



Delft University of Technology

Conversion of time-varying Stokes coefficients into mass anomalies at the Earth's surface considering the Earth's oblateness

Ditmar, Pavel

DOI

[10.1007/s00190-018-1128-0](https://doi.org/10.1007/s00190-018-1128-0)

Publication date

2018

Document Version

Final published version

Published in

Journal of Geodesy

Citation (APA)

Ditmar, P. (2018). 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. <https://doi.org/10.1007/s00190-018-1128-0>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Conversion of time-varying Stokes coefficients into mass anomalies at the Earth's surface considering the Earth's oblateness

Pavel Ditmar¹

Received: 1 July 2017 / Accepted: 6 February 2018 / Published online: 19 February 2018
© The Author(s) 2018. This article is an open access publication

Abstract

Time-varying Stokes coefficients estimated from GRACE satellite data are routinely converted into mass anomalies at the Earth's surface with the expression proposed for that purpose by Wahr et al. (J Geophys Res 103(B12):30,205–30,229, 1998). However, the results obtained with it represent mass transport at the spherical surface of 6378 km radius. We show that the accuracy of such conversion may be insufficient, especially if the target area is located in a polar region and the signal-to-noise ratio is high. For instance, the peak values of mean linear trends in 2003–2015 estimated over Greenland and Amundsen Sea embayment of West Antarctica may be underestimated in this way by about 15%. As a solution, we propose an updated expression for the conversion of Stokes coefficients into mass anomalies. This expression is based on the assumptions that: (i) mass transport takes place at the reference ellipsoid and (ii) at each point of interest, the ellipsoidal surface is approximated by the sphere with a radius equal to the current radial distance from the Earth's center ("locally spherical approximation"). The updated expression is nearly as simple as the traditionally used one but reduces the inaccuracies of the conversion procedure by an order of magnitude. In addition, we remind the reader that the conversion expressions are defined in spherical (geocentric) coordinates. We demonstrate that the difference between mass anomalies computed in spherical and ellipsoidal (geodetic) coordinates may not be negligible, so that a conversion of geodetic colatitudes into geocentric ones should not be omitted.

Keywords Stokes coefficients · Spherical harmonics · Time-varying gravity · Mass transport · GRACE · GRACE Follow-On

1 Introduction

Since the launch of Gravity Recovery and Climate Experiment (GRACE) satellite mission in 2002 (<http://www.csr.utexas.edu/grace>), satellite gravimetry has become one of the key tools to study large-scale mass transport in the Earth's system. Mass transport estimates based on GRACE data have been successfully applied in numerous studies of the solid Earth, cryosphere, oceans, and continental water resources (for an overview, see, e.g., Wouters et al. 2014). To produce these estimates, one exploits information about temporal gravity variations sensed by GRACE satellites.

Newton's law of gravitational attraction allows temporal gravity field variations to be computed uniquely, as soon as mass anomalies are given (see, e.g., Chao et al. 1987). Such a relationship is used, for instance, to esti-

mate rapid gravity field variations caused by non-tidal mass transport in the ocean and atmosphere. These estimates are distributed as a so-called Atmosphere and Ocean De-aliasing product AOD1B (Dobslaw et al. 2016), which is needed, along with other background force models, for a preprocessing of GRACE data. This is because rapid mass transport signals cannot be properly interpreted and play a role of additional source of noise in GRACE data. To properly process mass re-distribution in the atmosphere, it is essential to take into account its vertical structure (Swenson and Wahr 2002; Flechtner 2007), as well as the oblateness of the Earth (Forootan et al. 2013). All these considerations are taken into account in the production of the latest release (RL06) of the AOD1B product (Dobslaw et al. 2016).

The inverse problem—conversion of temporal gravity field variations into mass anomalies—is more difficult. It cannot be solved uniquely without additional assumptions about the mass transport. A commonly used assumption is that mass transport takes place at a sphere of a certain radius a . Then, the Stokes coefficients describing the time-varying gravity field can be uniquely related to Fourier coefficients of

✉ Pavel Ditmar
p.g.ditmar@tudelft.nl

¹ Delft University of Technology, Stevinweg 1, 2628, CN, Delft, The Netherlands

mass anomalies, provided that the corresponding base functions are defined as surface spherical harmonics (Chao et al. 1987). Wahr et al. (1998) proposed to apply such an approach to recovery of mass anomalies from GRACE data, and this has been done since then in hundreds of studies.

Unfortunately, the overwhelming majority of authors ignore the fact that the actual Earth's surface substantially deviates from the sphere. If the radius a is defined as the mean equatorial radius or the semimajor axis of the reference ellipsoid (which is typically the case), these deviations become particularly large in polar areas. Obviously, such deviations may result in some distortions in mass anomaly estimates (Chao 2005, 2016). However, they are not quantified so far in the context of GRACE-based estimates, the temporal resolution of which is limited to about 400 km (in terms of wavelengths).

Furthermore, the commonly used expressions for a conversion of Stokes coefficients into mass anomalies refer to *spherical* (geocentric) coordinates of a point, including the geocentric colatitude θ . However, only a few authors define the exploited coordinates properly. In most of cases, θ is defined as just "the colatitude." We consider this as an indication that the step of converting standard ellipsoidal (geodetic) geographical coordinates into spherical ones is likely omitted in many studies, especially if they are conducted by non-geodesists. Though the difference between geocentric and geodetic colatitudes is relatively minor ($\lesssim 11.5' \approx 21$ km), it might be improper to ignore it in all cases.

The primary goals of this publication are: (i) to demonstrate that mass anomaly estimates produced with the expressions from (Wahr et al. 1998) may not be sufficiently accurate (in particular, this concerns long-term mass losses in polar areas); (ii) to show that a simple modification of them may increase the accuracy by an order of magnitude; and (iii) to demonstrate that an (erroneous) interpretation of spherical coordinate θ as a geodetic colatitude may have a non-negligible effect onto the estimated mass anomalies.

It is important to stress that our discussion is limited to a recovery of a 2-D mass re-distribution. In practice, this means that our focus is on hydrology, ice sheets, and oceans. The corresponding mass variations take place at the Earth's surface or just below it (at a depth not exceeding in most cases a few hundreds of meters). Then, our assumption that mass transport is confined to a thin layer near the Earth's surface is fully justified, and a quantification of it in terms of equivalent water heights is physically meaningful.

Mass transport in the solid Earth definitely occurs deeper, spanning a much larger range of depths. For instance, hypocenters of large earthquakes are typically located at the depth of a few tens of km, whereas glacial isostatic adjustment (GIA) takes place in the asthenosphere, the top boundary of which is located in most places at the depth of 100–200 km. Therefore, the techniques discussed in this

paper are not applicable to mass transport in the solid Earth. A more extended discussion of limitations associated with a recovery of 3-D mass transport processes from GRACE data can be found in Chao (2016).

