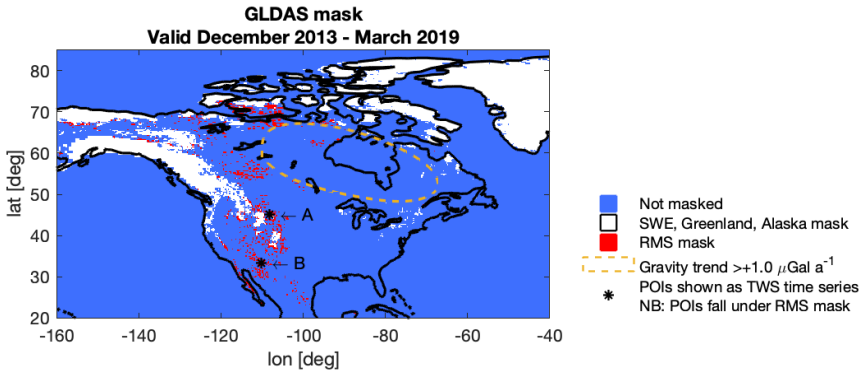



Figure 3.3: Mask that we apply to the GLDAS grid data before computing the GLDAS TWS trend. White grid points are masked because of a non-zero SWE in GLDAS in 90% of the epochs in aforementioned period, or because they are in Alaska or Greenland. Red points are masked because their RMS in 2016 is greater than 3 times their RMS in 2014. Points that are masked in both the SWE and the RMS mask are white. Blue points are not masked. Dashed orange line indicated the area where the gravity trend derived from GRACE CSR RL06 ($l_{max}$ = 13) >+1.0$\mu$Gal a$^{-1}$. Asterisks indicate locations of the time series on the top row of Figure 3.4. NB: the grid points of the POIs fall under the RMS mask (obscured by asterisk markers in the figure). We recompute the mask for each analysis to correspond to the time frame of the satellite data that is being corrected. This particular mask applies to the period December 2013 to March 2019.

online data display tool *Giovanni*[2], ruling out processing errors on our side. The top row of Figure 3.4 shows two examples of such signal jumps. The geographical locations of these time series are shown as Points Of Interest (POIs) marked by black asterisks in Figure 3.3. The bottom row of Figure 3.4 shows time series of nearby grid points that show no jump, which suggests the jumps are non-geophysical as such a substantial increase in SWC would surely affect neighboring grid points. Note how the jumps are of a magnitude multiple times the pre-2016 variability of the TWS estimate, and how the variability of the TWS estimate increases substantially after the 2016 jump. Given alongside the GLDAS TWS estimates in Figure 3.4 are the mass change estimates from the GRACE CSR RL06m mascons. We choose the mascons for this Figure and not the spherical harmonics solutions as the mascons have been corrected for GIA (by subtracting the model of Peltier et al. [61]), which means they should correlate better to the GLDAS TWS estimates. Note how the mass observations as given by the mascons do not indicate any notable jumps. This further argues that the TWS jumps in GLDAS-NOAH are unlikely to correlate to real geophysical mass redistribution.

Other than finding out that these jumps are in the SWC component of GLDAS, we have been unable to find the source of this error. The Hydrological Sciences Laboratory at NASA Goddard Space Flight Center (GSFC), responsible for GLDAS-NOAH, acknowledged that the jumps were likely erroneous and identified that they are caused by a switch in precipitation forcing data on 31 October 2015[3]. GSFC is currently reprocessing GLDAS-NOAH v2.1 using updated precipitation data for the entire model. The results are expected in September 2019[4]. This means we are forced to apply an ad-hoc correction, i.e., a correction based on our observation of the available data. For each grid point, we compute the RMS of the year before the jump (2014) and the RMS of the year after the jump (2016). We now mask all grid points where $RMS_{2016} : RMS_{2014} > 3:1$, where 3 is an empirically determined coefficient. Like the SWE correction described in Section 3.2.2, we set the masked grid points equal to the arithmetic mean of the eight surrounding grid points. Using the fraction of the pre-jump and post-jump RMS could lead to irregularities in locations where the pre-jump RMS is close to zero, i.e., very dry areas. Here, a small increase in RMS across 2015 could lead to the point being labelled as faulty. However, we believe this shortcoming of this method will minimally affect our results, as our area of interest is not a dry area.

The mask that we have computed is shown in red in Figure 3.3. When comparing this mask to the location of the GIA-induced gravity trend, which is shown in, e.g., Figure 1.3 on page 4, note how most of the excluded grid points fall outside of the area that is dominated by the GIA signal. The dashed orange line in Figure 3.3 indicates the areas where the gravity trend derived from GRACE CSR RL06 ($l_{max}$ = 13) >+1.0$\mu$Gal a$^{-1}$. This means that any hydrology corrections directly located in the GIA region will be minimally affected by the mask. Mostly, the mask will reduce long-wave leakage.

## 3.3. Long-wavelength leakage and ringing

In this study we use a spectral representation of the Earth's gravity field. While this approach has many advantages in terms of computational efficiency and opportunities for insight into the spectral distribution of various signals, it also

[2] `http://giovanni.gsfc.nasa.gov/giovanni/`
[3] Personal communication between dr. Mark Tamisiea at CSR and dr. Matthew Rodell at GSFC.
[4] Ms. Hiroko Kato Beaudoing, personal communication.

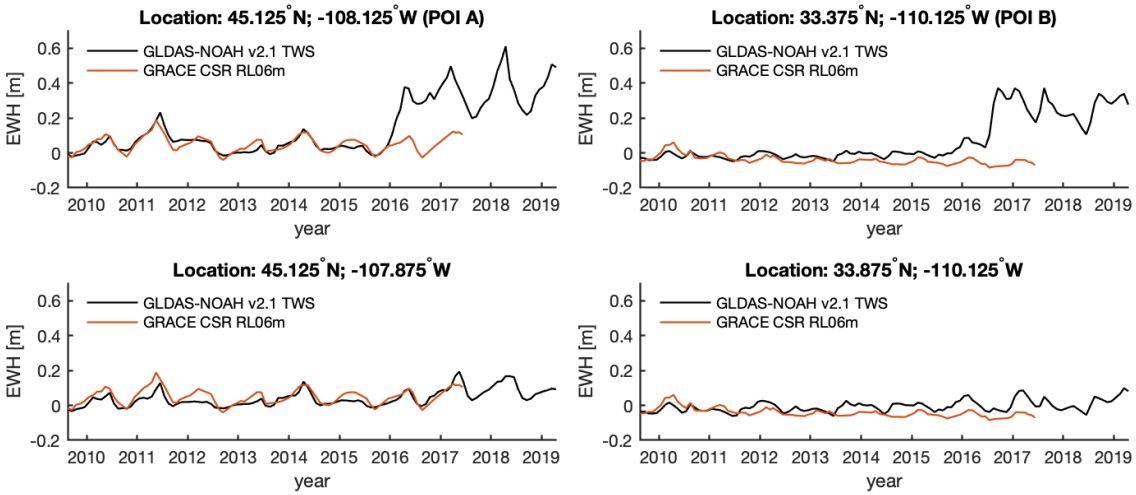



Figure 3.4: Time series of TWS results from GLDAS-NOAH v2.1 and mass change observations from the GRACE CSR RL06m mascons. Values are given as anomalies relative to a January 2004 to December 2009 mean baseline. Shown here on the top row are two arbitrarily selected grid points that show the jumps (also indicated as POI A and B on Figure 3.3). On the bottom row are two grid points located nearby the points of the top row, but that show no jump.

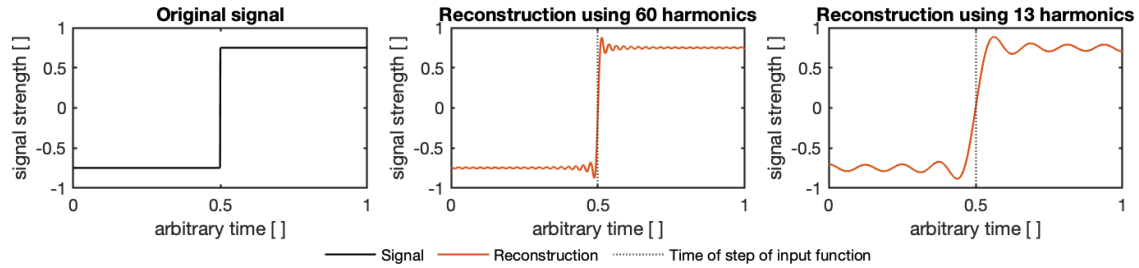



Figure 3.5: Left: arbitrary step function. Middle: reconstruction of the step function of the right panel using 60 harmonics. Right: reconstruction of the step function of the right panel using 13 harmonics. The dotted vertical lines in the middle and right panel indicate the time of the step of the input function.

has some disadvantages. The decomposition of a potential field into a set of spherical harmonics shares some similarities with a Fourier transform of a continuous function into a Fourier series of weighted sines and cosines. To better understand the particular limitations of spectral harmonics that we are trying to introduce in this section, we will therefore these limitations in the context of a 1-dimensional Fourier series first.

### 3.3.1. The Gibbs phenomenon

The Gibbs phenomenon is a general term used to describe the peculiar behavior of a Fourier series near a strong gradient. We illustrate this effect using an arbitrary step function. This function can be seen in the left panel of Figure 3.5. Next, we reconstruct the step function via a Fourier series, using amplitude-frequency pairs that we get from a Fourier transform of the step function. The middle panel in Figure 3.5 shows the reconstructed signal from a Fourier series of 60 harmonics, and the right panel shows the reconstructed signal when we use 13 harmonics. To fully recover the step function, one would need an infinite number of harmonics. In reality, the number of harmonics in a Fourier sum will always be finite. The Gibbs phenomenon is related to this discrepancy.

Two elements in Figure 3.5 are important for this study. First, note the vertical black line in the middle and right panel. This line indicates the location of the original step function. Especially in the right panel it is apparent that the reconstructed signal does not have the same infinite gradient as the step function. Rather, the gradient is drastically reduced. If we regard the area under the curve as the power of the signal, note how some power has leaked out to time steps before the time of the step function. Note the oscillations in the middle and right panel before and after the step at $t = 0.5$. This is another element of the Gibbs phenomenon. Note how, when we reduce the number of harmonics, the amplitude and the wavelength of the Gibbs oscillations increase.

Using 60 and 13 harmonics to reconstruct the signal in Figure 3.5 was not a completely arbitrary choice. Application of GRACE gravity fields to geophysical study typically happens with a maximum model degree of 60, albeit often with some Gaussian filtering. If we go back to Figure 2.2, we see that the noise in the Swarm fields starts rapidly increasing after degree 13. The authors responsible for the combined Swarm fields recommend to only use the fields up to degree 12-15 for geophysical study[5]. In other words, the mid panel of Figure 3.5 is indicative of the magnitude of the signal leakage and Gibbs ringing in a typical GRACE study, whereas the right panel shows this for a typical Swarm study. To not create a resolution discrepancy, we will bring GRACE down to the same maximum degree as Swarm once we compare the two, which means the ringing and leakage in the expansions of both the GRACE- and Swarm-derived Stokes coefficients will be substantial.

### 3.3.2. Applied correction

Even though there will most likely not be a gradient in the Earth's gravity field as steep as the step function in Figure 3.5, signal leakage has always been a matter of concern in GRACE application studies [112]. Figure 3.5 also suggests that unwanted ringing effects will likely be even more pronounced in our study than in a typical GRACE study. As we are mostly working in the gravity trend domain, ice mass loss is a matter of concern as it induces substantial gravity trends. Additionally, locations with a GIA-induced gravity trend are often close to areas of ice mass loss, which makes sense as GIA itself is caused by growth and decay of ice sheets. Both of these characteristics of the gravity trend induced by the Greenland ice mass loss can also be seen in Figure 1.3. As was explained in Section 3.2.2, a secondary area of our interest due to ice mass loss is Alaska and western Canada. Both the gravity trends induced by ice mass loss in Greenland and the Alaska area can be seen in the GRACE- and Swarm-derived gravity trend maps in Figure 1.2. Remember that our GLDAS-based correction excludes Greenland, as the results of GLDAS-NOAH are not accurate there, and how we masked Alaska to make room for the correction proposed here.

For the goal of this study, the comparison between two satellite-derived gravity trends, one could argue that the influences of long-wavelength leakage cancel out once we subtract the two trend results from each other. However, the ice mass loss in Greenland has recently slowed down [14, 52, 64], which means there is a difference between the ice mass loss in Greenland during the GRACE period and the ice mass loss in the Swarm period. This means we must account for Greenland signal interference in our comparison between GRACE and Swarm, because the difference in ice mass loss between the two periods should translate into a different influence on the observed GIA-induced gravity rate. Alaska is of secondary interest, as it has also experienced accelerating ice mass loss in the last decades [3] but is further away from our region of interest.

## Greenland

Various correction methods have been proposed to correct GRACE-derived GIA observations for ice mass loss in Greenland, either based on the GRACE data itself [5], independent data sources such as altimetry [33, 90], or filtering [31, 75]. Corrections computed independently from GRACE data have our preference, as we know in this case that we are not correcting for any other components in the GRACE observations of the Greenland area (i.e., the same ringing and leakage effects described in this section). However, just like we saw for the hydrology correction in the last section, the fact that we use very recent Swarm data (up to March 2019) greatly constrains our choices of usable independent data sets.

Sørensen et al. [90] showed that satellite altimetry-based Greenland mass change estimates correlate best to GRACE estimates, but these are not available for the Swarm time frame. This means we have to settle for potentially less accurate mass loss estimates, but keep in mind that eventually we will be looking at the differential trend between the GRACE and Swarm observations, and that any constant error therefore has very limited influence on our final results. We use results from Mouginot et al. [52], who employ a survey of thickness, surface elevation, velocity, and surface mass balance of 260 glaciers in Greenland to calculate the total Greenland mass balance. The online supplement of Mouginot et al. [52] gives the mass balance rate estimates per year, from which we derive -261.9 ±57 GT $a^{-1}$ for the GRACE period and -219.7 ±57 GT $a^{-1}$ for the Swarm period (defining the periods as 2002-2017, and 2014-2018, respectively). As was mentioned earlier in this section, only the differential mass loss rate is relevant for our study, as we are trying to isolate a congruent GIA signal from GRACE and Swarm, and not trying to isolate the most accurate signal in the absolute sense. For this reason, we will not verify these mass loss rates further.

Next, we apply the method of Sørensen et al. [90] to convert this mass balance rate estimate into a gravitational trend, by placing a uniform mass rate signal on the Greenland area and subsequently decomposing this into spherical harmonics. The methodology for spherical harmonics decomposition is outlined in Section A.4. The area where we place the mass rate signal is shown in the left panel of Figure 3.6. Note how this area correlates to the area that we set to 0 in our GLDAS-based hydrology correction (see Figure 3.3). The final corrections, and their effects on the trend estimates of the satellite data sets, will be shown in Section 4.7.

[5] Dr. João de Teixeira da Encarnação, personal communication.

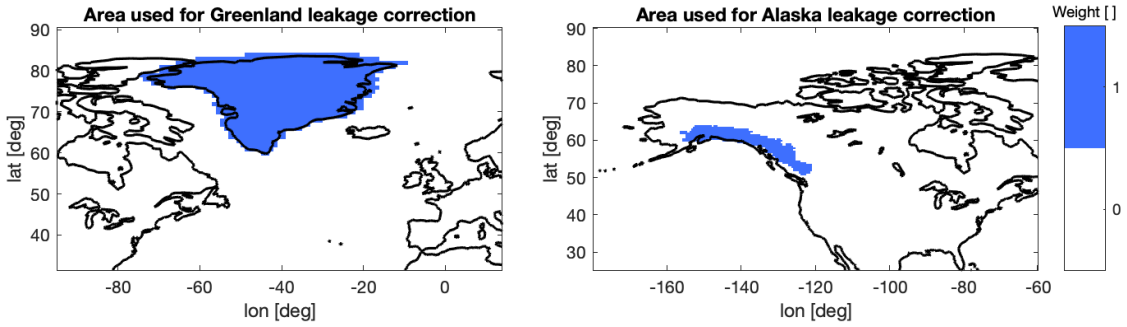



Figure 3.6: Earth surface area used for the long-wavelength leakage correction as described by Sørensen et al. [90].

### Alaska and western Canada

Acquiring recent mass balance estimates of Alaska and western Canada (hereafter: *Alaska*) proved even harder than for Greenland, and for that reason we will not apply the same methodology as for the Greenland correction. Instead, we propose a correction based on the GRACE mascon solutions. Root et al. [75] uses a similar approach, but with different mascon solutions. Mascons are attractive for this purpose, as they are less affected by unwanted effects of a spherical harmonics reconstruction such as leakage or ringing. We use the GRACE CSR RL06m mascons solutions [80, 82], which are distributed as a 0.25° grid. We derive the trend via least squares regression of a bias, trend, and (semi-)annual component, weighted via the methods of Section 2.6.1. Next, we isolate all gravity trends <0 $\mu$Gal a$^{-1}$ between latitude and longitude bounds [45°N, 70°N] and [-160°W, -118°W]. The right panel of Figure 3.6 shows the area we use. Note that this is exactly the same area that we masked in GLDAS, as described in Section 3.2.2. We compute the degree 719 spherical harmonic decomposition of the trend grid (see Section A.4) and apply the same processing as we do the satellite-derived trends we are trying to correct (i.e., truncation, Gaussian filtering, etc.).

The question now arises how we should extrapolate this result to the Swarm period. We use the *Fluctuations of Glaciers* (FoG) Database of the World Glacier Monitoring Service (WGMS) [120] to estimate the change in Alaskan ice mass loss rate between the period 2002-2017 ($\dot{m}_1^{(AL)}$) and 2014-2018 ($\dot{m}_2^{(AL)}$). This data is publicly available via the WGMS website[6]. We find 65 mass balance records of the Gulkana and Eklutna glaciers between 2002 and 2018, and compare the means of the available yearly mass loss records of the period 2002-2017 and of the period 2013-2018. We use the fraction of these means to estimate $\dot{m}_1^{(AL)} : \dot{m}_2^{(AL)} = 1 : 1.4326$. Even though we must acknowledge that this is a very crude estimation method, this accelerating mass loss is in line findings of other authors [3, 40]. We use this factor to linearly scale the GRACE mascon-derived correction before we apply it to the Swarm-derived gravity trend.

The correct implementation of both the Greenland and the Alaska correction can be verified once they have been applied to the satellite-derived gravity trends, as there should be little residual trend left. This does not mean the residual trend is 0, as some real geophysical signal due to GIA can be expected [119]. In Sørensen et al. [90] we learn that the residual trend in Greenland typically is ca. $\pm < 0.5$ $\mu$Gal a$^{-1}$.

## 3.4. Reducing very low-degree coefficients

To further focus our research on the Laurentide gravity signal, we aim to reduce the influence of long-wavelength leakage from ice mass loss signals coming from Antarctica. We experiment with doing this via two ways. First, we construct a Gaussian filter that reduces the very low-degree coefficients, similar to the approach of Root et al. [75]. Second, we follow Sasgen et al. [79] who simply reduce the very low degree coefficients to 0. Note that these two approaches are essentially the same methods as described in Section 2.4 but applied in such a way that the low-degree coefficients are reduced or discarded. We define the high-pass Gaussian scaling coefficient, $s_{hp,g}$ and the high-pass truncation scaling coefficient $s_{hp,t}$ as:

$$s_{hp,g}(l) = 1 - s_g(l) \tag{3.2}$$

$$s_{hp,t}(l) = 1 - s_t(l) \tag{3.3}$$

where $s_g$ and $s_t$ are the low-pass Gaussian and truncation scaling factor, as defined by Equation 2.8 and Equation 2.10, respectively, that are both a function of degree $l$. To distinguish between low-pass and high-pass truncation and Gaussian filtering, respectively, we will use $l_{max}$ and $r_g$ for low pass processing, and $l_{min}$ and $r_{hp,g}$ for high-pass processing.

[6] http://doi.org/10.5904/wgms-fog-2018-11

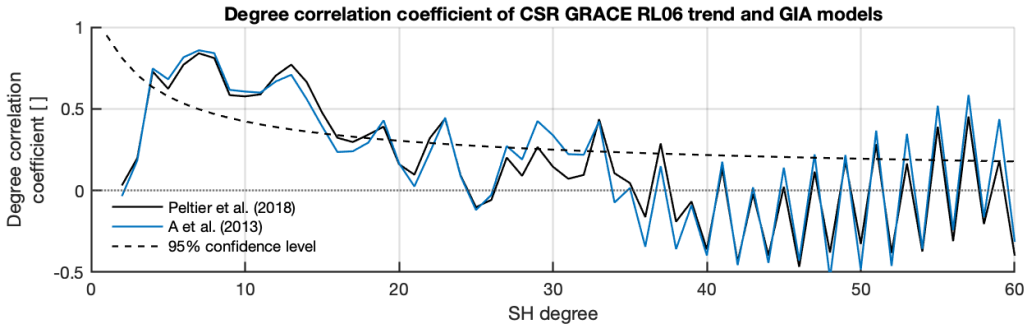



Figure 3.7: Degree correlation between GRACE CSR RL06 gravity trend (4/2002-6/2017) and gravity trend results of the ICE-6G_D GIA model of Peltier et al. [61] and the GIA model of A et al. [1]. Dashed line indicates the 95% confidence level.

Sasgen et al. [79] use the degree correlation between the GRACE-derived gravity trends and gravity trend results from a GIA model to motivate their choice of cutting all coefficients $l \leq 4$. Since we use newer GRACE data and a longer time series (CSR RL04 8/2002-8/2009 vs. CSR RL06 4/2002-6/2017), we have decided to repeat their degree correlation experiment. Additionally, we compare two GIA models to assure that any conclusions we derive are not driven by characteristics particular to any specific GIA model. We use the results of the ICE-6G_D GIA model of Peltier et al. [61] and the GIA model of A et al. [1]. We choose these model results because they have been shown to have good agreement with GRACE gravity observations but are computed independently from any GRACE data. ICE-6G_D and the model of A et al. [1] are also used for the GIA corrections to the GRACE CSR RL06m [80] and RL05m [82] mascon solutions, respectively. Gravity trend results of both GIA models are publicly available via the NASA PODAAC[7]. For conciseness sake, we will not cover the methodology behind the GIA models, but we refer the reader to Peltier et al. [58, 59, 61] and A et al. [1]. Like we did for the satellite-derived gravity fields (see Section 2.1.5), we set the degree 1 and $C_{2,0}$ coefficients in the GIA models results to 0. We subtract the corrections for ice mass loss in Greenland and Alaska outlined in Section 3.3 from the GRACE-derived gravity trends. Finally, we mask the Southern Hemisphere in the GIA model and the GRACE trend results, to focus our analysis on the Northern Hemisphere where our GIA region of interest is located.

Figure 3.7 shows the results of this experiment. Both GIA models yield a similar pattern of degree correlation w.r.t. GRACE. The correlation is <0.25 for degrees 2 and 3. The correlation between the GRACE-derived trend and ICE-6G_D is statistically significant for degrees 4-16, and for the model of A et al. [1] this is the case for degrees 4-15. For some degrees higher than 16 resp. 15 there is sporadic significant correlation. The erratic behavior of the degree correlation seen in the higher degrees (ca. $l > 35$) can be explained due to the fact that masking the Southern Hemisphere leads to strong gradients in the spatial expansion of the GIA model results and the Stokes coefficient trends, which induces ringing (see also Section 3.3).

We are most interested in the apparently uncorrelated low degrees. We believe this is likely caused by long-wavelength leakage (largely) induced by mass transport near the West Antarctic Ice Sheet, as this is a substantial trend signal (see, e.g., Figure 1.1) that we had not masked in this analysis. In this location, both ice mass loss and GIA affect the gravity trend. The GIA component (here: the GIA model) is positive, but the net gravity trend (i.e., the GRACE observations) is negative due to the contribution of ice mass loss. The Antarctic ice mass loss process is not stationary [89]. This means any leakage that influences the observations of the GIA-induced trend are not stationary either. Because the goal of this study is comparing two data sets that cover different time frames, any non-stationary corrections are especially important. For now we conclude that we will experiment with high-pass cut-offs at $l_{min} = 3, 4, 5$, where $l_{min}$ is the lowest degree harmonic not to be truncated. We convert to filter radii via Equation 2.9, which yields 3340, 2505, and 2004 km, respectively.

## 3.5. Analysis of the differential trend

In the final part of our comparison between GRACE and Swarm, we quantitatively assess the respective observations of the gravity trend in North America. For conciseness sake, we name the misfit between the GRACE- and Swarm-derived trend of Stokes coefficient or grid point $i$ the differential trend $\Delta$ of $i$, that is defined as:

$$\Delta_i = \hat{x}_{t,i}^{(GRACE)} - \hat{x}_{t,i}^{(Swarm)} \tag{3.4}$$

where $\hat{x}_{t,i}$ are the least squares estimates of the trend of $i$ derived from GRACE and Swarm, respectively.

[7] http://podaac-tools.jpl.nasa.gov/drive/

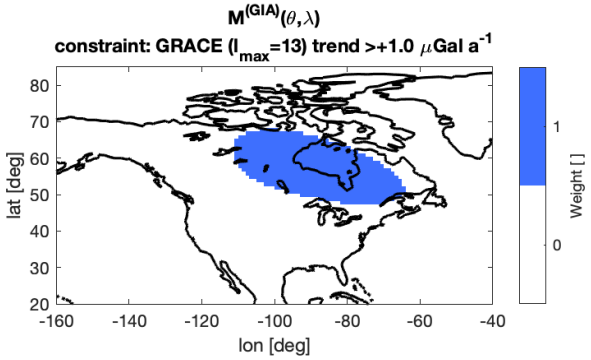



Figure 3.8: Mask used to isolate the positive trend anomaly around Hudson Bay, which is dominated by a GIA-induced trend. Mask is computed by identifying all areas in the North American continent with a >+1.0 $\mu$Gal a$^{-1}$ gravity trend as computed from GRACE CSR RL06 ($l_{max} = 13$), respectively.

### 3.5.1. Latitude-weighted GIA-region mean and RMS

To further aid comparison, we reduce the GRACE- and Swarm-derived trend, and any other derivatives, to a single number at selected points in the report. We use the latitude-weighted mean of the observations of the positive trend anomaly in the Hudson Bay area. To achieve optimal comparability of the results of this experiment, we need to define exactly which area we define as *dominated by a GIA-induced trend*. For this we take the GRACE CSR RL06 geopotential solutions spanning 4/2002-6/2017, truncate to $l_{max} = 13$ (will be motivated in Section 4.3), and determine the area on the North American continent where the expanded trend of the Stokes coefficients is >+1.0 $\mu$Gal a$^{-1}$. Figure 3.8 shows this region, that is also used by masking function $M^{(GIA)}(\theta,\lambda)$. We will always use the same mask, i.e., this exact definition of the *GIA-region*. Even when we look at Swarm data (or any other data), we use the area bounded by a >+1.0 $\mu$Gal a$^{-1}$ gravity trend as computed from aforementioned GRACE data. We will use the superscript $(\ )^{(GIA)}$ to indicate the latitude-weighted mean of the area in Figure 3.8, which may be computed from various gravity field derivatives. For an arbitrary function defined on the sphere $z(\theta,\lambda)$, $z^{(GIA)}$ and $z(\theta,\lambda)$ are thus related as follows:

$$z^{(GIA)} = \frac{\overset{\Omega}{\sum}\left[M^{(GIA)}(\theta,\lambda)\cdot\cos(\theta)\cdot z(\theta,\lambda)\right]}{\overset{\Omega}{\sum}\left[M^{(GIA)}(\theta,\lambda)\cdot\cos(\theta)\right]} \tag{3.5}$$

where $M^{(GIA)}(\theta,\lambda)$ is the masking function shown in Figure 3.8 and $\overset{\Omega}{\sum}$ indicates the summation over the entire sphere:

$$\overset{\Omega}{\sum} = \sum_{\lambda=-180°}^{180°}\sum_{\theta=-90°}^{90°} \tag{3.6}$$

A case of Equation 3.8 that we will use often is the latitude-weighted mean gravity trend in the GIA-region. We will denote this metric as $\dot{g}^{(GIA)}$. An exception to the method of Equation 3.8 applies when we want to reduce the outcome of Equation 3.4 or any other set of residuals to a single number. Here, we use the RMS, and not the mean, as highly variable residuals centered around zero would yield a zero mean, which could lead to the false conclusion that there is no discrepancy between the two metrics at hand. Instead, the RMS would correlate to the variance of said residuals and would likely yield a non-zero number in the case of variable residuals. Using $\Delta$ as an example, we write:

$$\Delta^{(GIA)} = \sqrt{\frac{\overset{\Omega}{\sum}\left[M^{(GIA)}(\theta,\lambda)\cdot\cos(\theta)\cdot\Delta_i^2\right]}{\overset{\Omega}{\sum}\left[M^{(GIA)}(\theta,\lambda)\cdot\cos(\theta)\right]}} \quad \text{for } i = 0,1,2,...,k \tag{3.7}$$

where all parameters are as defined earlier.

### 3.5.2. Normalizing the GRACE - Swarm misfit

In Equation 2.31 we saw the $\chi^2$-statistic, that normalized the residuals of a regression w.r.t. the uncertainty of the observations. In Section 2.5.3 we used the $\chi^2$-statistic as part of a larger analysis, but by itself the $\chi^2$-statistic can also be used to assess the discrepancy between the GRACE- and Swarm-derived gravity trends. For observations with a Gaussian

uncertainty distribution, a $\chi^2$ of 1 suggests that the misfit is within a single standard error $\sigma$, and $\chi^2 = 4$ suggests that the misfit is within $2\sigma$. We will use the methods outlined in Section 2.6 to assess the distributions of the errors in the various satellite-derived gravity fields.

In GIA study, the mean $\chi^2$-value is often minimized in an optimization process constraining a GIA model via observations of, e.g., relative sea level (RSL), $\dot{J}_2$, GPS-derived surface elevation [56, 98, 106]. The authors use the test to assess the difference between the GIA observation and the result from a GIA model. We employ it in a similar way, but instead of a GIA model, we take the GRACE-derived trend as *truth*. To assess this assumption, we will first compare the propagated trend uncertainty from GRACE and Swarm in Section 4.2. For now, we hypothesize that the GRACE uncertainty will be substantially lower than the Swarm uncertainty. Building on Equation 2.31, we compute the $\chi^2$-statistic of grid point or Stokes coefficient $i$ via [63]:

$$\chi_i^2 = \left( \frac{\hat{x}_{t,i}^{(GRACE)} - \hat{x}_{t,i}^{(Swarm)}}{\sigma_i} \right)^2 \tag{3.8}$$

where $\hat{x}_{t,i}$ is the least-squares estimate of the trend at grid point $i$ derived from the Swarm or the GRACE data. We hypothesize that the errors of Swarm are at least an order or magnitude bigger than the errors of GRACE, therefore we take for $\sigma_i$ the estimated uncertainty of $\hat{x}_{t,i}^{(Swarm)}$.

For observations from potentially dissimilar Gaussian distributions, the $\chi^2$-operation reduces the probability distribution of all observations to the same standard normal distribution, i.e., $N(0,1)$ (see Equation 2.40). For independent observations, this means the mean of $n$ number of $\chi^2$-values is distributed according to a $\chi^2$-distribution with $n$ degrees of freedom. This allows one to assign a confidence interval to the ensemble of observed data. We choose not to apply this method as the grid points in the maps of the Stokes coefficients expansions are not independent data points. Instead, we will show figures of the spatial distribution of the $\chi^2$-statistic, which allows us to address the distribution of the GRACE - Swarm trend misfit, in light of the uncertainties of those trend estimates.

## 3.6. Validation via GRACE-FO

In the final set of experiments of this study, we validate our findings via data of the recently launched successor of GRACE, GRACE-Follow On. The epochs of the GRACE-FO geopotential solutions are between June 2018 to April 2019, which places them at the far end of the Swarm fields. In the first section, we will attempt to derive gravity trends from the GRACE-FO data. Next, we will validate our results regarding the continuity between GRACE and Swarm in three ways. First, we will compare time series of modeled gravity change at specific grid points, using the GRACE-FO results from the first section. This allows us to assess how the three series of geopotential solutions compare independently from each other. Next, we will compare spatial maps of gravity trends derived from individual missions fields to gravity trends derived from sequences combining GRACE, Swarm, and GRACE-FO fields. Finally, we will append the GRACE data with either Swarm or GRACE-FO observations, and assess how this affects trends estimated from the combined time series w.r.t. trend combined from only GRACE. This latter experiment will help us understand the misfit between GRACE and Swarm better. If the comparison GRACE - GRACE-FO yields a similar misfit, this argues the misfit of GRACE and Swarm was caused by real geophysical signal and not noise. Note that all analyses in this chapter, either of GRACE, Swarm, or GRACE-FO data, will be limited to the spatial resolution to which Swarm is sensitive (ca. 1500 km). We leave the analysis of the GRACE-FO fields at resolutions beyond the senstivity of Swarm for future work.

### 3.6.1. GRACE-FO gravity trends

As the GRACE-FO data set is very short (9 mo), the chances that we will be able to successfully fit a (semi-)annual component are low. This could affect the results we get for other components of our regression model, where we are most interested in the trend. In an attempt to mitigate this issue, we test two additional regression approaches that reduce the number of independent variables in the model that we fit to the GRACE-FO data. We test three cases: one where we use only GRACE-FO data, one where we use the phase information from GRACE, and one where we use phase and amplitude information from GRACE. The first case applies the methods of Section 2.5 and will therefore not be covered here.

The second case can be explained as inserting the GRACE phase estimates in 2.11 before fitting the model to the GRACE-FO observations. This method is based on the assumption that the phases of the geophysical processes that induce the semi-annual and annual periodic mass change have not changed between the GRACE and GRACE-FO time periods. We rewrite Equation 2.11 as:

$$f(t) = \sum_{i=1}^{n} \left[ A_i \cdot t^{(i-1)} \right] + \sum_{j=1}^{m} \left[ B_j \sin \frac{2\pi}{T_j} \left( t - \hat{\omega}_j^{(GRACE)} \right) \right] + \varepsilon \tag{3.9}$$

where $\hat{\omega}_j^{(GRACE)}$ is the least square estimate of the phase of the periodic component $j$, as derived from the GRACE data (see Equation 2.21 to Equation 2.25), and all other parameters are as defined earlier.

For the third method, we use the estimated amplitude and phase that we get from applying the least squares regression to the GRACE data, to extrapolate these components into the GRACE-FO time frame. Next, we subtract the periodic components from the GRACE-FO observations, and fit a model with only a bias and a trend term (i.e., choosing $i = 2$ and $j = 0$ in Equation 2.11) via least squares regression. This method is based on the assumption that the phase *and* amplitude of the mass change-inducing geophysical processes has not changed between the GRACE and GRACE-FO time periods, which intuitively feels less realistic than the assumption we did for the first case.

### 3.6.2. Successive time series

In this first validation experiment, we expand the results of the regression of the GRACE, Swarm, and GRACE-FO gravity fields into the spatial domain, and compute time series of the gravity variation at specific points on the map where we know the GIA-induced trend is substantial. Ideally, the GRACE, Swarm, and GRACE-FO gravity observations should all lie on the same line here. We compare the observations of GRACE, Swarm, and GRACE-FO to the trend line that we compute via the regression.

While we agree that this is a rather trivial experiment, it does illustrate the central idea of the study, the GIA gravity signal acting as a mutual baseline for multiple satellite missions, very well. Additionally, results from this experiment are easy to understand for the less experienced reader, as we are not working in the domain of the first time derivative (i.e., the trend) anymore, but simply in the time domain. This may add to the understanding of the rest of the report.

### 3.6.3. Adding Swarm to the combined GRACE time series

In this second validation experiment we compare trends derived from different (combinations of) satellite data. First, we will compare the trends derived from the Swarm fields to trends computed from a combination of GRACE and GRACE-FO, where we trim the latter combination to the start and end dates of the Swarm fields. We will use a weighted regression with weights derived via the methods of Section 2.6.1 for both the Swarm fields and the combined GRACE / GRACE-FO fields.

Next, we will compare the combination GRACE / GRACE-FO to the combination GRACE / Swarm / GRACE-FO. We will do this for both an unweighted and two weighted cases. Our hypothesis is that the uncertainties of the Swarm observations will be substantially higher than the GRACE and GRACE-FO observations, and therefore it is likely that the first weighted case based on the weighing methods in Section 2.6.1 will yield a result where the Swarm observations contribute little to none to the results. For this reason, we will also experiment with a second weighing scheme that only accounts for the time evolution of the uncertainties. We do this by scaling the results from the methods of Section 2.6.1 in such a way that the time domain mean uncertainty equals unity. We will denote this method as using *trivial weights*. Note that, for the Swarm fields, we set the mean of the errors to the period after May 2015, as Van den IJssel et al. [104] suggests the accuracy of the Swarm-derived fields is non-nominal before that period due to influences from increased ionospheric activity.

In this experiment we do not look at the Swarm or GRACE-FO data by itself, but rather how adding the observations to the GRACE time series affects the trend results. As explained in Section 1.2, the goal of GRACE-FO is to continue the GRACE observations, and the reason why we are interested in Swarm is to cover the gap between the end of GRACE and the start of GRACE-FO. This makes this experiment very interesting, because it directly assesses if the Swarm observations could be a valuable addition to the existing time series of mass change. Unfortunately, we must immediately add the comment that correctly combining GRACE, Swarm, and GRACE-FO will likely require a more intricate weighing scheme than what we described above to optimally use all available information from Swarm. We have decided that this combination strategy is beyond the scope of this study, which means the results of this study should be regarded as tentative first steps in this research area.

### 3.6.4. Differential trend of combined time series w.r.t. GRACE

In the third and final validation experiment we assess the differential trend between a GRACE-only time series, and a time series that consists of either GRACE and Swarm data, or GRACE and GRACE-FO data. We rewrite Equation 3.4 as:

$$\Delta_i = \hat{x}_{t,i}^{(Comb)} - \hat{x}_{t,i}^{(GRACE)} \tag{3.10}$$

where $\hat{x}_{t,i}^{Comb}$ is now the least squares estimate of the trend of Stokes coefficient or grid point $i$, as computed from either one of the combined time series.

As there are 63 Swarm epochs, but only 9 for GRACE-FO, we must first account for this discrepancy before appending

the time series. A weighing scheme would not make sense, as the Swarm observations cover a different time period than the GRACE-FO observations. Mass change that occurred between August 2017 and June 2018 is simply not observed by the combined GRACE - GRACE-FO time series but will have an effect on any metric we derive from a GRACE - Swarm combination. Because of this, we have decided to only use 9 Swarm epochs for this experiment, that we have chosen so that they line up as well as possible with the GRACE-FO fields. We use June 2018 to March 2019 but leave out September 2018.

We will be looking for two specific things in the expanded maps of the differential trends. First, we are interested in the residual observations around Hudson Bay. If our assumptions of a stationary GIA-induced gravity trend and inter-annual hydrology variations canceling out over time are correct, there should be no differential signal here. Second, we will compare other artefacts of the differential trend map. Here, we are interested in whether Swarm observes the various geophysical signals in a way similar to GRACE-FO.

# 4

# Results

Similarly to the methodology chapters, this results chapter starts with the outcomes of more general experiments that we have applied to the GRACE- and Swarm-derived gravity fields. Results from these experiments are then used to determine the optimal configuration of the following analyses, which in the end aim towards deriving the best possible trend observations from the GRACE and Swarm geopotential solutions. This not only happens in the forward manner, but also in an iterative fashion. Note that this report is restricted to describing the final results. Most notably, the maximum degree configuration regarding gravity field truncation is derived in Section 4.3, but was already used in sections before that (and even sporadically in the previous chapters). We aim to refer the reader to subsequent sections as little as possible to increase the readability of the report.

This chapter starts by assessing the influence of the tidal aliases in the geopotential solutions to finalize the definition of our regression model (Equation 2.11). Section 4.2 investigates the observation uncertainties of the GRACE and Swarm solution and propagates the observation uncertainties into gravity trend uncertainties. Section 4.3 shows the gravity trends we derive from GRACE and Swarm after reducing the influence from high-degree noise in the Swarm solutions. Section 4.4 investigates the effect of the data set length on the gravity trend estimates. At this point, we have answered research sub-question Q.1. Next, Sections 4.5 to 4.8 first investigate the variability of the gravity trend through time, i.e., the magnitude of the components other than the linear GIA-induced trend, and then present the results of correcting for these signal dilutions. Now, we have answered research sub-question Q.2. Then, Section 4.9 contains the analysis of the discrepancy between the GRACE- and Swarm-derived trends. This section answers research sub-question Q.3 and can be regarded as a quantitative summary of the entire results chapter. Finally, Section 4.10 shows the validation of our results and answers research sub-question Q.4.

Definitions of acronyms are given in the List of Acronyms on page xiii. Symbol definitions and units, and values used for constants can be found in the List of Symbols on page xv.