The structure of this paper is as follows. In Sect. 2, we present a general expression for the computation of mass anomalies at the Earth's surface, as well as possible simplifications of that expression, depending on the assumption about the Earth's surface shape. That section concludes by a comparison of the proposed simplifications using real GRACE data, which allows us to identify the expression that is the most appropriate in practice. In Sect. 3, we address the issue of the proper definition of the coordinate θ . Section 4 is left for a discussion and conclusions.

2 Estimation of mass anomalies under different assumptions about the Earth surface geometry

2.1 Theory

2.1.1 General information

Information delivered by GRACE is usually provided to the Earth science community in the form of monthly sets of Stokes coefficients. Those coefficients represent the mean value of the Earth's gravitation potential $U(r, \theta, \phi, t)$ within a given month. The subtraction of a long-term mean value from each coefficient results in a time series of its temporal variations. Those variations can be linked to temporal variations $\Delta U(r, \theta, \phi)$ of the gravitational potential (see, e.g., Heiskanen and Moritz 1967, Eqs. 2–39):

$$\Delta U(r, \theta, \phi) = \frac{GM_E}{a} \sum_{l=1}^{L_{\max}} \left(\frac{a}{r}\right)^{l+1} \sum_{m=0}^l [\Delta \bar{C}_{lm} \bar{P}_{lm}(\theta) \cos m\phi + \Delta \bar{S}_{lm} \bar{P}_{lm}(\theta) \sin m\phi], \quad (1)$$

where (r, θ, ϕ) are spherical coordinates of a given point (radial distance, colatitude, and longitude) in the terrestrial reference frame; G is the universal gravitational constant; M_E is the Earth's mass; a is the Earth's radius (a more specific definition of this parameter is addressed below); l and m are the spherical harmonic degree and order, respectively; L_{\max} is a model-specific maximum degree; ΔC_{lm} and ΔS_{lm} are temporal variations of Stokes coefficients (by definition, $\Delta S_{lm} = 0$ for $m = 0$); and $\bar{P}_{l,m}$ are normalized associated Legendre functions. The argument of time is omitted in Eq. (1) for the sake of brevity. The summation does not contain the degree-0 term, since variations in the total mass of the Earth (which are described by that term) are negligible.

The exact definition of the Earth’s radius a is a matter of convention. Classically, this parameter is defined as the semimajor axis of the reference ellipsoid or the mean Earth’s equatorial radius: $a \sim 6,378,136$ m (Heiskanen and Moritz 1967). GRACE data products are also provided in line with this definition. On the other hand, many publications define a , explicitly or implicitly, as the mean radius of the entire Earth: $a = 6371$ km (e.g., Chao et al. 1987; Wahr et al. 1998; Swenson and Wahr 2002). Strictly speaking, this means that an application of the expressions derived in those publications requires the corresponding rescaling of the GRACE-based Stokes coefficients. To the best of the author’s knowledge, however, this is never done in practice. Apparently, the impact of this rescaling is assumed to be minor. In any case, in all the derivations presented below, a is defined as the semimajor axis of the reference ellipsoid.

To make the further derivations simpler, we rewrite Eq. (1) in a more compact form:

$$\Delta U(r, \theta, \phi) = \frac{GM_E}{a} \sum_{l=1}^{L_{\max}} \left(\frac{a}{r}\right)^{l+1} \sum_{m=-l}^l \Delta C_{lm} \bar{Y}_{lm}(\theta, \phi). \tag{2}$$

In this notation, spherical harmonic order m runs from $-l$ to l and \bar{Y}_{lm} are 4π -normalized surface spherical harmonics:

$$\bar{Y}_{lm}(\theta, \phi) = \bar{P}_{l,|m|}(\cos \theta) \begin{cases} \cos m\phi & (m \geq 0) \\ \sin(-m\phi) & (m < 0). \end{cases} \tag{3}$$

The temporal variations ΔC_{lm} of Stokes coefficients in Eq. (2) are related to the traditionally considered ones as

$$\Delta C_{lm} = \begin{cases} \Delta \bar{C}_{lm} & (m \geq 0) \\ \Delta \bar{S}_{lm} & (m < 0). \end{cases} \tag{4}$$

In the further derivations, we make use of the publication by Swenson and Wahr (2002) as a starting point. We also use the same notation, when possible.

The general expression that connects temporal variations of density $\Delta\rho(r, \theta, \phi)$ at/inside the Earth with temporal variations of Stokes coefficients is:

$$\Delta C_{lm} = \frac{a^2}{M_E} \frac{1}{(2l+1)} \iint_{\Omega} \Delta I_l(\theta, \phi) \bar{Y}_{lm}(\theta, \phi) d\Omega, \tag{5}$$

where integration covers the entire sphere:

$$\iint_{\Omega} d\Omega = \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta, \tag{6}$$

whereas $\Delta I_l(\theta, \phi)$ describes vertically integrated density variations:

$$\Delta I_l(\theta, \phi) = \int_0^{\text{top of atmos}} \left(\frac{r}{a}\right)^{l+2} \Delta\rho(r, \theta, \phi) dr. \tag{7}$$

Equation (5) is virtually equivalent to Eq. (2) in Swenson and Wahr (2002), with the exception that we watch the difference between the mean Earth’s radius and the equatorial one.

As it is explained in Sect. 1, we assume that mass transport takes place in a thin layer at the Earth surface. As such, it can be represented by a single mass layer, so that $\Delta I_l(\theta, \phi)$ can be approximated as

$$\begin{aligned} \Delta I_l(\theta, \phi) &= \int_{\text{thin layer}} \left(\frac{r_s(\theta, \phi)}{a}\right)^{l+2} \Delta\rho(r, \theta, \phi) dr \\ &\approx \left(1 + \frac{\xi + h}{a}\right)^{l+2} \Delta\sigma(\theta, \phi) \end{aligned} \tag{8}$$

[cf. Eq. (7) in Swenson and Wahr (2002)]. In Eq. (8), $\Delta\sigma(\theta, \phi)$ are variations of surface density (i.e., mass variations per unit area) and $r_s(\theta, \phi)$ is the function describing the shape of the Earth’s surface that can be represented with a high accuracy as

$$r_s(\theta, \phi) = a + \xi(\theta, \phi) + h(\theta, \phi) \tag{9}$$

with $\xi(\theta, \phi)$ being the height of the geoid above the sphere of radius a , whereas $h(\theta, \phi)$ is the orthometric height of the Earth’s surface topography [cf. Eq. (5) in Swenson and Wahr (2002)]. In the expressions below, the arguments in the functions $\xi(\theta, \phi)$ and $h(\theta, \phi)$ will not be explicitly written for the sake of brevity.

Variations of surface density can be related to the mass anomalies in terms of equivalent water heights (EWH) $\Delta H_w(\theta, \phi)$ in terms of equivalent water heights (EWH) with a simple formula

$$\Delta H_w(\theta, \phi) = \frac{\Delta\sigma(\theta, \phi)}{\rho_w}, \text{ where } \rho_w \text{ is water density.} \tag{10}$$

Equations (5) and (8) describe the link between the temporal variations of surface density and the temporal variations of Stokes coefficients under the assumption that the Earth is a rigid body. In practice, solid Earth reacts elastically to changes in the load on the Earth’s surface (see, e.g., Boy and Chao 2005). Hence, actual variations in the gravitational potential comprise both the direct effect of mass transport and the elastic deformation of the Earth (deformation potential). In order to take this into account, additional scaling factors $(1 + k_l)$ are introduced, where k_l are load Love numbers (Wahr et al. 1998), so that Eq. (8) turns into

$$\Delta I_l(\theta, \phi) \approx (1 + k_l) \left(1 + \frac{\xi + h}{a}\right)^{l+2} \Delta\sigma(\theta, \phi). \tag{11}$$

The function $\Delta\sigma(\theta, \phi)$, as any other function of coordinates (θ, ϕ) , can be represented in terms of the spherical harmonic expansion. After retaining the spherical harmonic degrees consistently with Eq. (1), we have:

$$\Delta\sigma(\theta, \phi) = a\rho_w \sum_{l'=1}^{L_{\max}} \sum_{m'=-l'}^{l'} \Delta\tilde{C}_{l'm'} \bar{Y}_{l'm'}(\theta, \phi), \tag{12}$$

where $\Delta\tilde{C}_{l'm'}$ are Fourier coefficients, which can be computed on the basis of the Stokes coefficients ΔC_{lm} . To that end, we insert Eqs. (12) and (11) into Eq. (5), which yields:

$$\begin{aligned} \Delta C_{lm} &= \frac{a^3 \rho_w}{M_E} \frac{(1+k_l)}{(2l+1)} \iint_{\Omega} \bar{Y}_{lm}(\theta, \phi) \left(1 + \frac{\xi+h}{a}\right)^{l+2} \\ &\times \sum_{l'=1}^{L_{\max}} \sum_{m'=-l'}^{l'} \Delta\tilde{C}_{l'm'} \bar{Y}_{l'm'}(\theta, \phi) \, d\Omega. \end{aligned} \tag{13}$$

By interchanging the order of the summation and integration, we readily obtain:

$$\Delta C_{lm} = \frac{4\pi a^3 \rho_w}{M_E} \frac{(1+k_l)}{(2l+1)} \sum_{l'=1}^{L_{\max}} \sum_{m'=-l'}^{l'} B_{l,m,l',m'} \Delta\tilde{C}_{l'm'} \tag{14}$$

with

$$\begin{aligned} B_{l,m,l',m'} &= \frac{1}{4\pi} \iint_{\Omega} \bar{Y}_{lm}(\theta, \phi) \bar{Y}_{l'm'}(\theta, \phi) \\ &\times \left(1 + \frac{\xi+h}{a}\right)^{l+2} \, d\Omega. \end{aligned} \tag{15}$$

Equation (14) represents a system of linear equations with constant coefficients $B_{l,m,l',m'}$, which form a square matrix. By solving this system, one can transform the Stokes coefficients into the coefficients $\Delta\tilde{C}_{l'm'}$. The system can be simplified further under some assumptions about the geometry of the Earth’s surface, as discussed below.

2.1.2 Spherical Earth approximation (radius = 6378 km)

Let us assume that the Earth is the sphere of radius a , i.e., $r_s(\theta, \phi) = a$, $\xi(\theta, \phi) = 0$, and $h(\theta, \phi) = 0$. Then, in view of the fact that the surface spherical harmonics form an orthogonal set on a sphere, the matrix formed by coefficients $B_{l,m,l',m'}$ turns into the unit one:

$$B_{l,m,l',m'} = \delta_{l,l'} \delta_{m,m'}, \tag{16}$$

where $\delta_{i,k}$ is the Kronecker delta. Furthermore, the Earth’s mass can be related to its mean density ρ_E :

$$M_E = \frac{4}{3} \pi a_E^3 \rho_E, \tag{17}$$

where a_E is the mean Earth’s radius. Then, one can insert Eqs. (17) and (16) into (14). If the difference between a_E^3 and a^3 is neglected (which is of the order of 0.3%), this readily results in:

$$\Delta C_{lm} = \frac{3\rho_w}{\rho_E} \frac{(1+k_l)}{(2l+1)} \Delta\tilde{C}_{lm}. \tag{18}$$

Thus, the computation of the coefficients $\Delta\tilde{C}_{lm}$ reduces to a scaling of the Stokes coefficients. A combination of Eqs. (18), (12), and (10) yields the well-known expression proposed for GRACE data processing by Wahr et al. (1998):

$$\Delta H_w(\theta, \phi) = \frac{a}{3} \frac{\rho_E}{\rho_w} \sum_{l=1}^{L_{\max}} \sum_{m=-l}^l \frac{(2l+1)}{(1+k_l)} \Delta C_{lm} \bar{Y}_{lm}(\theta, \phi). \tag{19}$$

2.1.3 Ellipsoidal Earth approximation

Let the Earth’s surface geometry be an ellipsoid of rotation. Then, $h(\theta, \phi) = 0$ and $\xi(\theta, \phi) = \zeta(\theta)$, where $\zeta(\theta)$ is the height of the ellipsoid above the sphere of radius a . The radial distance $r(\theta)$ of the points at the ellipsoidal surface is given by:

$$r(\theta) = a + \zeta(\theta) = a \frac{1-f}{\sqrt{1-e^2 \sin^2 \theta}}, \tag{20}$$

where f is the ellipsoid flattening (WGS84 value: $f = 1/298.2572$) and e is eccentricity ($e^2 = 2f - f^2$). Then, Eq. (15) turns into

$$\begin{aligned} B_{l,m,l',m'} &= \frac{1}{4\pi} \iint_{\Omega} \bar{Y}_{lm}(\theta, \phi) \bar{Y}_{l'm'}(\theta, \phi) \\ &\times \left(\frac{1-f}{\sqrt{1-e^2 \sin^2 \theta}}\right)^{l+2} \, d\Omega. \end{aligned} \tag{21}$$

Furthermore, the trigonometric functions $\sin m\phi$ and $\cos m\phi$ form an orthogonal set in the interval $[0; 2\pi]$:

$$\begin{cases} \int_0^{2\pi} \sin m\phi \cdot \cos m'\phi \, d\phi = 0 \\ \int_0^{2\pi} \cos m\phi \cdot \cos m'\phi \, d\phi = \int_0^{2\pi} \sin m\phi \cdot \sin m'\phi \, d\phi \\ = \pi(1 + \delta_{m,0}) \cdot \delta_{m,m'}. \end{cases} \tag{22}$$

In view of Eq. (6), this allows the expression for the elements $B_{l,m,l',m'}$ to be simplified to:

$$\begin{aligned}
 B_{l,m,l',m'} &= \frac{(1 + \delta_{m,0}) \cdot \delta_{m,m'}}{4} \\
 &\times \int_0^\pi \bar{P}_{l,|m|}(\cos \theta) \cdot \bar{P}_{l',|m'|}(\cos \theta) \sin \theta \\
 &\times \left(\frac{1-f}{\sqrt{1-e^2 \sin^2 \theta}} \right)^{l+2} d\theta \\
 &= \frac{(1 + \delta_{m,0}) \cdot \delta_{m,m'}}{4} \\
 &\times \int_{-1}^1 \bar{P}_{l,|m|}(x) \cdot \bar{P}_{l',|m'|}(x) \\
 &\times \left(\frac{1-f}{\sqrt{1-e^2(1-x^2)}} \right)^{l+2} dx. \tag{23}
 \end{aligned}$$

Thus, the system of linear equations given by Eq. (14) becomes block-diagonal and can be solved with ease.