## 4.1. Influence of tidal aliases

The $S_2$, $K_2$, and $K_1$ tides have been shown to alias into the GRACE solutions [67], with periods of 161, 1400, and 2800 days, respectively. Steffen et al. [94] suggest that not including a periodic component into the model of the least squares regression could lead to an incomplete cycle of that component affecting the trend estimate, as Section 2.5.3 also explained. Other authors, e.g., Sasgen et al. [79], disregard the tidal aliases altogether, citing that their overall influence on the trend estimate should be minimal. In this section we will assess the significance of the tidal aliases, and their influence on the trends derived from GRACE and Swarm. There are four reasons why it is hard to predict if the same aliasing will affect Swarm. First, remember how Section 1.2 explained how the Swarm solutions that we used are a combination of four independently computed solutions. The individual Swarm solutions do not all use the same tidal models, which makes it likely that tidal aliasing does not affect the individual models in the exact same way. Second, the weights of the individual solutions differ per epoch. Third, the Swarm time series could be too short (5.3 y) to retrieve a multi-year component. The $K_2$ and $S_2$ alias periods are ca. 3.8 y and 7.7 y, respectively. Finally, the altitude of the Swarm satellites is different from GRACE, leading to different ground track repeat patterns (NB: aforementioned periods have been computed to correspond to the tidal aliases in GRACE).

We use an F-test to determine if there is a significant periodic component in each Stokes coefficient time series of both the GRACE and the Swarm models. We use a 95% confidence level. As our current hypothesis is that the Swarm fields contain

| Geopotential solution | | Period of component | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Name** | $l_{max}$ | **1 y** | **0.5 y** | **161 d** | **1400 d** | **2800 d** | **1 y + 0.5 y** | **Tidal aliases*** |
| GRACE CSR RL06 | 60 | 91% | 49% | 44% | 53% | 69% | 91% | 82% |
| GRACE CSR RL06 | 15 | 98% | 61% | 27% | 49% | 73% | 100% | 94% |
| Swarm | 40 | 31% | 24% | 25% | 29% | 37% | 28% | 34% |
| Swarm | 15 | 57% | 29% | 23% | 39% | 44% | 55% | 51% |

NB: these are *not* percentages related to confidence intervals (see text or figure caption)!

Table 4.1: Percentage of Stokes coefficients, compared to total number of coefficients, that have significant periodic component at a 95% confidence level. In all F-tests, the restricted model has no periodic components, except for the *Tidal aliases** column. Here, a model with (semi-)annual components and all three tidal aliases is compared to a model with only (semi-)annual components.

mostly noise above ca. degree 15, we execute this test for both the full GRACE and Swarm solutions (i.e., $l_{max} = 60$ and $l_{max} = 40$, respectively), and a version truncated to degree 15, as we will likely reduce the higher-degree coefficients in a later stage to reduce the influences of high-degree noise in Swarm.

Table 4.1 presents the percentage of the total number of Stokes coefficients that have a significant periodic component. Note that the percentages do not refer to confidence levels. For GRACE, we see that including a single periodic component generally does not lead to more coefficients having a statistically improved fit, with the exception of the annual component. This makes sense, as many of the mass transport inducing processes related to climate follow an annual pattern. Including both the annual and semi-annual component leads to a significant fit in the largest number of coefficients, especially when we truncate the fields. We find that the significant aliases are spectrally distributed (not shown) in a way that reminds us of the GRACE uncertainty distribution: mostly the coefficients in the lower orders have significant tidal aliases. Few sectoral coefficients have significant periodic components. We suspect that increased noise level in this part of the coefficient spectrum is the reason for this.

For Swarm, there are two important observations. First, we see that in the full fields, none of the (combinations of) components are significant in more than half of the coefficients. When we truncate to degree 15, we see that a model with only an annual component allows for a significantly improved fit in slightly more (57% vs. 55%) coefficients than a model with both the annual and semi-annual component. We also see that for slightly more than half (51%) of the coefficients in the truncated field, adding all three tidal aliases leads to a significantly improved model fit w.r.t. a model with only (semi-)annual components. When we look at the spectral distribution of which coefficients yield a significantly improved model fit when including any (combination of) periodic component(s), this again seems to inversely correlate to degree number, and thus noise levels. This leads us to conclude that at least part of the discrepancy between GRACE and Swarm in the number of coefficients that have significant tidal components is due to higher noise levels in Swarm.

Table 4.1 might tell us something about the significance of the periodic component, but it does not tell us anything about the (magnitude of the) influence on our signal of interest: the gravity trend in Laurentia. To assess this, we compare maps of the gravity trend, one derived using a model with only a bias, trend, and (semi-)annual component, and the other also including the tidal aliases. We use the GRACE data for this, as we know most Stokes coefficients of this data set have a significant component with a period equal to those of the tidal aliases.

The results are in Figure 4.1. We see that the positive trend anomaly around Hudson Bay is practically unaffected by the in- or exclusion of the tidal aliases. The residual trend between the top two panels is negligible. The RMS of the trend residual in the area where the GIA-induced trend is dominant, $\Delta^{(GIA)}$, (red ellipse in bottom panels in Figure 4.9, see Section 3.5) equals $3.776 \times 10{-}4$ $\mu$Gal a$^{-1}$. This result, the fact that the F-test did not seem to provide conclusive evidence for the Swarm fields, and the fact that the Swarm fields are likely too short to retrieve very long-period components, lead us to conclude that we will not include the tidal aliases in our fitted model. This means we also not correct for tidal influences in our trend variability experiment described in Section 3.1. Even though Table 4.1 showed that including only an annual component leads to a significantly better model fit in more coefficients than including (semi-)annual components, we decide to use the same regression model for GRACE and Swarm to achieve an equal comparison between the two. We define our model as:

$$f(t) = \sum_{i=1}^{2} \left[ A_i \cdot t^{(i-1)} \right] + \sum_{j=1}^{2} \left[ B_j \sin\left(\frac{2\pi}{T_j} t\right) + C_j \cos\left(\frac{2\pi}{T_j} t\right) \right] + \varepsilon \tag{4.1}$$

with $T_1 = 1$ y and $T_2 = 0.5$ y and all other parameters as defined in Section 2.5.

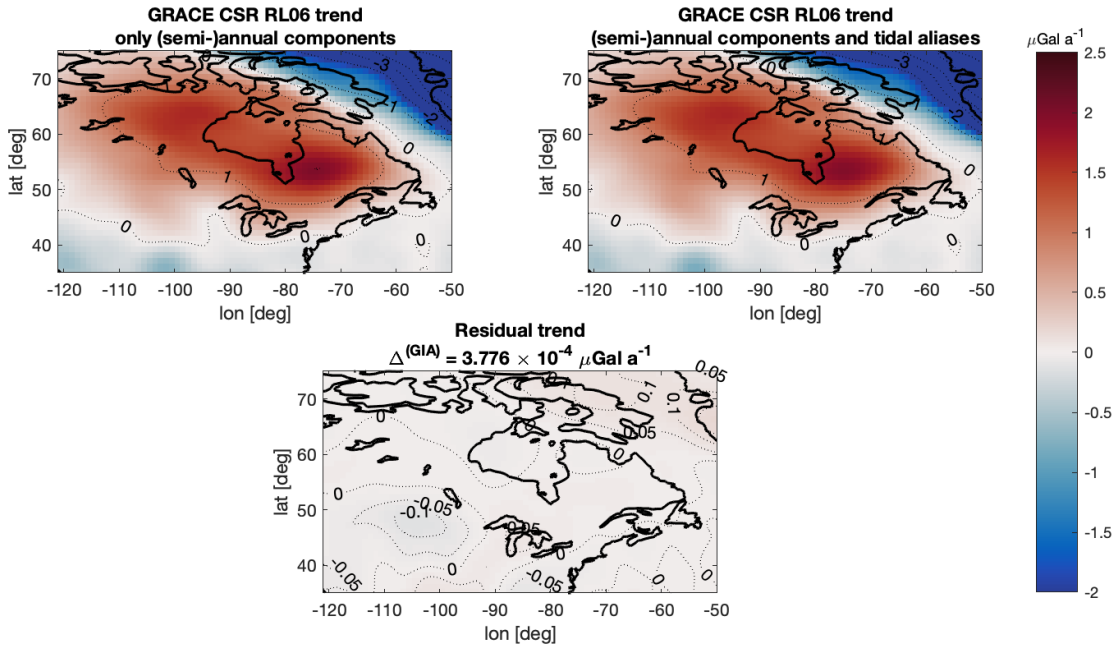



Figure 4.1: Trend derived from GRACE CSR RL06 via least squares regression of a bias, trend, and (semi-)annual component (left), and via regression of a bias, trend, (semi-)annual component, and the $S_2$, $K_2$, and $K_1$ tidal aliases (right). Bottom panel shows the residual between the top two panels. Only processing applied to the fields is a Gaussian filter with a 300 km half width, to reduce noise. NB: $\Delta^{(GIA)}$ is defined in Equation 3.7.

## 4.2. Error estimates

In this section, we will look at our trend error computations for Swarm and GRACE. First, we will compare the magnitude and spatial distribution of the trend error, as this is the signal component we are most interested in the end. Next, we will address the probability distribution of the observation errors, as normal error distribution is a prerequisite for some of the methods that we will be applying in the next chapter. Finally, we will address the temporal distribution of the observation errors, as this is an important metric in the weighted least squares method that we use to estimate the trend.

### 4.2.1. Spatial and spectral distribution

Figure 4.2 shows the spatial distribution of the trend errors. We compute this by first finding the standard deviations of the Stokes coefficient time series via the method of Section 2.6.1, and then propagating this into an uncertainty estimate for the trend via the methods of Section 2.5.2. The errors in GRACE and Swarm are multiple orders of magnitude apart and thus the spatial and spectral plots in this section do not use the same color scale. We truncate the fields at $l_{max} = 13$. We will motivate this choice in the section hereafter.

When we look at the errors for the truncated GRACE fields, the band near the equator is reminiscent of what we have seen in the $l_{max} = 60$ GRACE fields in Figure 2.5 on page 18. The anomaly near the poles does seem a little counter-intuitive, as the GRACE mission had virtually no polar gap ($i = 89.5°$ [99]), and the measurement density is highest at the poles. This anomaly is not visible in the maps of the error of the degree 60 fields in Figure 2.5.

We believe that the answer can be found in the spectral distribution of the errors (left panel of Figure 4.3). We see that the zonal coefficients up to degree 10, especially $C_{3,0}$ and $C_{4,0}$, have relatively high errors. This could explain the high-latitude artefact, that is spread across all longitudes. A look at the spectral distribution of the errors of the degree 60 fields (not shown) suggests that errors of the higher degree coefficients in these solutions compensate the errors of the lower degree zonal coefficients, which could explain the fact why there is no high-latitude artefact in Figure 2.5. Figure 4.4 shows the spatial distribution of the GRACE and Swarm errors when we exclude the zonal coefficients from the analysis. We see that the latitude-dependency of the errors reverses in the higher latitudes, which is more in line with what we see in literature [32, 114]. This suggest that in lower-resolution GRACE and Swarm solutions the uncertainties of the zonal coefficients indeed dominate the spatial distribution. These errors excluding the zonal coefficients were only computed to complement the analyses in this section, and are not used anywhere in this study.

It is believed that this relatively low sensitivity to the low degree zonal coefficients is due to the GRACE orbital geometry

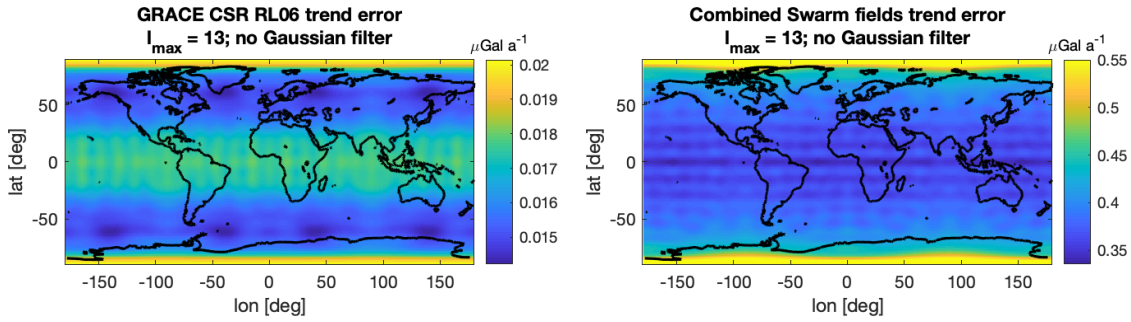



Figure 4.2: Spatial distribution of the errors in GRACE CSR RL06 and the Swarm fields, as computed via the method outlined in Section 2.6.1. NB: as the errors are more than an order of magnitude apart, and as the goal of this figure is to illustrate the spatial distribution of the errors, the panels use a different color scale.

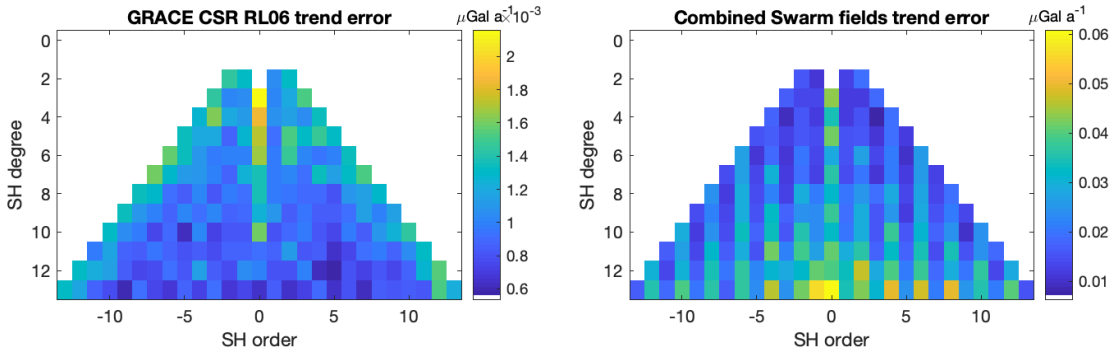



Figure 4.3: Spectral distribution of the errors in GRACE CSR RL06 and the Swarm fields, as computed via the method outlined in Section 2.6.1. NB: as the errors are more than an order of magnitude apart, and as the goal of this figure is to illustrate the spectral distribution of the errors, the panels use a different color scale. NB: degree 1 and $C_{2,0}$ coefficients are set to 0 (as in all analyses in this study).

and short separation (ca. 200 km [99]) between the satellites [16]. For this reason, many authors replace the GRACE-derived $C_{2,0}$ with results from SLR observations [18]. Section 2.1.5 describes why we choose to remove $C_{2,0}$ from our analyses altogether.

Swarm shows a number of differences w.r.t. GRACE. First, there is no band around the equator. We believe that this is due to the difference in orbital geometry ($i = 87.4°$ for Swarm [35]) and the better anisotropy of the measurement system. The GRACE KBR only measures the range-rate in the along-track direction. The GPS position observations of the Swarm satellites are isotropic. Together, this leads to better (relative) sensitivity of Swarm to the sectoral coefficients.

Secondly, we see vertical artefacts in the spectral distribution of the Swarm error, which is shown in the right panel of Figure 4.3. It seems that the Stokes coefficients with $m = 0$ or an even number for $m$ have a higher errors. We observe this pattern across the entire spectrum (also up to $l_{max} = 40$, not shown here). We propose that this characteristic of the Swarm fields is an interesting topic for future research, as it could indicate that there is a correlation in the Swarm errors. This could provide an opportunity for improved anisotropic processing of the Swarm fields.

Similarly to GRACE, we observe an increase in error near the poles in the spatial domain, which translates into increased errors in the zonal coefficients in the spectral domain. However, unlike what we saw for GRACE, the spatial distribution of the Swarm errors shown in Figure 4.2 holds when we increase $l_{max}$. Van den IJssel et al. [104] suggest that sensitivity of the Swarm GPS receivers to ionospheric activity degrade the position observations near the geomagnetic poles. We will further cover this sensitivity to increased ionospheric activity, and its effect on the Swarm errors, in Section 4.2.4.

### 4.2.2. Probability distribution

Section 2.6.2 outlined how we use a histogram of the normalized residuals to assess their distribution. We do this because it allows us to assign confidence intervals to the satellite gravity observations and consequently the computed trends, as well as investigate if we may apply certain statistical tests (see Section 2.5.3, Section 3.5). Here, we will first compare

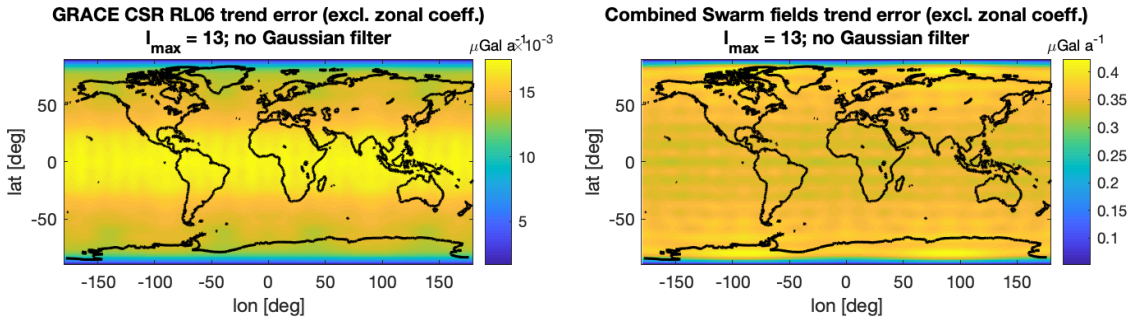



Figure 4.4: Same as Figure 4.2, but after reducing all zonal ($m = 0$) coefficients to 0. NB: as the errors are more than an order of magnitude apart, and as the goal of this figure is to illustrate the spatial distribution of the errors, the panels use a different color scale. NB: the errors values shown here are shown only to illustrate our conclusion regarding the dominance of the low-degree zonal coefficients in the error spectrum when truncating the fields to $l_{max} = 13$ (see text) and are not used anywhere in this study.

the probability distributions of the errors computed via the method of Wahr et al. [114] and the errors computed via our method. We argue in Section 2.6.1 that the method of Wahr et al. [114] fails to account for real inter-annual signals, and therefore overestimates the error. Additionally, this would imply that the resulting errors are not Gaussian distributed. Conversely, this should mean that our method, which does account for long-term signals, yields errors that are (closer to) normally distributed. The difference between the verification we did in Section 2.6.2 and what we will show here is that we use the CSR RL06 GRACE data here. Next, we will compute and assess the residuals of the Swarm fields, after applying our method.

Figure 4.5 presents the results of this experiment. When we compare the residuals from our method to those from the method of Wahr et al. [113], we find that our method leads to residuals that are closer to a normal distribution. This suggests that the additional steps in our method, a moving mean and periodic components estimates from a moving window, are indeed removing more real geophysical signal than the simpler model of Wahr et al. [114]. Nonetheless, even the residuals from our method are not an exact fit to the plotted Gaussian distribution with identical mean and variance. This suggests that either the more intricate error estimation method that we considered does not remove all real geophysical signal, or that the underlying distribution of the GRACE errors is non-Gaussian. The first argument could be motivated by the fact that the periodic hydrology cycles on Earth do not exactly follow a sinusoidal shape, but rather a more skewed periodic function [73]. Since we attempt to model this signal via a weighted combination of sines and cosines, i.e. a model that does follow a sinusoidal shape, it can be expected that this model fails to capture all signal.

The observant reader might have noticed a discrepancy between the results shown in the left panel of Figure 4.5 and the verification of the method of Wahr et al. [113] in Figure 2.6, page 19. Note that in the verification, we used a far shorter time series of GRACE data, so the effect of failing to account for long-term components in the Stokes coefficient time series was not as pronounced there. Nonetheless, the spiking artefacts we see around the middle of the graphs, where the histogram locally extends above the fitted normal distribution, were also seen when we recomputed the results of Wahr et al. [114] (and seen in the paper itself), albeit to a lesser extent. When we look at the residuals of the Swarm fields, we see that they are even closer to being normally distributed than the GRACE residuals. This is an encouraging result for the coming chapters of this report, where we will apply the inferential statistics to further analyze the continuity between the observations GRACE and Swarm of the GIA-induced gravity trend.

### 4.2.3. Errors computed in the spatial domain

We can further motivate the conclusion that we can use the RMSE as standard deviation of the Swarm-derived Stokes coefficients, by comparing the results of aforementioned experiment to results from repeating this experiment in the spatial domain, i.e., applying the method of Section 2.6.1 to time series of gravity observations at a particular map location (instead of time series of Stokes coefficients). Wahr et al. [114] showed that for GRACE this comparison suggests that the correlation of the GRACE errors (which manifests itself as the spatial distribution shown in Figure 4.2) reduced the reliability of the confidence intervals of this method. A simple way to show this is to compare Figure 4.2 to a map of the RMS of the post-fit residuals computed in the spatial domain. Figure 4.6 shows such a map, and we can immediately see that for GRACE (left panel) the spatial distribution of the errors is radically different. We see no band-shaped artefact around the equator, and instead see error values that are concentrated in location where we expect substantial mass variability (e.g., Amazon). On the contrary, we also see that the Swarm-derived errors are much more evenly spaced across the map. It is clear the spatial distribution of the Swarm errors in Figure 4.6 is not as homogeneous as what we see in Figure

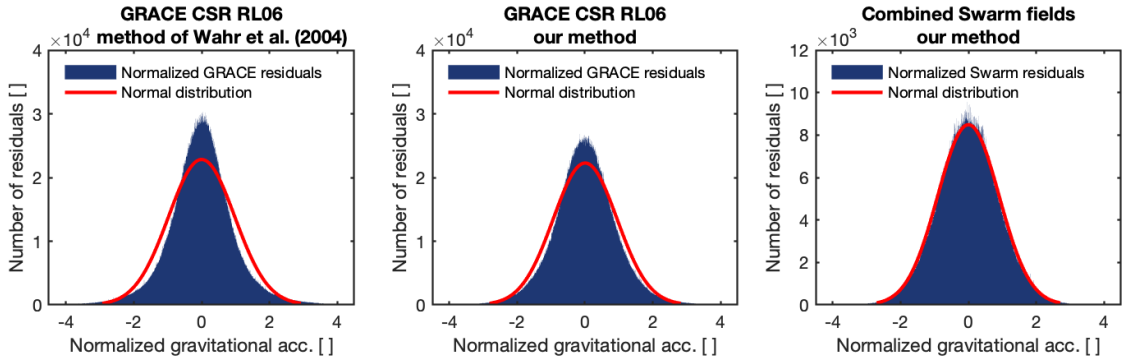



Figure 4.5: Left panel: histogram of normalized GRACE CSR RL06 residuals, computed via method of Wahr et al. [113]. Middle panel: histogram of normalized GRACE CSR RL06 residuals, computed via our method (see Section 2.6.1 on page 15). Right panel: histogram of normalized residuals of the Swarm fields, computed via our method.

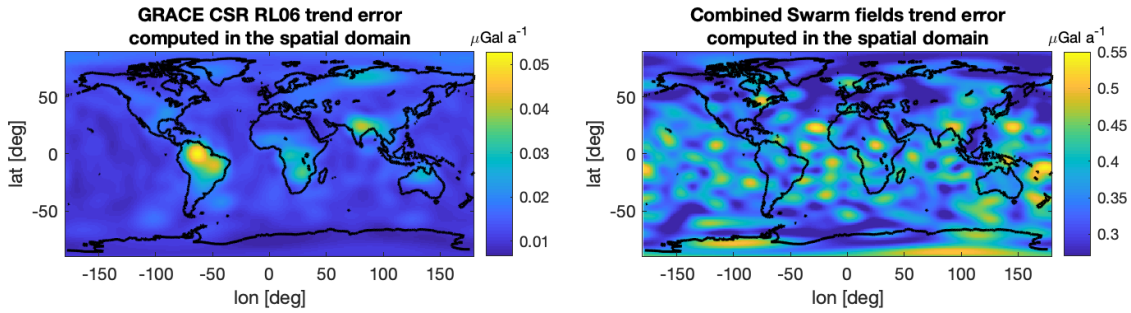



Figure 4.6: Spatial distribution of the errors in GRACE CSR RL06 (left panel) and the Swarm fields (right panel), as computed via the method outlined in Section 2.6.1, this time applied in the spatial domain. NB: as the errors are more than an order of magnitude apart, and as the goal of this figure is to illustrate the spatial distribution of the errors, the panels use a different color scale. Additionally, note that for the same reason the color scales in this Figure are also slightly different from those in Figure 4.2.

4.2, but then again we cannot expect to fully recover the underlying RMS in the case of doing this in the spatial domain, as we have only 64 observations per grid point. Furthermore, we see that the magnitude of the Swarm errors computed in the spatial domain does not correlate to areas of high mass redistribution variability as much as GRACE does (i.e., Amazon does not show an error peak for Swarm). This suggests that the sensitivity of the Swarm measurement system to mass redistribution not well captured by the model we described in Section 2.6.1 (i.e., 1-y moving mean, periodicals from 2-y window) is much lower than for GRACE.

To further substantiate this claim, we look at histograms of the normalized post-fit residuals again, this time computed by applying our error estimated method of Section 2.6.1 in the spatial domain. True mass signals are smaller in magnitude over the oceans than over land (see, e.g., Figure 1.1 on page 2), and therefore we do this analysis separately for the grid points that are located on land and those that are located over the oceans. We define the land area as all grid points within the coast line shown as a black line in, e.g., Figure 4.6. Some shallow water bodies still contain substantial signal [75], and therefore we define the limit of the ocean area mask 800 km offshore from aforementioned coastlines.

The results are shown in Figure 4.7. When we look at GRACE, we see a substantial difference in how the two histograms fit the Gaussian distribution. While both histograms appear to be slightly skewed to the left, the fit to the Gaussian distribution of the ocean area residuals is much closer. The skewness we see here also showed in the histograms in Figure 4.5. The spiking artefacts that we identified there indicated small-magnitude errors occurred more often than they would if the errors where exactly Gaussian distributed. Going back to Figure 4.7, we see that both the land area and ocean area residuals of Swarm are closer to being normally distributed. The land area histogram shows a slight skewness in a similar location to where the GRACE residuals showed such a feature, but the overall shape is still considerably closer to the Gaussian distribution. This suggests that the Swarm fields also have, albeit limited, sensitivity to inter-annual land area signals, likely hydrology variations. Also, this means our method of using the RMSE as standard deviation is more valid for the Swarm-derived fields.

The conclusion we find is equal parts encouraging and discouraging: the Swarm errors might be substantially larger than

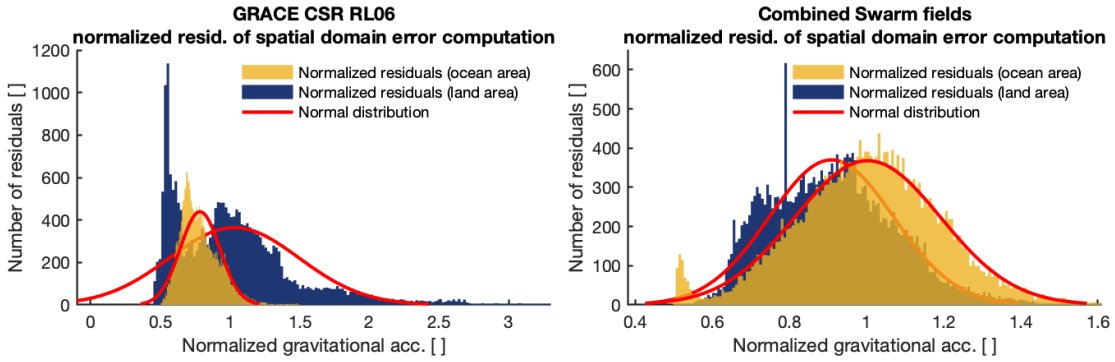



Figure 4.7: Histogram of normalized GRACE CSR RL06 residuals, computed via our method Section 2.6.1, but this time applied in the spatial domain, of GRACE CSR RL06 (left panel) and the Swarm fields (right panel).

the GRACE errors, but at least they are normally distributed. This in turn means that it is unlikely that we will be able to retrieve small-scale geophysical signal from the Swarm data, but conversely it also means that we can assign reliable confidence intervals to the signals that we do find and that inferential statistical tests are valid.

### 4.2.4. Temporal distribution

Following the approach of Wahr et al. [114], we assess the time-dependency of the observation errors of our gravity fields by expanding the residuals of each particular epoch via Equation 2.3, and then computing the latitude-weighted global mean of each epoch. Note that in this case we use the residuals of only this particular epoch, and not the RMS of the residuals. Figure 4.8 shows the temporal distribution of the GRACE and the Swarm errors. The top plot is shown to give an impression of the differences between the GRACE and the Swarm errors. Note that the x-axis of the top plot is set to start in 2009, to increase the readability of the plot. The middle and bottom graph of Figure 4.8 show the time evolution of the GRACE and Swarm errors relative to their mean. When we use the errors in a weighted least squares regressions, it is the relative magnitude of the errors that weights the individual observations. A higher relative error will lead to a lower weight (see also Section 2.5.2).

The GRACE errors (best seen in middle plot of Figure 4.8) mostly stay within a ±50% band of the mean error. The inter-epoch variability of the GRACE error can for a great part be explained by different ground track distributions. The GRACE error seems to increase substantially after ca. the second half of 2016. This is obviously correlated to the fact that one of the GRACE accelerometers was turned off due to power issues in September 2016 [101]. The solutions produced thereafter use data from only one accelerometer for correcting the non-conservative forces [101]. As this affects roughly one sixteenth of the total epochs, we believe the influence of this error increase is minimal. The accelerometer was briefly turned back on in May 2017, allowing for the last two-accelerometer solution, as is clearly visible in Figure 4.8.

The bottom plot of Figure 4.8 shows that the errors in the Swarm field vary through time more substantially than their GRACE counterparts. The Swarm errors in the beginning of the data set are roughly an order of magnitude larger than the GRACE errors. Most authors agree that this is caused by increased ionospheric plasma irregularities during this time, which are known to lead to satellite GPS signal loss (*GPS blackouts*). This error propagates into the derived gravity fields [24, 86, 122]. Van den IJssel et al. [104] allow us to suggest that improved GPS receiver settings aboard the Swarm satellites after 5 May 2015 also help improve the gravity solutions.

If one regards the period after ca. early 2016 as nominal performance, the error at the beginning of the time series are more than twice the amplitude of the nominal errors. Moreover, almost half the Swarm time series (start to ca. 1/2016) have errors that are larger than the errors in the period after 2016. This suggests that weighing the observations is particularly important when performing a regression of the Swarm data. We assess this by comparing trend results of an unweighted and weighted least squares regression of the Swarm data. We truncated the fields at $l_{max} = 13$. We will motivate this choice in the section hereafter.

Figure 4.9 shows the results of this comparison. It appears that weighing the observations in the least squares regression has a negligible impact on the trend result from GRACE but brings the results from Swarm substantially closer to GRACE. The RMS of the trend residual in the area where the GIA-induced trend is dominant, $\Delta^{(GIA)}$, (see Section 3.5) equals 0.2667 $\mu$Gal a$^{-1}$ for the unweighted regression and 0.1189 $\mu$Gal a$^{-1}$ for the weighted regression. We conclude that applying a weighted least squares regression, and not an unweighted one, is particularly important for the Swarm-derived fields, as their errors vary substantially through time (i.e., the observations show heteroskedasticity). Unless noted otherwise, all regressions hereafter are computed via weighted least squares regression, using errors computed via the method in

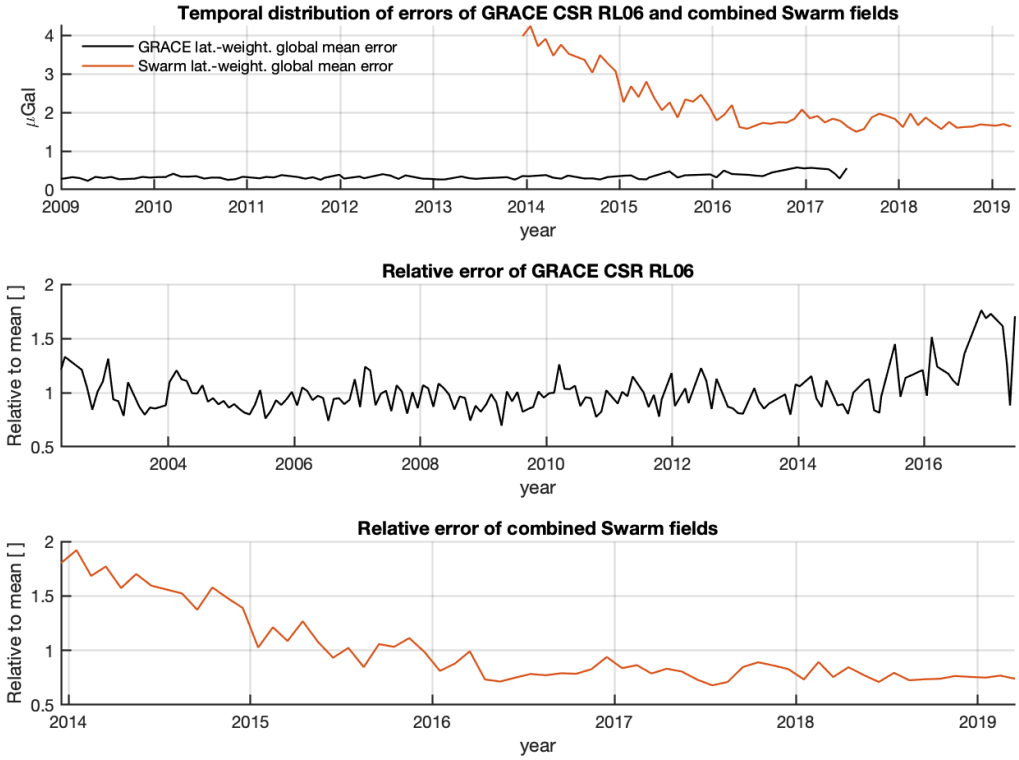



Figure 4.8: Temporal distribution of errors of the GRACE CSR RL06 fields and the Swarm fields. Markers in the top plot indicate individual fields, to illustrate the data gaps in GRACE. This figure shows latitude-weighted global means of the errors as computed via the method in Section 2.6.1. The middle and bottom panel show the same data as the top panel, only relative to their respective means. NB: time axis in the top plot is set to start in 2009 to increase the readability of the figure - the middle panel shows the full GRACE time series.

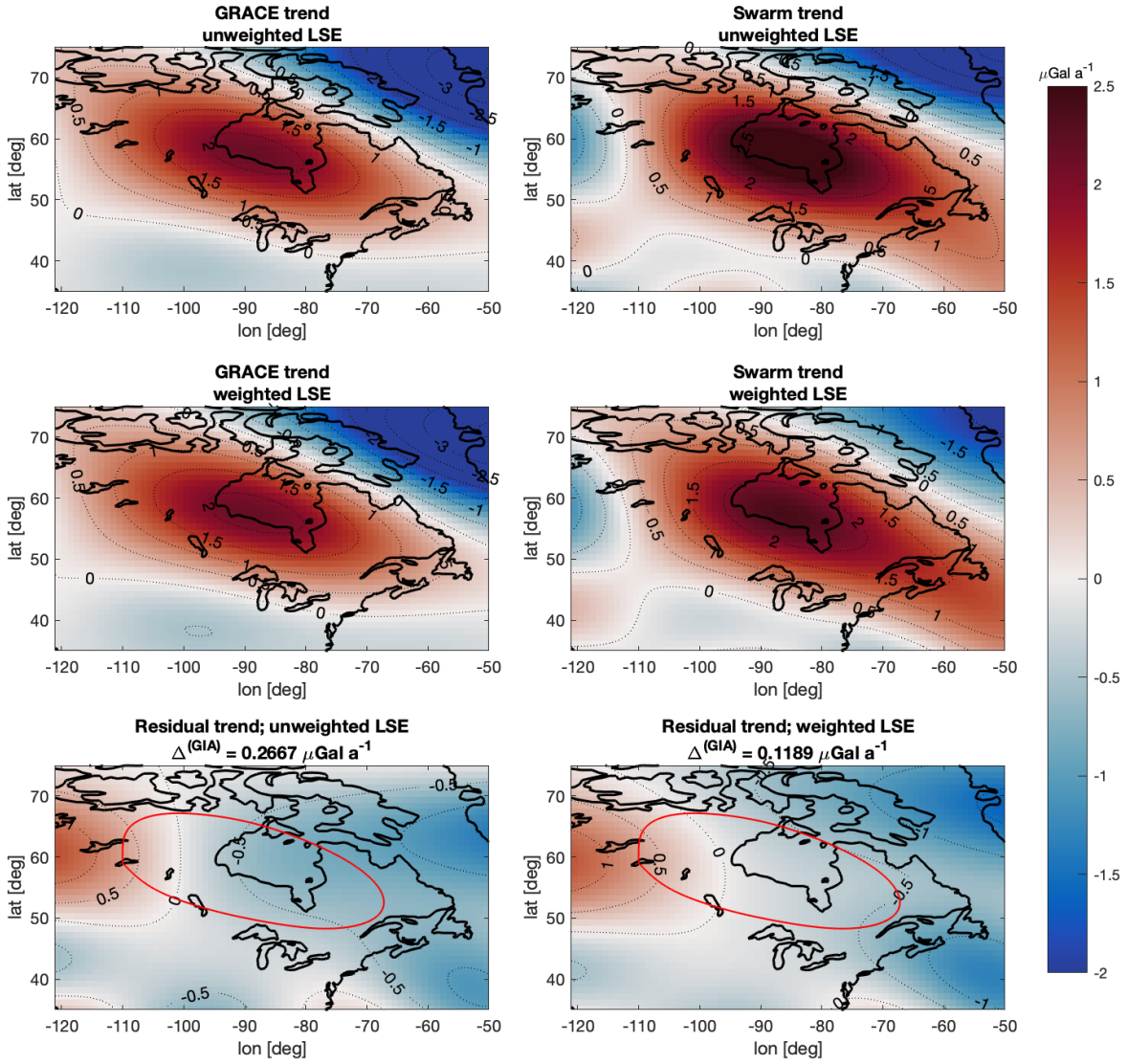



Figure 4.9: Trends computed from GRACE CSR RL06 (left panels), and Swarm fields (right panels). All fields are truncated to $l_{max} = 13$. The top row uses an unweighted least squares approach, and the middle row uses a weighted least squares approach, with weights derived via the method in Section 2.6.1. The bottom row shows the residual trend after subtracting the Swarm trend from the GRACE trend for unweighted regression (bottom left) vs. weighted regression (bottom right). NB: $\Delta^{(GIA)}$ is defined in Equation 3.7. NB: red ellipse indicates the area where the GIA-induced trend is dominant, see Section 3.5.

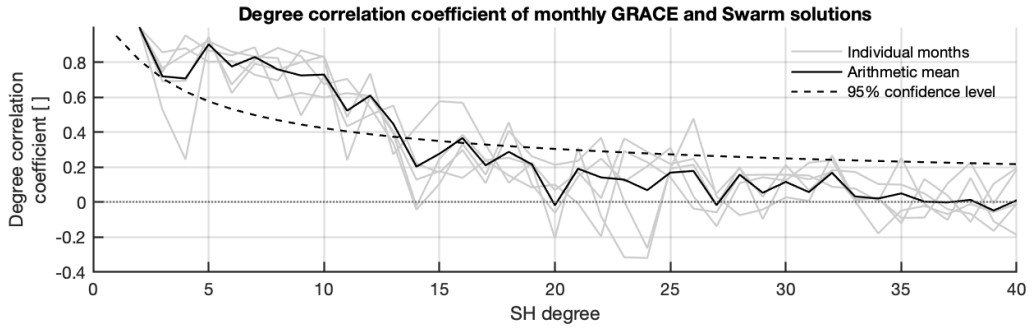



Figure 4.10: Degree correlation between GRACE CSR RL06 and the Swarm fields. Solid black line indicates the arithmetic mean of the degree correlation of the months July, August and September 2015, February and July 2016, and May 2017, and grey lines indicate the individual months. Dashed line indicates the 95% confidence level.

Section 2.6.1.

## 4.3. Noise reduction

The following section shows and discusses results of the noise reduction approaches described in Section 2.4. First, we will do another generalized comparison of the GRACE- and Swarm-derived gravity fields to find a general starting point for the configuration of the noise reduction routines. Then, we simplify the problem to a few key metrics to derive tentative conclusions about the effects of truncating or Gaussian filtering the fields. Finally, we will dive deeper into the full spatial representations of the truncated and filtered geopotential solutions.

### 4.3.1. Choosing a starting point

From Figure 2.2 in Section 2.4 showed that simply expanding the Swarm-derived Stokes coefficients via Equation 2.3 and plotting the results on a map will most likely not yield workable results, as the noise in the higher degrees overpowers the true signal in the lower degrees. Before examining the fields, we need to reduce the influence of high-degree noise, which we will do via either truncation or Gaussian filtering.