2.1.4 Spherical Earth approximation (arbitrary radius)

Finally, one can also approximate the Earth’s surface with a sphere of an arbitrary radius: $r_s(\theta, \phi) = a + \Delta r$ (i.e., $\xi(\theta, \phi) = \Delta r$, $h(\theta, \phi) = 0$), with Δr being an arbitrary value, provided that $\Delta r > -a$. In that case, the expression for the elements $B_{l,m,l',m'}$ simplifies to:

$$B_{l,m,l',m'} = \delta_{l,l'} \delta_{m,m'} \left(1 + \frac{\Delta r}{a} \right)^{l+2}, \tag{24}$$

so that the link between the coefficients $\Delta \tilde{C}_{lm}$ and the Stokes coefficients is:

$$\Delta C_{lm} = \frac{4\pi a^3 \rho_w}{M_E} \frac{(1 + k_l)}{(2l + 1)} \left(1 + \frac{\Delta r}{a} \right)^{l+2} \Delta \tilde{C}_{lm}. \tag{25}$$

The parameter Δr can be chosen for a given target region such that the difference between the actual radial distance of the points at the Earth’s surface and the Earth’s equatorial radius a is taken into account.

One step further is a “locally spherical” approximation. With this, we mean that the parameter Δr can be chosen not as a single constant for the entire target region, but as a value dependent on the location of the current point where the mass anomaly is computed. Let us assume, for example, that mass transport takes place at the ellipsoid, i.e., $\Delta r = \Delta r(\theta) = \zeta(\theta)$. Then, a combination of Eqs. (25), (20), (12), and (10) yields the following expression for the mass anomalies in terms of EWH:

$$\begin{aligned}
 \Delta H_w(\theta, \phi) &= \frac{M_E}{4\pi a^2 \rho_w} \sum_{l=1}^{L_{\max}} \sum_{m=-l}^l \frac{(2l + 1)}{(1 + k_l)} \Delta C_{lm} \\
 &\times \left(\frac{\sqrt{1-e^2 \sin^2 \theta}}{1-f} \right)^{l+2} \bar{Y}_{lm}(\theta, \phi) \tag{26}
 \end{aligned}$$

or, in view of Eq. (17),

$$\begin{aligned}
 \Delta H_w(\theta, \phi) &= \frac{a}{3} \left(\frac{a_E}{a} \right)^3 \frac{\rho_E}{\rho_w} \sum_{l=1}^{L_{\max}} \sum_{m=-l}^l \frac{(2l + 1)}{(1 + k_l)} \Delta C_{lm} \\
 &\times \left(\frac{\sqrt{1-e^2 \sin^2 \theta}}{1-f} \right)^{l+2} \bar{Y}_{lm}(\theta, \phi). \tag{27}
 \end{aligned}$$

This expression resembles Eq. (19), but contains an additional degree-dependent scaling factor $\left(\frac{a_E}{a}\right)^3 \cdot \left(\frac{\sqrt{1-e^2 \sin^2 \theta}}{1-f}\right)^{l+2}$.

The most appropriate choice of the Earth’s geometry approximation is discussed below.

2.2 Selection of the most appropriate approximation of the Earth’s surface geometry

In the previous section, we presented a number of alternative expressions to convert temporal variations of Stokes coefficients into mass anomalies. The complexity of the associated computations depends on the adopted approximation of the Earth’s shape. In order to identify the most appropriate computational scheme, we consider the computation of mass anomalies from real GRACE data.

2.2.1 Data

In this study, we use monthly gravity field solutions in 2003–2015 produced at the Center for Space Research (University of Texas at Austin) (Bettadpur 2012). For a few months, the solutions are absent. Furthermore, we ignored the solutions that were not limited to a specific calendar month. As such, 135 monthly solutions in total are exploited. Each of these solutions is formed by a set of Stokes coefficients complete to degree 96. No filtering is applied.

It is expected that mass anomalies at relatively short spatial scales are particularly sensitive to the deviations of the assumed Earth’s surface geometry from the actual one. Therefore, it makes sense to focus on the scenarios where the high-frequency signal in GRACE data is strong, whereas the noise level is low. In view of this, we use the time series of Stokes coefficients to estimate the mean rate of linear mass change (i.e., linear trend) in the entire time interval 2003–2015. In this way, random noise in the coefficients is largely suppressed. The linear trend is co-estimated with a bias and

a quadratic term, as well as with annual and semiannual (co-)sinusoidal variations [see, e.g., Eq. (15) in Siemes et al. (2013), for more detail].

2.2.2 Computation of mass anomalies at the actual Earth's surface

As a reference, we convert the temporal variations of Stokes coefficients into mass anomalies using Eqs. (14) and (15) explicitly. To define the elevations $h(\theta, \phi)$ of the Earth's surface, we use the Global Land One-kilometer Base Elevation (GLOBE) digital elevation model (GLOBE Task Team et al. 1999). For inland areas, the model provides terrain elevations above the mean sea level. The oceans are flagged; we set the elevations there equal to zero. The geoid heights $\xi(\theta, \phi)$ above the sphere of radius a are approximated by Eq. (20) for the ellipsoid (thus, we ignore the differences between the reference ellipsoid and geoid as relatively minor, $\lesssim 100\text{m}$).

In general, the computed mass change rates are heavily contaminated by random noise, which is not surprising in view of the absence of filtering (not shown). Nevertheless, noise in polar areas is relatively low and allows one to clearly see strong signals over the territory of Greenland and the Amundsen Sea embayment of West Antarctica (Fig. 1). The strongest negative trends are observed at the Jakobshavn Isbræ (West Greenland) and at the Pine Island glacier (West Antarctica), see Fig. 1 and Table 1. A rapid ice mass loss at both locations is also revealed by other observation techniques (e.g., Groh et al. 2014; Mouginot et al. 2014).

We explain the low noise level in Fig. 1, in spite of the absence of any filtering, by a combination of several factors. First, using 13 years of data allows random noise to be largely averaged out, as it is already mentioned above. Second, a large density of GRACE ground tracks, as well as the intersection of ascending and descending tracks at relatively large angles, ensures a good coverage of polar areas, which reduces random noise further. Third, Greenland and West Antarctica are notorious for the presence of very strong negative trends due to a rapid ice mass loss. Ironically, these locations are far away from the equator, so that deviations of the actual Earth's surface from the sphere of radius $a \approx 6378\text{ km}$ are large there. Thus, the presented areas can be considered as the “worst-case scenarios” for Eq. (19), which is traditionally used to convert Stokes coefficients into mass anomalies. Therefore, our further analysis is limited to the two geographical areas shown in Fig. 1.

2.2.3 Computation of mass anomalies at the sphere of 6378-km radius

Conversion of Stokes coefficients into mass anomalies with the commonly used Eq. (19) reveals large differences

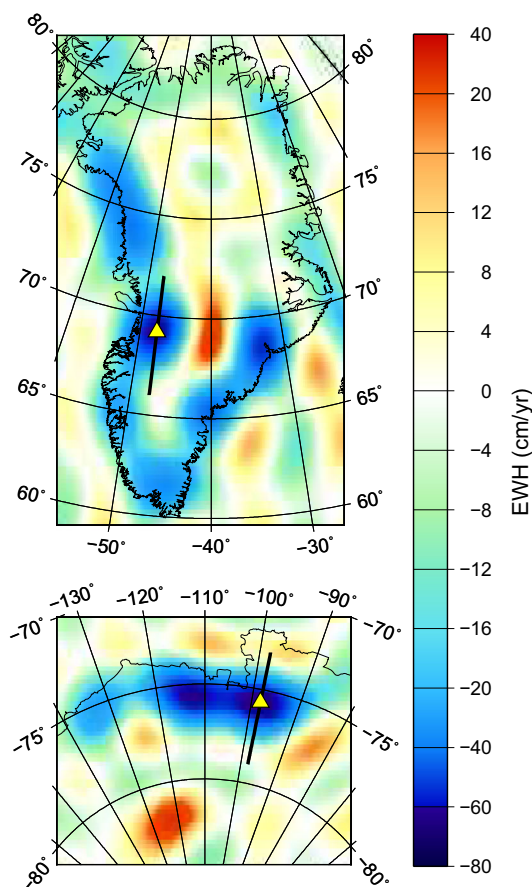


Fig. 1 Mean rate of linear mass change in Greenland (top) and Amundsen Sea embayment of West Antarctica (bottom) in 2003–2015, in terms of EWH. The estimates are based on GRACE Release-5 monthly solutions produced at CSR. The mass anomalies are computed at the actual Earth's surface, using GLOBE digital elevation model. Yellow triangles denote the locations of the peak signal: the Jakobshavn Isbræ in Greenland and the Pine Island glacier in West Antarctica. Thick black lines denote the profiles addressed in Fig. 5

from the values computed at the actual Earth's surface (see Fig. 2a). The observed differences are clearly anti-correlated with the total signal shown in Fig. 1. In other words, the traditionally used expression damps the recovered signals. This is not surprising, since the polar Earth's radius is 21 km smaller than the equatorial one, so that actual mass transport in the considered case takes place about 20 km further away from GRACE satellites than it is implicitly assumed in the traditional computations. The largest differences reach about 10 cm/year (Fig. 2a, Table 1).

In order to better quantify the observed signal damping, we also represent the difference $D(\theta, \phi)$ between the two variants of mass change rates in percentages, for which purpose the following expression is used:

$$D(\theta, \phi) = \frac{\left(\dot{H}_w^{(\text{appr})}(\theta, \phi) - \dot{H}_w^{(\text{ref})}(\theta, \phi) \right)}{\left| \dot{H}_w^{(\text{ref})}(\theta, \phi) \right|} \times 100, \quad (28)$$

Table 1 Mass change rates (in terms of EWH) computed at the actual Earth’s surface, as well as the errors introduced by different assumptions about its geometry. The considered points are the locations of the peak signal

	Jakobshavn Isbrae (Greenland)	Pine Island glacier (West Antarctica)
Peak signal location		
Longitude	47°W	98°W
Latitude	69°N	76°S
Signal	− 64.0 cm/year	− 68.6 cm/year
Error of the spherical approximation ($R = 6378$ km)	9.2 cm/year (14.4%)	9.7 cm/year (14.2%)
Error of the ellipsoidal approximation	− 1.0 cm/year (− 1.5%)	− 0.4 cm/year (− 0.6%)
Error of the (ellipsoid-based) locally spherical approximation	− 1.2 cm/year (− 1.8%)	− 0.6 cm/year (− 0.8%)
Difference between the locally spherical and ellipsoidal approximation	− 0.2 cm/year (− 0.3%)	− 0.2 cm/year (− 0.2%)

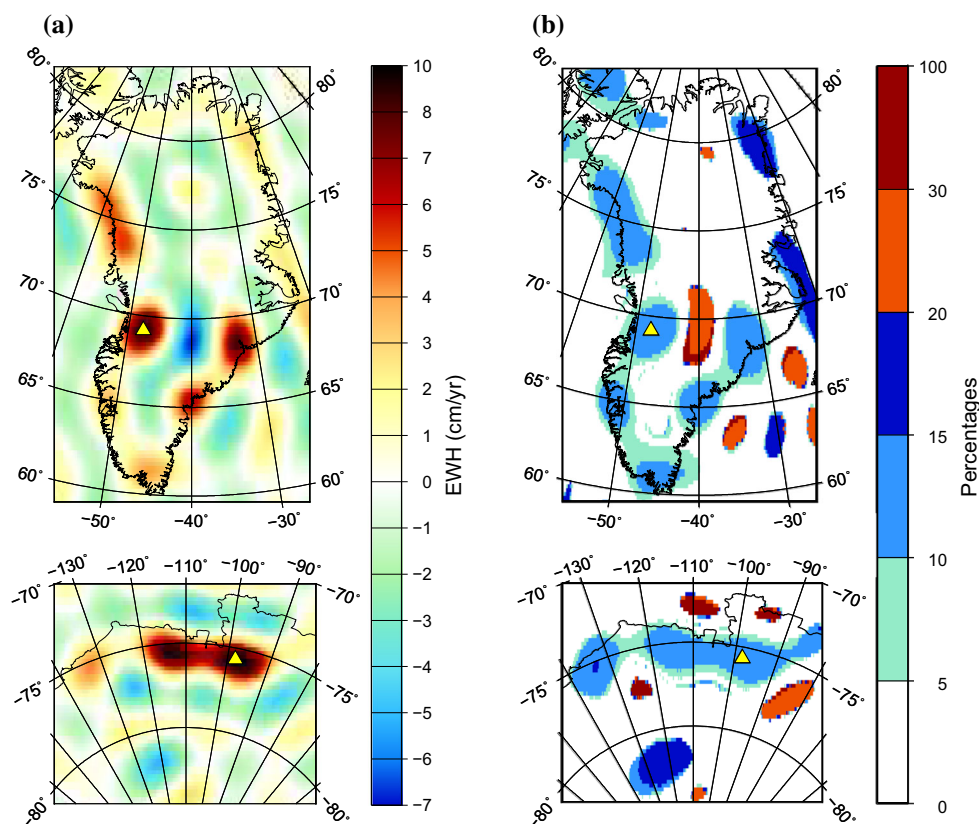


Fig. 2 Difference (left) and the absolute value of relative difference in percentages (right) between the mass change rates computed at the sphere of radius $a = 6,378,136$ m and those computed at the actual