To complement the analysis in Figure 2.1 on page 10, we look at the degree correlation between a selection of GRACE and Swarm epochs to determine a starting point for the values of the truncation and Gaussian filtering parameters. The epochs are selected on basis of three criteria aimed at finding an equal comparison of nominal performance. First, the epochs of GRACE and Swarm must be derived from data collected during the same time period. Especially near the end of the GRACE mission the fields do not always cover exactly 1 month [102]. Second, we want months with normal levels of ionospheric activity, as increased ionospheric activity leads to a deterioration of the Swarm-derived fields [e.g. 86, 122]. Finally, we select months where GRACE had two functioning accelerometers (i.e., before September 2016 or May 2017), as the single-accelerometer fields show a slight deterioration [102]. We have chosen July, August and September 2015, February and July 2016, and May 2017. Similar to the approach in Section 2.4, we subtract the GGM05C static gravity model [70] before computing the degree correlation to focus our analysis on the time-variable part of the geopotential solutions.

Figure 4.10 shows a plot of the degree correlation between GRACE CSR RL06 and the Swarm fields of these 5 months. As the goal of this experiment is not to study individual months, but to get a feeling of the correlation between GRACE and Swarm as a whole, we are primarily interested in the mean degree correlation, which is shown in black. We see that the correlation is above 0.7 up to degree 10, but decreases sharply after degree 12. The correlation is significant at a 95% confidence level through degrees 13. From degree 19 on the correlation is under 0.2. We conclude that we will likely find good agreement between GRACE and Swarm through degree 13. This is in line with the conclusion of other authors [10, 24].

Note that there is no perfect test to determine to what degree the Stokes coefficient of a particular field contain true signal, and at what degree they are mostly noise. However, from this experiment we learn that the coefficients above degree 20 are likely to be dominated by noise. Together with conclusions we derived from Figure 2.2 we can now argue that the region between degree 10 - 20 are worth examining in terms of Gaussian filter and truncation configurations.

### 4.3.2. Signal-to-noise ratio

The most intuitive way of assessing the effects of truncation and/or Gaussian filtering on the trends derived from the gravity field time series would be to make spatial maps of the trend in the expanded Stokes coefficients, and compare those visually. However, it is hard to precisely *quantify* the effects via comparison of spatial maps. For this reason, we will first reduce the maps of the expanded Stokes coefficient trends to two key metrics: the latitude-weighted mean trend error and the latitude weighted gravity trend, both computed in the region that is shown in Figure 3.8 (Section 3.5 described how the limits of this region are found). We have chosen these metrics as the purpose of the truncation or Gaussian filtering is to reduce the noise, while leaving the true geophysical signal untouched as much as possible. Reducing each map of the results to two numbers greatly aids the comparison. We compute the trend via weighted least squares regression of the model in Equation 4.1, with weights as computed via the method in Section 2.6.1. We use all available data, i.e., April 2002 to June 2016 for GRACE and December 2013 to March 2019 for Swarm.

We compute the two aforementioned metrics from the GRACE and Swarm fields truncated to a maximum degree that decreases from 20 to 8. The red lines Figure 4.11 show the results of these experiments, where the top left plot shows the latitude-weighted mean gravity trend in the GIA-region (i.e., $\dot{\bar{g}}^{(GIA)}$ as defined in Section 3.5), and top right plots shows the latitude-weighted mean trend error. Swarm is shown as a solid line and GRACE is shown as a dash-dotted line. We have also included the gravity trend from the ICE-6G_D GIA model of Peltier et al. [61] in the analysis (shown as dashed lines). For conciseness sake we include only one GIA model. Next, we repeat the experiment, but apply a Gaussian filter instead of truncation. We choose filter half-widths that that correspond to the maximum degrees chosen for the last experiment (20 to 8). Equation 2.9 describes how the degree number and filter half width are related. The black lines Figure 4.11 show the results of these Gaussian filtering experiments. Finally, both bottom plots in Figure 4.11 show the signal-to-noise ratio (SNR) of the maximum trend in the GIA region, which is defined as the square of the signal (i.e., top right plot squared), divided by the square of the error (i.e., top left plot squared). As the SNR results span ca. 5 orders of magnitude they are spread across two graphs, where the left panel shows Swarm and the right panel shows GRACE. Again, the red line indicates the SNR for various $l_{max}$ of the fields, and the black line indicates various Gaussian filter half widths. Note that the red axes in Figure 4.11, indicating the truncation degree, are linear but the black axes indicating the Gaussian filter radius correspond to the red axes via the inverse relation in Equation 2.9.

All graphs in Figure 4.11 should be interpreted carefully, as we have reduced a complete expansion of a series of Stokes coefficients to just a few data points. For this reason we only aim to derive a tentative direction, and not a definitive conclusion, in terms of truncation and filtering configuration for our next section, where we will look at the complete spatial map.

The GRACE-derived $\dot{\bar{g}}^{(GIA)}$ (top left in Figure 4.11) is barely affected by truncation until $l_{max} = 12$, after which we see an accelerating decrease in amplitude. The trends from the GIA model show a similar behavior, albeit with a less abrupt drop. The signal attenuation of the Gaussian filter (the distance between the black and red line can be interpreted as such) is clearly visible in the satellite data, but appears to be limited in the GIA model case. As the filter radius increases, $\dot{\bar{g}}^{(GIA)}$ decreases at an increasing rate. For Swarm, we see a similar behavior as far as the Gaussian filter. When we look at truncation of the Swarm fields, we see a rather erratic behavior of $\dot{\bar{g}}^{(GIA)}$ for truncation with $l_{max} > 15$. We argue that this is caused by spurious trend results due to increased noise levels in the higher degrees. The amplitude of the Swarm-derived trend is always higher than for GRACE, and is closest to GRACE for maximum degrees 11, 12, and 15. The flat red line in the top left plot suggest that degrees > 12 add little to the trend derived from the GRACE fields, although the results of the GIA model suggest that this limit lies slightly higher, as ca. degree 16. For the Swarm fields the opposite is true, which suggests that this increase in magnitude of the trend derived from the Swarm fields is caused by noise in the degrees > 12.

The top right panel in Figure 4.11 shows that the trend error is ca. two orders of magnitude larger for Swarm than for GRACE for $l_{max} = 20$ or a corresponding Gaussian filter, and that this decreases to ca. one order of magnitude for $l_{max} = 8$ or a corresponding Gaussian filter. For GRACE, the trend error shows similar response to lowering the maximum degree truncation or increasing the half width of the Gaussian filter show a similar decrease in trend error as the field is truncated to a lower maximum degree or if the filter half-width is increased. Gaussian filtering appear to lead to a lower trend error than truncation at corresponding $r_{Gauss}$ and $l_{max}$. For Swarm, truncation leads to a lower trend error than Gaussian filtering for $l_{max} > 10$. is a more effective means of noise reduction than Gaussian filtering until degree 10. As the error in the Swarm fields is substantially higher than the errors in the GRACE fields, and as we are expecting to use the fields at a higher spatial resolution than $l_{max} = 10$, we conclude that this experiment suggests that truncation is a more appropriate method of noise reduction for this study.

The highest SNR (bottom plots in Figure 4.11) differs for both data sets and both noise reduction methods. For truncation, the GRACE SNR maximum is at $l_{max} = 12$, whereas for Swarm this is at $l_{max} = 9$. For the Gaussian filter, the Swarm SNR increased close to exponentially for increasing filter half width. For GRACE, there is a maximum SNR for 770 km. Truncation yields a SNR that is ca. one order of magnitude higher for both data sets, for all $l_{max}$ or corresponding $r_{Gauss}$. We conclude that truncation is likely to be the more appropriate choice for this study, as heavy noise reduction is required

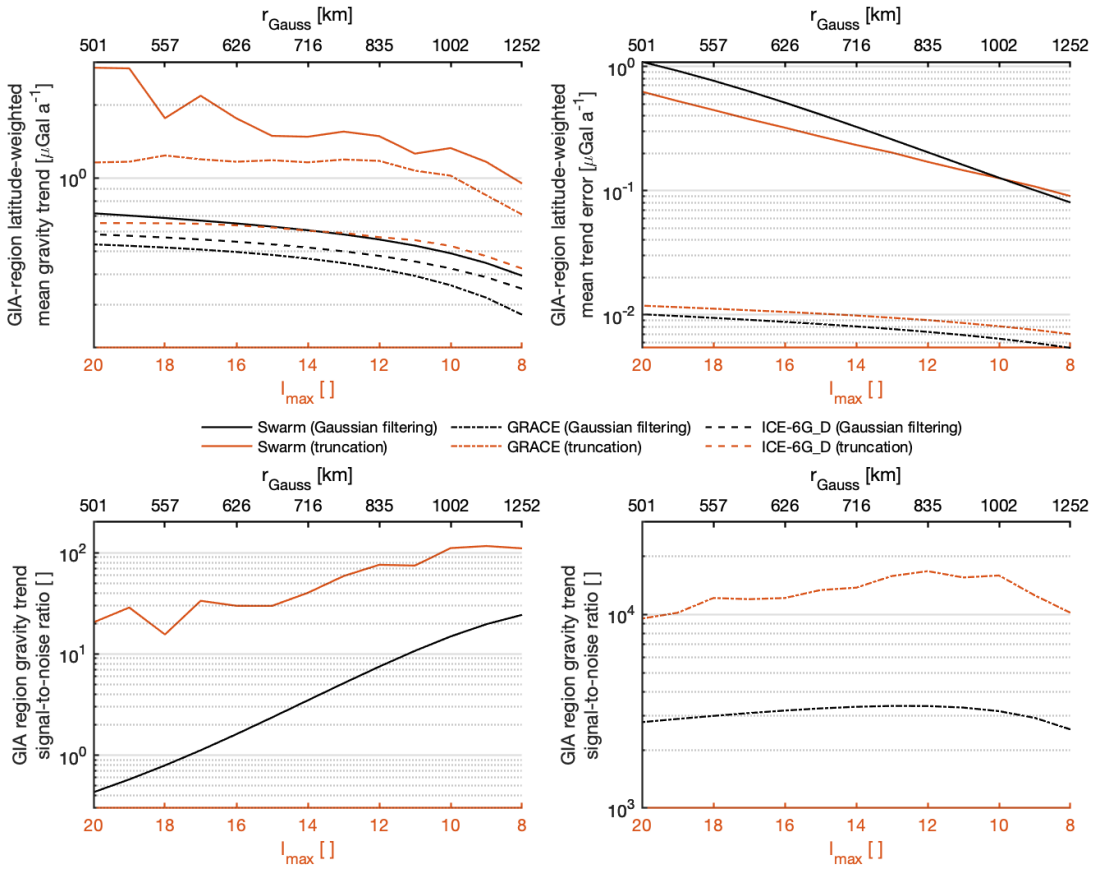



Figure 4.11: Top left: the effect of truncation (red line) and Gaussian filtering (black line) on the latitude-weighted mean gravity trend in the Laurentian GIA region, computed from the Swarm gravity fields, GRACE CSR RL06, or the ICE-6G_D GIA model of Peltier et al. [61]. Top right: the effect of truncation (red line) and Gaussian filtering (black line) on the latitude-weighted mean of the trend error in the Laurentian GIA region, computed from the Swarm gravity fields and GRACE CSR RL06. Bottom: signal-to-noise ratio of maximum gravity trend in GIA region for various truncation and Gaussian filtering configurations of the Swarm fields (bottom left) and GRACE CSR RL06 (bottom right).

to make the Swarm fields usable for geophysical study. This does mean that the spatial resolution of the fields is reduced, and that any conclusion that we find regarding a continuity between GRACE and Swarm will be limited to the coefficients not truncated.

### 4.3.3. Spatial comparison

In Section 4.3.2 we simplified the problem at hand to two metrics, one concerning the error and one concerning the signal of interest, and deduced that truncating the fields to ca. degree 12-15 should allow us to optimally reduce gravity field noise. Before making a final decision regarding noise reduction method and parameter configuration, we will compare spatial maps of the Stokes coefficient trends derived from the GRACE and Swarm gravity field in this section. We will focus on the latitude and longitude domains [35°N, 75°N] and [-120°W, -50°W], as we know the North American GIA-induced gravity trend is most pronounced here. The data and trend computation method are all as outlined in the first paragraph of Section 4.3.2.

The spatial maps are shown in Figure 4.12. Moving from $l_{max} = 15$ to $l_{max} = 12$, we see that there is only minimal change in the spatial distribution and magnitude of the gravity trend observed by GRACE. The only observable change is a slight increase in the north-south dimension of the positive trend anomaly around Hudson Bay, and a decrease in its east-west dimension. For Swarm, we see much more effects of decreasing $l_{max}$ from 15 to 12. For $l_{max} = 15$ there are two peaks in the positive trend anomaly, one on each side of Hudson Bay. This pattern in somewhat reminiscent of the *double bulls eye* pattern of the Laurentian GIA-induced gravity trend that is observed in GRACE solutions truncated to a higher degree [76] or in results from GIA models [61]. The differences in spatial patterns of the $l_{max} = 13$ and $l_{max} = 12$ are minimal, both in magnitude and in spatial distribution. The latitude-weighted RMS of the residual trend between GRACE and Swarm, $\Delta^{(GIA)}$, (see Section 3.5) equals 0.1226 $\mu$Gal a$^{-1}$, 0.1567 $\mu$Gal a$^{-1}$, 0.1189 $\mu$Gal a$^{-1}$, and 0.2101 $\mu$Gal a$^{-1}$ for respectively $l_{max} = 15, 14, 13, 12$. One could be inclined to think this means $l_{max} = 15$ yields almost as good as an agreement between GRACE and Swarm as $l_{max} = 13$, but Figure 4.12 clearly shows a discrepancy in spatial distribution of the trend between GRACE and Swarm for this truncation setting. The fact that the same setting yields a low $\Delta^{(GIA)}$ illustrates how limiting an analysis to a single number can lead to false conclusions, i.e., why we choose to show both the spatial maps and a quantitative analysis such as the $\Delta^{(GIA)}$ or the results in Section 4.3.2. We conclude that $l_{max} = 13$ is the optimal setting to achieve agreement between the spatial distributions of the gravity trend observations of the two satellite missions, and truncate as little information from the field as possible.

Section 4.3.2 shows truncation leads to a higher SNR in both GRACE and Swarm, compared to Gaussian filtering. Nonetheless, this increased SNR comes at the price of a reduced spatial resolution, and as such Gaussian filtering remains a popular noise reduction approach in the scientific community. For completeness sake we also present the results of applying a Gaussian filter to the Swarm-derived trends. First, we truncate the GRACE fields to degree 40, to match the Swarm fields. The filter half width settings are chosen so that they roughly correspond to our choices for $l_{max}$ in Figure 4.12, but also cover filter half widths common in literature [24, 46]. We decide on 500, 668, 716, and 835 km, which corresponds to degrees 20, 15, 14, 12, respectively. Finally, we compute the gravity trend via the same method as in the truncation experiment.

Figure 4.13 shows the gravity trends from GRACE and Swarm after applying the Gaussian filter. For GRACE, we see that the signal attenuation of Gaussian filtering becomes very apparent when we compare Figure 4.12 and Figure 4.13. $\Delta^{(GIA)}$ equals 2.763 $\mu$Gal a$^{-1}$, 0.2671 $\mu$Gal a$^{-1}$, 0.1864 $\mu$Gal a$^{-1}$, and 0.1299 $\mu$Gal a$^{-1}$ for respectively 500, 668, 716, and 835 km half width settings. The $\Delta^{(GIA)}$ for $r_{Gauss} = 835$ km comes close to the $\Delta^{(GIA)}$ for $l_{max} = 13$ we found in the last paragraph (= 0.1189 $\mu$Gal a$^{-1}$). Nonetheless, the clear discrepancy between GRACE and Swarm in spatial distribution of the trend shown in Figure 4.13 leads us to conclude that the Swarm trend results are still substantially affected by high-degree noise, which makes Gaussian filtering a less appropriate approach for this study than truncation.

Finally, we conclude by taking into account the findings so far and by comparing the GRACE- and Swarm-derived trend in Figure 4.12 and Figure 4.13 that truncation yields better agreement between the trend observations of GRACE and Swarm. We keep our earlier conclusion of applying truncation to $l_{max} = 13$ as a method of noise reduction. We apply the same processing to GRACE and Swarm, even though the GRACE data contains less noise than the Swarm data, to achieve an equal comparison.

## 4.4. Comparing Swarm to a longer time series of GPS-derived solutions

Before attempting to isolate the GIA-induced component from the gravity trends that we derived in Section 4.3, we will apply the methodology used to far to a second time series of GPS-derived gravity fields. We use the GeoQ fields that were introduced in Section C.

We motivate our substitution of the GeoQ fields for the Swarm fields in this experiment by studying the degree correlation

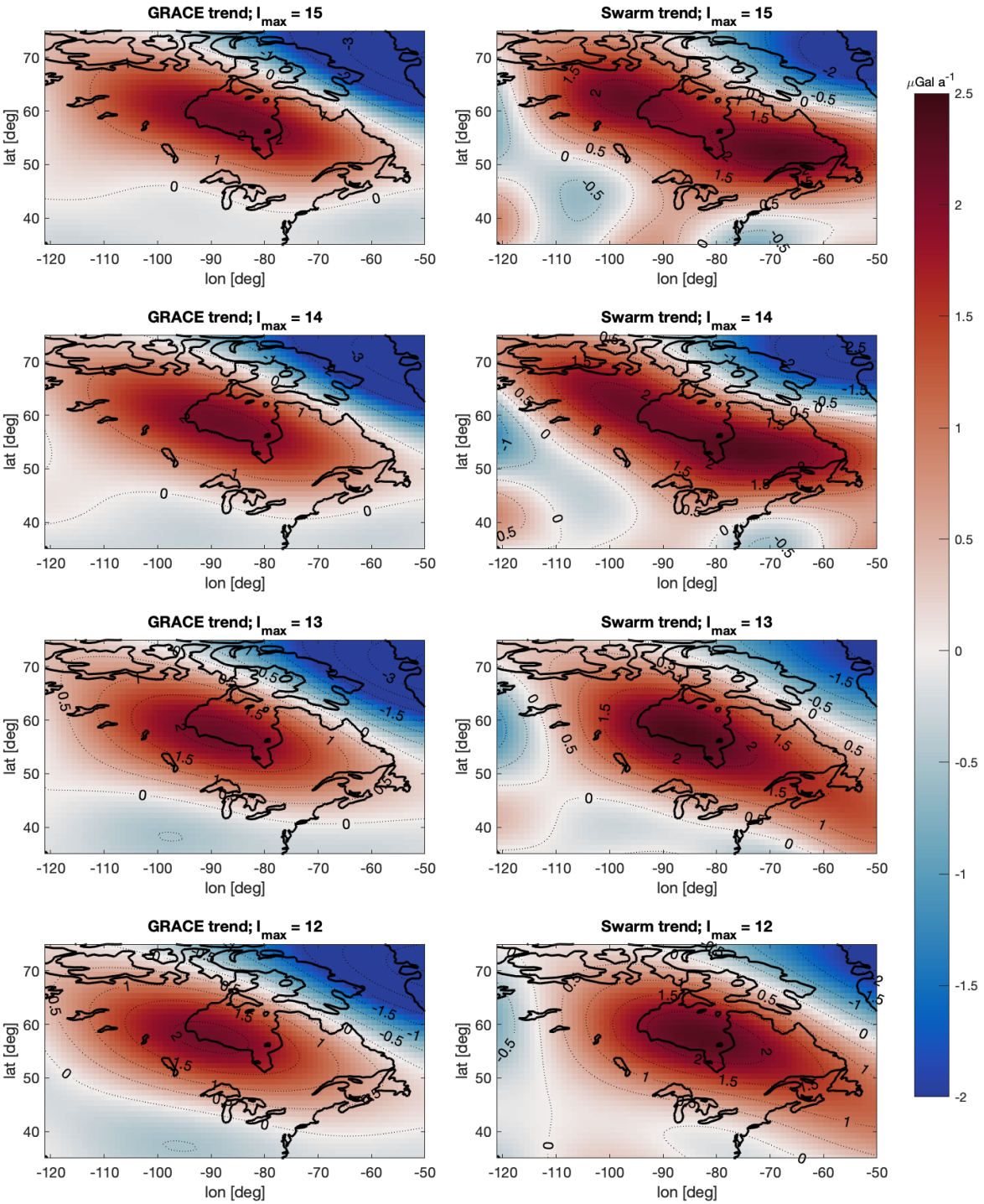



Figure 4.12: Gravity trend computed from GRACE CSR RL06 (covering 4/2002 - 6/2017) or Swarm fields (covering 12/2013 - 3/2019) via weighted least squares regression of a bias, trend, and (semi-)annual component, for fields truncated to a maximum degree of 15, 14, 13, or 12.

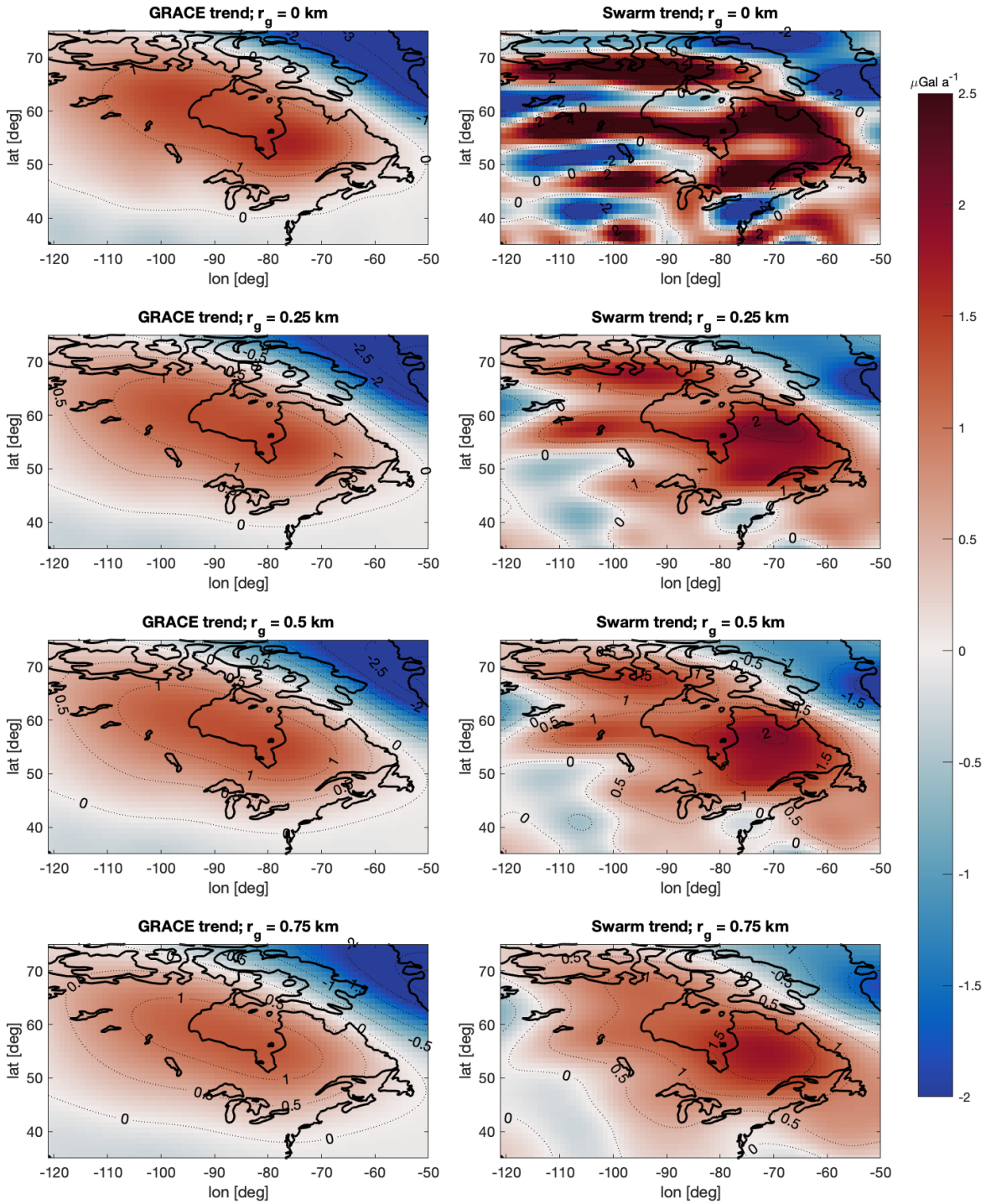



Figure 4.13: Gravity trend computed GRACE CSR RL06 (covering 4/2002 - 6/2017) or Swarm fields (covering 12/2013 - 3/2019) via least squares regression of a bias, trend, and (semi-)annual component, for fields filtered with a Gaussian filter with a 500, 668, 716, or 835 km radius.

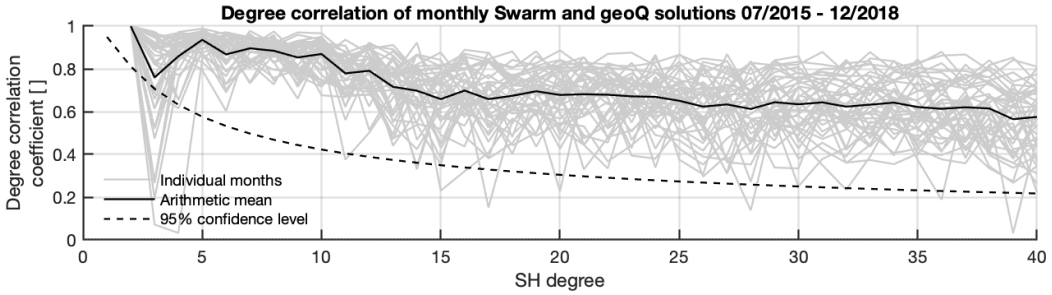



Figure 4.14: Degree correlation between the Swarm fields and the GeoQ fields. black line indicates the arithmetic mean of the degree correlation of the months July to December 2017, and grey lines indicate the individual months. Dashed line indicates the 95% confidence level.

of the time-variable part of the gravity fields again (i.e., degree correlation after removing a static model). The experiment shown in Figure 4.10, which compared GRACE and Swarm, followed a similar approach. For the comparison of Swarm and GeoQ in this section we again take the mean correlation coefficients of months where we determine that the fields show nominal performance. Here, we apply the additional constraint of choosing months where the GeoQ fields are free of GRACE GPS position observations. We use all months from July 2017 to December 2018. Figure 4.14 shows that for degrees 2-10 the correlation between Swarm and GeoQ is high (>0.8). The only exception is degree 3, where the decreased correlations likely is due to the SLR information in the lower degree coefficients of the GeoQ fields. For degrees 11-13, the correlation decreases but remains significant throughout the entire degree range. Be careful in comparing Figure 4.14 to Figure 4.10, as the first compares Swarm and GeoQ (which we argue and show are very similar as Swarm dominates the GeoQ epochs that we use), whereas the second compares GRACE and Swarm (two independent data sets). We conclude that the GeoQ and Swarm field are sufficiently similar to allow substitution of the GeoQ fields for the Swarm fields for this experiment. This conclusion needs the qualification that we have based it on only a part of the GeoQ solution, even though we will be using the entire data set for this experiment.

### 4.4.1. GeoQ-derived observations of the GIA-induced gravity trend

The following section shows gravity trends derived from the GeoQ fields. We use the regression model described in Section 4.1 with uncertainties computed via the method described in Section 2.6.1. The GeoQ uncertainties are very similar to Swarm when it comes to their spatial, spectral, and probability distribution. We provide a brief analysis of the spatial, spectral, and probability distribution of the GeoQ uncertainties in Appendix C. We use all 192 available monthly epochs of the GeoQ time series, which means the GeoQ time series now contains GRACE GPS position observations and GRACE accelerometer data (see also Section C). This, and the fact that GeoQ and the combined Swarm fields use different processing approaches (again see Section C) means we should be careful in extending any conclusion we derive from the GeoQ fields to GRACE. Similar to what was shown for GRACE and Swarm in Figure 4.12, we examine the gravity trend in North America for various $l_{max}$. For conciseness sake we will not show the results of various Gaussian filter configurations, but note that Weigelt et al. [117] recommends using a 750 km half width. In the second part of this section we repeat this analysis for a subset of the GeoQ time series that is equal in length to Swarm.

Figure 4.15 compares the GeoQ-derived trends to GRACE-derived trends for equal, arbitrarily chosen $l_{max}$. Compared to GRACE (i.e., in Equation 3.7 replace Swarm with GeoQ), $\Delta^{(GIA)}$ equals 0.0567, 0.0119, 0.0142, and 0.0125 $\mu$Gal a$^{-1}$, for respectively $l_{max} = 25, 22, 19, 16$. Aiming to be as consequent as possible we still use the same area to compute the RMS (i.e., Figure 3.8), even though we now analyze higher spatial resolution solutions. Note from Figure 4.15 and the $\Delta^{(GIA)}$ how, even for $l_{max} = 22$, both the spatial distribution and the amplitude of the gravity trend shows good agreement with GRACE. For $l_{max} = 25$ the trend peak east of Hudson Bay appear to be overestimated by GeoQ w.r.t. GRACE. Of course, a word of warning is warranted for the interpretation of these results as the two data sets in this experiment are not independent. Additionally, the data from which we derived the trends in Figure 4.15 covers mostly the same time period (4/2002 - 6/2017 for GRACE, 1/2003 - 12/2018 for GeoQ), which means any non-stationary inter-annual variability of the gravity field (e.g., hydrology) will affect both trend estimates in mostly the same way.

Figure 4.16 shows the trend estimates for the subset of GeoQ fields equal to the length of the Swarm fields. We use September 2013 through December 2018. As expected, we had to apply more rigorous truncation to reduce the influence from noise and find agreement between GeoQ and GRACE. The $l_{max}$ at which we can recognize the positive trend anomaly around Hudson Bay are similar to the results for Swarm. Note that Figure 4.12 shows GRACE- and Swarm-derived for the same $l_{max}$. The GeoQ trends differ more from GRACE than Swarm does, both when it comes to the magnitude of the trend component and the spatial distribution thereof. Note how the positive trend anomaly around Hudson Bay appears

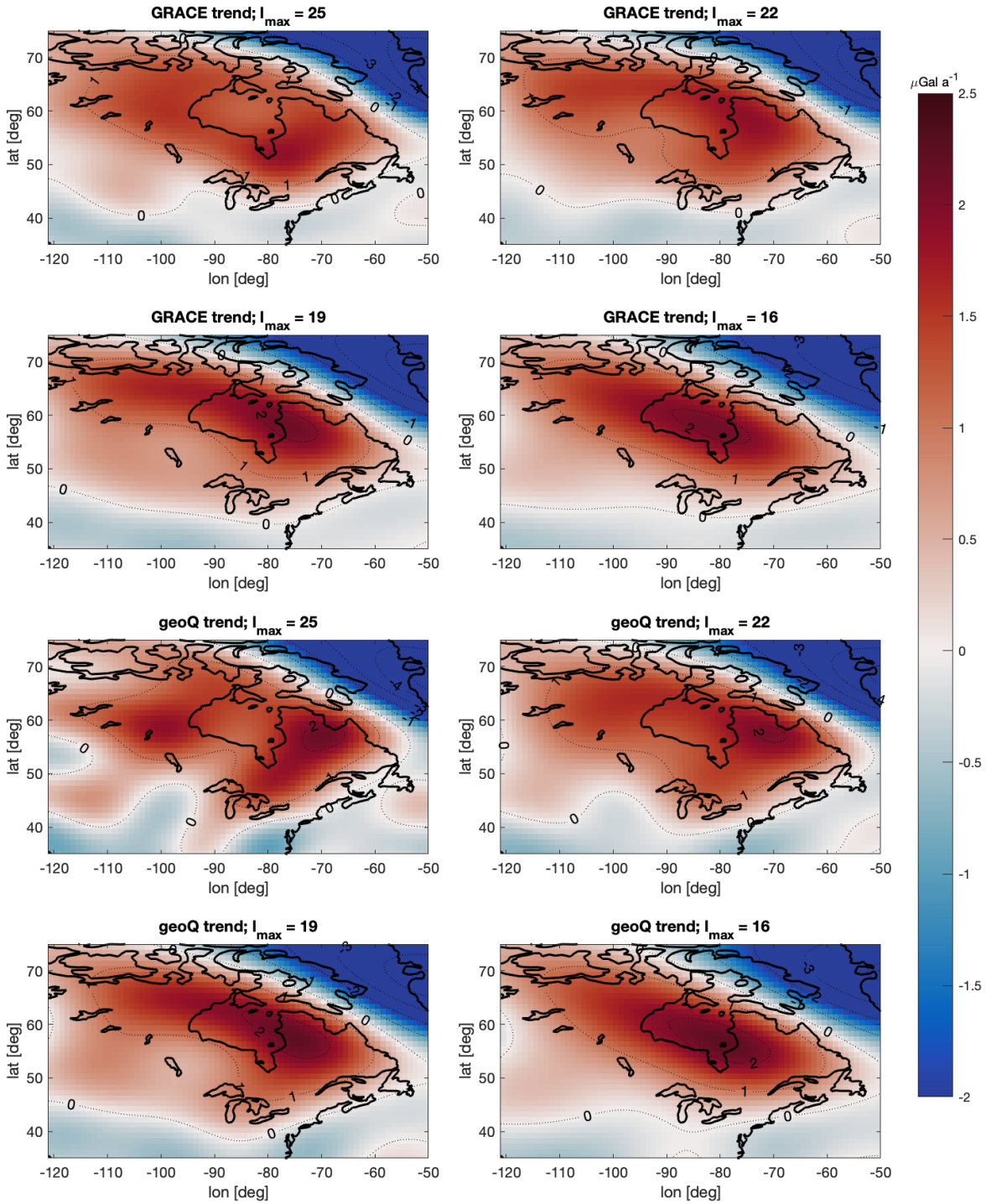



Figure 4.15: Gravity trend computed from GRACE CSR RL06 (covering 4/2002 - 6/2017) or the GeoQ multi-satellite fields (covering 1/2003 - 12/2018) via weighted least squares regression of a bias, trend, and (semi-)annual component, for fields truncated at various maximum degrees.

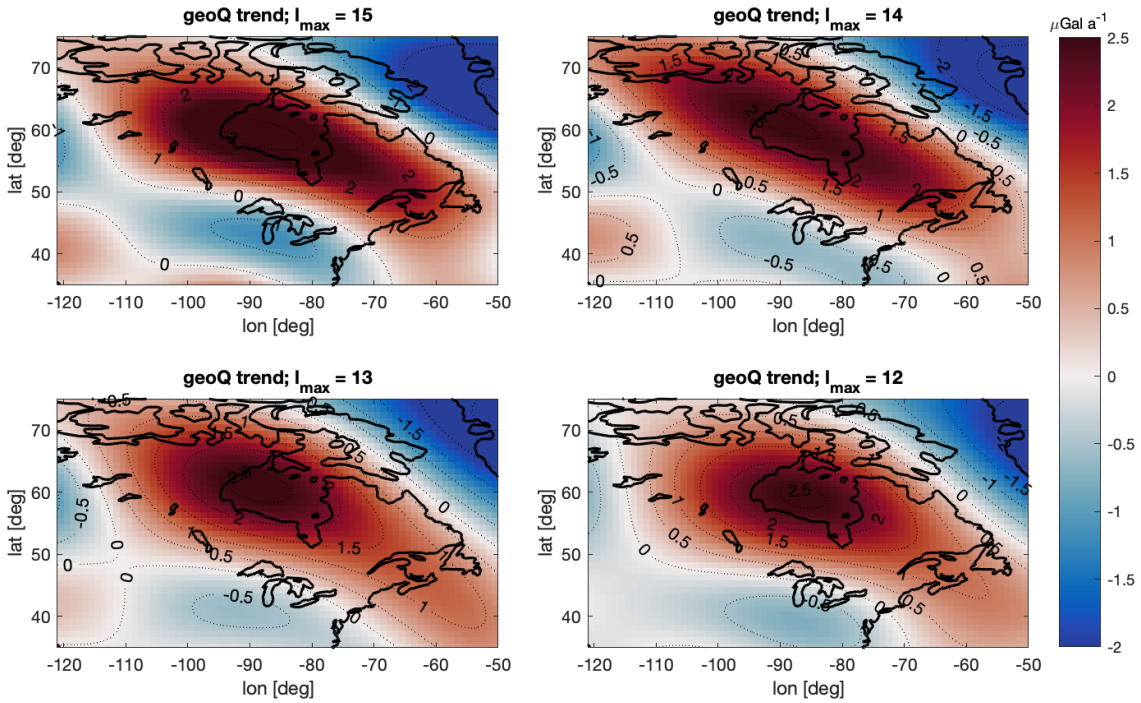



Figure 4.16: Gravity trend computed the GeoQ multi-satellite fields (covering 9/2013 - 12/2018) via weighted least squares regression of a bias, trend, and (semi-)annual component, for fields truncated at various maximum degrees. NB: Figure 4.12 shows GRACE- and Swarm-derived trends for identical processing steps.

to be shifted to the north-east w.r.t. GRACE or Swarm. For respectively $l_{max} = 15, 14, 13, 12$, we compute $\Delta^{(GIA)}$ (again w.r.t. GRACE data 4/2002-6/2017, i.e., the comparison performed for Swarm in Section 4.3.3). These values are listed in Table 4.2. Also given in this table are the values for $\dot{g}^{(GIA)}$ and the maximum gravity trend $\dot{g}$ in the GIA-region (as in Figure 3.8). First, note how $\Delta^{(GIA)}$ is larger for GeoQ, except when $l_{max} = 12$. Also, note how the values for $\dot{g}^{(GIA)}$ of GeoQ and Swarm are rather close, especially for $l_{max} = 14, 13$. We argue that this is only because the peak of the GeoQ-observed trend has shifted and we still use the same definition of the GIA-region. The fact that the maximum $\dot{g}$ (bottom row Table 4.2) are all substantially higher than those of GRACE supports this explanation. For every $l_{max}$ shown in the table the maximum $\dot{g}$ of Swarm is closer to GRACE than GeoQ.

In summary, we have seen that a 64 mo subset of the GeoQ fields yields roughly the same trend results as the Swarm fields, albeit that Swarm is closer to GRACE in trend amplitude and spatial distribution. However, when we use the full 192 mo GeoQ data set, we find that we can derive trends that align with GRACE at much higher spatial resolutions, up to degree 22 or ca. 900 km. We conclude that trends derived from a longer (i.e., 192 mo vs. 64 mo) GeoQ time series agree with GRACE up to a higher spatial resolution. This suggests the influence of noise (in the higher degrees) reduces for longer time series. When we extend this conclusion to Swarm, we can make the careful prediction that for a future, longer time series of combined Swarm fields the maximum usable spatial resolution for gravity trends (that we now find

| | **GeoQ** (9/2013-12/2018) | | | | **Swarm** (12/2013-3/2019) | | | | **GRACE** (4/2002-6/2017) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $l_{max}$ | 15 | 14 | 13 | 12 | 15 | 14 | 13 | 12 | 15 | 14 | 13 | 12 |
| $\Delta^{(GIA)}$ w.r.t. GRACE | 0.3581 | 0.2089 | 0.1803 | 0.1646 | 0.1226 | 0.1567 | 0.1189 | 0.2101 | n/a | n/a | n/a | n/a |
| $\dot{g}^{(GIA)}$ | 1.987 | 1.751 | 1.649 | 1.659 | 1.716 | 1.761 | 1.691 | 1.584 | 1.528 | 1.523 | 1.533 | 1.510 |
| Maximum $\dot{g}$ in GIA-region | 3.175 | 2.566 | 2.617 | 2.555 | 2.330 | 2.333 | 2.398 | 2.327 | 2.169 | 2.142 | 2.158 | 2.133 |

NB: all units are $\mu$Gal a$^{-1}$.

Table 4.2: Amplitude metrics of the Hudson Bay gravity trend computed from GeoQ, Swarm, and GRACE. NB: $\Delta^{(GIA)}$ and $\dot{g}^{(GIA)}$ are as defined in Section 3.5.

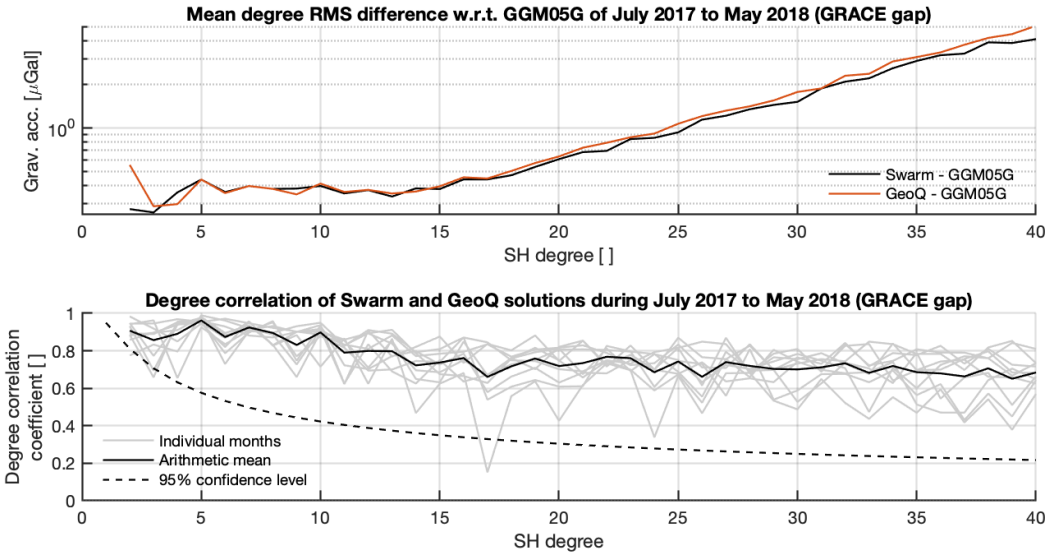



Figure 4.17: Degree RMS difference of the Swarm fields (black line) and the GeoQ multi-satellite fields (red line) w.r.t. the GGM05G static gravity model during the GRACE gap (July 2017 to May 2018).

equals $l_{max} = 13 \cong$ ca.1500 km) will likely increase.