Earth’s surface. The relative differences are shown only at the locations where the signal exceeds in absolute value 10 cm/year. For more details, see the caption of Fig. 1 and the main text

where $\dot{H}_w^{(appr)}$ is the mass change rate computed with the approximate formula and $\dot{H}_w^{(ref)}$ is the reference mass change rate. The results are presented in Fig. 2b. We only show the observed differences if the reference signal at a given point exceeds in absolute value 10 cm/year. In this way, we mask out the large relative differences that are caused by a small value in the denominator in Eq. (28). In spite of that, the observed relative differences are, in general, quite large. At the locations of strongest negative trends, they are of the order

of 15% (Table 1), whereas at some other locations they are even larger, reaching 20% and more (Fig. 2b). One may argue that the largest relative differences may be associated with noise (e.g., over the ocean and, perhaps, over the inner part of Greenland). Mass anomaly estimates there must be filtered anyway, so that the errors introduced by the spherical Earth approximation are not critical. Nevertheless, even if we limit the discussion only to the coastal areas, the signal damping

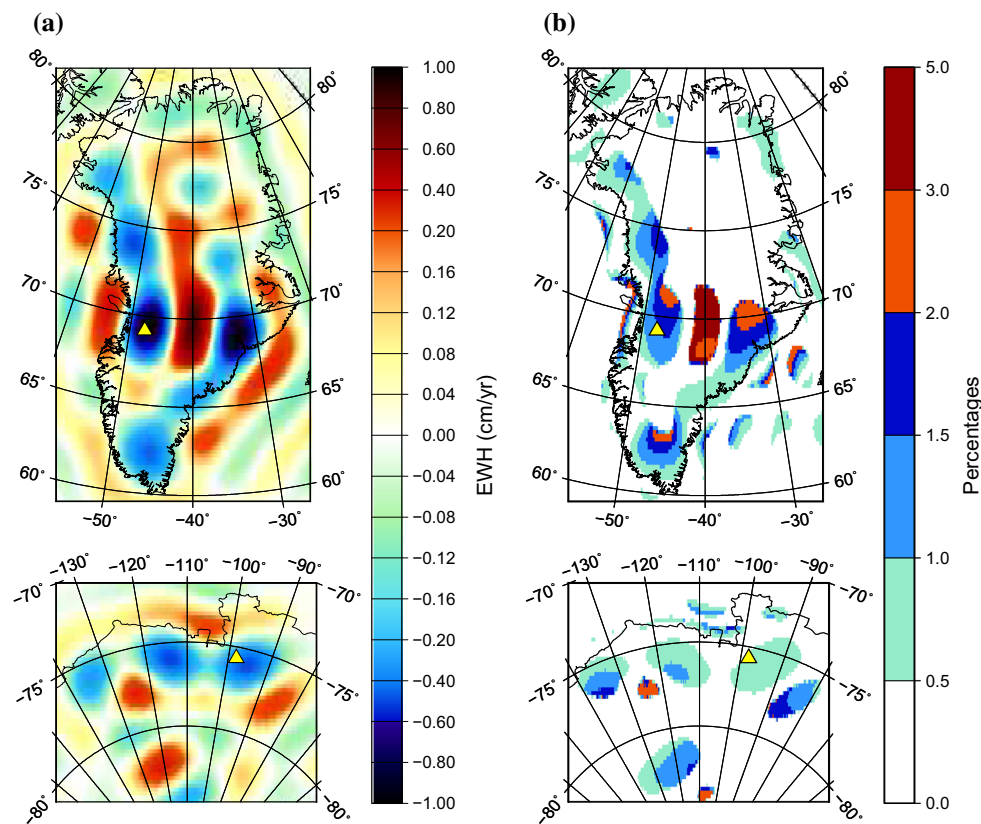


Fig. 3 Same as Fig. 2, but the compared mass change rates are computed at the reference ellipsoid and at the actual Earth's surface

caused by the considered approximation is at the level of up to 10–15%, which is hardly acceptable.

2.2.4 Computation of mass anomalies at the reference ellipsoid

Next, we compute mass anomalies under the assumption that mass transport takes place at the reference ellipsoid. To that end, we invert the block-diagonal matrix given by Eq. (23). The resulting mass change rates are an order of magnitude closer to those computed at the actual Earth's surface, as compared to those produced under the assumption that the Earth is a sphere of 6378 km radius. The differences do not exceed 1 cm/year (Fig. 3a, Table 1). The relative differences are within 1.5% at the locations of the peak signal and typically stay within the 3% limit elsewhere (Table 1, Fig. 3b). Thus, the approximation of the Earth surface geometry with the reference ellipsoid may improve the conversion accuracy by an order of magnitude, as compared to the traditional approximation with the sphere of radius $a = 6378$ km.

In addition, it is worth noticing that the spatial pattern of the observed differences shows a clear positive correlation with the signal itself (cf. Figs. 1 and 3a). In other words, the recovered signal is sharper than the actual one. Obviously, this is due to the fact that the surface of the reference ellipsoid

(which is close to the sea level) is a few km further away from the GRACE satellites than the actual surface of ice sheets. Finally, relatively large differences are observed in the ocean areas at the western and southeastern coasts of Greenland. We relate them to a strong gradient of mass anomalies at the Greenland coasts. In view of a limited spectrum of the function $\Delta H_w(\theta, \phi)$, ringing artefacts associated with the Gibbs phenomenon must occur in those areas. As soon as recovered signal becomes sharper, these artefacts become more pronounced.

2.2.5 Computation of mass anomalies at the reference ellipsoid using the locally spherical approximation

Finally, we compute mass anomalies at the reference ellipsoid using the locally spherical approximation. That is, the spherical Earth's surface expression is used in the computations, but the radius of the sphere is latitude-dependent. At each latitude, it is set equal to the distance between the reference ellipsoid and the center of the Earth, cf. Eq. (26). The obtained linear trend estimates are surprisingly close to those produced with the explicit procedure addressed in the previous section. The differences between the results do not exceed 0.4 cm/year; see Fig. 4a. The relative differences stay at the level of at most 1–2% (Fig. 4b). Thus, the locally