### 4.4.2. Comparing Swarm and GeoQ during the GRACE gap

Much of the current interest in GPS-derived geopotential solutions is due to fact that they are likely to be (part of) the solution to the GRACE gap. However, it is very likely that the majority of those interested in an Earth mass transport product will resort to the GRACE fields for the period April 2002 to June 2017, and (in the near future) to GRACE-FO fields for the period after June 2018. To a certain extent, this makes the bulk of the GPS-derived geopotential solutions obsolete in the near future. Of course, there are some research areas (e.g, very low degree Stokes coefficients) where observations from GPS (or SLR) seem to perform better than the ll-SST results of GRACE, and there will always be studies with a specific research interest in GPS-derived geopotential solutions, but we will leave such highly specific applications out of account for now.

As such, it really only makes sense to compare the Swarm fields and the multi-satellite solutions during the specific time period of the GRACE gap (July 2017 to May 2018). Additionally, this yields a fairer comparison between the two as GeoQ does not contain any position observations or accelerometer data of the GRACE satellites. In light of the research goal of Weigelt et al. [116, 117] of computing the highest quality geopotential solutions free of ll-SST information including GRACE GPS and accelerometer data makes sense, but it does give a somewhat biased view to anyone expecting that solutions of this quality will be available when no GRACE data is available.

The top panel of Figure 4.17 shows the degree difference of the Swarm fields and the GeoQ multi-satellite fields w.r.t. the GGM05G static gravity fields model during the GRACE gap. This is similar for what we showed for GRACE and Swarm in Figure 2.2, but this time we take the arithmetic mean over all the epochs in the GRACE gap (July 2017 - May 2018). To illustrate the fact that the GeoQ and Swarm solutions contain similar but not identical information, Figure 4.17 also shows the degree correlation for the GRACE gap epochs. The degree difference discrepancies between Swarm and GeoQ between degrees 2 to 6 are hard to interpret due to the fact that the GeoQ solutions also contain information derived from SLR observations in those degrees. Similarly, the differences we see for degrees 7 to ca. 14 can both be caused by increased noise or by an increased magnitude in the time-variable part of the gravity field that Swarm and GeoQ both observe. We know those degrees contain predominantly real geophysical signal for both solutions, but we do not know which solution is more accurate. Instead, the only valid conclusion we can draw from the degree RMS graph in Figure 4.17 concerns the higher degrees, where we see a small but consistent difference between Swarm and GeoQ that suggests the GeoQ solutions contain more noise in those degrees. Of course, this conclusion needs the qualification that most of these degrees have a signal-to-noise ratio below unity [23], which means the signal is unlikely to be usable for geophysical studies. Nonetheless, the consistently lower noise does suggest that the method of combining Swarm fields derived via various inversion procedures is a more effective approach than using position observations of other satellites than the Swarm satellites and only using the accelerations approach.

In conclusion, the results of the last section suggests that the length of a (GPS-derived) time series of geopotential solutions has a substantial effect on the gravity trends that can be derived. Velicogna et al. [110] suggest the same applies to GRACE, which the authors found when using GRACE data to quantify the acceleration of high-latitude ice loss processes. Results from our experiments suggest that longer time series provide an opportunity to derive credible trends up to a higher spatial resolution. Also, we confirm the findings of Arnold et al. [4], who suggest including position observations from other satellites into the solutions helps mitigate the errors due to the Swarm system's sensitivity to ionospheric plasma gradients. Nonetheless, we have also shown that in the period where there are no other geopotential products available that can offer the spatial resolution of GPS-derived fields (i.e., the GRACE gap), the Swarm solutions show a small but consistent improvement in high degree noise w.r.t. the GeoQ solutions. Any added benefit that the GeoQ fields offer due to the information from SLR observations in the very low degree coefficients can be extended to Swarm by applying a similar methodology to combine hl-SST and SLR observations in a future study. Additionally, we suggests combining the method de Teixeira da Encarnação et al. [23] (i.e., applying four different inversion procedures and combining solutions) and the method of Weigelt et al. [116] (i.e., including other orbiters than Swarm and combining with SLR) as focus of a second future study. We end this section by declaring that in the current configuration the Swarm solutions pose the better solution to bridge the GRACE gap.

## 4.5. Trend variability

The following section covers the results of the trend variability experiment described in Section 3.1. We will first look at the GRACE data by itself. As this data is substantially less noisy (see, e.g., Figure 2.2 or 4.11) than the Swarm-derived fields and has a longer time coverage, it provides us the best chance of deriving any conclusions about the true trend variability in the long term. We will then attempt to extend these conclusions to the Swarm data by comparing GRACE to the GeoQ fields again. We use the GeoQ fields and not the Swarm fields as the length of the entire time series greatly affects our trend variability experiment. For example, if we use a window length of 5 years, the GRACE or GeoQ data would yield ca. 3 non-overlapping subsets, whereas 5-year subsets from Swarm would approximately comprise the entire Swarm data set. Figure 4.9 showed the importance of using weighted regression, especially in the case of the Swarm data. For the trend variability experiments we therefore also use weighted least squares regression. The weights are derived once for the entire time series, as described in Section 2.6.1.

### 4.5.1. Variability of mean GIA-region trend

For a more concise analysis, we will first apply the trend variability experiment and compute the latitude-weighted mean trend of the area in Figure 3.8 (NB: same area as used for Figure 4.11 or $\Delta^{(GIA)}$). We use a 3-, 5-, or 7.5-year window length. Note that this means having 5, 3, and 2 independent (i.e., free of overlap) time windows, respectively. Also note how 3 years corresponds to the experiment of Van der Wal et al. [106] and 5 years corresponds roughly to the length of the Swarm data set. Figure 4.18 shows the results. The top panel shows the results of using GRACE data and $l_{max} = 60$, the middle panels shows GRACE data and $l_{max} = 13$ and the bottom panel shows GeoQ data and $l_{max} = 13$. To make sure our results are not affected by short-wavelength noise in the $l_{max} = 60$ case, we apply a 300 km radius Gaussian filter. For equal comparison we do the same to the other two panels. The data points indicated by the lines are time stamped with the mean of the time stamps of the earliest and latest satellite epochs used in that particular window. E.g., for a continuous data set of monthly epochs, a data point on the 5-year window line marked mapped to 1 January 2010 in Figure 4.18 would be computed from satellite data of 1 July 2007 through 1 June 2012. For the same data set, a data point of the 7.5-year window line mapped to 1 September 2014 in Figure 4.18 would be computed from satellite data of 1 January 2010 through 1 June 2017. Note how the time axes of the three plots are equal. Aforementioned time stamping convention thus means that GRACE and GeoQ data points that are on the same vertical line in Figure 4.18 use the same start and end date for the window from which the mean trend is computed for the particular data point.

The most striking conclusion from Figure 4.18 is that the trend variability is very clearly a function of the window length, and that a longer window length reduced the trend variability considerably for all data sets and spatial resolutions. Even though we use weighted averages and not trend maxima, we find similar differences between the smallest and largest trend as Van der Wal et al. [106] for a 3-year window. Interestingly, this difference is almost equal for the GRACE $l_{max} = 60$ and the $l_{max} = 13$ case (0.756 $\mu$Gal a$^{-1}$ and 0.736 $\mu$Gal a$^{-1}$, respectively). In fact, when we move to longer window lengths, we find that the difference between the smallest and the largest trend in any window is slightly higher for the $l_{max} = 13$ case. For a 5-year window we find 0.235 $\mu$Gal a$^{-1}$ vs. 0.329 $\mu$Gal a$^{-1}$ and for a 7.5-year window we find 0.196 $\mu$Gal a$^{-1}$ vs. 0.207 $\mu$Gal a$^{-1}$, both using $l_{max} = 60$ and $l_{max} = 13$, respectively. For the $l_{max} = 13$ GeoQ data these numbers are substantially higher at 1.62 $\mu$Gal a$^{-1}$, 0.776 $\mu$Gal a$^{-1}$, and 0.410 $\mu$Gal a$^{-1}$ for respectively a 3-, 5-, and 7.5-year window.

Note how the lines of the middle and bottom panels in Figure 4.18 appear to have a positive slope, especially in the 5- and 7.5-year window case. This would suggest that the gravity trend in the Hudson Bay area is accelerating (or at least the trend of the coefficients up to degree 13), which would undermine our idea of extrapolating the GRACE observations

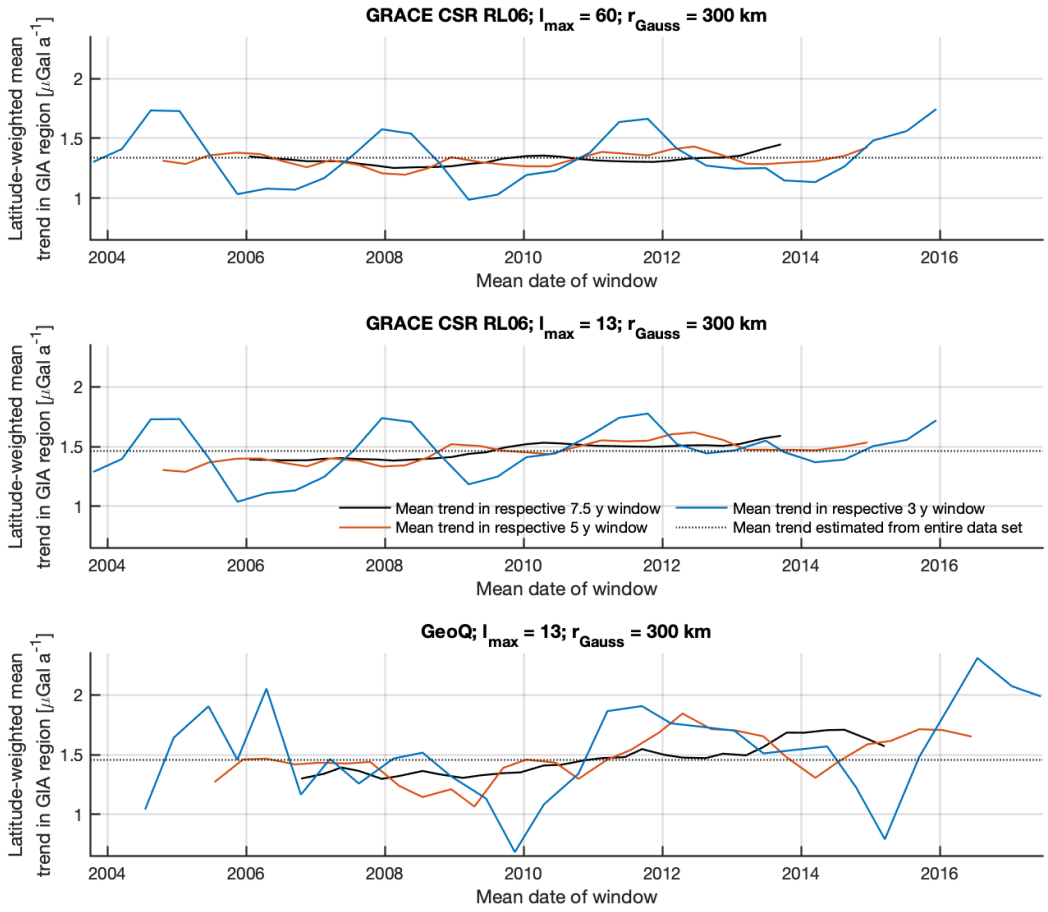



Figure 4.18: Time evolution of latitude-weighted mean gravity trend in Hudson Bay area as derived from GRACE CSR RL06, ($l_{max}$ = 60; top panel and $l_{max}$ = 13; middle panel) and GeoQ ($l_{max}$ = 13; bottom panel). In all cases a 300 km radius Gaussian filter was applied. All panels and all lines use the same area to compute the mean trend from, that is limited by a >+1.0 $\mu$Gal a$^{-1}$ gravity trend in the entire GRACE CSR RL06 data set with $l_{max}$ = 13, i.e., the area in Figure 3.8.

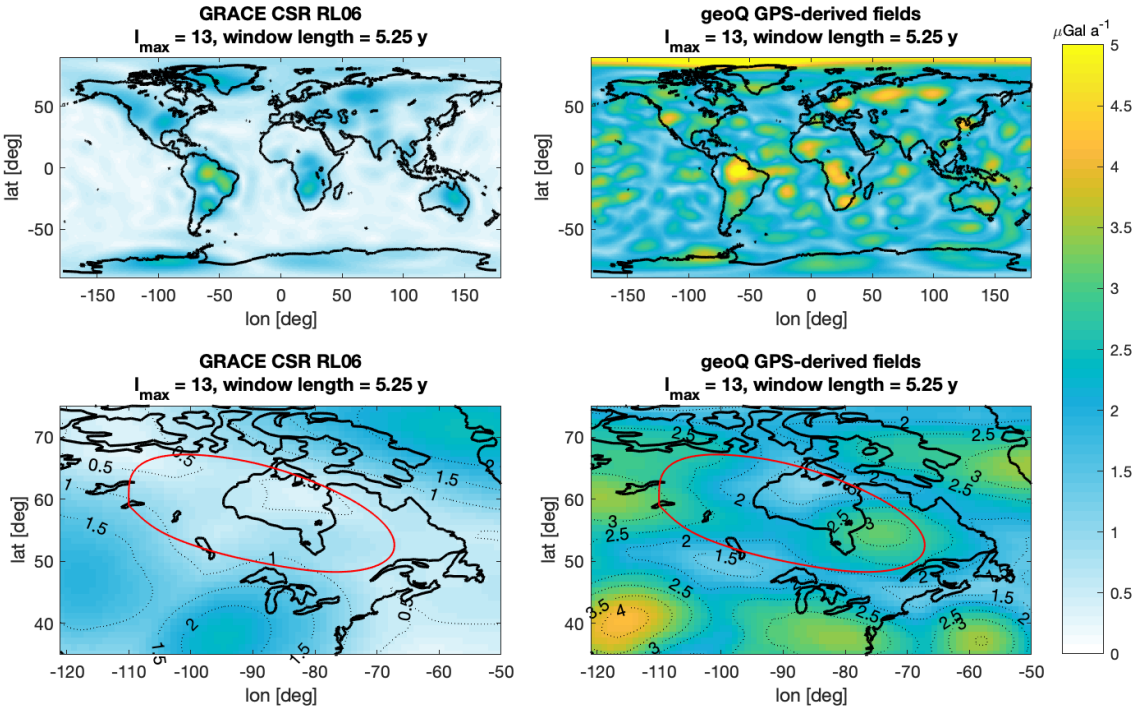



Figure 4.19: Trend variability analysis of the GRACE CSR RL06 and the GeoQ GPS-derived satellite fields. The fields are truncated to $l_{max}$ = 13, and we use a moving window with a length of 5.25 years to estimate the trends. Red ellipse indicates the area where the GIA-induced gravity trend is dominant, see Section 3.5. NB: color scale is not the same as in Figure 4.20, as increasing the window length decreases trend variability.

of this trend into the Swarm time period. Also, for the first half of the data set, the 3-year window line appears to follow a periodic behavior with a period of ca. 3 year. Taking into account the context of this study and the limited amount of available data, we argue that both these conclusions need strong qualification. First, the trend variability seen in the 3- and 5-year window plots, regardless of $l_{max}$, is likely well within the uncertainty limits of the Swarm data. In that case, any time-variability of the gravity trend would be unlikely to lead to a statistically significant discrepancy between GRACE- and Swarm-derived observations. Second, we argue that the data set at hand is too short to derive any conclusions about the long-term variability (be it periodicity or acceleration) of data derivatives such as the trend. To derive such conclusions, longer time series are needed. Third, keep in mind that the windows we use to compute the trends from in this experiment are highly overlapping. This makes it hard to apply any inferential analysis that would allow us to derive conclusions about the underlying processes that hold beyond the time span of the current data at hand. We conclude that the trend variability is negatively correlated to the length of the data set, and that the spatial resolution of the data has limited effect of the trend variability in this experiment where we use the latitude-weighted mean trends.

### 4.5.2. Comparing GRACE to GPS-derived gravity fields

In the second part of this section we assess the trend variability in the spatial domain, as described in Section 3.1. We trim the GeoQ fields to the length of GRACE, but they do not completely overlap due to the later start date of the raw GeoQ data (GRACE: 4/2002-6/2017 vs. GeoQ: 1/2003-3/2018). To align with how we plan to use the Swarm-derived fields in this research, we have truncated all fields to $l_{max}$ = 13. We have chosen a window length from which we estimate the trends, roughly equal to the length of the Swarm time series (5.25 y). The left panels of Figure 4.19 show the results for GRACE. Note that the areas of increased trend variability correlate to areas of known substantial inter-annual variation in mass change. We see increased trend variability in the Amazon, where the year-to-year amplitude of the annual periodic gravity component substantially varies [83]. Note the effects of non-stationary ice melt in Greenland [52] and Antarctica [89]. Closer to our area of interest, we see influence from inter-annual surface hydrology variations in the North American continent [106]. Interestingly, these variations seem to be minimal in the area where there is a substantial GIA-induced gravity trend (red ellipse shows limits of Section 3.5). For the GeoQ fields (right panels of Figure 4.19) the trend variability is substantially higher than for GRACE across the map. We interpret this as an effect of the increased noise levels in the GeoQ fields (NB: $l_{max}$ = 13 in this experiment). In fact, the noise levels in GeoQ are so high that they obscure the trend variability due to surface hydrology variations in North America that we see GRACE. This leads us to conclude that

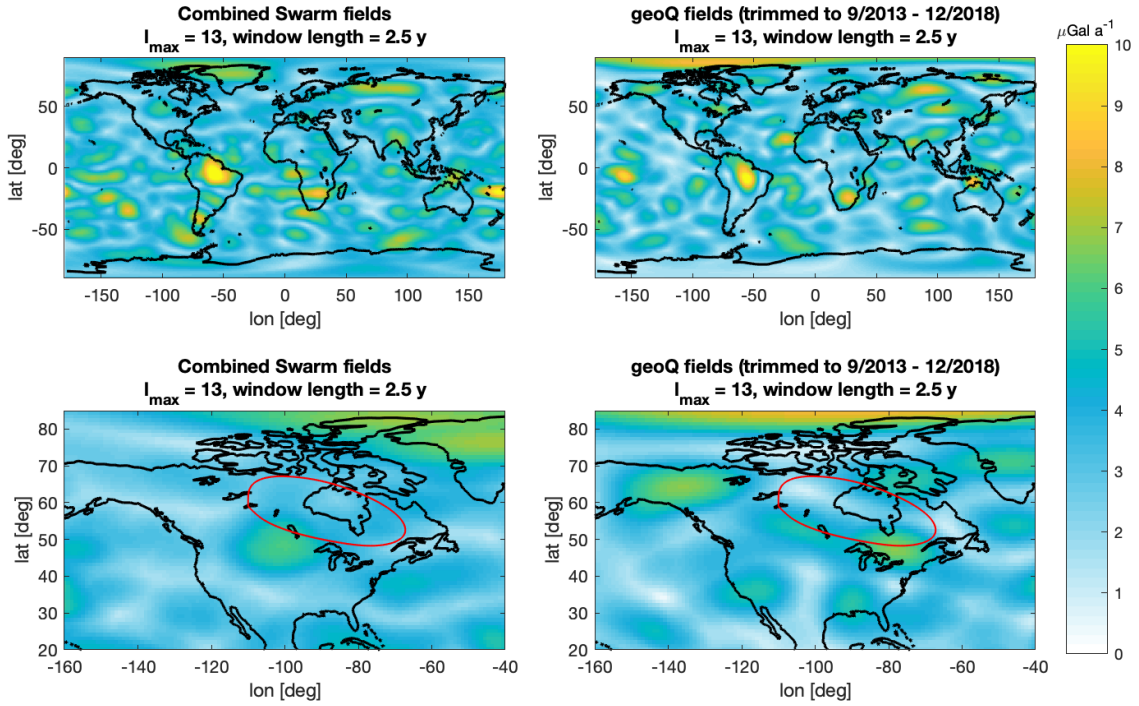



Figure 4.20: Trend variability analysis of the Swarm fields and the GeoQ GPS-derived satellite fields. We trimmed the GeoQ fields to start at September 2012 and end in December 2017, to match the length of the Swarm fields. The fields are truncated to $l_{max} = 13$, and we use a moving window with a length of 2.5 years to estimate the trends. NB: color scale is not the same as in Figure 4.19, as decreasing the window length increases trend variability. Red ellipse indicates the area where the GIA-induced gravity trend is dominant, see Section 3.5.

inter-annual hydrology variations will likely be retrievable from GPS-derived gravity fields to only a limited degree.

For our chosen configuration of $l_{max} = 13$ and a window length of 5.25 years, the variability of the trend signal around Hudson Bay derived from GRACE is between 0.5-1 $\mu$Gal a$^{-1}$. The latitude-weighted mean variability of the area in the red ellipse equals 0.6878 $\mu$Gal a$^{-1}$. For GeoQ this equals 2.258 $\mu$Gal a$^{-1}$, which is very close to the amplitude of the GIA-induced trend of ca. 2 $\mu$Gal a$^{-1}$. Both the GRACE and GeoQ results seem high w.r.t. the results shown in Figure 4.18, but keep in mind that we compared latitude-weighted mean values in that analysis, whereas we look at the difference between the minimum and maximum of all windows here. To put this into perspective, we compute the latitude-weighted mean of the ocean areas (defined as 1000 km offshore from coastlines in Figure 4.20) of the GRACE result, which equals 0.3368 $\mu$Gal a$^{-1}$. Comparing this to the RMS of the ocean area trend of $8.640 \times 10{-}4$ $\mu$Gal a$^{-1}$ (computed from the entire GRACE time series) suggests that our experiment likely yields overestimated results. We argue that this is due to the fact that we use the difference between the maximum and minimum trend as our result, which is likely very susceptible to outliers even with the ad-hoc correction we applied (see Section 3.1). We conclude that this experiment gives interesting insights into the spatial distribution of the trend variability, but likely overestimates its amplitude.

To better extend our conclusions to the Swarm fields, we repeat the experiment of Figure 4.19 using the Swarm fields and an overlapping subset of the GeoQ fields. Due to the shorter length of the data at hands, we use a window length of 2.5 y (this yields 2 independent subsets). Again, we truncate the fields to $l_{max} = 13$. Figure 4.20 shows that the trend variability of Swarm and GeoQ is comparable, both in magnitude and in its spatial distribution. The latitude-weighted mean is 3.974 $\mu$Gal a$^{-1}$ for Swarm and 3.494 $\mu$Gal a$^{-1}$ for GeoQ. When compute this from only the ocean areas (again defined as 1000 km offshore), we find 3.981 $\mu$Gal a$^{-1}$ vs. 3.474 $\mu$Gal a$^{-1}$. We conclude that we should be careful in extending the GeoQ-derived conclusions to the Swarm fields, as the Swarm fields appear to show a slightly higher trend variability.

## 4.6. Correcting for inter-annual hydrology variation

Section 3.2 described how inter-annual hydrology variations can obscure a GIA-induced gravity trend, and how we plan to correct for this via results of the GLDAS-NOAH v2.1 model [74]. This section shows the corrections that we compute and the GRACE- and Swarm-derived trends after applying the corrections. Figure 4.21 shows that the trends in TWS, as

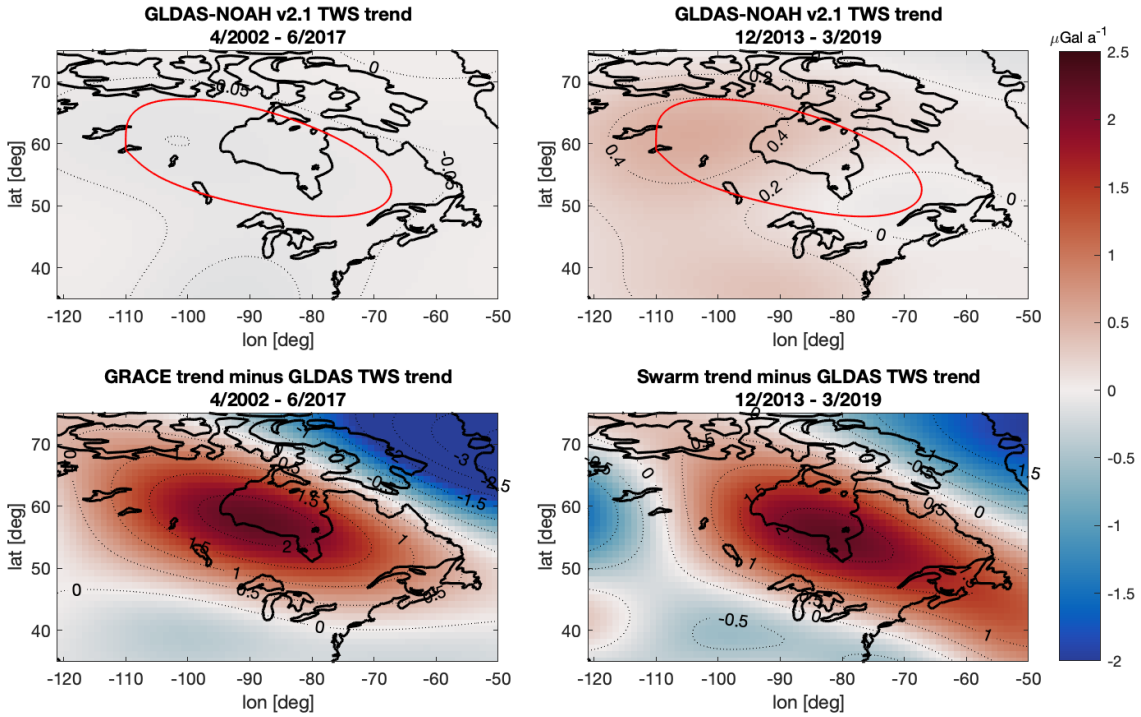



Figure 4.21: Top row: trends in the TWS, as computed from a spherical harmonics decomposition ($l_{max}$ = 13) of GLDAS-NOAH results, for the period covering the GRACE mission (left), and the Swarm mission (right). Values in the contour lines correspond to the gravity rate in $\mu$Gal a$^{-1}$. Red ellipse indicates the area where the GIA-induced gravity trend is dominant, see Section 3.5. Bottom row: gravity trend in GRACE CSR RL06 and the Swarm fields, both truncated to $l_{max}$ = 13, after subtracting GLDAS trends on the top row.

computed from GLDAS, are very low in the GRACE period. The GIA-region latitude-weighted mean, $\dot{g}^{(GIA)}$, (see Equation 3.5) equals -0.07860 $\mu$Gal a$^{-1}$. We suspect two factors influence the near-zero results in the top left plot of Figure 4.21. First, these results argue in favor of a theory that was already postulated by, e.g., Van der Wal et al. [106] and Roy [76]: most inter-annual hydrology variations reduce to zero over longer periods of time. Secondly, we argue that the low spatial resolution that we use reduces the overall amplitude of the correction. Computing the GLDAS correction for a $l_{max} = 60$ GRACE trend (not shown) yields a $\dot{g}^{(GIA)}$ of -0.09493 $\mu$Gal a$^{-1}$. Comparing the bottom left panel in Figure 4.21 to the corresponding panel in Figure 4.12, we see that the influence of the GLDAS-based hydrology correction on the GRACE-derived trend is minimal.

The GLDAS-derived TWS trends in the Swarm period (Figure 4.21 top right) are larger in amplitude than what we saw for GRACE. $\dot{g}^{(GIA)}$ equals 0.2423 $\mu$Gal a$^{-1}$ (0.3212 $\mu$Gal a$^{-1}$ in a hypothetical $l_{max}$ = 60 case). The higher amplitude of the TWS trend in the Swarm period w.r.t. GRACE makes sense in light of the first argument stated in the last paragraph. Comparing Figure 4.21 to Figure 4.12, we see that the GLDAS-based correction does make a difference for the Swarm-derived trend. We see that the positive trend anomaly centered around Hudson Bay is slightly decreased in magnitude. Additionally, a negative trend anomaly at the edge of the map (ca. 60°N; -120°W) seems to increase in magnitude. However, $\Delta^{(GIA)}$ (see Equation 3.7) increases to 0.2707 $\mu$Gal a$^{-1}$, compared to 0.1189 $\mu$Gal a$^{-1}$ without the GLDAS correction. While this is a discouraging result, it is in line with findings of various other authors who concluded that model-based TWS trends are often not suited to isolate a GIA-induced trend [60, 77, 92, 106].

When interpreting the results of the GLDAS-based gravity trend correction, we account for the fact that our empirical RMS-based correction for the irregularities in GLDAS TWS estimates described in Section 3.2.3 may fail to filter out all erroneous grid points. In this case, the TWS trends in the Swarm period would be more affected than those in GRACE, as the signal jumps all occurred around the middle of the Swarm period. This provides an additional explanation as to why GLDAS-derived trends are especially hard to interpret in this study.

## 4.7. Correcting for leakage and ringing

In Section 3.3 explained how reconstructing a signal from a finite number of harmonics introduces ringing into the solution when the signal has large gradients. The same section explained how ice mass loss will induce such undulations

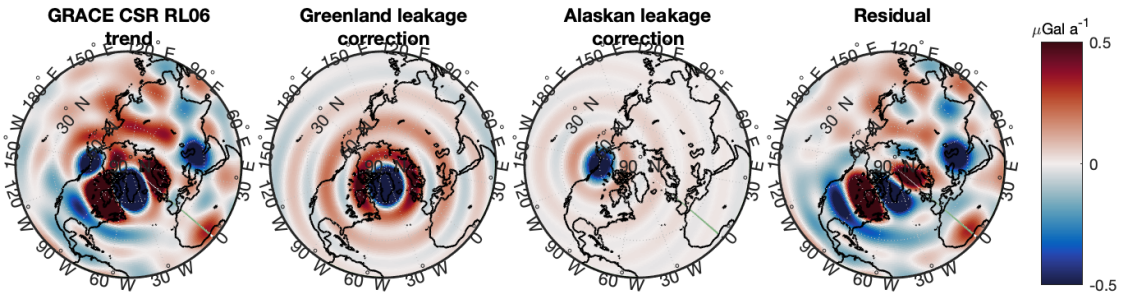



Figure 4.22: Far-left: gravity trend (4/2002 - 6/2017) computed from the GRACE CSR RL06 spherical harmonics solutions, with $l_{max}$ = 13. Middle-left: our proposed correction for the Greenland ice loss, computed from a mass rate estimate of Mouginot et al. [52] and using the implementation of Sørensen et al. [90]. Middle-right: results of our mascon-based correction for Alaskan ice loss. Far-right: residual GRACE trend after subtracting both corrections. NB: the color scale of these plots has been adjusted to highlight the ringing effects.

in the gravity trend domain. Figure 4.22 shows the ringing artefacts induced by ice mass loss. The far-left panel shows the trend in the GRACE CSR RL06 spherical harmonics solutions, truncated to $l_{max}$ = 13 as this is how we will be using the Swarm data. We do not show this for the Swarm data as the ringing will be obscured by noise in those fields. By far the strongest signal in the trend map of the GRACE data is the negative anomaly associated with ice melt in Greenland. Note how the ringing artefacts form concentric circles around the Greenland trend anomaly. The positive circular artefact closest to Greenland itself passes right through the middle of our area of interest around Hudson Bay. This suggests the trend value we find in that area is an overestimation of the true gravity trend, as the results shown in the far-left panel of Figure 4.22 are the sum of the ringing artefacts and the true signal. The middle-left panel of Figure 4.22 shows the gravity trend computed from the estimate of mass loss during the GRACE period of Mouginot et al. [52] via the method of Sørensen et al. [90]. Note how the spatial pattern of concentric circles agrees well with the left plot. The middle-right panel of Figure 4.22 shows the correction we apply for long-wavelength leakage from the ice mass loss signal from Alaska. Again, note the spatial pattern of concentric circles. Since the (spatial) magnitude of the Alaskan signal is smaller, it is harder to discern the ringing artefacts induced by the Alaskan trend gradient from the GRACE data. The far-right panel shows the residual trend after subtracting both corrections. Note how the rings around the Greenland ice mass loss signal have reduced.

Figure 4.23 shows the amplitude and spatial distribution of the leakage corrections in our area of interest around Hudson Bay. We see that the Greenland correction is ca. 0.2 $\mu$Gal a$^{-1}$ in the area where the GIA-induced trend in dominant (red line in Figure 4.23, indicates limit of area shown in Figure 3.8). The magnitude of the Alaska correction is roughly an order of magnitude lower. Compare the third row in Figure 4.23 to the $l_{max}$ = 13 panels in Figure 4.12 to see the difference between applying and not applying both the Greenland and Alaska correction. We see that the GRACE trend decreases the most, going from $\dot{g}^{(GIA)}$ equals 1.533 $\mu$Gal a$^{-1}$ to 1.186 $\mu$Gal a$^{-1}$ (no correction vs. both corrections). For Swarm, this equals 1.691 $\mu$Gal a$^{-1}$ and 1.372 $\mu$Gal a$^{-1}$, respectively ($l_{max}$ = 13). Even though the amplitude of the correction is small, it reduces the agreement between GRACE and Swarm in both trend amplitude and trend spatial distribution.

The bottom row of Figure 4.23, which shows the residual GRACE and Swarm trend after applying both corrections, verifies that we have successfully removed most of the signal trend in Greenland and Alaska. The residual Alaskan signal in GRACE and the Greenland signal in Swarm $\leq$ 0.5 $\mu$Gal a$^{-1}$. Conversely, it appears that there is more residual signal in Greenland in the GRACE-derived gravity trend near the Greenland south-west coast, which could be explained by the homogeneity of the correction (our modeled Greenland ice mass loss does not vary across the area shown on the right in Figure 3.6). We know the true ice mass loss on the west coast has a higher amplitude than on the other parts of Greenland during the GRACE period [52]. For Swarm, the residual trend could be related to noise in the gravity fields. If we look south of Alaska, we see an alternating pattern of artefacts over the oceans in the Swarm map, which are likely noise. The residual Alaskan signal that we see in the Swarm map seems to coincide with a trough of this alternating pattern.

When comparing the corrected GRACE- and Swarm-derived trends, only the amplitude of the differential correction will be relevant. The $\Delta^{(GIA)}$ (see Equation 3.7) of the Greenland and Alaska corrections are very small at $1.230 \times 10^{-3}$ $\mu$Gal a$^{-1}$ and $5.579 \times 10^{-4}$ $\mu$Gal a$^{-1}$, respectively. This leads us to conclude that ringing will likely have a negligible impact on any discrepancy between GRACE- and Swarm-derived trends.

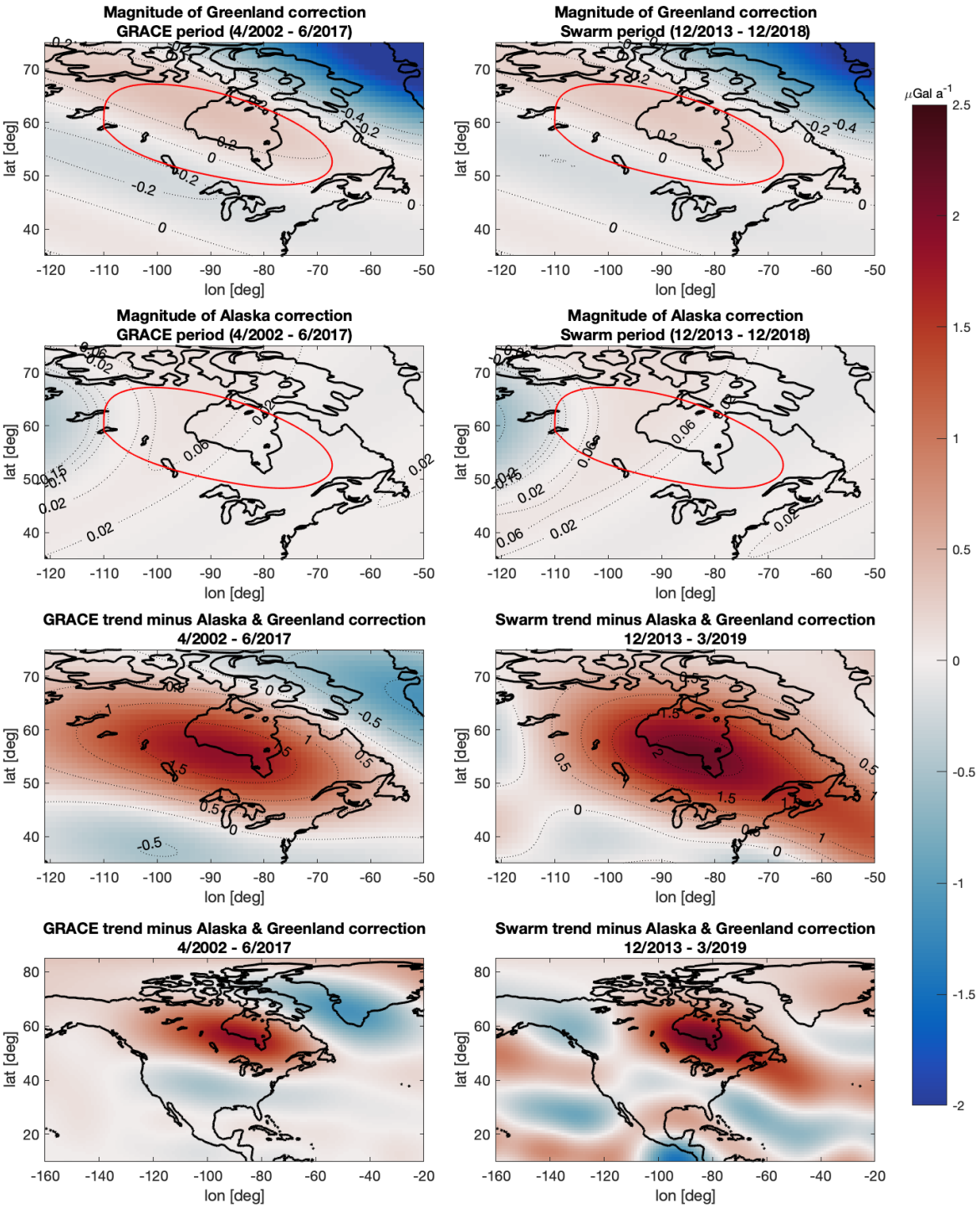



Figure 4.23: Magnitude of the ringing induced in the Hudson Bay area by mass change in Greenland (top panels) and Alaska (second row), as computed from our correction methods, and the corrected GRACE- and Swarm-derived gravity trends (third and bottom row). NB: the third row is a close-up of the bottom row. Red ellipse indicates the area where the GIA-induced gravity trend is dominant, see Section 3.5. NB: Figure 4.12 shows the $l_{max} = 13$ trends without the corrections applied.

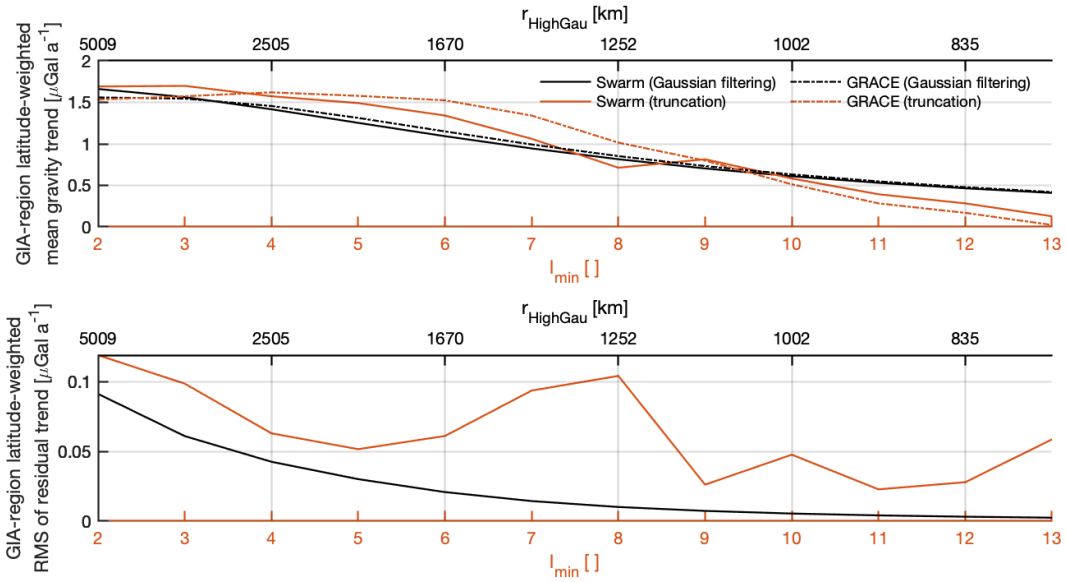



Figure 4.24: Caption

## 4.8. Reducing very low degree coefficients

Section 3.7 discussed the degree correlation of the GRACE-derived gravity trend and the gravity trend as computed by the ICE-6G_D GIA model of Peltier et al. [61] and the GIA model of A et al. [1]. Degrees ≤ 3 yielded no significant correlation between GRACE trends and the GIA models. Sasgen et al. [79] suggests this is because of the ice melt in Antarctica, and more importantly, that this long-wavelength signal affects the trend estimates in Laurentia.

This section presents the results of truncating the very low-degree coefficients or applying a high-pass Gaussian filter, the latter as in Root et al. [75]. These methodology of this section is very much like the methods outlined in Section 2.4 (corresponding to results in Section 4.3), only now applied to the very low degrees instead of the high degrees. Note that the reason why we apply the truncation or Gaussian filtering does differ between Section 2.4 and this section. In Section 2.4 we aim to reduce the influence from high-degree noise, whereas in this section we aim to reduce the influence of non-GIA signals. From Figure 3.7 we determine the truncation setting $l_{min} = 4$, where $l_{min}$ refers to the lowest degree harmonic that is not truncated. Before showing the trend maps, we will first look at a generalized case like in Figure 4.11. We will again show the latitude-weighted mean trend of the Hudson Bay area (i.e., $\dot{\bar{g}}^{(GIA)}$, but instead of the error and SNR we will plot the value for $\Delta^{(GIA)}$ (see Equation 3.7).

Figure 4.24 shows the results of this experiment. Note how, generally speaking, reducing more coefficients (either by truncation or filtering) tends to reduce the $\Delta^{(GIA)}$. However, this is a trivial solution as it reduces the amplitude of the entire trend signal. Also, one could be accused of cherry-picking if one chooses the cutoff point based on an analysis like Figure 4.24. As such, we decide to aim for a conservative setting of which very low degrees to reduce. Rather than basing our analysis on the results in Figure 4.24, we go back to Figure 3.7 and keep the truncation setting $l_{min} = 4$. Note that this is one degree lower (i.e., we keep more coefficients) than what Sasgen et al. [79] used.

We decide not to further pursue high-pass Gaussian filtering. Root et al. [75] successfully used this technique in the Barents Sea, but note that the GIA-induced gravity trend there is concentrated in a much smaller area and of a much smaller amplitude. From Root et al. [75, Figure 1a, Figure 1b] we estimate the positive GIA-induced trend anomaly in the Barents Sea to be ca. <500 km in diameter, while this is >2000 km for the Hudson Bay trend anomaly (see, e.g., Figure 1.3). This makes it likely that our GIA-induced signal extends to lower SH degrees than the signal in the Barents Sea. For amplitude, Root et al. [75] cites typical values of 0.1-0.2 $\mu$Gal a$^{-1}$ whereas we find >2 $\mu$Gal a$^{-1}$ (see, e.g., Figure 4.12). This makes it likely that our signal will dominate obscuring signals that would have to be filtered to retrieve the Barents Sea GIA-induced trend. Finally, we argue that truncating instead of Gaussian filtering would aid the validity of any conclusion we derive regarding the GRACE - Swarm continuity, as truncation leaves the remaining coefficients completely untouched.

Figure 4.25 compares GRACE and Swarm trends derived from degrees 2-13 (we never use degree 1, see Section 2.1.5) to trend derived from degrees 4-14. Note how high-pass truncating reduced the amplitude discrepancy of the observed trends. $\Delta^{(GIA)}$ reduces from 0.1189 $\mu$Gal a$^{-1}$ to 0.06290 $\mu$Gal a$^{-1}$. This is caused by both an increase in the amplitude of the GRACE trend and a decrease in the Swarm (when comparing $l_{min} = 2$ to $l_{min} = 4$). We find $\dot{\bar{g}}^{(GIA)}$ values of

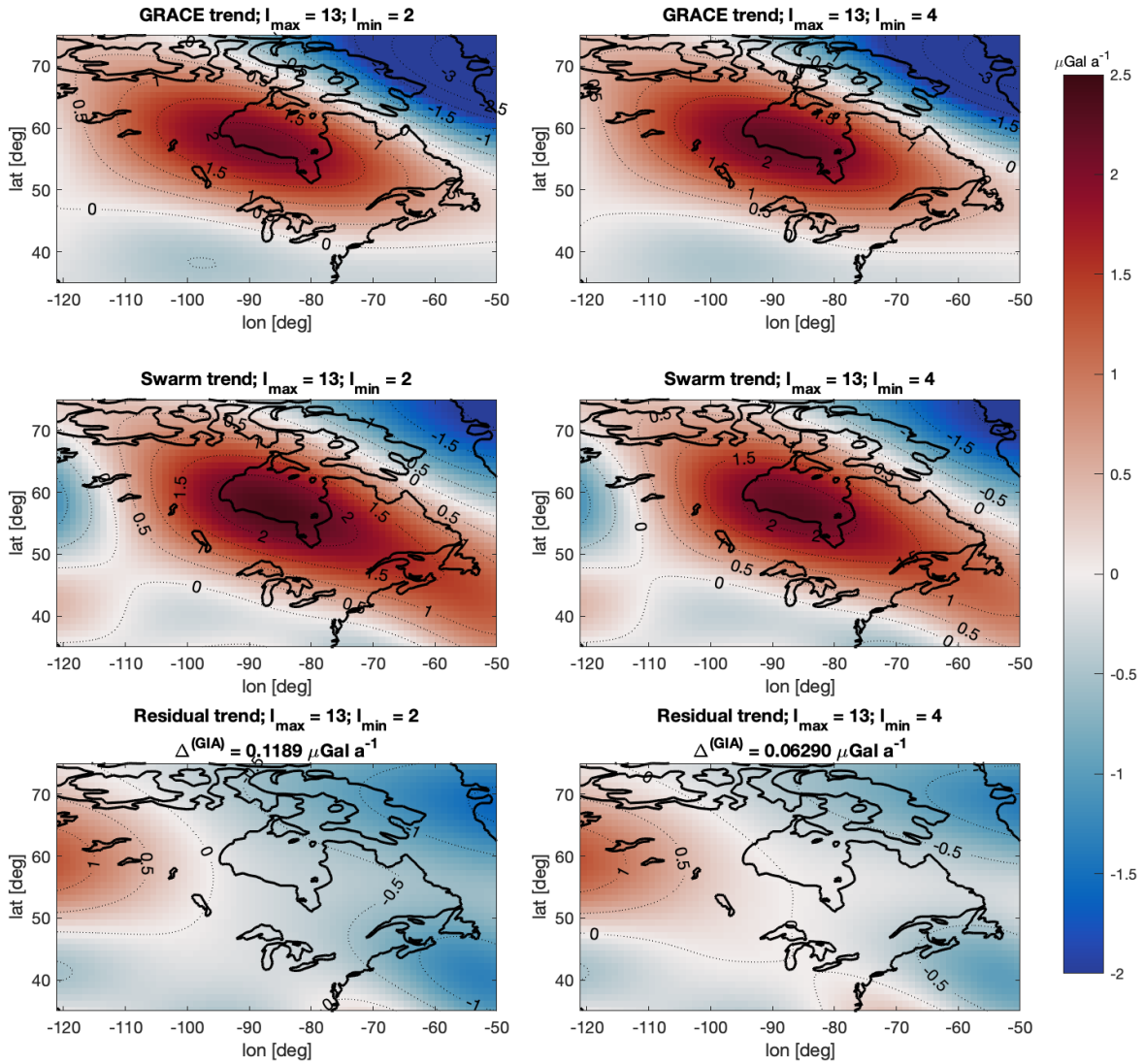



Figure 4.25: Left column: GRACE- and Swarm-derived gravity trend and residuals for $l_{max}$ = 13 and $l_{min}$ = 2 (NB: degree 1 in not used in this entire study, see Section 2.1.5). Right column: GRACE- and Swarm-derived gravity trend and residuals for $l_{max}$ = 13 and $l_{min}$ = 4. NB: $l_{min}$ denotes the lowest degree *not* reduced to zero.

1.5331 and 1.6911 $\mu$Gal a$^{-1}$ (GRACE vs. Swarm for $l_{min}$ = 2), and 1.6203 and 1.5730 $\mu$Gal a$^{-1}$ (GRACE vs. Swarm for $l_{min}$ = 4). Postulating that GRACE and Swarm observe the exact same GIA-induced trend, this would suggest that any long-wavelength leakage in degree 2 and 3 had a different amplitude in the GRACE period than in the Swarm period.

## 4.9. Analysis of the misfit between GRACE and Swarm

Section 4.3 explained why we argue that truncation to $l_{max}$ = 13 is the optimal noise reduction approach for this study. Section 4.8 presented the results of truncating the degree 2 and 3 coefficients from the GRACE and Swarm fields. This decreased the $\Delta^{(GIA)}$ to 0.06290 $\mu$Gal a$^{-1}$, where we concluded that this is due to the fact that the influence from (non-stationary) long-wavelength leakage is reduced. In this section we present the results of a quantitative analysis of $\Delta^{(GIA)}$ in light of the observation uncertainties found in Section 4.2. This analysis based on the $\chi^2$ methodology as presented in Section 3.5.2, and will show in what areas the discrepancy between GRACE and Swarm is statistically significant, where we define significant as >2 Swarm trend standard errors.

We will execute this analysis for four scenarios. First, we take the $l_{max}$ = 13, $l_{min}$ = 4 trend results. Second, we apply the GLDAS-based hydrology correction of Section 4.6. Third, we will apply the undulations correction of Section 4.7. Fourth,

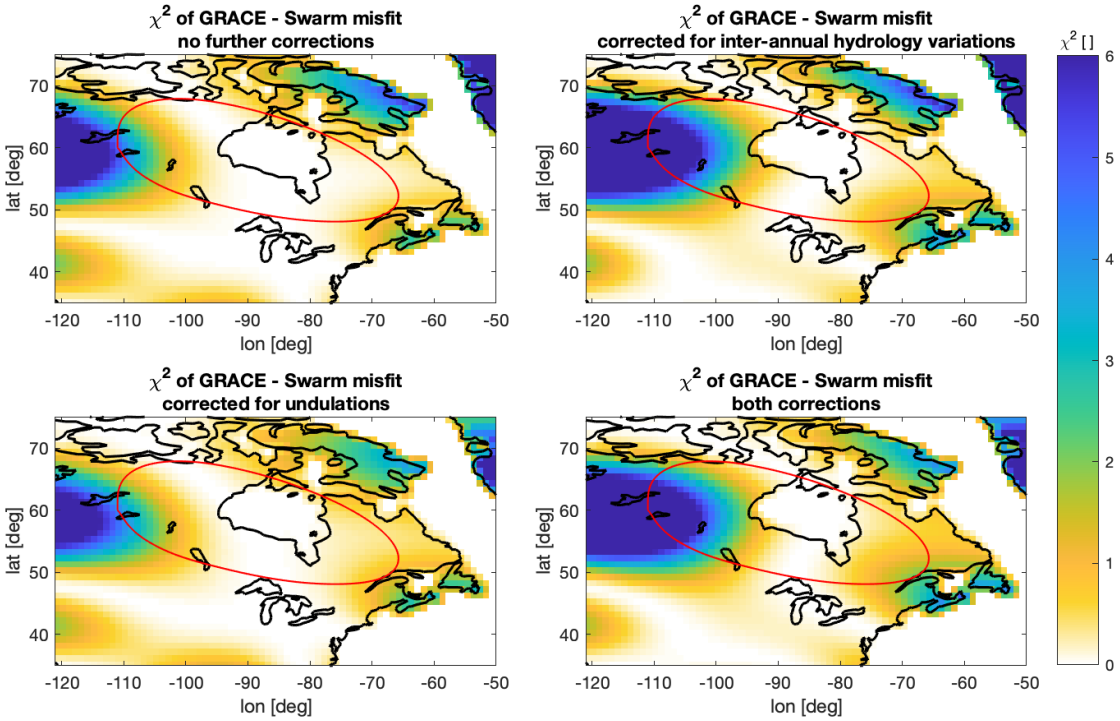



Figure 4.26: $\chi^2$ statistic of the misfit between the trends computed from GRACE CSR RL06 and the Swarm fields, both truncated to $l_{max} = 13$ and $l_{min} = 4$. Ocean areas are masked to increase the readability of the plots. The top left panel has no further correction applied. The top right panel shows the same, but with a correction for inter-annual hydrology variations based on the GLDAS-NOAH v2.1 model (see Section 3.2 and 4.6). The bottom left panel shows the same as the top left, but with a correction for undulation effects based on the RL06m mascon solutions (see Section 3.3 and 4.7). The bottom right shows the same as the top left, but with both corrections applied. NB: red line indicates the area where the GIA-induced gravity trend is dominant (see Section 3.5).

we apply both corrections. We have decided to use the errors of Swarm for $\sigma$ in Equation 3.8. Section 4.2 shows that the errors of GRACE are ca. a factor 20 lower. We assume the error of the ringing correction of Section 4.7 to be negligible, as the magnitude of this correction is also negligible. For the GLDAS-based correction, Van der Wal et al. [106] cites a mean error of the GLDAS-derived gravity trend of 0.13 $\mu$Gal a$^{-1}$. We are unable to recompute this error for our study as we do not have access to other hydrology models that cover the Swarm data set (Van der Wal et al. [106] bases the error estimate on the differences between various hydrology models). Secondly, we choose not to copy the result of Van der Wal et al. [106] as that study likely used an older version of GLDAS. The authors do not give a version number, but the publication is from 2008, when GLDAS-NOAH v2.1 was not yet available. We decide to neglect the GLDAS uncertainty, but acknowledge that this likely leads to an underestimation of the true error.

Note that in the following section, we will apply the high-pass truncation setting of Section 4.8 (i.e., $l_{min} = 4$) to all gravity solutions, but also to aforementioned corrections. The corrections in Section 4.6 and Section 4.7 did not use a high-pass correction (apart from taking out degree 1 and $C_{2,0}$). For conciseness sake will not show the recomputed corrections or how they affect the gravity trend, but only focus on the $\chi^2$-statistic (which, of course, is computed from the gravity trend, see Equation 3.8).

Figure 4.26 shows the spatial distribution of the $\chi^2$-statistic for the four scenarios. Note how even without the hydrology or leakage correction the agreement between GRACE and Swarm is striking (NB: constraints in this panel are $l_{max} = 13$ and $l_{min} = 4$). In the area where the GRACE-derived trend is >+1 $\mu$Gal a$^{-1}$, the latitude-weighted mean $\chi^2$ (i.e., $(\chi^2)^{(GIA)}$ equals 0.4259, indicating that the mean misfit is below one standard error (SE) of the Swarm trend (NB: $\chi^2 = 1 \widehat{=} 1$ SD, $\chi^2 = 4 \widehat{=} 2$ SD, see Section 3.5). The top right panel in Figure 4.26 shows the results of applying the GLDAS correction but not the undulations correction. Note how the misfit west and south of Hudson Bay increases, but the misfit north of Hudson Bay slightly decreases. $(\chi^2)^{(GIA)}$ increases to 1.604, indicating the misfit is more than one Swarm SE but less than two Swarm SE. Note that the substantial increase in misfit west of Hudson Bay seems to correlate to the areas where we had identified irregularities in the GLDAS SWC time series (see Section 3.2.3, Figure 3.3). The bottom left panel of Figure 4.26 shows the results of the undulations correction. In Section 4.7 we had already found that the magnitude of this correction in negligible (ca. four orders of magnitude lower than the GIA gravity trend). As such, we find that is hard to tell exactly where the misfit increased or decreased, except for a small misfit decrease north of Hudson Bay. $(\chi^2)^{(GIA)}$

slightly increases w.r.t. using no corrections to 0.4633. Finally, it seems that applying both corrections (bottom right panel Figure 4.26 results in a misfit increase ($(\chi^2)^{(GIA)}$ = 1.628, almost fully attributable to GLDAS.

We conclude, even without any corrections, there is a strong similarity in both magnitude and spatial distribution of the GRACE- and Swarm-derived gravity trends in around Hudson Bay. The mean $\chi^2$ values of the region where the GIA-induced trend is dominant indicates that the misfit is well below our confidence level of two Swarm SE. In fact, only 6 of the 641 (0.9%) of the grid points that we identify as GIA-region (see Figure 3.8) yield $\chi^2 \geq 4$, i.e., a statistically significant misfit. These points are all located in the area west of Hudson Bay.

Applying either the hydrology correction, the leakage correction, or both, deteriorates this agreement. We argue that this is likely caused by errors in the data that we have used for the corrections. The hydrology correction is based on GLDAS, even though Section 3.2.3 shows that these data contain irregularities. Additionally, various authors suggest no current hydrology model is sufficiently accurate to correct GRACE-derived gravity trends [60, 77, 92, 106]. NASA GSFC aims to publish reprocessed GLDAS results in September 2019[1], and we argue that recomputing the GLDAS correction with the reprocessed results could lead to a decrease in GRACE - Swarm misfit. The undulations correction is based on Greenland mass balance estimates based on a survey of glacier thickness, surface elevation, velocity, and surface mass balance [52], even though Sørensen et al. [90] suggested satellite altimetry estimates correlate to GRACE better. This suggests we should repeat this experiment when altimetry-derived mass balance estimates are available for the Swarm time frame.

## 4.10. Validation of results

In this final section of the results chapter of this study, we first assess the trend results from the GRACE-FO geopotential solutions individually. Next, we execute three experiments using the GRACE-FO data to verify the conclusions we derived so far from comparing GRACE and Swarm and assess the benefits of adding Swarm information to the combined GRACE and GRACE-FO time series.

### 4.10.1. GRACE-FO gravity trends

This section shows gravity trends derived from GRACE-FO data, using either the least squares regression method of Section 2.5, or either one of our approaches using GRACE data presented in Section 3.6.1. Figure 4.27 shows the gravity trends. Note that the color scale in this Figure is twice the range of color scales in most other trend maps in this report. We see that either one of our methods greatly reduces the amplitude of the expanded trends. Adding more (i.e., amplitude and phase vs. only phase) GRACE data decreases the amplitude of the derived trends shown in Figure 4.27. We also see that none of the expansions show a clear positive trend anomaly around Hudson Bay - the place where we would expect a GIA-induced gravity trend. Of course, taking into account the amplitude of the trends that we do see, this could be due to the fact that the GIA signal is obscured by TWS fluctuations.

We know that the time-variable gravity signals observed by GRACE-FO over the oceans should be substantially lower than over land [114]. Comparing the latitude-weighted RMS of the ocean vs. the land area yields 2.329 vs. 7.344 $\mu$Gal a$^{-1}$, 1.144 vs. 4.664 $\mu$Gal a$^{-1}$, and 0.5839 vs. 3.382 $\mu$Gal a$^{-1}$, for respectively the top left, top right, and bottom plot in Figure 4.27. We define the land area by the coastlines of Figure 4.27 and the ocean area as >1000 km offshore. The difference between land and ocean area suggests that (at least part of) the trend results are caused by real geophysical signal. As the ocean area RMS and land area RMS both decrease when we include more GRACE data, we argue that the limited amount of available GRACE-FO data yields to an overestimation of the trend, and that including GRACE information in the regression reduces the magnitude of this overestimation. We will therefore use the GRACE-FO trends derived using GRACE phase and amplitude information in our successive time series presented in Section 4.10.2 (i.e., bottom plot of Figure 4.27).

We attempt to both further explain the gravity trends we retrieve from GRACE-FO and potentially isolate the GIA-induced trend by subtracting TWS estimates from GLDAS-NOAH v2.1 (see Section 3.2.1). Again, we test three approaches similar to what we applied to the GRACE-FO data. In the first case, we derive the TWS trend from only the June 2018 to April 2019 GLDAS data (GRACE-FO time frame; we exclude days not in the GRACE-FO solutions). The second case uses the phase estimate from the GLDAS TWS results of January 2017 to April 2019, but computes the trend from only the data in the GRACE-FO time frame. This start date was chosen as it minimizes influences from the irregularities identified in Section 3.2.3. The third case subtracts (semi-)annual component estimates from the January 2017 to April 2019 GLDAS data, then computes the TWS trend of the results in the GRACE-FO time frame.

Figure 4.28 shows the GLDAS TWS results of July 2018 to April 2019 derived via the three aforementioned methods in the left column. Where the GRACE-FO trends (Figure 4.27) decreased in amplitude when we moved from simple least squares regression to a method using external phase or amplitude data, the GLDAS TWS trends appear to show different behavior.

[1] Ms. Hiroko Kato Beaudoing, personal communication.

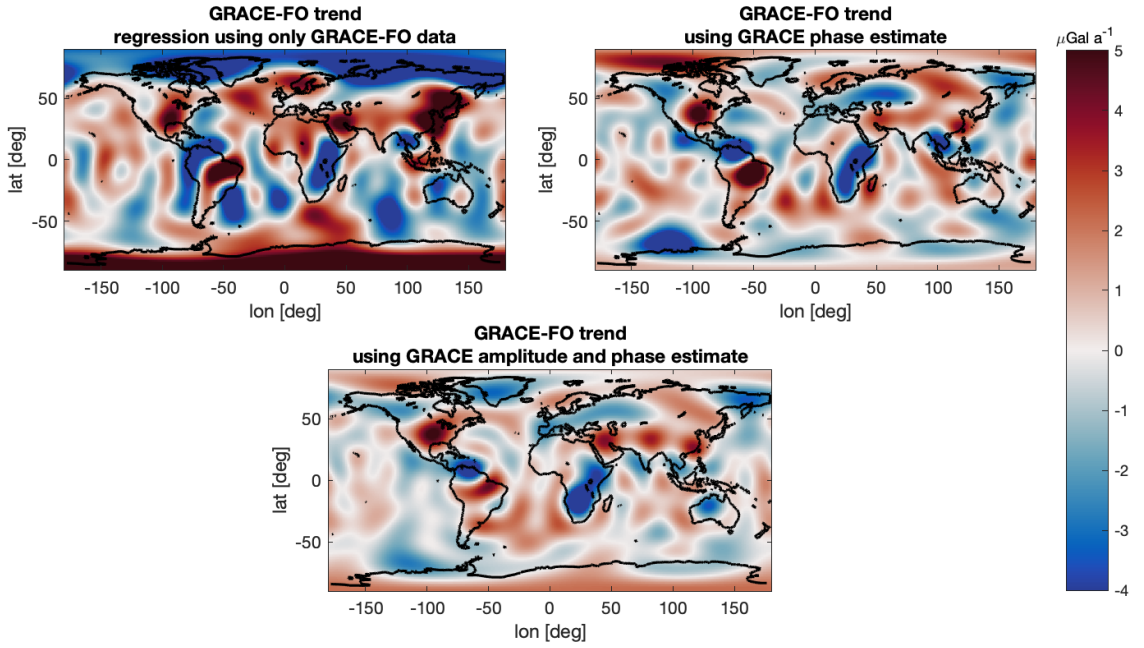



Figure 4.27: Gravity trend computed from mean of the CSR, JPL, and GFZ RL06 GRACE-FO fields, using the least squares regression method of Section 2.5 (top left), using the regression model which includes the GRACE phase estimates (top right), and after subtracting the (semi-)annual components estimated from GRACE data (bottom). NB: color scale is twice the range of most of the other figures showing gravity trends.

Especially in the case where we use phase data of an earlier period, we see that the ocean areas yield spurious results (NB: GLDAS TWS model data equals 0 for all ocean areas). The right column shows the GRACE-FO trends after subtracting the GLDAS TWS trends. We compare the latitude-weighted RMS of the ocean vs. the land area again and find 2.307 vs. 6.693 $\mu$Gal a$^{-1}$, 1.193 vs. 2.896 $\mu$Gal a$^{-1}$, and 0.5314 vs. 1.714 $\mu$Gal a$^{-1}$, for respectively the top, middle, and bottom plot in the right column of Figure 4.28. We interpret the fact that the ocean and land RMS values approach each other when we include more GRACE data as a confirmation that we are indeed reducing the effect of spurious trend results by including the GRACE data. Nonetheless, the fact that there is still a substantial RMS difference between the land and ocean area illustrates that we have not removed all land signal or that some results are still of spurious origin. Note that the land area latitude-weighted RMS of all GIA-induced gravity trends (as computed from ICE-6G_D [61]) equals 0.02540 $\mu$Gal a$^{-1}$, and is thus far too small to explain the observed difference between the land and ocean area RMS in the GRACE-FO trends. We conclude that GLDAS likely removed part of the TWS trends, therefore makes the GIA-signal more pronounced in the gravity trend, but that there is still residual trend signal that we cannot explain. Nonetheless, we decide that we will repeat the experiment described in Section 4.10.2 with a GLDAS-corrected GRACE-FO derived trend.

## 4.10.2. Successive time series

In this section we will look at independently derived gravitational acceleration results of GRACE, Swarm, and GRACE-FO at different points of interest (POIs) in the region on interest. We choose points that are evenly spaced across the GIA-dominated positive gravity trend anomaly around Hudson Bay. Figure 4.29 shows the POIs, as well as the gravity trends derived from GRACE and Swarm, using $l_{max}$ = 13 and $l_{min}$ = 4. Note that these are the same gravity trends as shows on the second row of Figure 4.25, only now with the added information about the location of our POIs. Also note how POIs A, B, C, and D are approximately evenly spaced along the northwest-southeast axis of the trend anomaly, and POIs E, F, G, and H are spaced along the northeast-southwest axis.

The left column of Figure 4.30 shows the observations of GRACE, Swarm, and GRACE-FO of the gravitational acceleration at the POIs. All gravity fields are truncated to $l_{max}$ = 13 and $l_{min}$ = 4. To increase the readability of the graphs, we have centered all gravity fields to the mean of the Swarm fields and trimmed the GRACE data to start in 2012. Note how the Swarm derived-time series show substantially greater variability than the GRACE- or GRACE-FO derived time series. Also note how this variability changes through time - especially the first two years show increased variability. This illustrates the heteroskedasticity of the Swarm observations, which can be explained by the sensitivity of the Swarm measurement system to increased ionospheric activity, as described in Section 4.2.4. It is hard to derive sound conclusions from only 9 data points, but the GRACE-FO observations are where we would expect them if we roughly extrapolated GRACE. Ad-

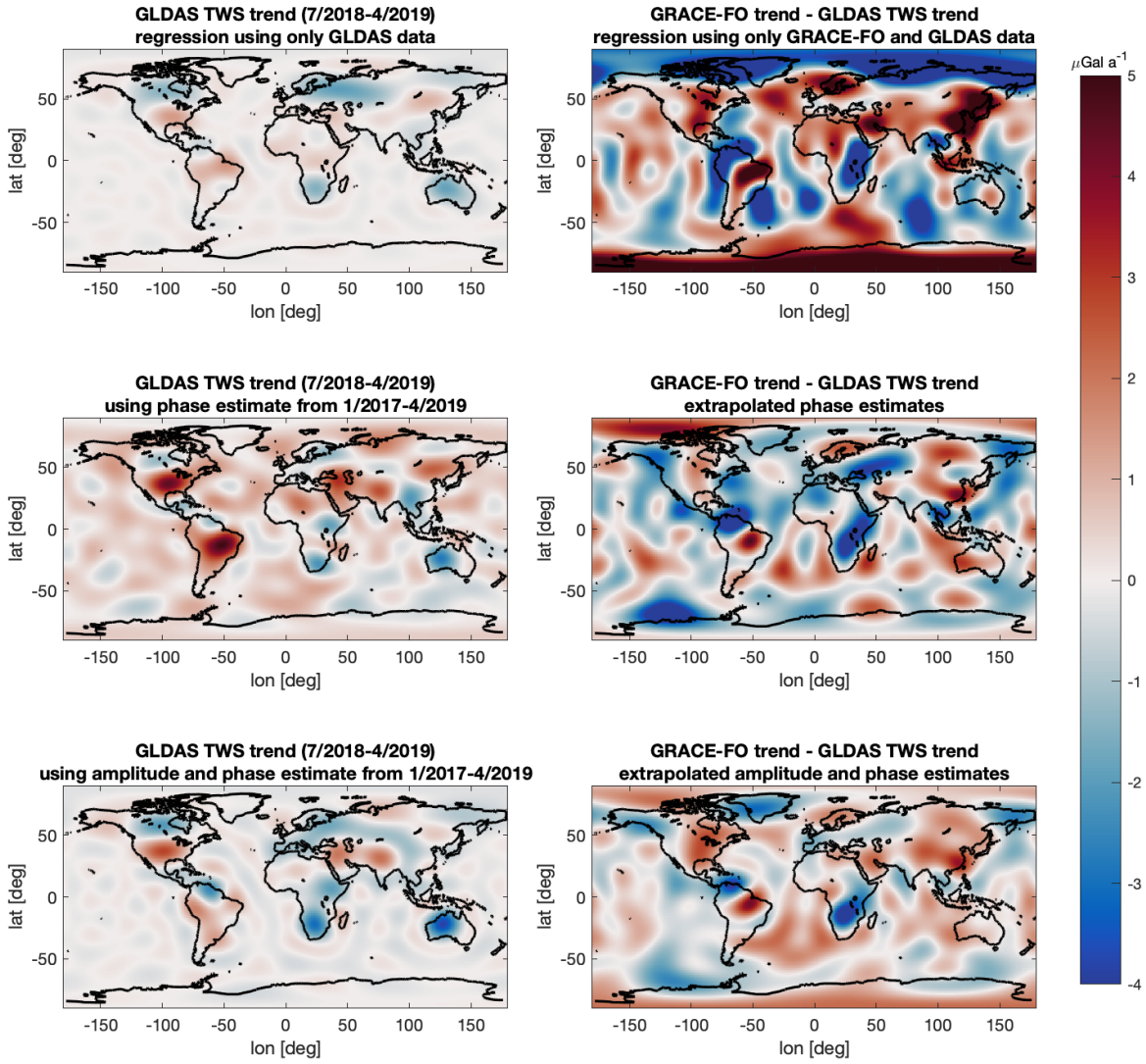



Figure 4.28: Left column: TWS trend computed from GLDAS in GRACE-FO time frame, using the least squares regression method of Section 2.5 (top), using a regression model which includes the GLDAS-NOAH phase estimates since January 2017 (middle), and after subtracting the (semi-)annual components estimated from GLDAS data since January 2017 (bottom). Right column: GRACE-FO-derived trends of Figure 4.27 minus GLDAS-derived TWS trends on the left side of this figure. NB: color scale is twice the range of most of the other figures showing gravity trends.

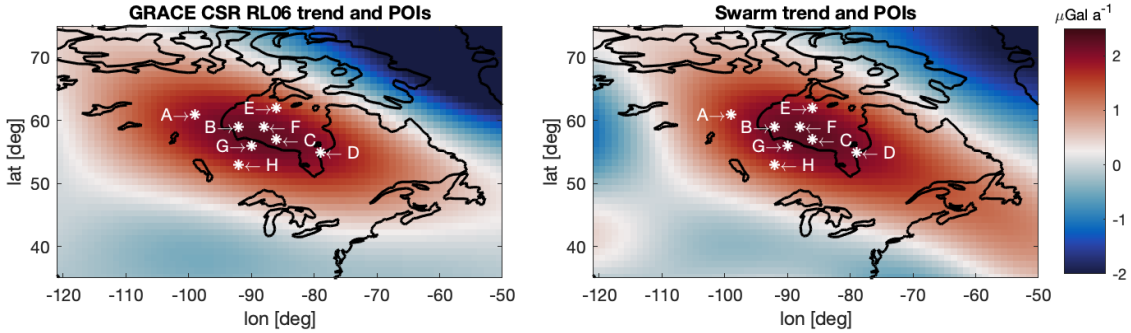



Figure 4.29: Gravity trends derived from GRACE CSR RL06 and the Swarm fields, with $l_{max}$ = 13 and $l_{min}$ = 4. White marks indicate the locations of the 8 POIs A to H. NB: this figure shows the same gravity trends as the second row of Figure 4.25, only now with the added information about the location of our POIs.

ditionally, the GRACE-FO observations seem to be located around the mean of the Swarm observations, generally. This illustrates well how, even from 9 months of GRACE-FO data, we can see that Swarm and GRACE-FO observe roughly the same signal, albeit with higher noise levels for Swarm. Also note how, for POI A and B, the gradient of the Swarm observations seem to decrease after ca. early 2018, and how GRACE-FO seems to follow this pattern.

The right column of Figure 4.30 shows trend lines derived from GRACE, Swarm, and GRACE-FO for the gravitational acceleration at the four POIs. The lines were computed from a least-squared regression of the full data sets, using a model consisting of a bias, trend, and (semi-)annual component. For GRACE-FO, only a bias and trend were fitted, after subtracting GRACE-derived (semi-)annual components (see also Section 3.6.2). The dotted lines indicate ±2 standard deviations of the observations w.r.t the trend line, where we derived the values for the standard deviations via the methods described Section 2.6.1. Comparing the trend lines, the GRACE-FO derived trends are obviously not on the same line as GRACE and Swarm. However, as less than a year's worth of data is available, we argue that it was always unlikely that we would be successful in deriving a credible trend from GRACE-FO. Nonetheless, we do see that the Swarm- and GRACE-FO derived lines connect reasonably well, given that the gradient of the GRACE-FO line is off. More encouraging is the comparison of the GRACE- and Swarm-derived trend lines. We see that for all four POIs, the lines line up (almost) perfectly. POIs A and D seem to have a small bias discrepancy, where Swarm is below GRACE in both graphs. In POIs B and C we see that there is a small discrepancy between the gradients of the trend lines, where is appears that the gradient of the Swarm line is slightly higher in both graphs. Nonetheless, we see that the trend lines are well within the error margins, regardless if we use the Swarm or the much narrower GRACE margins. We conclude that the observations of the GIA-induced trend show near perfect continuity between GRACE and Swarm. For GRACE-FO, we conclude that there is currently not enough data available to successfully derive a gravity trend.

Finally, we repeat the experiment, but revise the GRACE-FO trends for inter-annual TWS variation via GLDAS-NOAH. Figure 4.32 shows the results of this experiment. We see that, compared to Figures 4.30 and 4.31, the GRACE-FO-derived trend is closer to what we see in GRACE and Swarm. Nonetheless, the agreement between GRACE and GRACE-FO and Swarm and GRACE-FO is nowhere near as close as the agreement between GRACE and Swarm. We conclude that, while the TWS correction improves the agreement between GRACE-FO and the other two missions to some extent, the results are still sub-optimal due to the scarcely available GRACE-FO data.

### 4.10.3. Adding Swarm to the combined GRACE time series

This section shows trend maps derived from various (combinations of) satellite data, as outlined in Section 3.6.3. Figure 4.33 compares the gravity trend derived from the Swarm field and from the combination GRACE / GRACE-FO. We have truncated all fields to $l_{max} = 13$ to limit our analysis to the resolution that Swarm is sensitive to, and we find the gravity trend via weighted least squares regression using weights derived via the method of Section 2.6.1.

In Figure 4.33 we see that the Swarm fields yield substantially higher trend amplitudes over the oceans. We interpret this as an effect of the increased noise levels in the Swarm fields. Apart from that, the spatial distribution and magnitude of the Swarm and combined GRACE / GRACE-FO trend seems to correlate reasonably well. A notable exception is east Antarctica, where the negative trend anomaly in Swarm is larger in spatial size and amplitude. Similarly, the amplitudes of the gravity trends in South America seems higher in Swarm. Keep in mind that South America has highly variable inter-annual gravity signals (see, e.g., Figure 4.19), and that Rignot et al. [72] recently identified east Antarctica as an area of accelerating ice mass loss. Both these reasons, of course along with the increased noise levels in Swarm, could explain the discrepancy between Swarm and GRACE / GRACE-FO that we see here.

Next, we compare the combination GRACE / GRACE-FO and GRACE / Swarm / GRACE-FO. We do this for an unweighted case, a case using weights based on the RMSE (i.e., Section 2.6.1). In Section 3.6.3 we predicted that using the RMSE weighs would substantially reduce the influence on the Swarm data as the absolute uncertainty of the Swarm observations is ca. a factor 20 higher than for GRACE (see, e.g., Figure 4.2). For this reason, we include a third case where we only take the heteroskedasticity of the observations into account and reduce the mean of the errors to unity (the trivial weights). For Swarm, we compute this mean from the period after May 2015, as we had identified the period before that as abnormally high in uncertainty. The resulting weights have a standard deviation of 0.2433 for GRACE, 0.4358 for Swarm, and 0.3042 for GRACE-FO (NB: these are dimensionless numbers). This once again illustrates the heteroscedasticity of Swarm.

The latitude-weighted global RMS of the residual trend between the combined GRACE / GRACE-FO and the GRACE / Swarm / GRACE-FO data sets confirms this, and is equal to 0.0520, 0.0086, and 0.0534 $\mu$Gal a$^{-1}$, for respectively the unweighted and two weighted cases described above. Note how all residual RMS values are also very low, and for that reason we will not show the absolute trend maps in the section (they look almost identical), but only the residual maps. Figure 4.34 shows these residual maps. Note how the unweighted case and the trivial case are almost identical, this makes sense as the only substantial difference here is that we have down-weighed the high-uncertainty Swarm observations before May 2015. The RMSE-based weighs yield a vastly different result. Note how the influence of adding Swarm to the GRACE - GRACE-FO time series now is ca. one order of magnitude smaller. Also, the areas where the trend is influences by

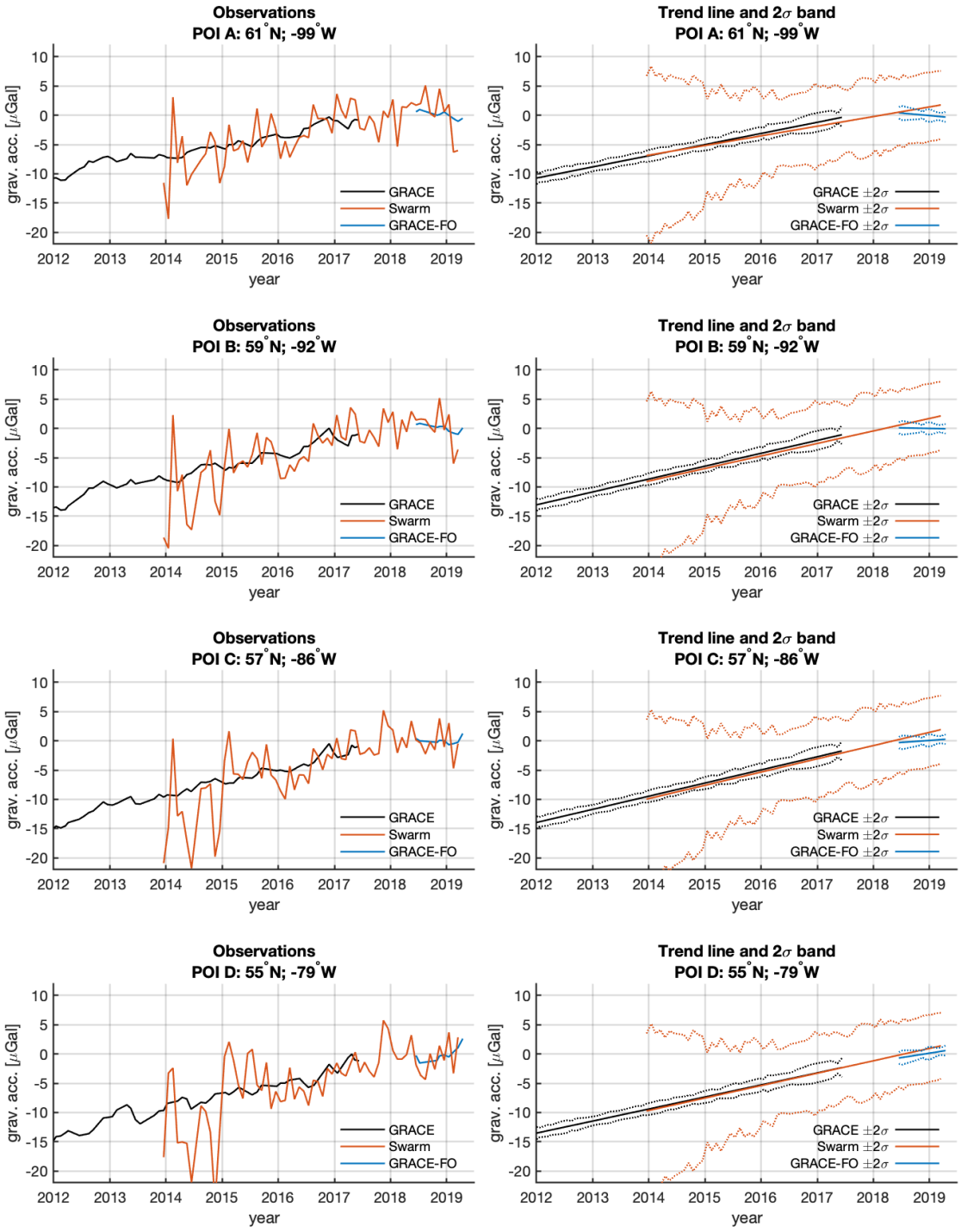



Figure 4.30: Left column: time series of gravitational acceleration at POIs A to D, as observed by GRACE, Swarm, and GRACE-FO. For all fields, $l_{max} = 13$ and $l_{min} = 4$. Results are normalized to the mean of the Swarm fields. Right column: trend line, as computed from a least-squared regression of a bias, trend, and (semi-)annual component at POIs A to D. For GRACE-FO, only a bias and trend were fitted, after subtracting GRACE-derived (semi-)annual components. Dotted lines indicate ±2 standard deviations of the observations w.r.t the trend line. NB: the time axis is set to start in 2012 to increase the readability of the plots.