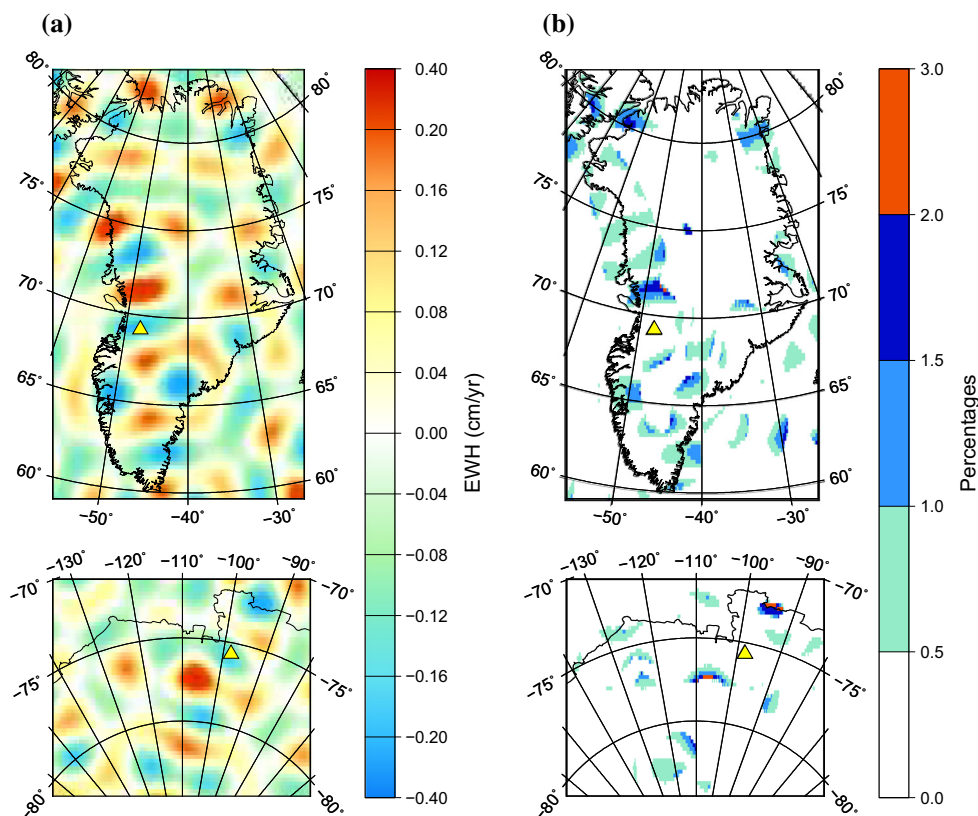


Fig. 4 Same as Fig. 2, but the compared mass change rates are computed at the reference ellipsoid: using a locally spherical approximation and explicitly

spherical approximation of the reference ellipsoid ensures an almost the same conversion accuracy as the usage of the reference ellipsoid explicitly. On the other hand, the conversion based on the locally spherical approximation is easier to implement in practice, since it does not require solving any systems of linear equations.

3 Ellipsoidal versus spherical coordinates

In this section, we discuss the difference between mass anomalies computed in spherical (geocentric) and ellipsoidal (geodetic) coordinates. The maximum difference between geodetic and geocentric (co-)latitudes is observed near the 45° latitudes, reaching approximately 11.5 arc-minutes or 21 km. In the polar areas, which are considered in our examples, this difference is smaller: of the order of 10–15 km or even less. One may argue that such differences must be negligible in the estimation of mass anomalies from GRACE data, since the spatial resolution of those estimates is a few hundreds of km. To demonstrate that such a statement may be unfair, we consider the meridional profiles that cross the Jakobshavn Isbræ in Greenland and the Pine Island glacier in West Antarctic (Fig. 5). The linear trend estimates along

these profiles are presented as functions of both geodetic and geocentric latitude (Fig. 5). This figure clearly shows that the accuracy of locating a sharp signal may by far exceed the spatial resolution of GRACE data. As such, the difference between the estimates in spherical and ellipsoidal coordinates is clearly visible. Thus, we believe that the conversion of geodetic colatitudes into geocentric ones must not be omitted in GRACE data processing.

4 Discussion and conclusions

To convert time-varying Stokes coefficients into mass anomalies at the Earth's surface, geoscientists routinely use Eq. (19) or its equivalents. However, the results obtained with this expression represent mass transport at the spherical surface of 6378 km radius. In this study, we show that the accuracy of such a conversion may be insufficient, especially if the target area is located in a polar region and the signal-to-noise ratio is high. For instance, the mean linear trends in 2003–2015 estimated over Greenland and Amundsen Sea embayment of West Antarctica may be underestimated in this way by 10–15% or even more. Such an error may definitely exceed the noise level of current mass transport estimates. Moreover,

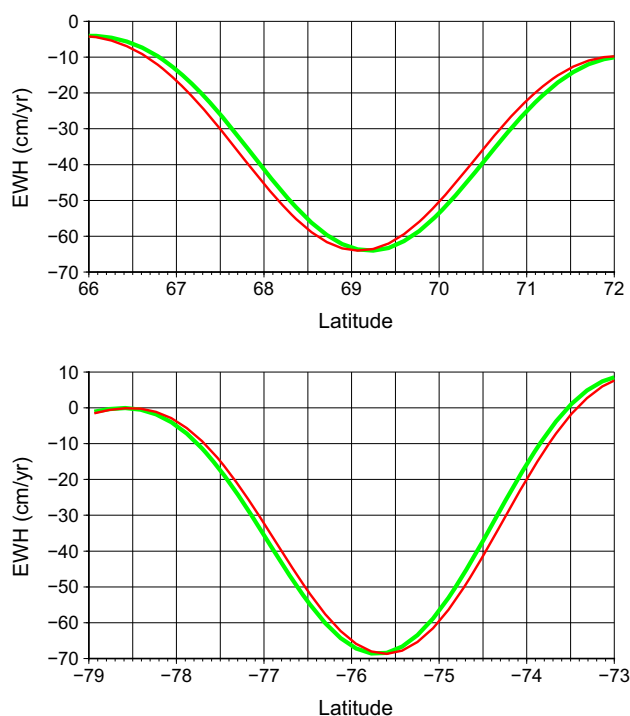


Fig. 5 Mass change rates computed along the meridional profiles crossing the Jakobshavn Isbræ in Greenland (top) and the Pine Island glacier in West Antarctica (bottom); the exact location of the profiles is shown in Fig. 1. The computed values are presented as functions of geocentric latitude (in red) and geodetic latitude (in green). In both cases, it is assumed that mass transport takes place at the actual Earth's surface, which is described by the GLOBE digital elevation model

there are no doubts that the accuracy and spatial resolution of future estimates will increase further, so that a limited accuracy of the conversion based on Eq. (19) will likely become in the future even less tolerable. There are several reasons to expect improvements in the mass anomaly estimation in the foreseeable future: (i) an ongoing progress in the techniques for satellite gravity data processing; (ii) a continuously increasing duration of mass anomaly time series, which leads to a high accuracy of mean estimates for the entire available time interval (including mean linear trends); and (iii) the forthcoming launch of the GRACE Follow-On (GFO) satellite mission (<https://gracefo.jpl.nasa.gov>). GFO satellites will be equipped with a laser interferometer, allowing for an order-of-magnitude increase in the accuracy of inter-satellite ranging. Though the increase in the ranging accuracy may not result in a proportional increase in the accuracy of estimated mass anomalies in general (Flechtner et al. 2016), the level-1 data will definitely become cleaner at high frequencies, since ranging noise at those frequencies is dominant (Flury et al. 2008; Ditmar et al. 2012). Then, this reduction in noise level will likely have a positive effect onto the estimates of mass anomalies at small spatial scales, which are particularly vulnerable if the Earth's geometry is defined inaccurately.