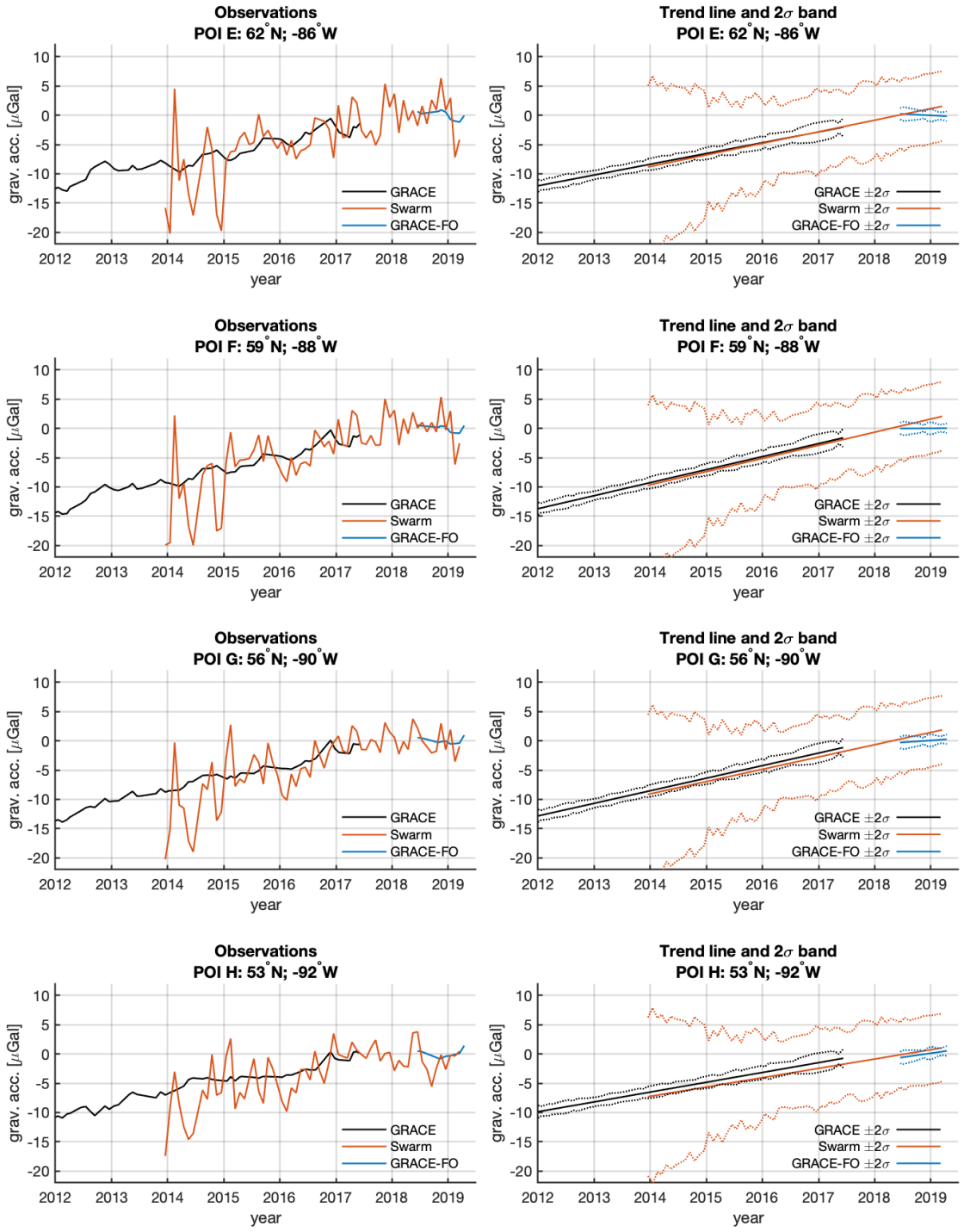



Figure 4.31: Same as Figure 4.30, but for POIs E to H.

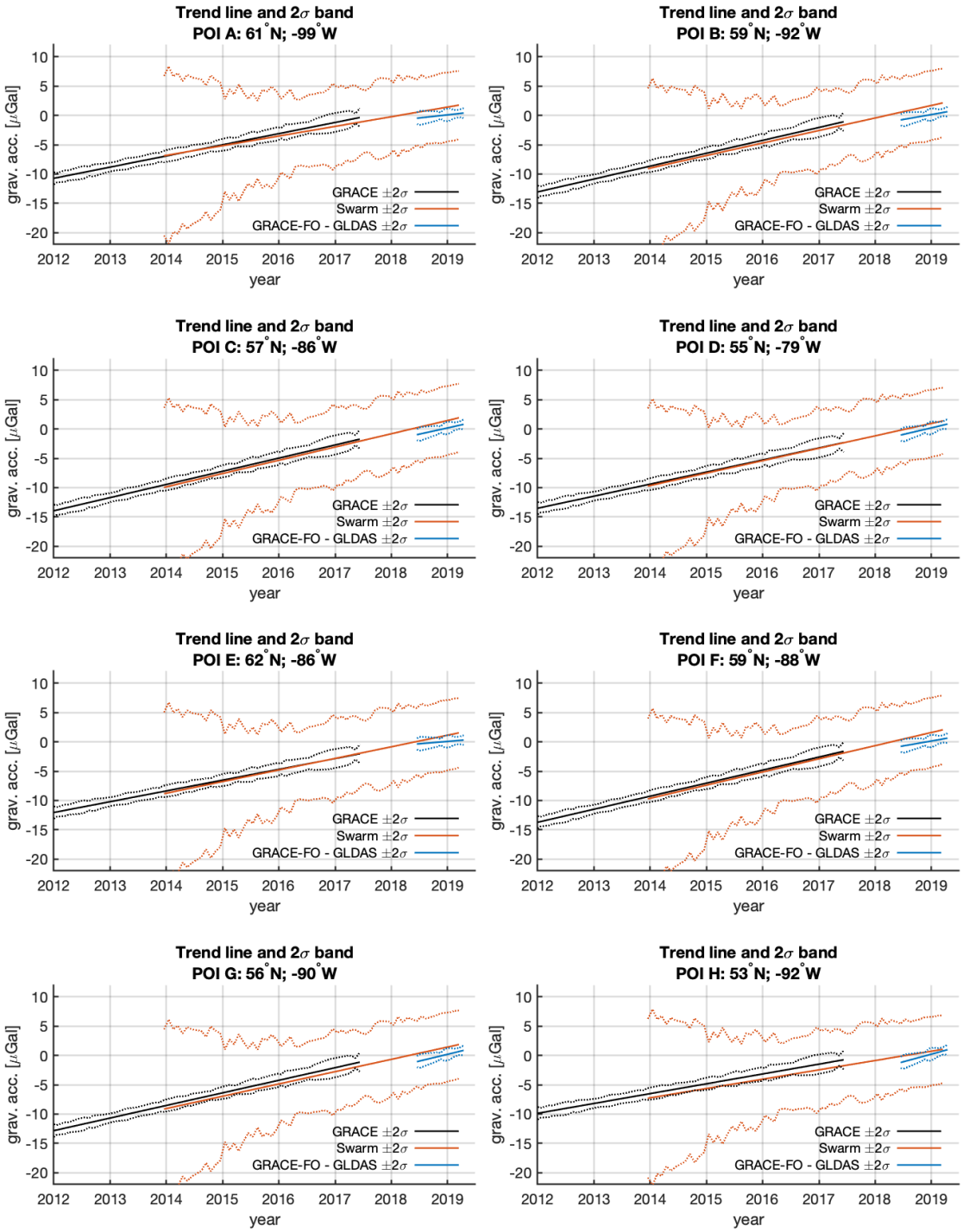



Figure 4.32: Same as the right columns of Figure 4.30 and Figure 4.31, but after correcting the GRACE-FO-derived trends via GLDAS-NOAH, as described in Section 3.6.2.

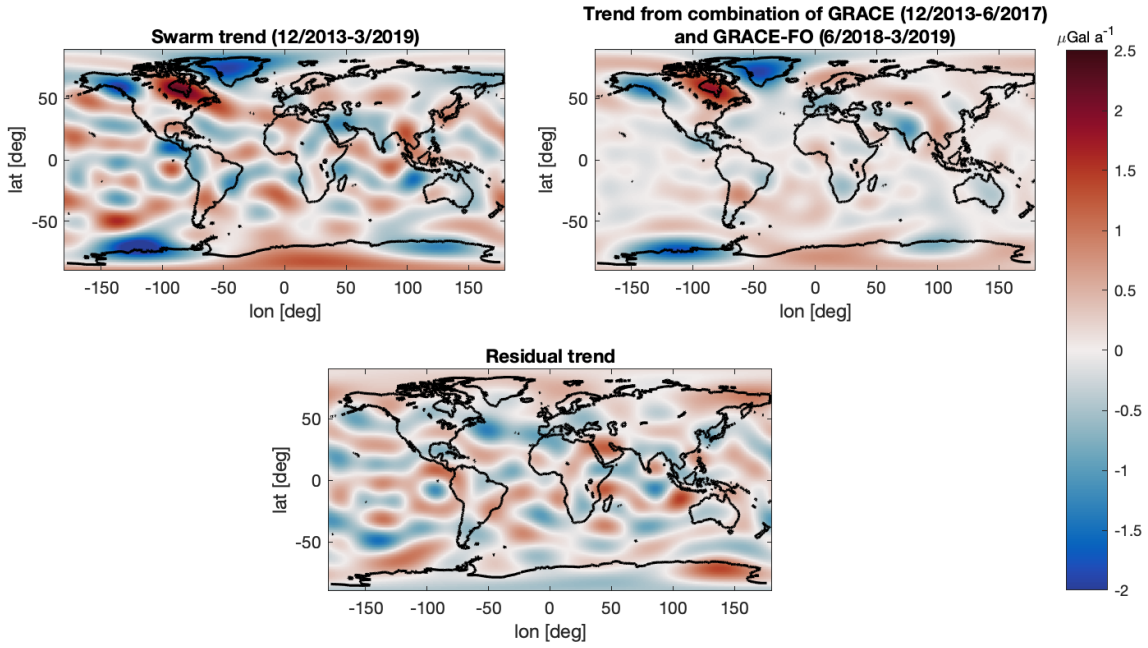



Figure 4.33: Gravity trend derived from the combined Swarm fields (12/2013-3/2019; top left), and the combination GRACE / GRACE-FO (trimmed to also 12/2013-3/2019; top left). All fields are truncated to $l_{max}$ = 13 and we apply a weighted regression using weights derived via the method of Section 2.6.1.

the Swarm data seem to be much less localized. Note the peak over the Amazon, where we know inter-annual variability of the gravity signal is high, and where it is therefore likely that Swarm adds real observations that were mission from GRACE - GRACE-FO due to the GRACE gap.

We conclude this experiment by arguing that we have found the first tentative signs that Swarm can add relevant information to (trend results of) the GRACE - GRACE-FO time series. It appears that weighing the observations in crucial in extracting all relevant information from Swarm, and therefore we recommend this as a topic for future study.

### 4.10.4. Differential trend of combined time series w.r.t. GRACE

Section 4.10.2 that the observations of GRACE and Swarm of the GIA-induced gravity trend in North America showed a striking continuity. Extracting a trend from the short GRACE-FO time series proved difficult, and for that reason GRACE-FO was largely left out of the discussion in the Section 4.10.2. The following experiment is more suited for the short coverage of GRACE-FO. Since we are appending the Swarm and the GRACE-FO data to the existing time series, we could theoretically even assess the influence of adding a single epoch. For conciseness sake, we will focus on the entire available data set at once, and compare it to a similarly trimmed Swarm-derived time series, as described in Section 3.6.4.

The results of this experiment are shown in Figure 4.35. Immediately, we see that the combination of GRACE and Swarm yield a noisier result than the combination of GRACE and GRACE-FO. However, we can also identify various artefacts in the maps that are likely to be real geophysical signals. As differential results are harder to interpret than an absolute ones, we advise the reader to briefly revisit Equation 3.10 before studying Figure 4.35. Additionally, we have printed an interpretation aid in Table 4.3.

The most striking artefact in Figure 4.35 is the strong negative structure in the Greenland area. GRACE measured a strong mass rate here, but studies after the end of the GRACE mission suggest this rate has decelerated since [e.g. 51, 52]. The fact that we see a negative residual rate of a rate that was negative in the first place, confirms that the rate must have decelerated. This relation is illustrated by the seventh column in Table 4.3. Other artefacts shared by the GRACE-Swarm and GRACE-GRACE-FO combination are the positive differential trend anomaly in India, which seems to be surrounded by areas of a negative differential trend, and positive differential trend structures in Australia, Antarctica (ca. 90°E), and western Canada. In the latter case, note that this positive trend anomaly seems to correlate to an area of increased GRACE - Swarm misfit (best shown in Figure 4.26). This positive differential trend anomaly implies that the gravity trend after the GRACE-period increased (see columns 3, 4, 6 of Table 4.3). Accordingly, this suggests that the GRACE-Swarm misfit over the North American continent was caused by real geophysical signal. Finally, note how the area around Hudson Bay, where the GIA-induced positive trend anomaly is the strongest, remains constant across both maps. This suggests that

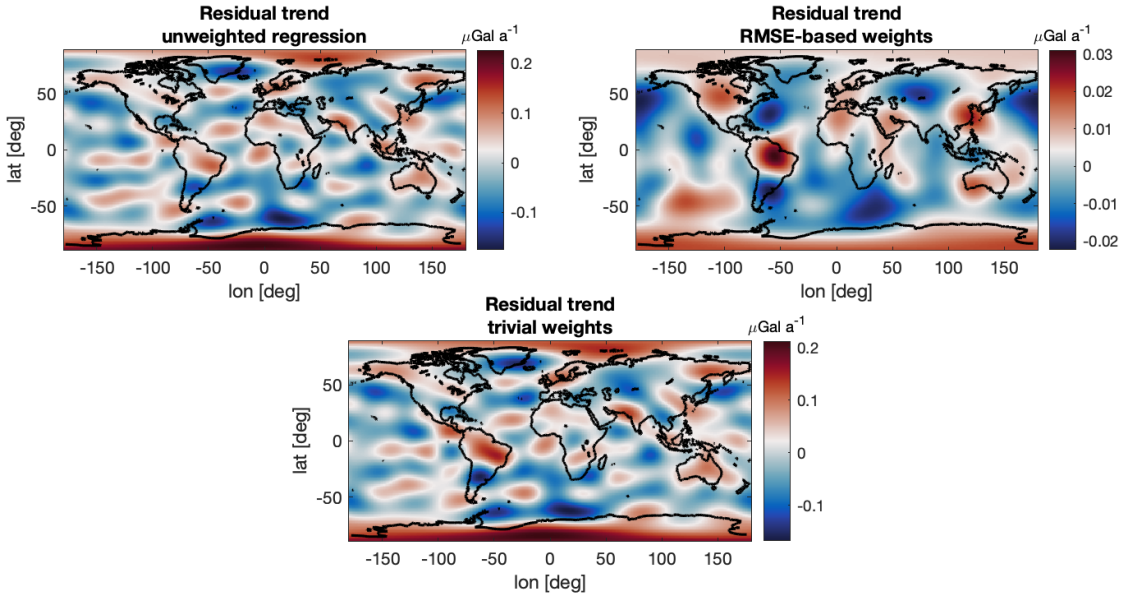



Figure 4.34: Residuals of the gravity trend derived from GRACE / GRACE-FO and the gravity trend derived from GRACE / Swarm / GRACE-FO. We use all available data and three different weighing approaches (see text). NB: color scale is different per panel to best illustrate the spatial distribution of the residuals (magnitude can be compared via RMS values in text).

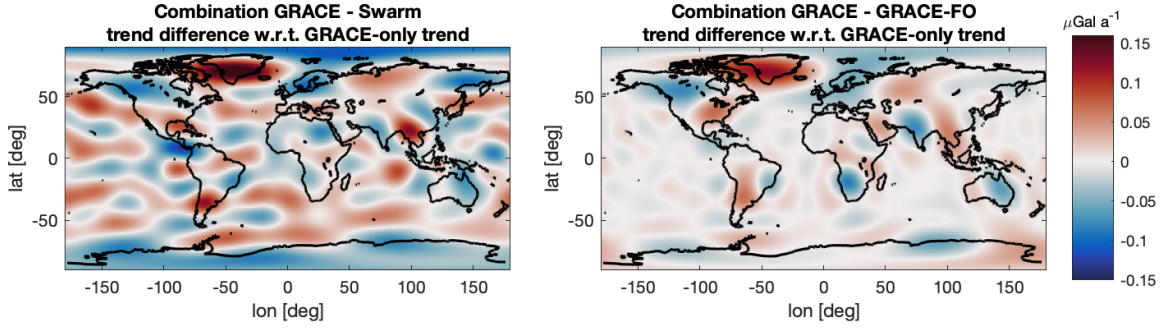



Figure 4.35: Differential trend between a time series of GRACE and Swarm, and GRACE (left), and differential trend between a time series of GRACE and GRACE-FO, and GRACE (right).

the GRACE observations of this trend have converged to the underlying secular GIA-induced signal, and that influences for inter-annual hydrology signals have canceled out at this observation series length.

We finish this chapter by stating that this last GRACE-FO-based experiment successfully validated our conclusions concerning the capabilities of the Swarm measurement system to observe gravity trends. Additionally, we have shown that adding Swarm observations to a time series of GRACE observations yields similar results to appending GRACE with GRACE-FO. This result strongly argues that Swarm has the potential to make significant contributions in the bridging of the GRACE gap.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Trend in combined time series | 0 | 0 | 0 | + | + | + | - | - | - |
| Trend in only GRACE | 0 | + | - | 0 | + | - | 0 | + | - |
| **Differential trend** | 0 | - | + | + | 0 | + + | - | - - | 0 |

Table 4.3: Aid to interpret the sign of the differential trend shown in Figure 4.35. Top row indicates a zero, positive or negative trend computed from the time series combining GRACE and Swarm or GRACE and GRACE-FO data, middle row indicates the sign of the trend computed from only GRACE, and bottom row given the sign of the resultant differential trend.

# 5

# Conclusion and Recommendations

The final chapter of this report first concludes the results presented in the last chapter. We return to the research question and sub-questions posed in Section 1.3, and decide if we have been able to answer them. Next, we build on these conclusion to provide recommendations for future studies concerning the use of Swarm-derived geopotential solutions to bridge the GRACE gap.

NB: *typical results* or *typical values* in this chapter refer to the latitude weighted mean (or RMS; will be explicitly noted) of all signal in the area limited by a GRACE-derived gravity trend of >+1.0 $\mu$Gal a$^{-1}$, as shown in Figure 3.8 (i.e., $\dot{g}^{(GIA)}$ and $\Delta^{(GIA)}$ in the main text).

## 5.1. Conclusions

Following the success of the pioneering GRACE mission, there was a strong call from the scientific community to continue the time series of mass change observations that this mission started. A successor mission, GRACE-FO, was launched in 2018, which leaves a data gap of roughly one year between GRACE and GRACE-FO. The optimal method to fill the *GRACE gap* is currently a topic of debate, with many authors arguing for geopotential solutions derived from GPS position observations of the ESA Swarm mission. In this study we aim to use the gravitational trend induced by GIA to show a continuity between time series of mass change inferred from GRACE and Swarm satellite data. Past studies using GIA observations (relative sea level change, GPS-derived surface deformation, [time-variable] gravity, etc.) to constrain Earth rheology found a relatively high mantle viscosity for the area around Hudson Bay [92, 119]. This suggests GIA-induced gravity trends can be approximated as linear on a multi-decadal timescale, which means GRACE ('02-'17) and Swarm ('13, active) should observe a similar gravity trend in the area. Such a continuity would validate similar sensitivity to mass change trends of the GRACE and Swarm measurements systems, which is a key requirement for Swarm to successfully bridge the gap between GRACE and GRACE-FO.

We extracted the gravity trend from the satellite data via least squares regression. Steffen et al. [94] warned that a failure to account for the $S_2$, $K_1$, and $K_2$ tidal aliases in the fitted model could affect the computed trends. In Section 4.1 we show via an F-test that significant periodic components with periods corresponding to aforementioned aliases exist in 94% of the GRACE Stokes coefficients and in 51% of the Swarm Stokes coefficients up to degree 15. Nonetheless, we also show that including said aliases in the regression model has a negligible effect on the gravity trend estimated from GRACE (Figure 4.1). The short length of the Swarm data set (64 mo) w.r.t. the periods of the aliases (longest: $K_1 \sim 92$ mo) decreases the chance of successfully isolating these components from the Swarm data. This, together with the fact that the GRACE trends was virtually unaffected by the inclusion of the tidal aliases into the regression model, made us decide to neglect the aliases altogether. Our final regression model consisted of a bias, trend, and (semi-)annual component, as given in Equation 4.1.

A sound quantification of the observation uncertainties is vital if we want to successfully propagate to an uncertainty of the estimated trend parameter, which is required for a quantitative assessment of the discrepancy in trend observations of GRACE and Swarm. We combined work of Wahr et al. [114], Van der Wal et al. [106], and Velicogna and Wahr [109], and constructed a method to use the residual RMS (the RMSE) as uncertainty estimate. We found the RMSE by subtracting a 1-year moving mean, and periodic components from a 2-year moving window (see Equation 2.33) from time series of Stokes coefficients and taking the RMS of the residuals. Section 4.2 shows our analysis of the uncertainties of the GRACE CSR RL06 and the Swarm fields. In the spatial domain, the GRACE errors show a familiar distribution with a band shaped

artifact around the equator [114] and the Swarm errors are much more homogeneously distributed. Typical values for the $1\sigma$ uncertainty of the gravity trend in the Hudson Bay area are 0.0153 $\mu$Gal a$^{-1}$ for GRACE and 0.400 $\mu$Gal a$^{-1}$ for Swarm (using $l_{max}$ = 13). In the spectral domain, the relative increase in uncertainty of the sectoral GRACE coefficients was also a familiar result. Swarm yielded a peculiar spectral pattern where the Stokes coefficients from even orders appear to have a relative increase in RMSE. The normalized residuals from Swarm follow a Gaussian distribution and for GRACE they are close to doing to. Finally, Figure 4.8 shows that the Swarm observation uncertainty varies substantially through time. Figure 4.9 shows that not accounting for this heteroscedasticity in the regression (i.e., using unweighted vs. using weighted regression) leads to an overestimate of the Swarm-derived gravity trend.

Where the primary source of observations for the GRACE gravity fields are micron-level range rate measurements of the inter-satellite KBR system, these are centimeter-level GPS position observations for the Swarm fields. Understandably, this leads to substantially more noise in the Swarm fields, especially in the higher degrees. We experiment with both Gaussian filtering and truncation of the maximum degree to reduce the influence of the high degree noise. Section 4.3 shows how we aim to reduce noise as much as possible, whilst not losing (too much) signal amplitude of the gravity trend. We found that Gaussian filtering leads to substantial attenuation of the gravity trend, and decide that truncation at $l_{max}$ = 13 provides the optimal trade off between aforementioned criteria. This yields typical values for the gravity trend in the Hudson Bay area of 1.533 $\mu$Gal a$^{-1}$ for GRACE and ca. 1.691 $\mu$Gal a$^{-1}$ for Swarm. The typical residual RMS equals 0.1189 $\mu$Gal a$^{-1}$. This, together with the findings of the last three paragraphs, answers research sub-question Q.1.

The observed mass change at a location is the net sum of all mass change in that particular vertical column. This means the GIA-induced gravity trend is obscured by other mass change processes. Section 4.5 shows the results of an experiment where we aim to compute the net magnitude of all the influence that dilute the GIA-induced gravity trend. Interestingly, the area where this magnitude is at its minimum seems to coincide with the area where the positive gravity trend anomaly in North America is at its maximum. In that area, we found a difference of <0.5 $\mu$Gal a$^{-1}$ for trends derived from 5-year windows of the GRACE data. Isolating the GIA-induced component from gravity trend observations requires correcting these observations for mass change due to inter-annual surface hydrology variations. The near real-time coverage of the Swarm-derived gravity fields restricts the hydrology models we can use. We settle on GLDAS-NOAH v2.1. During pre-processing, we mask the Greenland area and Alaska area as shown in Figure 3.6. Additionally, we mask areas under permanent snow cover. Section 3.2.3 described irregularities in the GLDAS SWC that we identified during the first iteration of this study. The behavior described in that section also exists in the GLDAS source data, and thus we are forced to apply an ad-hoc correction. We choose to mask all grid points where the RMS of the TWS in 2016 equals 3 times or more the RMS of the 2014 TWS. Section 4.6 described the results of correcting the GRACE- and Swarm-derived gravity trends via GLDAS. The typical magnitude of the correction in the Hudson Bay area is -0.07860 $\mu$Gal a$^{-1}$ for GRACE and +0.2423 $\mu$Gal a$^{-1}$ for Swarm. Note that we use all available data in both cases, which suggests hydrology influences reduce in amplitude for longer time scales. Applying the correction decreases the agreement between GRACE and Swarm of the Hudson Bay area gravity trend observations. For the GLDAS-based hydrology correction, we conclude that this is either due to the irregularities in GLDAS that we had identified earlier, or due to the known low agreement between GRACE mass observations and any hydrology model [83]. For the ice mass loss corrections, we conclude that we have used sub-optimal data to base our mass rate estimates on (glacier survey studies mostly based on in-situ velocity measurements and airborne data). Sørensen et al. [90] suggest satellite altimetry-derived mass estimates correlate best to GRACE, but these are not (yet) available for the Swarm period.

As the application of the GRACE- and Swarm-derived gravity fields in this study happens at a drastically lower maximum degree w.r.t. typical GRACE studies (13 vs. 60), leakage and ringing effects are amplified. Section 3.3 schematically illustrated this. Furthermore, many geophysical processes at the origin of the leakage and ringing effects (mainly ice mass loss) are not stationary across the GRACE and Swarm periods, which means they could lead to an increase in misfit between the two. We apply separate corrections for leakage and/or ringing induced by ice mass loss in Greenland, and ice mass loss in Alaska. For Greenland, we apply the leakage correction of Sørensen et al. [90] with a mass loss estimate from Mouginot et al. [52]. For Alaska, we subtracted the Alaskan gravity trend isolated from the CSR RL06m mascons, that we scale via glacier mass balance estimates [120] in the case of Swarm. We find that the magnitude of the corrections is negligible. The amplitude of the differential correction (which is most relevant when comparing GRACE and Swarm) equals $1.230 \times 10^{-3}$ $\mu$Gal a$^{-1}$ and $5.579 \times 10^{-4}$ $\mu$Gal a$^{-1}$ for respectively Greenland and Alaska. Further long-wavelength leakage can theoretically be expected from Antarctic ice mass loss [79]. Figure 3.7 shows the degree correlation between the Northern Hemisphere results of the GRACE-derived gravity trend (after applying the Greenland and Alaska corrections) and the results of the ICE-6G_D GIA model of Peltier et al. [61]. We found no significant correlation for degrees below 4. Section 4.8 shows the results of high-pass Gaussian filtering and truncation, where we concluded that truncating all coefficients below degree 4 yields a substantially increased agreement between GRACE and Swarm. Typical values for the gravity trend in the Hudson Bay area are now ca. 1.6203 $\mu$Gal a$^{-1}$ for GRACE and ca. 1.5730 $\mu$Gal a$^{-1}$ for Swarm, with a residual RMS of 0.06290 $\mu$Gal a$^{-1}$. Figure 4.25 shows how the agreement in spatial distribution of the gravity trend is striking. This answers research sub-question Q.2.

In Section C we introduced a second GPS-derived time series of gravity fields, to infer conclusions about the influence of the data set length on the derived gravity trends. We use the GeoQ fields of Weigelt et al. [116, 117], that are based

on kinematic orbits of 27 different satellites inverted via the accelerations approach [9]. Section 2.1.4 outlined why we did not focus on these fields in the first place. Figure 4.15 shows how trends derived from 192 months of GeoQ fields show good agreement with GRACE up to degree 22. However, Figure 4.16 shows how subset of the GeoQ fields of a length comparable to the Swarm fields (64 mo) yields an overestimation of the gravity trend around Hudson Bay area w.r.t. to GRACE. Finally, we show that the high-degree noise in the GeoQ fields is higher than the noise in the Swarm fields for the epochs in the GRACE gap (July 2017 to May 2018). We conclude that the Swarm fields yield result closer to GRACE for the 64 mo subset, and that it is therefore likely that the Swarm-derived gravity fields will allow for higher spatial resolution trends to be derived when the data set length grows in the future.

We use a $\chi^2$-test to assess the discrepancy between the GRACE- and Swarm-derived gravity trends w.r.t. the trend uncertainty we found earlier. Section 4.9 shows and discussed the results of this experiment, where we found that all of the GRACE - Swarm misfit of the >+1$\mu$Gal a$^{-1}$ gravity trend anomaly around Hudson Bay is well within the 1$\sigma$ bound of the Swarm trend results, with a typical $\chi^2$ value of 0.4259. This answers research sub-question Q.3. The uncertainty of the GRACE-derived trend observation is neglected here as it is more than an order of magnitude smaller than its Swarm-derived counterpart. The only region where the GRACE - Swarm misfit exceeds the 1$\sigma$ bound is near the Canadian Rocky Mountains. Here, Swarm observes a positive trend anomaly, whereas GRACE does not. We hypothesize that this signal is induced by inter-annual variations in continental water storage. We are unable to reduce the misfit in this area via a GLDAS-derived correction, but it should be noted that aforementioned SWC irregularities that we identified in GLDAS are prevalent in this area and that some of this area is masked out in out pre-processing of GLDAS due to permanent snow cover. We conclude that we have been successful in identifying a continuation between GRACE and Swarm of observations of the GIA-induced gravity trend in North America. This suggests that both systems have a similar sensitivity to gravity trends at the spatial scales at which the Swarm-derived gravity fields contain mostly real signal, which is ca. 1500 km. Also, note that we have only looked at the highest amplitude GIA signal on Earth, and not at gravity trends with a lower amplitude. As such, we must add that strictly taken our conclusion is limited to high amplitude (>+1.0 $\mu$Gal a$^{-1}$) gravity trends. This answers our main research question.

We end the report by validating our recommendations via the recently GRACE-FO fields in Section 4.10. We do this in two steps. First, we study time series of modeled mass change at specific locations carefully chosen to give a cross section of the positive gravity trend anomaly near Hudson Bay, that are independently regressed from GRACE, Swarm, or GRACE-FO data. As the GRACE-FO data set is too short (9 mo) to successfully retrieve a (semi-)annual component, we subtract GRACE-derived results before fitting a bias and trend term to the GRACE-FO data, as was described in Section 3.6. We found that the time series of mass change of GRACE and Swarm line up near perfectly for all grid points. The GRACE-FO-derived time series yield less satisfactory results, likely due to the very limited amount of available data. The latitude-weighted ocean area RMS of the GRACE-FO trends equals 0.5839 $\mu$Gal a$^{-1}$ (NB: for a trend derived from a longer GRACE time series this number equals $8.640 \times 10{-}4$ $\mu$Gal a$^{-1}$). We repeat the experiment after correcting GRACE-FO for TWS variations via GLDAS, which yields trends in the Hudson Bay area that are positive, but still far from GRACE observations in amplitude and spatial distribution.

In the second experiment we investigate the effect of adding Swarm observations to the GRACE / GRACE-FO combined time series. Section 4.2 showed the GRACE and Swarm errors are more than an order or magnitude apart, and thus correct weighing of the observation is a key issue in combining these data. We experiment with an unweighted case, our RMSE-based error estimates, and a third case where we set the RMSE errors to a mean of 0 (to only take into account the relative time evolution of the observations uncertainties of the three data sets). The first and third case both yield noisy results, but the RMSE-based case finds that Swarm adds gravity trend information in areas where we expect substantial inter-annual gravity variability during the GRACE gap (e.g., the Amazon). Here, it should be added that the influence from Swarm on the GRACE / GRACE-FO combinations was of a very low amplitude (max. $\pm$0.3 $\mu$Gal a$^{-1}$).

In the third experiment we look at the differential trend between GRACE and a time series of either chronologically sequenced GRACE and Swarm fields, or sequenced GRACE and GRACE-FO field. To achieve equal comparison we use only the last 9 Swarm epochs in the first case. We found that appending the GRACE time series with Swarm observation yields roughly the same results as doing so with GRACE-FO observations, albeit at a higher noise level. The deceleration of Greenland ice mass loss, acceleration of Alaskan ice mass loss, and ground water depletion in Australia are just three examples of geophysical processes after the GRACE period that can be resolved once the data is combined with Swarm (or GRACE-FO). Additionally, aforementioned misfit near the Canadian Rocky Mountains is also observed in the Swarm data. This suggests that this misfit was caused by real geophysical signal, which argues that the Swarm measurement system can be used to infer conclusions about terrestrial water storage. We conclude that, in general, results from GRACE-FO align well with results from Swarm. This validates our earlier conclusion: at spatial scales up to ca. 1500 km, the GRACE and Swarm measurements systems have similar sensitivity to gravity trends, which further establishes geopotential observations derived from Swarm satellite data as the solution to bridge the data gap between the GRACE and GRACE-FO satellite missions. The conclusions derived from these three validation experiments answer research sub-question Q.4.

## 5.2. Recommendations for future work

### 5.2.1. Trend continuity in higher degrees

We have shown that the main obstacle in extracting geophysical signals from GPS-derived gravity fields, whether or not to compare with GRACE, is noise in the values for the Stokes coefficients. This noise (partially) originates in the errors in the GPS observations of the satellite's location, which propagate into the resulting estimates of the Stokes coefficients. GPS data reprocessing is a subject of current debate, and could be applicable to anyone trying to increase the quality of the Swarm-derived fields.

In this study, we have applied generic methods to reduce this noise. If we look back on the past two decades of GRACE gravity observations, we see a substantial increase in both processing strategies and processing complexity. Anisotropic filtering, as briefly mentioned in Section 2.4, was successful when applied to the GRACE data in the right configuration. We make an attempt at anisotropic filtering in Appendix B, applying a filter that is based on the statistical significance of the trend component in a particular Stokes coefficient time series. An interesting starting point for more advanced processing of the Swarm fields could be the vertical artefacts in the spectral distributions of the trend, as shown in Figure 4.3. We conclude by recommending that any user of the Swarm data can get a head start carefully studying relevant lessons learned from GRACE.

We have found that the length of the available data set also influences how high-degree noise propagates into the trend results. In Section 4.5 and Section 4.4 we saw that deriving a trend from a longer time series reduced influences from both noise and other inter-annual mass transport processes. In Section 4.4.2 we saw that we were able to successfully derive trends from coefficients up to ca. degree 22 from the 192-month GeoQ gravity fields, whereas in Section 4.3 we saw that for the 64-month Swarm this was only possible up to ca. degree 13. Nonetheless, when we performed a limited analysis of the Swarm and GeoQ solutions during the GRACE gap in Section 4.4.2, and found that the Swarm solutions show less high-degree noise. Together, these findings suggest that it is likely that we will be able to derive higher resolution conclusion from the Swarm fields in the future, when there is more data available. It would be interesting to repeat the experiment of this study at that point, and see if our conclusion of a GIA-induced gravity trend observation that continues between GRACE and Swarm still holds.

### 5.2.2. Combining Swarm observations with other data

Various authors have shown that time series of Stokes coefficients derived from SLR position observations of Earth orbiters can add valuable information to the low degrees of (Swarm-derived) gravity fields [e.g. 51, 116, 117] and that increasing the number of used orbiters increases the quality of the results [e.g. 11, 12, 49]. Additionally, a limited amount of studies have looked at using DORIS-observations for this [97]. This suggests that the optimal geopotential solution to bridge the GRACE gap uses both Swarm and (most likely) SLR information. SLR for the low degrees, and Swarm for the moderate to higher degrees. The reader should keep in mind that our conclusion regarding the continuity between GRACE and Swarm is not valid for degrees 1-3, as we had truncated these (see Section 4.8). We motivated this decision by our hypothesis that these degrees would be affected by leakage and undulations caused by ice mass loss in Greenland and Antarctica (see Figure 3.7) which, as they are not stationary mass transport processes, affect GRACE and Swarm differently. However, there is always the option that the decrease in discrepancy between the trends derived from GRACE and Swarm after high-pass truncation was caused by inaccurate low-degree Stokes coefficients in either data set. After all, the GRACE measurements system has already been shown to have limited sensitivity to the very low degree coefficients [e.g. 17, 43].

Sun et al. [95] shows potential in using deep learning via convoluted neural networks to extend time series of the mismatch between GRACE observations and the GLDAS-NOAH land surface hydrology model. In this study, we have shown a continuity between the lower to moderate degrees of the Swarm fields ($l_{max}$ = 13), and GRACE. The results of Section 4.4 suggested that this resolution can likely be increased once longer Swarm time series are available. We have also shown that for the higher degrees (ca. $l \geq 22$), noise dominates the GPS-derived fields which means we cannot isolate any geophysical signal from the observations. We believe the work of Sun et al. [95] shows potential for complimenting the Swarm- and SLR-derived Stokes coefficients in those higher degrees. Sun et al. [95] have shown that machine learning-based propagation of GRACE observations is possible in the spatial domain. As we have credible observations in the lower degrees, it would be interesting to see if these can be included in the extrapolation process. Some qualification of this methodology is required, as it would lead to results that are based on past data, instead of new observations. This would likely decrease the credibility of the resulting products in the eyes of researchers interested in using those products for geophysical study.

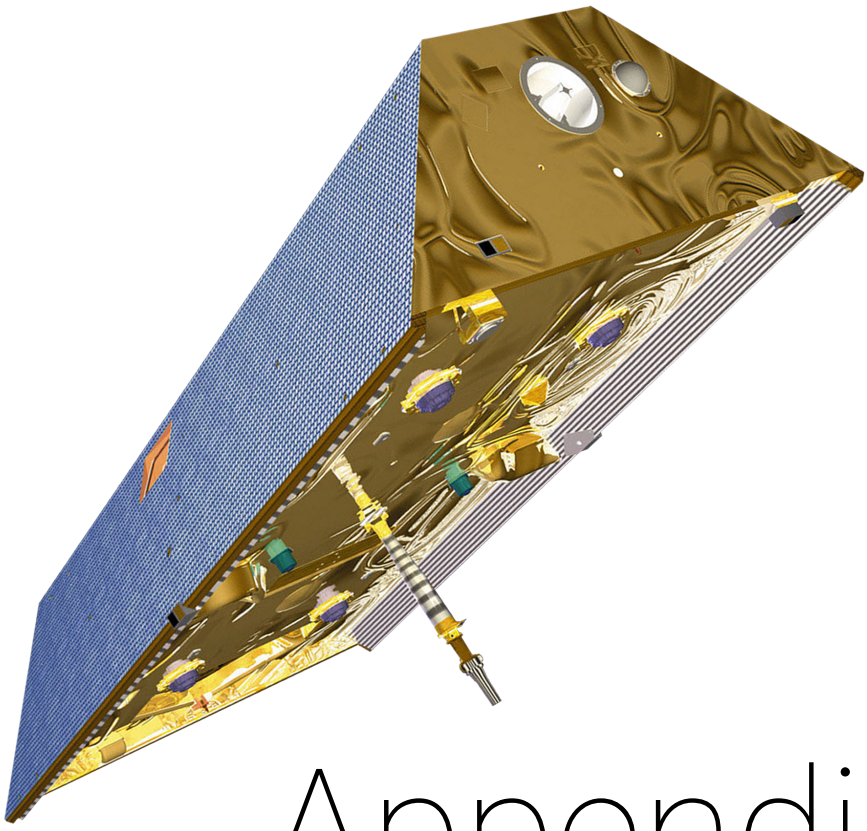
### 5.2.3. Using GIA as a baseline for GRACE-FO

Section 4.5 shows that the area where the positive gravity trend anomaly in North America is most dominant (i.e., the ellipse in Figure 3.8) seems to coincide with the area that shows minimal variation in gravity trends derived from different time frames. This argues that influences from other inter-annual mass transport processes mostly cancel out over time, and that the gravity trend around Hudson Bay is linear on a multi-decadal time scale. We used these findings to illustrate that there is a continuity between GRACE and Swarm. This finding could very well be extended to GRACE-FO, which suggests that the GRACE-derived gravity trend in that area could be used as a *baseline* for the GRACE-FO data products. This would mean a real geophysical signal can be used to cross-calibrate the two missions, which is an attractive option as this approach is close to most applications of the GRACE and GRACE-FO solutions. Using GIA as a baseline for GRACE-FO is a research topic currently under review of various GRACE-FO science teams[1]

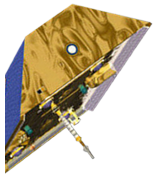
### 5.2.4. Weighing GRACE / Swarm / GFO

Section 4.2 showed that the GRACE and Swarm observation uncertainties are multiple orders of magnitude apart. This means weighing the observations will be a crucial but non-trivial step in extracting as much information from the data sets as possible when they are combined with each other and with GRACE-FO. In Section 4.10 we attempted to combine all three data sets. It appeared that weighing via the RMSE (Section 2.6.1) yields results that correspond to expectations (e.g., concentration of inter-annual gravity variability in the Amazon). Nonetheless, we believe that a more thorough study into the optimal weighing approach for the GRACE / Swarm / GRACE-FO combination could be beneficial in extracting the most amount of available data.

[1] Dr. Byron Tapley, personal communication.



# Appendices



# A

# Auxiliary parameters and functions

## A.1. Love numbers

The Love numbers are dimensionless parameters that define the rigidity of a (planetary) body. We use numbers for the PREM Earth model as given in Wahr et al. [112], which are printed below. Love numbers for degrees not in this table are found through linear interpolation.

| **Degree** $l$ [ ] | **Love number** $k_l$ [ ] | **Degree** $l$ [ ] | **Love number** $k_l$ [ ] |
|---|---|---|---|
| 0 | 0.000 | 12 | -0.064 |
| 1 | 0.027 | 15 | -0.058 |
| 2 | -0.303 | 20 | -0.051 |
| 3 | -0.194 | 30 | -0.040 |
| 4 | -0.132 | 40 | -0.033 |
| 5 | -0.104 | 50 | -0.027 |
| 6 | -0.089 | 70 | -0.020 |
| 7 | -0.081 | 100 | -0.014 |
| 8 | -0.076 | 150 | -0.010 |
| 9 | -0.072 | 200 | -0.007 |
| 10 | -0.069 | | |

Table A.1: Love numbers for the PREM Earth model [29] as given in Wahr et al. [112].

## A.2. Legendre functions and polynomials

The associated Legendre functions $P_{l,m}(\sin\lambda)$ of degree $l$ and order $m$, are defined as [37]:

$$P_{l,m}(x) = (1-x^2)^{m/2}\frac{d^m P_l(x)}{dx^m} \tag{A.1}$$

with the *Legendre Polynomials* defined as:

$$P_l(x) = \frac{1}{(-2)^l\, l!}\frac{d^l}{dx^l}(1-x^2)^l \tag{A.2}$$

Various recursive methods exist for computationally efficient evaluation of the Legendre Functions. We will not cover those here but refer the reader to, e.g., Rapp [66].

## A.3. Normalization conventions

For numerical reasons, the data products used in this thesis are distributed as normalized Stokes coefficients. We define the normalization factor, $N_{l,m}$, the fully-normalized Legendre function, $\bar{P}_{l,m}(x)$, and for completeness sake the normal-

ized spherical harmonic coefficients, $\bar{C}_{l,m}$ and $\bar{S}_{l,m}$, as [62]:

$$N_{l,m} = \left[(2n+1)(2-\delta_{0m})\frac{(l-m)!}{(l+m)!}\right]^{1/2} \tag{A.3}$$

$$\bar{P}_{l,m}(x) = N_{l,m}P_{l,m}(x) \tag{A.4}$$

$$C_{l,m} = N_{l,m}\bar{C}_{l,m} \tag{A.5}$$

$$S_{l,m} = N_{l,m}\bar{S}_{l,m} \tag{A.6}$$

$$\tag{A.7}$$

with Kronecker delta $\delta_{0m}$ defined as:

$$\delta_{0m} = \begin{cases} 1, & \text{if } m = 0 \\ 0, & \text{if } m \neq 0 \end{cases} \tag{A.8}$$

Note how Equation 2.3 on page 9 yields the same result, as long as one uses either all normalized, or all non-normalized Stokes coefficients and Legendre functions.

## A.4. Spectral decomposition of a function on a sphere

Let $f(\theta,\lambda)$ be a function that is defined on a sphere, using longitude $\lambda$ and latitude $\theta$. Surface spherical harmonics form a complete orthogonal set on a sphere, which means $f(\theta,\lambda)$ can be uniquely represented as a Fourier series of surface harmonics [37]:

$$f(\theta,\lambda) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}\bar{C}_{l,m}\bar{Y}_{l,m}(\theta,\lambda) \tag{A.9}$$

where $\bar{Y}_{l,m}$ is the normalized surface spherical harmonic, that is defined as:

$$\bar{Y}_{l,m}(\theta,\lambda) = \bar{P}_{l,m}(\cos\theta)\begin{cases}\cos m\lambda, & \text{if } m \geq 0 \\ \sin|m|\lambda, & \text{if } m < 0\end{cases} \tag{A.10}$$

The normalized associated Legendre functions $\bar{P}_{l,|m|}$ are as shown in Section A.2. Next, we multiply both sides of Equation A.9 with normalized surface spherical harmonic, $\bar{Y}_{l',m'}(\theta,\lambda)$ and take the integral over the unit sphere. As the surface spherical harmonics are orthogonal, the inner product any $Y^{(1)}_{l,m}$ and $Y^{(2)}_{l,m}$ equals 0 unless $l_1 = l_2$ and $m_1 = m_2$. Finally, we use the normalization definition [62]:

$$||\bar{Y}_{l,m}||^2 = 4\pi \tag{A.11}$$

which yields:

$$\bar{C}_{l,m} = \frac{1}{4\pi}\iint_{\Omega} f(\theta,\lambda)\bar{Y}_{l,m}(\theta,\lambda)d\Omega \tag{A.12}$$

This relation can be evaluated in a computationally efficient way by dividing both sides into a real and an imaginary part, and employing a Fast Fourier Transform algorithm. This will not be covered here but we refer the reader to Driscoll and Healy [28]. For a more comprehensive description of spherical harmonics, we refer the reader to texts such as Heiskanen and Moritz [37].

## A.5. Spectral decomposition of a Dirac delta function

The derivation given in this section build on the relations given in Section A.4. Given the Dirac delta function $\delta_{\theta,\lambda}$, on a sphere at longitude $\lambda_\delta$ and latitude $\theta_\delta$, the normalized Stokes coefficients of the spherical harmonics decomposition are found by integrating the following equation over the unit sphere [37]:

$$\bar{C}^{\delta}_{l,m} = \frac{1}{4\pi}\iint_{\Omega}\delta_{\theta,\lambda}\bar{Y}_{l,m}(\theta,\lambda)d\Omega \tag{A.13}$$

as the integral of a Dirac delta is equal to unity, this simplifies to:

$$\bar{C}^{\delta}_{l,m} = \frac{1}{4\pi}\bar{Y}_{l,m}(\lambda_\delta,\theta_\delta) \tag{A.14}$$

where $\bar{Y}_{l,m}$ is a normalized surface spherical harmonic that is as defined earlier.

## A.6. Cumulative distribution functions

We use MATLAB's `tcdf` function to calculate the probability $p$ of observing T-test statistic $T$ and use the `tinv` function to compute the critical $T$ statistic, $T_\alpha$, for a given confidence level $\alpha$. The `tcdf` evaluates the cumulative distribution function below [2], and `tinv` is defined in terms of the inverse of this cumulative distribution function:

$$p = F_T(T|\nu) = \int_{-\infty}^{x} \frac{\Gamma\left[\frac{(\nu+1)}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]} \frac{1}{\sqrt{\nu\cdot\pi}} \frac{1}{\left[1+\frac{t^2}{\nu}\right]^{\frac{\nu+1}{2}}} dt \tag{A.15}$$

$$T_\alpha = F_T^{-1}(p|\nu) = \left\{T\colon F_T(T|\nu) = \alpha\right\} \tag{A.16}$$

where $\nu$ is the degrees of freedom and $\Gamma(x)$ is an ordinary Gamma function, which is defined as:

$$\Gamma(x) = (x-1)! \quad \text{for} \quad x > 0 \tag{A.17}$$

We calculate the probability $p$ of finding an F-test statistic within the interval $[0,\Phi]$ from a single observation via the cumulative distribution function [2]:

$$p = F_F(\Phi|\nu_1,\nu_2) = \int_0^{\Phi} \frac{\Gamma\left[\frac{(\nu_1+\nu_2)}{2}\right]}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \frac{t^{\frac{\nu_1-2}{2}}}{\left[1+\left(\frac{\nu_1}{\nu_2}\right)t\right]^{\frac{\nu_1+\nu_2}{2}}} dt \tag{A.18}$$

where $\nu_1$ is referred to as the *degrees of freedom in the numerator* and $\nu_2$ is referred to as the *degrees of freedom in the denominator*.

# B

# Case study: statistical filtering of Swarm-derived gravity trends

## Introduction

In Figure 4.3 on page 38 we saw how the spectral distribution of the Swarm errors shows a peculiar pattern of vertical artefacts. In this appendix, we attempt to adapt our noise reduction methods to this apparent heterogeneous error distribution, with the goal of deriving trends of a higher spatial resolution than we found possible with the earlier methodology of this study.

The correlation that we derived from Figure 4.3 was that the noise magnitude was higher in Stokes coefficients with an even order number. Intuitively, one might now argue that we should just filter these coefficient more aggressively than the unevenly numbered order. However, it is not the absolute magnitude of the error that warrants a particular amount of filtering, but the magnitude of the noise relative to the magnitude of the real signal. Assessing the magnitude of a signal in relation to its uncertainty is in a nutshell what statistics is about, and for that reason we will attempt to use the power of inferential statistics to anisotropically filter the Swarm gravity fields.

## Approach

In Section 2.6 we made the assumption that each time series of the satellite-derived Stokes coefficient observations consists of deterministic signal and stochastic noise, and that we can accurately describe the deterministic signal with a model consisting of a bias, trend, annual, and semi-annual component. However, when we fit such a model to the observations via least squares regression, both the signal and the noise affect our estimated model parameters. Consequently, the expansion of the estimated parameters into a spatial map of, e.g., the gravity trend shows artefacts due to real trends in the Stokes coefficients trends and due to noise. We reduce the risk of a Stokes coefficient trend caused by noise propagating into our final spatial map of the gravity trend, by computing that map from only those coefficients that have a statistically significant trend. Because the trend result is affected by the results of the other model components, we expand this method and test for significance for all model components, before including them in our final model.

This method was first applied to the annual component in the GRACE observations by Davis et al. [21], and later extended to the secular component by Davis et al. [20] and Sasgen et al. [78]. Since then, various authors have shown the benefits of statistical filtering, often combining the statistical filter with Gaussian filtering or another anisotropic destriping method, to study TWS trends [e.g. 36] or GIA [e.g. 79]. The authors suggest statistical filtering is effective at reducing noise (or striping in GRACE context), with less attenuation of real signal compared to Gaussian filtering.

It should be noted that over time other anisotropic processing routines have been proven to be more effective in reducing GRACE errors, e.g., regularization [e.g. 81]. However, those methods are not as trivial for us to apply to the Swarm data. For this reason, we choose to apply the method of Davis et al. [20]. This approach is different from the method of Sasgen et al. [78] or Sasgen et al. [79] in the sense that is uses statistical tests weighted by the standard deviation of the observations. Given the heteroscedasticity of the Swarm data, we believe this will be beneficial in our case.

We can describe the statistical filter as an extension of Equation 2.3, where we now add a term $M_{p,l,m}$, that is a function of degree $l$, order $m$, and parameter $p$, where the parameter is any one of the modeling parameters (e.g., trend, bias, annual

component, etc.). We write:

$$f^{(p)}(\theta,\lambda) = \sum_{l=0}^{\infty}\sum_{m=0}^{l} P_{l,m}(\sin\theta)\left(M_{l,m}^{(p)}\cdot\hat{C}_{l,m}^{(p)}\cos m\lambda + M_{l,m}^{(p)}\cdot\hat{S}_{l,m}^{(p)}\sin m\lambda\right) \tag{B.1}$$

where $f^{(p)}(r,\lambda,\theta)$ is the value for the particular modeling parameter $p$ at longitude $\lambda$ and latitude $\theta$. Note that $\hat{C}_{l,m}^{(p)}$ and $\hat{S}_{l,m}^{(p)}$ are *not* the Stokes coefficients from the gravity models. Rather, they are the least squares estimates of the particular parameter in the time series of Stokes coefficient $C_{l,m}$ and $S_{l,m}$, respectively. Note that this means that the statistical filter essentially means expanding only the Stokes coefficients that passed the significance test, and rejecting all other coefficients.

We compute the entries for $M_{p,l,m}$ separately per model parameter, per Stokes coefficient. We use a Student's T-test to determine the statistical significance of the bias and trend component, and an F-test to determine the significance of the periodic components. Both tests follow roughly the same outline. First, we define a null hypothesis that states our estimated parameters are not significant. In other words, the result we derived from our data can be attributed to random chance. We will try to reject the null hypothesis by proving its probability of occurring is low. For this we use a predetermined confidence level of 95%. This means there is a 5% probability of falsely concluding a parameter is significant. When we reject the null hypothesis, we accept the alternative hypothesis, which states our parameter estimate is significant (i.e., *reductio ad absurdum*). Note that this means we never *prove* our alternative hypothesis; we only show the null hypothesis is very unlikely.

Sasgen et al. [79] suggested a more restrictive version of the statistical filter, which the authors suggest is more effective at reducing noise. The authors performed the significance tests separately for the CSR, JPL, and GFZ GRACE geopotential solutions, and only the coefficients that passed the significance test in all three of the solutions were included in the final expansion. We will experiment with applying this methodology using the various components of the combined Swarm solutions: the AIUB, ASU, IfG, and OSU Swarm-derived geopotential solutions[1]. We implement this by expanding only those coefficients from the combined Swarm solutions that pass the significance test for each of the four individual solutions.

## Statistical tests

The Student's T-test is executed separately for the least-squares estimate of the bias and the trend parameter. However, it is defined the same for both parameters, which we will therefore both denote as $\hat{x}_{p,i}$ here. We define the hypotheses:

- $H_{T0}$: there is no significant bias or trend in the observations of Stokes coefficient $i$, i.e., $\hat{x}_{p,i} = 0$.
- $H_{Ta}$: there is a significant, non-zero bias or trend in the observations of Stokes coefficient $i$, i.e., $\hat{x}_{p,i} \neq 0$.

The first step is computing the *T-statistic* of $\hat{x}_{p,i}$, which is defined as follows:

$$T_{p,i} = \frac{\hat{x}_{p,i}}{SE_{\hat{x}_{p,i}}} \sim \tau_\nu \tag{B.2}$$

where $SE_{\hat{x}_{p,i}}$ is the standard error of $\hat{x}_{p,i}$. As the methods described in Section 2.6.1 were based on using the RMS of the post-fit residuals (i.e., the RMSE which is equal to the standard error) as an error measurement for our observations, we can take the corresponding entry in the parameter covariance matrix (see Equation 2.29) for $SE_{\hat{x}_{p,i}}$.

$T$ will follow a Student's T-distribution with $\nu$ degrees of freedom, as is denoted by $\sim \tau_\nu$ in Equation B.2. Just like in Section 2.5.3, we can therefore calculate the probability of $H_{F0}$ occurring given the current data via the cumulative distribution function of the F-distribution. We do this via MATLAB again, using the `tcdf` function [2]. The mathematical background of `tcdf` is also in Appendix A.6. We have predetermined our confidence interval to be 95%. Since the least-squares estimate of the bias or trend can be both lesser than or greater than 0, we use a two-tailed test. This means we reject $H_{T0}$ when $p > 0.975$ or $p < 0.025$.

The F-test for the periodic components follows the F-test described in Section 2.5.3 on page 14. We will therefore not repeat the theoretical background here. We compare a model with a bias and trend component to a model with a bias, trend, and annual component to determine the significance of the annual component. In a second, analogical F-test we do this for the semi-annual component. In earlier iterations of this research we included the $S_2$, $K_1$, and $K_2$ tidal aliases. However, in Section 4.1 we will show that these aliases, while being statistically significant, have a negligible impact on the estimated trends, and are therefore omitted from the least squares problem.

[1] These data are not publicly available, but were kindly provided by dr. João de Teixeira da Encarnação at the CSR in Austin, Texas.

## Assessing spatial distortion

Filters for (gravity) fields can be split into isotropic and anisotropic filters. Isotropic filters affect all Stokes coefficients of a particular order in the same way, whereas anisotropic filters discriminate between orders within a degree. An example of an isotropic filter is the Gaussian filter. Here, the scaling factor from the filter is only a function of the Stokes coefficient degree (see also Equation 2.8). Anisotropic filters have proven to be very useful in reducing the GRACE striping problem [e.g. 96].

The degree of a Stokes coefficient correlates with the spatial size of the field artefacts described by that coefficient. All order of a particular degree are needed to reconstruct every artefact at a particular spatial size. This means an isotropic filter, that does not filter every order equally, will affect the spherical harmonic model's capability to reconstruct certain spatial structures. This a problematic, because the goal of filtering is to suppress noise, but not alter the magnitude, let alone the spatial distribution, of real geophysical signals. This complicates the interpretation of the results of the statistical filter.

We independently assess the spatial effect of the statistical filters we compute from the GRACE and Swarm data via methods shown in, e.g., Davis et al. [20] or Sasgen et al. [79]. We create a synthetic field solution of the spectral decomposition of a Dirac delta function $\delta$ at longitude $\lambda_\delta$ and latitude $\theta_\delta$. The derivation of the Stokes coefficients of these field are given in Appendix A, Section A.5 on page 84. Next, we exclude the same coefficients from the spectral decomposition of our synthetic field, as the statistical filter removes from our gravity fields. Finally, we assess the expansion of the reduced synthetic field to learn about the influence of our statistical filter on the spatial signal distribution. As the anisotropic statistical filter is a function of the signal location, we place the Dirac delta function near our signal of interest, which is the GIA signal around the Hudson Bay.

## Verification

The most straightforward verification of this method would be to compare our results to the results in Davis et al. [20], in terms of exactly which coefficients we deem significant. Unfortunately the authors do not provide this information in their text. Sasgen et al. [79] do provide this information, the only difference being that their method uses unweighted statistical tests. For this reason, we repeat our computations with unweighted statistical tests, recompute Figure 4 in Sasgen et al. [79], and compare to our method using weighted tests. Implementing the unweighted test is done by replacing the standard errors $SE_{\hat{x}_{p,i}}$ in Equation B.2 with the corresponding entries from the aposteriori estimate of the covariance matrix:

$$P_p = \frac{\varepsilon^T \varepsilon}{\nu} \cdot \left(A^T A\right)^{-1} \tag{B.3}$$

where all parameters are as defined earlier.

We are careful to use the same data and time span as the authors used, which is the GRACE RL04 product from the German Research Centre for Geosciences (GFZ), spanning August 2002 to August 2009 [84]. The results of this are shown in the top left panel Figure B.1. Comparing this to Figure 4 in Sasgen et al. [79] shows an exact match, which leads us to conclude we have implemented the T-test part of statistical filter correctly. The T-test is the most important part of this filter, as it directly governs which coefficients are used to compute the trend map. The other parts of the filter routine (i.e., determining significance of annual component) have only a secondary influence, and therefore we do not execute any further verification.

The top right panel of Figure B.1 shows the results of applying the method of Davis et al. [20] (i.e., the method we propose to use) to the data Sasgen et al. [79] used. We have computed the errors in the GRACE GFZ RL04 fields with the method of Section 2.6.1. Note how this method passes more coefficients than the method of Sasgen et al. [79] (1090 vs. 843 out of possible 3721). Also note how the high-degree coefficients passed by Davis et al. [20] are more located towards the zonal coefficients and less is the sectoral domain, where the GRACE solutions are known to be dominated by noise [96].

The bottom panel of Figure B.1 compares degree RMS graphs of the trend computed from the raw GFZ RL04 data, this data after applying the method of Davis et al. [20], and this data after applying the method of Sasgen et al. [79]. We see that, with the exception of degrees 44, 45, and 47, the method of Davis et al. [20] is more effective at suppressing high-degree noise. Also note how the method of Davis et al. [20] is more effective at suppressing the degree RMS variance in the degrees past $l = 50$.

As a final measure of verification in the spatial domain, we compare the latitude-weighted RMS of the ocean area grid points of the expanded Stokes coefficient trends. We find a value of 0.44948 $\mu$Gal for the method of Sasgen et al. [79], and a value of 0.29518 $\mu$Gal for our method. This is substantially closer to the 0.11389 $\mu$Gal we get when we compute the same metric from the newer and lower noise CSR RL06 solutions. Finally, to test if we are not just repressing all signals, we compute the latitude-weighted mean of the trend in the GIA region (as in Figure 4.11). We find 1.1467 $\mu$Gal a$^{-1}$ (method

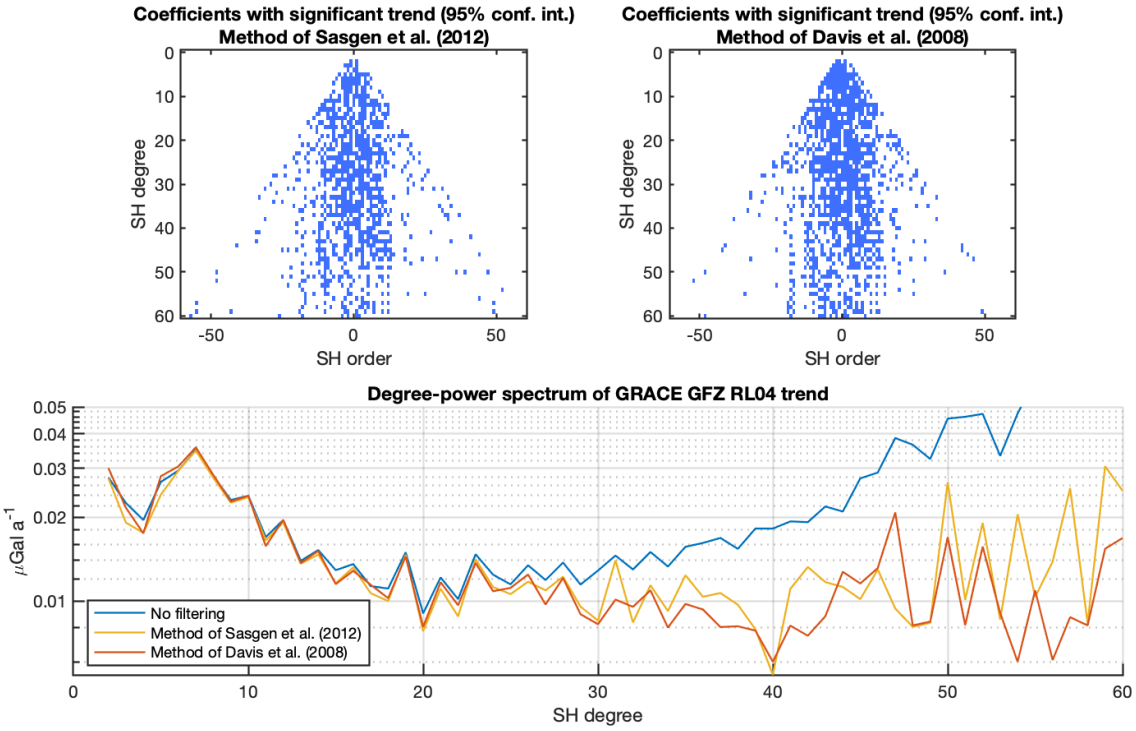



Figure B.1: Left: blue indicates statistical significance of the trend in time series of Stokes coefficients from GRACE GFZ RL04 spanning August 2002 to August 2009. Confidence interval is set at 95%. Shown here to verify our statistical filter. Compare this figure to Figure 4 in Sasgen et al. [79]. Right: same as left, but now including the uncertainty estimates of the Stokes coefficients in the statistical tests.

of Sasgen et al. [79]) and 1.1905 $\mu$Gal a$^{-1}$ (our method). We conclude that the method of Davis et al. [20] is likely to remove more noise, while leaving more GIA signal, compared to the method of Sasgen et al. [79].

# Results

Figure B.2 shows the coefficients in the GRACE CSR RL06 and the combined Swarm fields have a statistically significant trend in blue. For GRACE, we see that 1332 (79.5%) coefficients pass the test, and that the known spectral distribution of the coefficients without a significant trend component resembles the noise distribution: noise is more pronounced in the high-degree and high-order coefficients Swenson and Wahr [96]. Note that we have truncated the fields to $l_{max} = 40$ to match the Swarm fields.

The results for the combined Swarm fields as a whole are discouraging: only 124 coefficients have a significant trend (7.4%). We see that most of the coefficients passing the test are concentrated in the lower degrees, where we had already determined that noise levels are lower. Only looking at degrees up to 13, the number of coefficients passing the significance test increases to 26.3%, but remains low compared to 90.2% for GRACE. The rest of the passing coefficients indicate some spurious trend components that are arbitrarily spread across the spectrum, not indicating any correlation. When we look at the spatial distribution of the Swarm-derived trend after applying the statistical mask in Figure B.4, we see that the statistical filter is effective in reducing the magnitude discrepancy between the GRACE- and Swarm-observed trends for $_{max} = 13, 14, 15$. However, we also note that the statistical filter suppresses the Swarm trend result so much that it is now lower than the GRACE result. Regarding the positive trend anomaly around the Hudson Bay as a whole, we see that the statistical filter substantially affects its spatial distribution. The center seems to shift to the north-east, and its shape transforms from elliptical to round for the $_{max} = 13$ case, to an elongated shape for $_{max} = 14, 15$. Above this degree, the GRACE- and Swarm-derived statistically filtered trend clearly diverge.

In more restrictive case of the statistical filter, as described by Sasgen et al. [79], we only include coefficients in the final expansion that have passed the significance test for each individual solution. When we apply this to the four components of the combined Swarm solution, we find 169 (10.1%), 187 (11.2%), 166 (9.9%), and 196 (11.7%) coefficients with a significant trend in respectively the AUIB, ASU, IfG, and OSU solutions, as is shown in Figure B.5. Only 17 (1%) coefficients have a significant trend in all four solutions. Needless to say, expansion of these 17 coefficients does not lead to any reasonable results, and therefore this expansion is not shown. We conclude that the Swarm-derived fields contain too much noise to

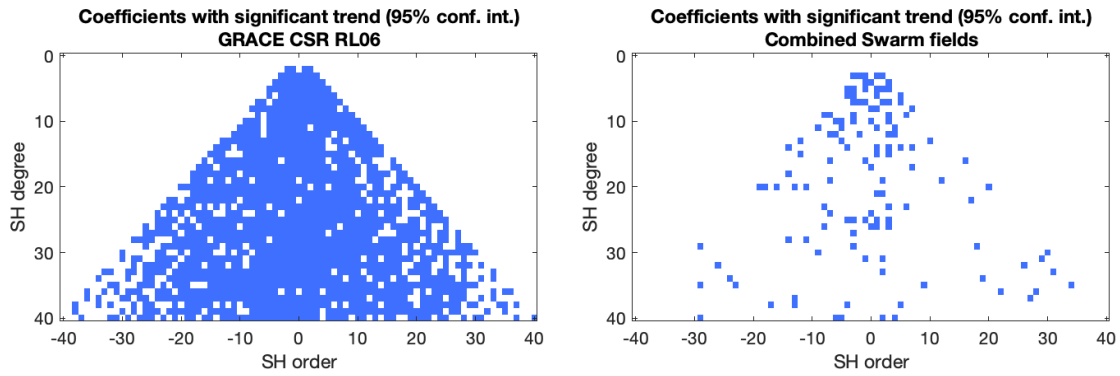



Figure B.2: Coefficients with a significant trend (95% confidence interval) in GRACE CSR RL06 (left) and the combined Swarm fields (right).

apply the more restrictive statistical filter of Sasgen et al. [79].

## Interpretation of results

In Section B we explained how anisotropic filtering can distort the spatial distribution of the reconstructed fields, and that interpretation of the results from the statistical filter is difficult for that reason. Here, we apply the statistical filters that we have computed from the gravity data (Figure B.2) to the synthetic fields consisting of a Dirac delta function at either the south-east coast of the Hudson Bay (58°N; 88°W), or on a point on the equator along the same meridian (0°N; 88°W). Further explanation of this experiment is in aforementioned section. As close to all GRACE coefficients up to degree 13 pass the statistical test, and Figure B.4 showed little to no impact of the statistical filter on the GRACE-derived trends, we will only execute this experiment for the Swarm-derived statistical filter.

Figure B.6 shows the result applying the Swarm-derived statistical filter to the spectral decomposition of the Dirac delta function. Both the Dirac function at the equator and the Dirac function near the Hudson Bay are spatially distorted after applying the filter. For the point at the equator, the east-west axis is elongated w.r.t. the north-south axis. The point near the Hudson Bay seems to have slightly increased in length in the north-south direction. This agrees with the observations of Davis et al. [20] and Sasgen et al. [79], who executed this experiment with a GRACE-derived mask.

Additionally, it seems like the function at the equator has shifted westward, and that the functions near the Hudson Bay has shifted to the northwest. This agrees with what we saw in Figure B.4, where the positive trend anomaly observed by Swarm around the Hudson Bay also shifted northeast after applying the spatial mask.

## Conclusion

The power of the statistical filter, when applied to GRACE, is that is suppresses the high-degree noise to such an extent that further processing is no longer needed [79]. For Swarm, we find that there is too much high degree noise to derive statistically significant trends. A very substantial part of the Swarm-derived coefficients is filtered out, with only sporadic coefficients past degree 13 passing the statistical tests. The statistical mask clearly indicates an error correlation between the high-degree-high-order coefficients of GRACE, but fails to show any correlation in the Swarm errors. Finally, the anisotropic nature of the statistical filter then leads to spatial distortion of the expanded Stokes coefficients, which makes the results harder to interpret. This greatly reduced the applicability of the statistical filter for geophysical study.

We conclude that, at time moment, we are unable to extract more geophysical signal from the Swarm-derived gravity fields using an anisotropic filter based on statistical significant of the trend component. Longer time series could make it easier to derive statistically significant parameters from the Swarm gravity fields, and therefore we do not rule out the statistical filter could be useful for application to the Swarm data in the future.

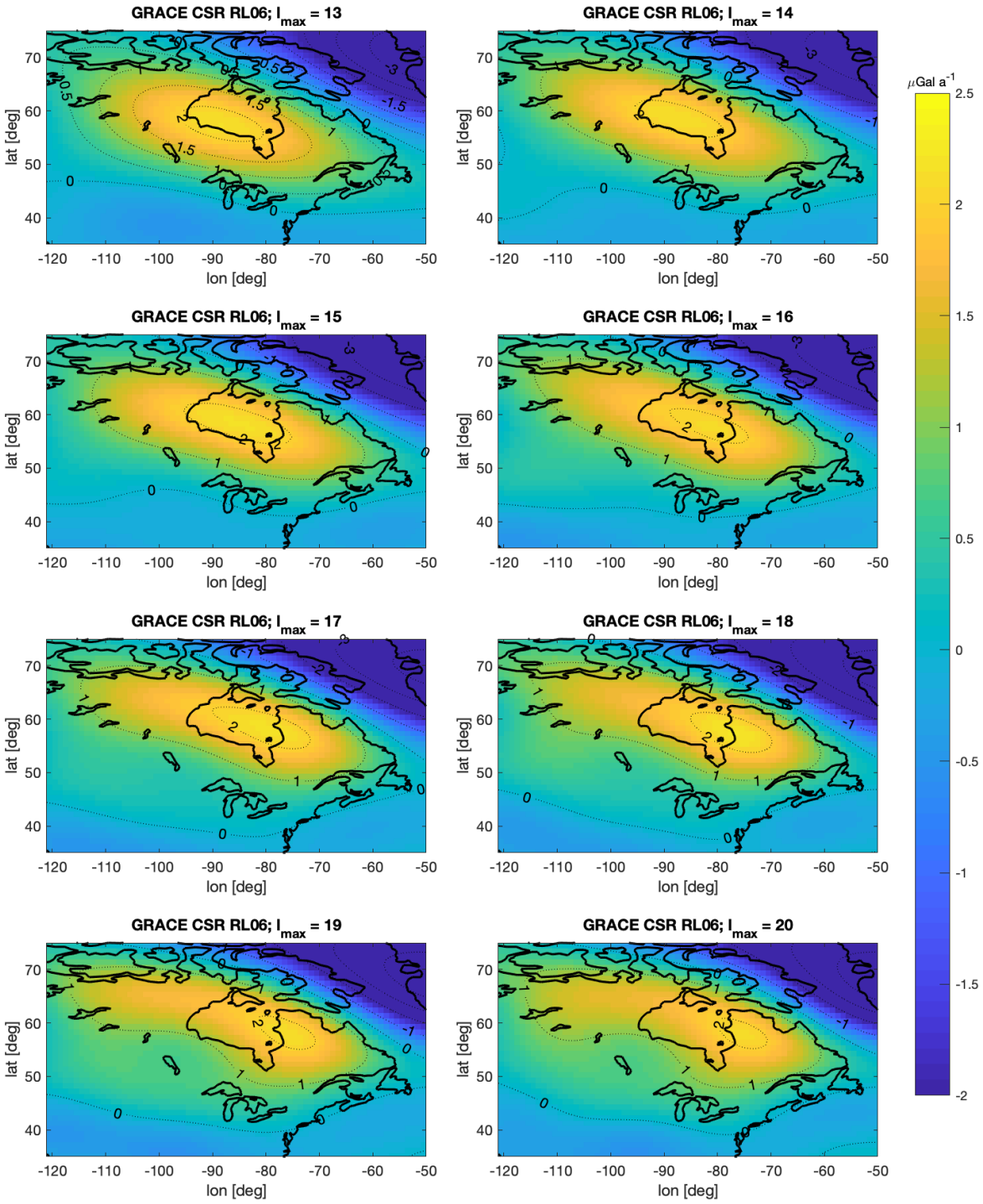



Figure B.3: Results of applying the statistical filter to the GRACE CSR RL06 fields truncated at various maximum degrees.

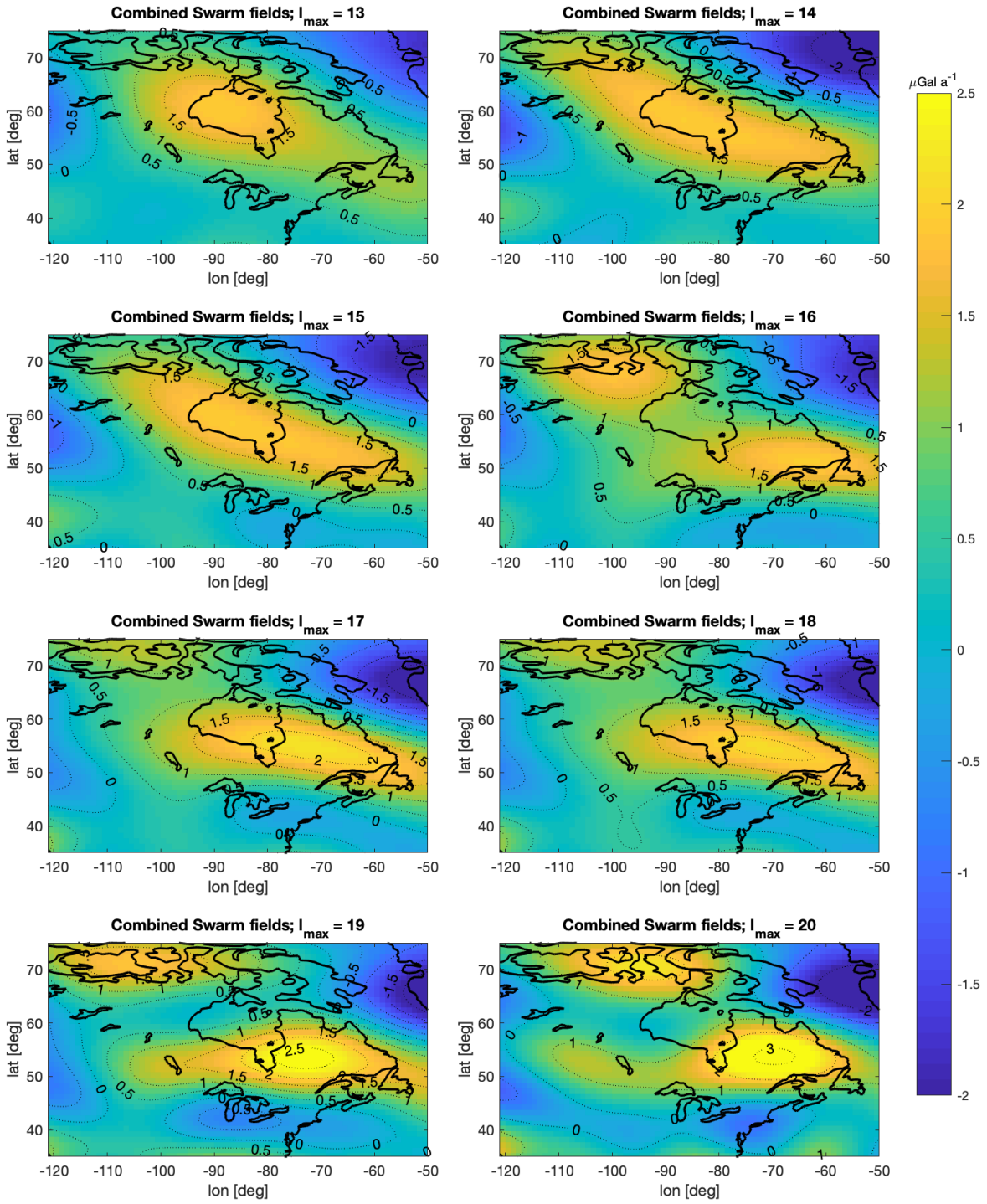



Figure B.4: Results of applying the statistical filter to the combined Swarm fields truncated at various maximum degrees.

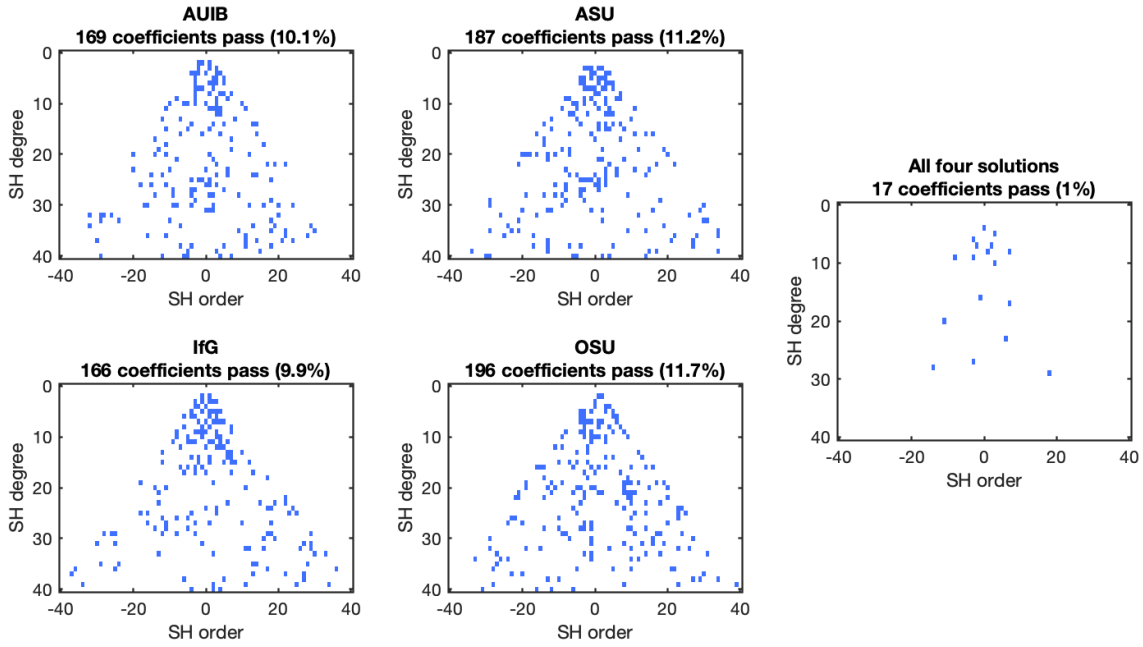



Figure B.5: Coefficients with a significant trend (95% confidence interval) in the individual components of the combined Swarm solutions (left and middle columns), and coefficients with a significant trend in all four solutions (right panel).

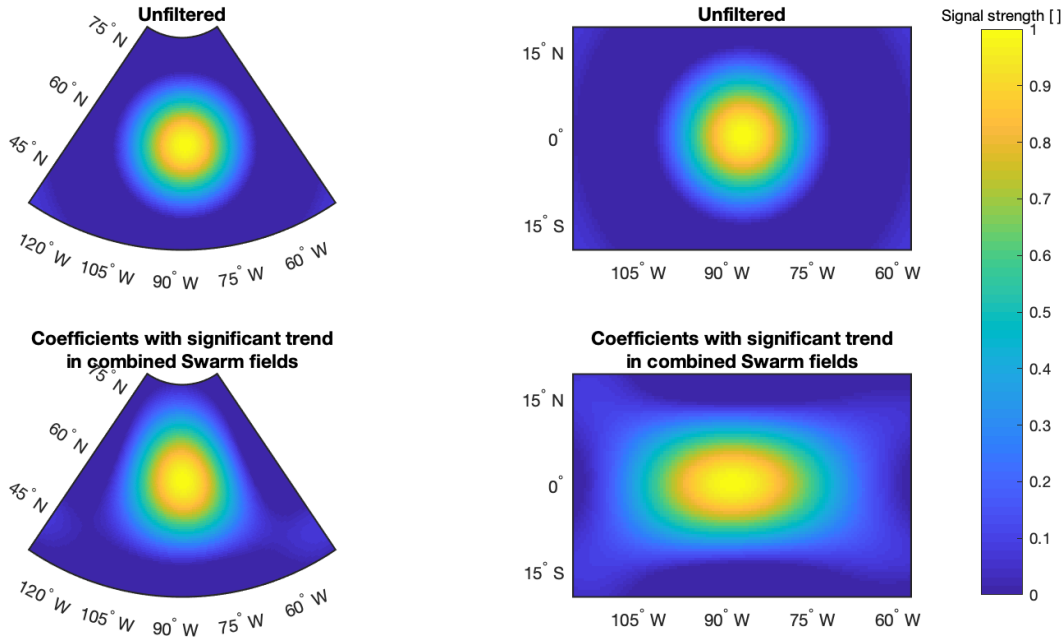



Figure B.6: Statistical filter computed from the trend estimates in the Stokes coefficients of the combined Swarm fields, applied to a Dirac delta function at 58°N; 88°W (left), and 0°N; 88°W (right).

# C

# Error analysis of the GeoQ solutions

In this appendix we will apply much of the same processing that we applied to the Swarm field in the main text of this thesis to the unfiltered GeoQ fields. Similar to Section , we do this for both the entire GeoQ data set, and a subset that we have trimmed to match the length of the combined Swarm fields of 64 months.

## Errors computed from Stokes coefficient RMSE

In this section we apply the methodology of Section 2.6 to the time series of GeoQ fields that is trimmed to the length of the combined Swarm solutions. We compute the spatial, spectral, and probability distributions of the observation uncertainties. First, we do this from RMSE results computed in the spectral domain, i.e., from Stokes coefficient time series. Second, we do the same but use RMSE results computed in the spatial domain, i.e., from time series of gravity change. For comparison with the GRACE- and Swarm-derived results in Section 4.2, we truncate the fields to $l_{max} = 13$.

Figure C.1 shows the results for the GeoQ fields trimmed to 64 months length, and Figure C.2 shows the results for the full GeoQ time series. When it comes to the spatial, spectral, or probability distribution we see very similar results for both the full and the trimmed data set to what we saw for the Swarm series in Figures 4.2, 4.3, and 4.5 (pp. 38 to 40). Spatially, we see a relatively homogeneously distributed errors with again a substantial increase near the poles. In Chapter 4 we explained this via the spectral distribution of the errors. As the spectral distribution of the errors in the Swarm-derived and multi-satellite fields are

## Errors computed from gravity RMSE

Here, we repeat the experiments of the last section, but apply the methods of Section 2.6 to gravity time series of a particular geographical location, instead of time series of Stokes coefficients observations.

Figure C.3 shows the results for the GeoQ fields trimmed to 64 months length, and Figure C.4 shows the results for the full GeoQ time series. For the trimmed series, we see a relatively homogeneous error distribution in the spatial domain. Nonetheless, the peaks in South America and India, along with the clearly non-Gaussian distribution of the land area residuals suggest that we have not been able to retrieve all of the geophysical signal from the observations. This is much like what we saw for GRACE in Section 4.2. For the full time series we see slightly different results. Spatially, the errors seem to increase for lower latitudes. This could be related to the fact that the ground track density, and thus the observation density, it higher at higher latitudes. Also, Figure C.4 seems to shows a relatively higher errors over the Southern Hemisphere. The probability distribution of the residuals of the full GeoQ data set also shows a peculiar pattern: the ocean area residuals diverge from the Gaussian distribution near the mean of the residuals. We have been unable to find a conclusive explanation for this behavior, but as the pattern is reminiscent of two overlapping normal distributions the origin is perhaps related to the fact that the GeoQ fields combine GPS and SLR observations.

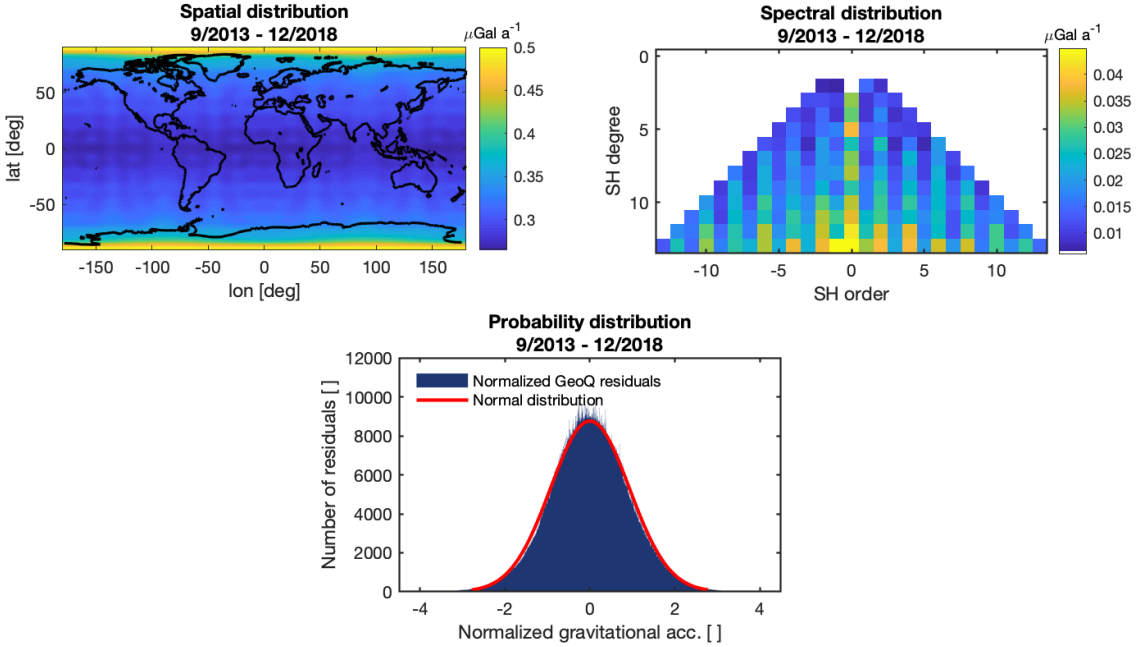



Figure C.1: Error analysis in the spectral domain for the GeoQ time series of GPS-derived geopotential fields, trimmed to September 2013 to December 2018. NB: color scales in this Figure do not match Figure C.2, nor any of the Figures in Section 4.2 where GRACE- and Swarm-derived error estimates are described.

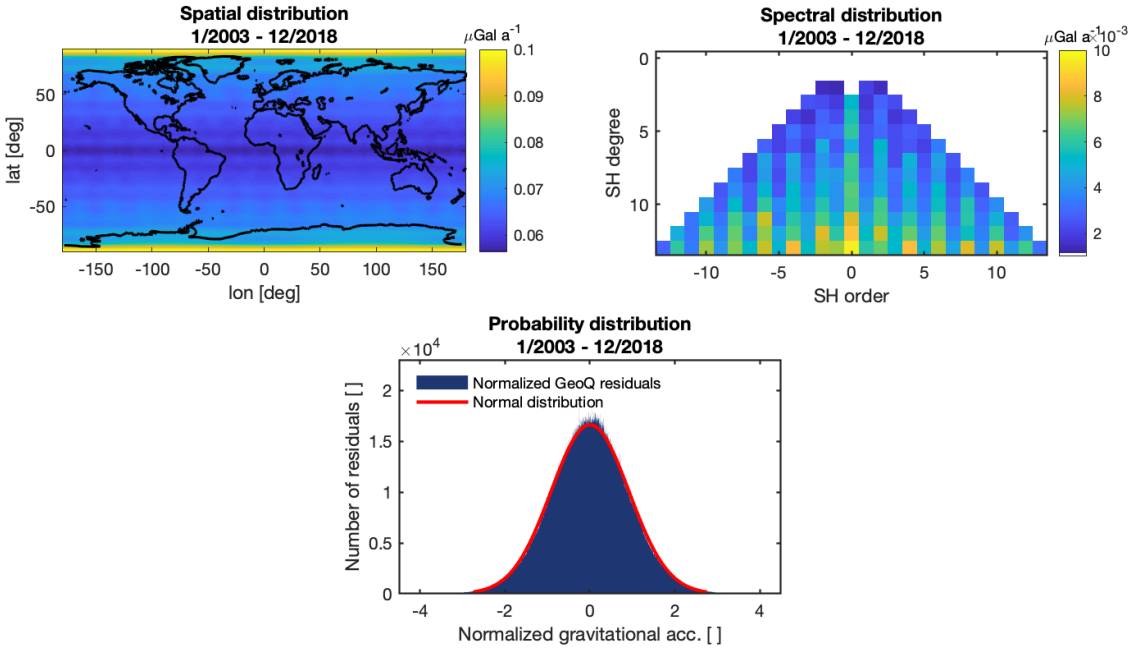



Figure C.2: Same as for Figure C.1, but using all available epochs in the GeoQ data set. NB: color scales in this Figure do not match Figure C.1, nor any of the Figures in Section 4.2 where GRACE- and Swarm-derived error estimates are described.

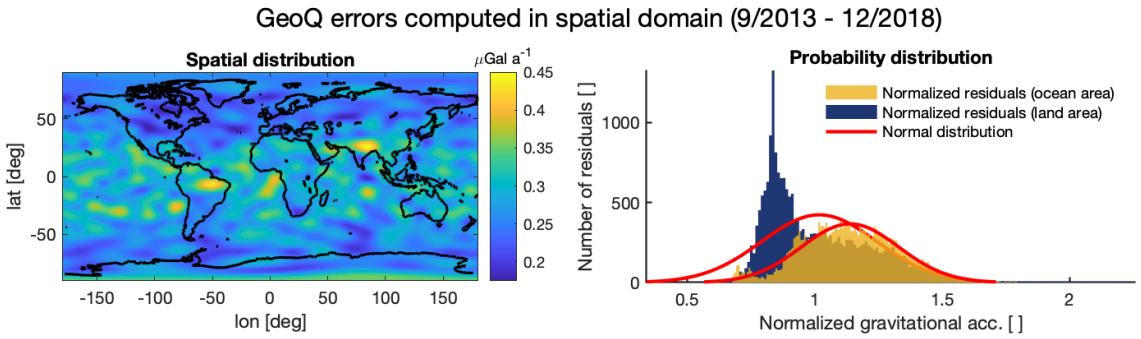



Figure C.3: Spatial distribution of the errors in the September 2013 to December 2018 subset of the GeoQ fields, as computed via the method outlined in Section 2.6.1, this time applied in the spatial domain. NB: color scale in this Figure does not match Figure C.4, nor any of the Figures in Section 4.2 where GRACE- and Swarm-derived error estimates are described.

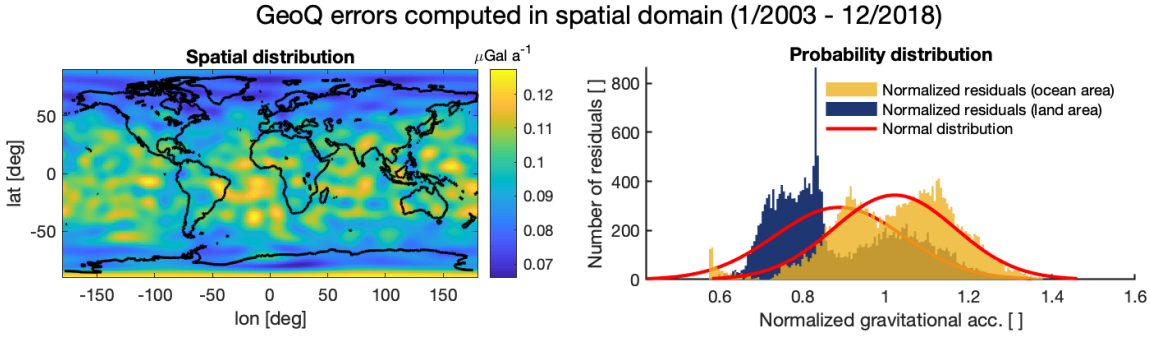



Figure C.4: Same as for Figure C.3, but using all available epochs in the GeoQ data set. NB: color scale in this Figure does not match Figure C.1, nor any of the Figures in Section 4.2 where GRACE- and Swarm-derived error estimates are described.

# Bibliography


[1] G. A, J. Wahr, and S. Zhong. Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: an application to Glacial Isostatic Adjustment in Antarctica and Canada. *Geophysical Journal International*, 192(2):557–572, 2013.

[2] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Government Printing Office, 1964.

[3] S. Adusumilli, A. Borsa, N. Lau, and W. Neely. Mass balance of Alaskan glaciers from GPS, GRACE, and satellite altimetry. In *AGU Fall Meeting Abstracts*, 2018.

[4] D. Arnold, U. Meyer, C. Dahle, and A. Jäggi. Combined Swarm/Sentinel Gravity Fields. In *Fourth Swarm Science Meeting*, 2017.

[5] O. Baur, M. Kuhn, and W. Featherstone. GRACE-derived ice-mass variations over Greenland by accounting for leakage effects. *Journal of Geophysical Research: Solid Earth*, 114(B6), 2009.

[6] R. Bernknopf, D. Brookshire, Y. Kuwayama, M. Macauley, M. Rodell, A. Thompson, P. Vail, and B. Zaitchik. The Value of Remotely Sensed Information: The Case of a GRACE-Enhanced Drought Severity Index. *Weather, Climate, and Society*, 10(1):187–203, 2018.

[7] S. Bettadpur. Level-2 Gravity Field Product User Handbook Rev 4.0. *Center for Space Research, The University of Texas at Austin*, 2018.

[8] G. Beutler, A. Jäggi, L. Mervart, and U. Meyer. The celestial mechanics approach: theoretical foundations. *Journal of Geodesy*, 84(10):605–624, 2010.

[9] A. Bezděk, J. Sebera, J. Klokočník, and J. Kostelecký. Gravity field models from kinematic orbits of CHAMP, GRACE and GOCE satellites. *Advances in Space Research*, 53(3):412–429, 2014.

[10] A. Bezděk, J. Sebera, J. de Teixeira da Encarnação, and J. Klokočník. Time-variable gravity fields derived from GPS tracking of Swarm. *Geophysical Journal International*, 205(3):1665–1669, 2016.

[11] M. Bloßfeld, S. Rudenko, A. Kehm, N. Panafidina, H. Müller, D. Angermann, U. Hugentobler, and M. Seitz. Consistent estimation of geodetic parameters from SLR satellite constellation measurements. *Journal of Geodesy*, 92(9): 1003–1021, 2018.

[12] M. Bloßfeld, V. Stefka, H. Müller, and M. Gerstl. Satellite laser ranging-a tool to realize GGOS? In *In: Rizos C., Willis P.(Eds.) IAG 150 Years, IAG Symposia*, volume 143, pages 540–547, 2015.

[13] Board, Space Studies and National Research Council. *Earth science and applications from space: national imperatives for the next decade and beyond*. National Academies Press, 2007.

[14] J. A. Bonin, D. P. Chambers, and M. Cheng. Using satellite laser ranging to measure ice mass change in Greenland and Antarctica. *The Cryosphere*, 12(1):71, 2018.

[15] D. P. Chambers, J. Wahr, and R. S. Nerem. Preliminary observations of global ocean mass variations with GRACE. *Geophysical Research Letters*, 31(13), 2004.

[16] J. Chen, M. Rodell, C. Wilson, and J. Famiglietti. Low degree spherical harmonic influences on Gravity Recovery and Climate Experiment (GRACE) water storage estimates. *Geophysical Research Letters*, 32(14), 2005.

[17] M. Cheng and J. Ries. The unexpected signal in GRACE estimates of $C_{20}$. *Journal of Geodesy*, 91(8):897–914, 2017.

[18] M. Cheng, B. Gunter, J. Ries, D. Chambers, and B. Tapley. Temporal variations in the Earth's gravity field from SLR and CHAMP GPS data. *3rd meeting of the International Gravity and Geoid Commission (IGGC)*, pages 424–431, 2002.

[19] C. Dahle, F. Flechtner, M. Murböck, G. Michalak, H. Neumayer, O. Abrykosov, A. Reinhold, and R. König. GRACE-FO Geopotential GSM Coefficients GFZ RL06. V. 6.0. GFZ Data Services, 2019. http://doi.org/10.5880/GFZ.GRACEFO_06_GSM.

[20] J. L. Davis, M. E. Tamisiea, P. Elósegui, J. Mitrovica, and E. M. Hill. A statistical filtering approach for Gravity Recovery and Climate Experiment (GRACE) gravity data. *Journal of Geophysical Research: Solid Earth*, 113(B4), 2008.

[21] J. Davis, P. Elósegui, J. Mitrovica, and M. Tamisiea. Climate-driven deformation of the solid Earth from GRACE and GPS. *Geophysical Research Letters*, 31(24), 2004.

[22] J. de Teixeira da Encarnação. *Next-generation satellite gravimetry for measuring mass transport in the Earth system*. PhD thesis, Delft University of Technology, 2015.

[23] J. de Teixeira da Encarnação, P. Visser, E. Doornbos, J. van den IJssel, X. Mao, E. Iorfida, D. Arnold, A. Jäggi, U. Meyer, A. Bezdek, J. Sebera, J. Klokocnik, M. Ellmer, T. Mayer-Gürr, S. Krauss, J. Guo, C. Zhang, C. Shum, and Y. Zhang. Complete 5-years time series of combined monthly gravity field models derived from Swarm GPS data. In *EGU2019-13412*, 2019.

[24] J. de Teixeira da Encarnação, D. Arnold, A. Bezděk, C. Dahle, E. Doornbos, J. van den IJssel, A. Jäggi, T. Mayer-Gürr, J. Sebera, P. Visser, and N. Zehentner. Gravity field models derived from Swarm GPS data. *Earth, Planets and Space*, 68(1):127, 2016. doi: 10.1186/s40623-016-0499-9.

[25] F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media, 2005.

[26] P. Ditmar. Conversion of time-varying Stokes coefficients into mass anomalies at the Earth's surface considering the Earth's oblateness. *Journal of Geodesy*, 92(12):1401–1412, 2018.

[27] P. Döll, M. Fritsche, A. Eicker, and H. M. Schmied. Seasonal water storage variations as impacted by water abstractions: comparing the output of a global hydrological model with GRACE and GPS observations. *Surveys in Geophysics*, 35(6):1311–1331, 2014.

[28] J. R. Driscoll and D. M. Healy. Computing Fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15(2):202–250, 1994.

[29] A. M. Dziewonski and D. L. Anderson. Preliminary reference Earth model. *Physics of the Earth and planetary interiors*, 25(4):297–356, 1981.

[30] K. M. Ellett, J. P. Walker, A. W. Western, and M. Rodell. A framework for assessing the potential of remote-sensed gravity to provide new insight on the hydrology of the Murray-Darling Basin. *Australasian Journal of Water Resources*, 10(2):125–138, 2006.

[31] H. H. Farahani, P. Ditmar, P. Inácio, O. Didova, B. Gunter, R. Klees, X. Guo, J. Guo, Y. Sun, X. Liu, et al. A high resolution model of linear trend in mass variations from DMT-2: Added value of accounting for coloured noise in GRACE data. *Journal of Geodynamics*, 103:12–25, 2017.

[32] W. Feng. GRAMAT: A comprehensive Matlab toolbox for estimating global mass variations from GRACE satellite data. *Earth Science Informatics*, pages 1–16, November 2018.

[33] R. Forsberg, L. Sørensen, and S. Simonsen. Greenland and Antarctica ice sheet mass changes and effects on global sea level. In *Integrative Study of the Mean Sea Level and Its Components*, pages 91–106. Springer, 2017.

[34] C. Fowler. *The Solid Earth: An Introduction to Global Geophysics*. Cambridge University Press, 2005.

[35] E. Friis-Christensen, H. Lühr, and G. Hulot. Swarm: A constellation to study the Earth's magnetic field. *Earth, planets and space*, 58(4):351–358, 2006.

[36] J. Guo, D. Mu, X. Liu, H. Yan, Z. Sun, and B. Guo. Water storage changes over the Tibetan Plateau revealed by GRACE mission. *Acta Geophysica*, 64(2):463–476, 2016.

[37] W. Heiskanen and H. Moritz. *Physical geodesy*. W.H. Freeman and Company, 1967.

[38] A. Jäggi, C. Dahle, D. Arnold, H. Bock, U. Meyer, G. Beutler, and J. van den IJssel. Swarm kinematic orbits and gravity fields from 18 months of GPS data. *Advances in Space Research*, 57(1):218–233, 2016.

[39] A. Jäggi, M. Weigelt, F. Flechtner, A. Güntner, T. Mayer-Gürr, S. Martinis, S. Bruinsma, J. Flury, S. Bourgogne, H. Steffen, et al. European Gravity Service for Improved Emergency Management (EGSIEM)—from concept to implementation. *Geophysical Journal International*, 218(3):1572–1590, 2019.

[40] S. Jin, T. Zhang, and F. Zou. Glacial density and GIA in Alaska estimated from ICESat, GPS and GRACE measurements. *Journal of Geophysical Research: Earth Surface*, 122(1):76–90, 2017.

[41] J. Kusche. Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models. *Journal of Geodesy*, 81(11):733–749, 2007.

[42] K. Lambeck, A. Purcell, and S. Zhao. The North American Late Wisconsin ice sheet and mantle viscosity from glacial rebound analyses. *Quaternary Science Reviews*, 158:172–210, 2017.

[43] J. M. Lemoine, S. Bruinsma, S. Loyer, R. Biancale, J. C. Marty, F. Perosanz, and G. Balmino. Temporal gravity field models inferred from GRACE data. *Advances in Space Research*, 39(10):1620–1629, 2007.

[44] E. W. Leuliette and L. Miller. Closing the sea level rise budget with altimetry, Argo, and GRACE. *Geophysical Research Letters*, 36(4), 2009.

[45] D. Lineberry. Celebrate With the World for One Night: Yuri's Night. *NASA Langley Center, The Researcher News*, 2009.

[46] C. Lück, J. Kusche, R. Rietbroek, and A. Löcher. Time-variable gravity fields and ocean mass change from 37 months of kinematic Swarm orbits. *Solid Earth*, 9:323–339, 2018.

[47] S. B. Luthcke, H. J. Zwally, W. Abdalati, D. Rowlands, R. Ray, R. Nerem, F. Lemoine, J. McCarthy, and D. Chinn. Recent Greenland ice mass loss by drainage system from satellite gravity observations. *Science*, 314(5803):1286–1289, 2006.

[48] Z. Martinec. The density contrast at the Mohorovičic discontinuity. *Geophysical Journal International*, 117(2): 539–544, 1994.

[49] K. Matsuo, T. Otsubo, M. Bloßfeld, and H. Müller. SLR contributions to time-variable gravity field and thermospheric neutral density estimates. In *ILRS Tech Workshop*, 2017.

[50] T. Mayer-Gürr. *Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2006.

[51] U. Meyer, K. Sosnica, D. Arnold, C. Dahle, D. Thaller, R. Dach, and A. Jäggi. SLR, GRACE and Swarm Gravity Field Determination and Combination. *Remote Sensing*, 11(8):956, 2019.

[52] J. Mouginot, E. Rignot, A. A. Bjørk, M. van den Broeke, R. Millan, M. Morlighem, B. Noël, B. Scheuchl, and M. Wood. Forty-six years of Greenland Ice Sheet mass balance from 1972 to 2018. *Proceedings of the National Academy of Sciences*, 116(19):9239–9244, 2019.

[53] NASA Jet Performance Laboratory. GRACE-FO monthly geopotential spherical harmonics JPL release 6.0, 2019. URL `https://podaac.jpl.nasa.gov/dataset/GRACEFO_L2_JPL_MONTHLY_0060`.

[54] National Academies of Sciences, Engineering, and Medicine. *Thriving on our changing planet: A decadal strategy for Earth observation from space*. National Academies Press, 2018.

[55] A. Paulson, S. Zhong, and J. Wahr. Inference of mantle viscosity from GRACE and relative sea level data. *Geophysical Journal International*, 171(2):497–508, 2007.

[56] A. Paulson, S. Zhong, and J. Wahr. Limitations on the inversion for mantle viscosity from postglacial rebound. *Geophysical Journal International*, 168(3):1195–1209, 2007.

[57] J. L. Peixoto. Incorrect F formulas in regression analysis. *The American Statistician*, 47(3):194–197, 1993.

[58] W. R. Peltier. The history of the Earth's rotation: Impacts of deep Earth physics and surface climate variability. *Treatise of Geophysics*, 9:221–279, 2015.

[59] W. R. Peltier, R. Drummond, and K. Roy. Comment on "Ocean mass from GRACE and glacial isostatic adjustment" by DP Chambers et al. *Journal of Geophysical Research: Solid Earth*, 117(B11), 2012.

[60] W. R. Peltier, D. F. Argus, and R. Drummond. Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model. *Journal of Geophysical Research: Solid Earth*, 120(1):450–487, 2015.

[61] W. R. Peltier, D. F. Argus, and R. Drummond. Comment on "An Assessment of the ICE-6G_C (VM5a) Glacial Isostatic Adjustment Model" by Purcell et al. *Journal of Geophysical Research: Solid Earth*, 123(2):2019–2028, 2018.

[62] G. Petit and B. Luzum. International Earth Rotation and Reference Systems Service (IERS) conventions (2010). Technical report, Bureau International des Poids et Mesures, 2010.

[63] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C++: The art of scientific computing*. Cambridge University Press, 1992.

[64] J. Ran, P. Ditmar, R. Klees, and H. Farahani. Statistically optimal estimation of Greenland Ice Sheet mass variations from GRACE monthly solutions using an improved mascon approach. *Journal of Geodesy*, 92(3):299–319, 2018.

[65] E. Rangelova, W. Van der Wal, A. Braun, M. G. Sideris, and P. Wu. Analysis of Gravity Recovery and Climate Experiment time-variable mass redistribution signals over North America by means of principal component analysis. *Journal of Geophysical Research: Earth Surface*, 112(3):1–12, 2007.

[66] R. H. Rapp. A Fortran program for the computation of gravimetric quantities from high degree spherical harmonic expansions. Technical report, The Ohio State University Department of Geodetic Science and Surveying, 1982.

[67] R. Ray and S. Luthcke. Tide model errors and GRACE gravimetry: towards a more realistic assessment. *Geophysical Journal International*, 167(3):1055–1059, 2006.

[68] R. D. Ray, D. Rowlands, and G. Egbert. Tidal models in a new era of satellite gravimetry. *Space Science Reviews*, 108 (1-2):271–282, 2003.

[69] A. S. Richey, B. F. Thomas, M.-H. Lo, J. T. Reager, J. S. Famiglietti, K. Voss, S. Swenson, and M. Rodell. Quantifying renewable groundwater stress with GRACE. *Water resources research*, 51(7):5217–5238, 2015.

[70] J. Ries, S. Bettadpur, R. Eanes, Z. Kang, U. Ko, C. McCullough, P. Nagel, N. Pie, S. Poole, T. Richter, et al. The Combination Global Gravity Model GGM05C. *Center for Space Research–The University of Texas at Austin, Texas, CSE-TM-16-01*, 2016.

[71] R. Rietbroek, M. Fritsche, C. Dahle, S.-E. Brunnabend, M. Behnisch, J. Kusche, F. Flechtner, J. Schröter, and R. Dietrich. Can GPS-derived surface loading bridge a GRACE mission gap? *Surveys in Geophysics*, 35(6):1267–1283, 2014.

[72] E. Rignot, J. Mouginot, B. Scheuchl, M. van den Broeke, M. J. van Wessem, and M. Morlighem. Four decades of Antarctic Ice Sheet mass balance from 1979–2017. *Proceedings of the National Academy of Sciences*, 116(4):1095–1103, 2019.

[73] M. Rodell, J. Famiglietti, D. Wiese, J. Reager, H. Beaudoing, F. Landerer, and M.-H. Lo. Emerging trends in global freshwater availability. *Nature*, 557(7707):651–659, 2018.

[74] M. Rodell, P. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, et al. The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85 (3):381–394, 2004.

[75] B. Root, L. Tarasov, and W. Van der Wal. GRACE gravity observations constrain Weichselian ice thickness in the Barents Sea. *Geophysical Research Letters*, 42(9):3313–3320, 2015.

[76] K. Roy. *High-quality constraints on the glacial isostatic adjustment process over North America: The ICE-7G_NA (VM7) model*. PhD thesis, University of Toronto, 2017.

[77] K. Roy and W. R. Peltier. Space-geodetic and water level gauge constraints on continental uplift and tilting over North America: regional convergence of the ICE-6G_C (VM5a/VM6) models. *Geophysical Journal International*, 210(2):1115–1142, 2017.

[78] I. Sasgen, Z. Martinec, and K. Fleming. Regional ice-mass changes and glacial-isostatic adjustment in Antarctica from GRACE. *Earth and Planetary Science Letters*, 264(3-4):391–401, 2007.

[79] I. Sasgen, V. Klemann, and Z. Martinec. Towards the inversion of GRACE gravity fields for present-day ice-mass changes and glacial-isostatic adjustment in North America and Greenland. *Journal of Geodynamics*, 59:49–63, 2012.

[80] H. Save. CSR GRACE RL06 Mascon Solutions, 2019. URL `https://doi.org/10.18738/T8/UN91VR`.

[81] H. Save, S. Bettadpur, and B. D. Tapley. Reducing errors in the GRACE gravity solutions using regularization. *Journal of Geodesy*, 86(9):695–711, 2012.

[82] H. Save, S. Bettadpur, and B. D. Tapley. High-resolution CSR GRACE RL05 mascons. *Journal of Geophysical Research: Solid Earth*, 121(10):7547–7569, 2016.

[83] B. R. Scanlon, Z. Zhang, H. Save, A. Y. Sun, H. M. Schmied, L. P. van Beek, D. N. Wiese, Y. Wada, D. Long, R. C. Reedy, L. Longuevergne, P. Döll, and M. F. Bierkens. Global models underestimate large decadal declining and rising water storage trends relative to GRACE satellite data. *Proceedings of the National Academy of Sciences*, 115 (6):E1080–E1089, 2018.

[84] R. Schmidt, F. Flechtner, U. Meyer, K.-H. Neumayer, C. Dahle, R. König, and J. Kusche. Hydrological signals observed by the GRACE satellites. *Surveys in Geophysics*, 29(4-5):319–334, 2008.

[85] R. Schmidt, F. Flechtner, U. Meyer, K.-H. Neumayer, C. Dahle, R. König, and J. Kusche. Hydrological signals observed by the GRACE satellites. *Surveys in Geophysics*, 29(4-5):319–334, 2008.

[86] L. F. Schreiter, D. Arnold, V. J. Sterken, and A. Jäggi. Mitigation of ionospheric signatures in Swarm GPS gravity field estimation using weighting strategies. *Annales geophysicae*, 37:111–127, 2019.

[87] K.-W. Seo, C. R. Wilson, J. Chen, and D. E. Waliser. GRACE's spatial aliasing error. *Geophysical Journal International*, 172(1):41–48, 2008.

[88] K. Shang, J. Guo, C. Shum, C. Dai, and J. Luo. GRACE time-variable gravity field recovery using an improved energy balance approach. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, 203(3):1773–1786, 2015.

[89] A. Shepherd, E. Ivins, E. Rignot, B. Smith, M. van den Broeke, I. Velicogna, P. Whitehouse, K. Briggs, I. Joughin, G. Krinner, et al. Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature*, 556:219–222, 2018.

[90] L. S. Sørensen, A. H. Jarosch, G. Aðalgeirsdóttir, V. R. Barletta, R. Forsberg, F. Pálsson, H. Björnsson, and T. Jóhannesson. The effect of signal leakage and glacial isostatic rebound on GRACE-derived ice mass changes in Iceland. *Geophysical Journal International*, 209(1):226–233, 2017.

[91] K. Sośnica, A. Jäggi, U. Meyer, D. Thaller, G. Beutler, D. Arnold, and R. Dach. Time variable Earth's gravity field from SLR satellites. *Journal of Geodesy*, 89(10):945–960, 2015.

[92] H. Steffen, H. Denker, and J. Müller. Glacial isostatic adjustment in Fennoscandia from GRACE data and comparison with geodynamical models. *Journal of Geodynamics*, 46(3-5):155–164, 2008.

[93] H. Steffen, O. Gitlein, H. Denker, J. Müller, and L. Timmen. Present rate of uplift in Fennoscandia from GRACE and absolute gravimetry. *Tectonophysics*, 474(1-2):69–77, 2009.

[94] H. Steffen, S. Petrovic, J. Müller, R. Schmidt, J. Wünsch, F. Barthelmes, and J. Kusche. Significance of secular trends of mass variations determined from GRACE solutions. *Journal of Geodynamics*, 48(3-5):157–165, 2009.

[95] A. Y. Sun, B. R. Scanlon, Z. Zhang, D. Walling, S. N. Bhanja, A. Mukherjee, and Z. Zhong. Combining Physically Based Modeling and Deep Learning for Fusing GRACE Satellite Data: Can We Learn From Mismatch? *Water Resources Research*, 55(2):1179–1195, 2019.

[96] S. Swenson and J. Wahr. Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters*, 33(8), 2006.

[97] M. J. Talpe, R. S. Nerem, E. Forootan, M. Schmidt, F. G. Lemoine, E. M. Enderlin, and F. W. Landerer. Ice mass change in Greenland and Antarctica between 1993 and 2013 from satellite gravity measurements. *Journal of Geodesy*, 91 (11):1283–1298, 2017.

[98] M. E. Tamisiea, J. X. Mitrovica, and J. L. Davis. GRACE Gravity Data Constrain Ancient Ice Geometries and Continental Dynamics over Laurentia. *Science*, 316(5826):881–883, 2007.

[99] B. D. Tapley, S. Bettadpur, M. Watkins, and C. Reigber. The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, 31(9):1–4, 2004.

[100] B. D. Tapley, S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins. GRACE measurements of mass variability in the Earth system. *Science*, 305(5683):503–505, 2004.

[101] B. D. Tapley, M. M. Watkins, F. Flechtner, C. Reigber, S. Bettadpur, M. Rodell, I. Sasgen, J. S. Famiglietti, F. W. Landerer, D. P. Chambers, J. T. Reager, A. S. Gardner, H. Save, E. R. Ivins, S. C. Swenson, C. Boening, C. Dahle, D. N. Wiese, H. Dobslaw, M. E. Tamisiea, and I. Velicogna. Contributions of GRACE to understanding climate change. *Nature Climate Change*, 9(5):358–369, 2019.

[102] The University Of Texas Center For Space Research. GRACE static field geopotential coefficients CSR release 6.0, 2018. URL `https://podaac.jpl.nasa.gov/dataset/GRACE_GSM_L2_GRAV_CSR_RL06`.

[103] The University of Texas Center for Space Research. GRACE-FO monthly geopotential spherical harmonics CSR release 6.0, 2019. URL `https://podaac.jpl.nasa.gov/dataset/GRACEFO_L2_CSR_MONTHLY_0060`.

[104] J. Van den IJssel, B. Forte, and O. Montenbruck. Impact of Swarm GPS receiver updates on POD performance. *Earth, Planets and Space*, 68(1):85, 2016.

[105] W. Van der Wal, E. Rangelova, M. Sideris, and P. Wu. Secular geoid rate from GRACE for vertical datum modernization. In *Gravity, Geoid and Earth Observation*, pages 611–617. Springer, 2010.

[106] W. Van der Wal, P. Wu, M. G. Sideris, and C. Shum. Use of GRACE determined secular gravity rates for glacial isostatic adjustment studies in North-America. *Journal of Geodynamics*, 46(3-5):144–154, 2008.

[107] W. van der Wal, E. Kurtenbach, J. Kusche, and B. Vermeersen. Radial and tangential gravity rates from GRACE in areas of glacial isostatic adjustment. *Geophysical Journal International*, 187(2):797–812, 2011.

[108] W. Van der Wal, P. L. Whitehouse, and E. J. Schrama. Effect of GIA models with 3D composite mantle viscosity on GRACE mass balance estimates for Antarctica. *Earth and Planetary Science Letters*, 414:134–143, 2015.

[109] I. Velicogna and J. Wahr. Time-variable gravity observations of ice sheet mass balance: Precision and limitations of the GRACE satellite data. *Geophysical Research Letters*, 40(12):3055–3063, 2013.

[110] I. Velicogna, T. Sutterley, and M. Van Den Broeke. Regional acceleration in ice mass loss from Greenland and Antarctica using GRACE time-variable gravity data. *Geophysical Research Letters*, 41(22):8130–8137, 2014.

[111] I. Velicogna. Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE. *Geophysical Research Letters*, 36(19), 2009.

[112] J. Wahr, M. Molenaar, and F. Bryan. Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE. *Journal of Geophysical Research: Solid Earth*, 103(B12):30205–30229, 1998.

[113] J. Wahr, S. Swenson, V. Zlotnicki, and I. Velicogna. Time-variable gravity from GRACE: First results. *Geophysical Research Letters*, 31(11):20–23, 2004.

[114] J. Wahr, S. Swenson, and I. Velicogna. Accuracy of GRACE mass estimates. *Geophysical Research Letters*, 33(6):1–5, 2006.

[115] L. Wang, J. L. Davis, E. M. Hill, and M. E. Tamisiea. Stochastic filtering for determining gravity variations for decade-long time series of GRACE gravity. *Journal of Geophysical Research: Solid Earth*, 121(4):2915–2931, 2016.

[116] M. Weigelt, A. Jäggi, U. Meyer, D. Arnold, A. Grahsl, K. Sosnica, C. Dahle, and F. Flechtner. HlSST and SLR-bridging the gap between GRACE and GRACE Follow-on. In *EGU General Assembly Conference Abstracts*, volume 19, page 14258, 2017.

[117] M. Weigelt, A. Jäggi, U. Meyer, D. Arnold, T. Mayer-Gürr, N. Zehentner, S. Krauss, B. Dutt Vishwakarma, B. Devaraju, H. Steffen, K. Sośnica, L. Ren, and S. Schön. Bridging the gap between GRACE and GRACE-Follow On. In *IUGG General Assembly Conference Abstracts*, 2019.

[118] H. White. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48(4):817–838, 1980.

[119] P. L. Whitehouse. Glacial isostatic adjustment modelling: historical perspectives, recent advances, and future directions. *Earth Surface Dynamics*, 6(2):401–429, 2018.

[120] World Glacier Monitoring Service (WGMS). Fluctuations of Glaciers Database, 2018. URL `http://dx.doi.org/10.5904/wgms-fog-2018-11`.

[121] X. Wu, J. Ray, and T. van Dam. Geocenter motion and its geodetic and geophysical implications. *Journal of Geodynamics*, 58:44–61, 2012.

[122] C. Xiong, C. Stolle, and H. Lühr. The Swarm satellite loss of GPS signal and its relation to ionospheric plasma irregularities. *Space Weather*, 14(8):563–577, 2016.

[123] D.-N. Yuan. GRACE Follow-On Level-2 Gravity Field Product User Handbook. *Jet Propulsion Laboratory, JPL D-103922*, 2019. URL `https://podaac-tools.jpl.nasa.gov/drive/files/allData/gracefo/docs/GRACE-FO_L2-UserHandbook_v1.0.pdf`.

[124] N. Zehentner and T. Mayer-Gürr. Precise orbit determination based on raw GPS measurements. *Journal of Geodesy*, 90(3):275–286, 2016.

[125] N. P. Zelensky, F. G. Lemoine, D. S. Chinn, S. Melachroinos, B. D. Beckley, J. W. Beall, and O. Bordyugov. Estimated SLR station position and network frame sensitivity to time-varying gravity. *Journal of Geodesy*, 88(6):517–537, 2014.

[126] X. Zhang, S. Jin, and X. Lu. Global Surface Mass Variations from Continuous GPS Observations and Satellite Altimetry Data. *Remote Sensing*, 9(10):1000, 2017.