As a solution, we propose an updated expression for the conversion of Stokes coefficients into mass anomalies. This expression is based on the assumptions that: (i) mass transport takes place at the Earth's surface that is approximated by the reference ellipsoid; (ii) at each point of interest, the Earth's surface is further approximated by the sphere with a radius equal to the current radial distance from the Earth's center ("locally spherical approximation"). The updated expression is nearly as simple as Eq. (19), but allows the inaccuracies associated with the conversion procedure to be reduced by an order of magnitude.

In addition, we demonstrate that it is advisable to convert geodetic (co-)latitudes into geocentric ones, when mass anomalies are computed. This is in spite of the fact that the shifts caused by this conversion are an order of magnitude smaller than the spatial resolution of GRACE-based estimates.

In "Appendix A," we summarize the recommended expressions for the conversion of Stokes coefficients into mass anomalies. Unlike in the main text, we use there the most traditional notation for the spherical harmonic expansion (when the orders start from 0 and not from $-l$), in order to facilitate the usage of the proposed expressions.

We would like to stress that the proposed conversion formula is particularly beneficial in the presence of strong signals in the range of high degrees, as it was the case in the considered examples. The absence of such signals makes the results much less sensitive to the assumption about the surface where mass transport takes place. For instance, the truncation of mass anomaly estimates at spherical harmonic degree 60 reduces the conversion errors 2–3 times, as compared to those presented in Fig. 2 (where the maximum degree is equal to 96).

In all the discussions so far, we assumed that the loading effects can be corrected just by introducing load Love numbers as additional scaling factors of the kind $(1 + k_l)$. For a non-spherical Earth, such a simple approach is, strictly speaking, incorrect. However, the loading effects manifest themselves mostly in the range of low degrees. This can be understood from the fact that the scaling factors $(1 + k_l)$ rapidly approach 1 as spherical harmonic degree increases. For instance, $(1 + k_l) > 0.96$ for degrees above 30 and $(1 + k_l) > 0.98$ for degrees above 70 (Wahr et al. 1998). Since the impact of the proposed conversion formula is mostly limited to the range of relatively high degrees, we believe that a simplified treatment of the loading effects is justified.

Finally, one may pose the questions whether mass anomalies can be uniquely restored considering that mass transport takes place at the Earth's surface of a complicated geometry. For the case of a spherical Earth, a unique recovery of mass anomalies is guaranteed by Eq. (18). This formula establishes a unique link between the Stokes coefficients (that describe variations of gravitational potential) and the Fourier coef-

ficients of mass anomalies. Moreover, this formula offers a practical way to make the corresponding conversion. By increasing the maximum spherical harmonic degree under consideration, one may, in principle, recover mass anomalies with an arbitrarily high spatial resolution. Thus, the spatial resolution of the results is fully defined by the spatial resolution of the exploited gravity field model. The situation with mass transport at the actual Earth's surface is more complicated. First of all, the unique conversion of time-varying Stokes coefficients into mass anomalies can only be guaranteed if the matrix composed of coefficients $B_{l,m,l',m'}$ (cf. Eq. 15) is invertible. In all the computations conducted in this study, this was indeed the case: this matrix was not only invertible, but also close to the unit one. However, it remains unclear if (or under what conditions) this matrix remains invertible in general. Furthermore, the presence of nonzero off-diagonal elements in this matrix implies that there is no unique link anymore between the spatial resolution of gravity field model and that of mass anomalies. For instance, high-frequency signals in terms of mass anomalies can map onto low-frequency signals in gravity field observations. If the former signals contain spherical degrees above L_{max} in a given data processing run, a proper conversion of Stokes coefficients into mass anomalies becomes impossible. In other words, a realistic Earth's geometry may result in a new type of high-frequency signal aliasing, which is absent when mass transport takes place at a spherical surface. A quantification of this effect and, if necessary, designing optimal schemes to mitigate it are the subjects of further studies.

Acknowledgements The author thanks the Center for Space Research at the University of Texas at Austin for providing the GRACE Level-2 data product. The author also thanks the GLOBE task team for providing the Global Land One-kilometer Base Elevation (GLOBE) digital elevation model. Roland Klees is thanked for valuable comments that allowed the author to improve the manuscript substantially. Finally, the author thanks two anonymous reviewers for their numerous suggestions, which resulted in a further improvement of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Recommended expressions to convert time-varying Stokes coefficients into mass anomalies at the Earth's surface

1. Identify the time interval of interest. Subtract the mean value of each Stokes coefficient in this time interval from the coefficient time series to produce temporal variations of Stokes coefficients ($\Delta C_{lm}, \Delta S_{lm}$).

2. For each point of interest, convert the geodetic colatitude θ_g into the geocentric colatitude θ_c :

$$\theta_c = \text{ATAN2}(\sin \theta_g, (1 - f)^2 \cos \theta_g), \tag{29}$$

where f is the flattening of the Earth's figure (1/298.2572) and

$$\text{ATAN2}(y, x) = \begin{cases} \arctan(y/x) & \text{if } x > 0 \\ \arctan(y/x) + \pi & \text{if } x < 0 \text{ and } y \geq 0 \\ \arctan(y/x) - \pi & \text{if } x < 0 \text{ and } y < 0 \\ \pi/2 & \text{if } x = 0 \text{ and } y > 0 \\ -\pi/2 & \text{if } x = 0 \text{ and } y < 0 \\ \text{undefined} & \text{if } x = 0 \text{ and } y = 0. \end{cases} \tag{30}$$

3. For each point of interest, compute the mass anomaly ΔH_w in terms of equivalent water heights using the following expression:

$$\begin{aligned} \Delta H_w(\theta_c, \phi) = & \frac{M_E}{4\pi a^2 \rho_w} \sum_{l=1}^{L_{max}} \frac{(2l+1)}{(1+k_l)} \\ & \times \left(\frac{\sqrt{1-e^2 \sin^2 \theta_c}}{1-f} \right)^{l+2} \\ & \times \left(\sum_{m=0}^l \Delta C_{lm} \bar{P}_{lm}(\cos \theta_c) \cos m\phi \right. \\ & \left. + \sum_{m=1}^l \Delta S_{lm} \bar{P}_{lm}(\cos \theta_c) \sin m\phi \right), \end{aligned} \tag{31}$$

where ϕ is the longitude of the current point; M_E is the Earth's mass (5.9722×10^{24} kg); a is the Earth's equatorial radius (6,378,136 m); ρ_w is water density (1000 kg/m³); L_{max} is the maximum spherical harmonic degree in the exploited GRACE data product; k_l are load Love numbers (Wahr et al. 1998); $e^2 = 2f - f^2$, and $\bar{P}_{lm}(x)$ are normalized associated Legendre functions of argument x .

References

Bettadpur SV (2012) Gravity Recovery and Climate Experiment, UTCSR level-2 processing standards document for level-2 product release 0005. GRACE 327-742 (CSR-GR-12-xx). Center for Space Research, University of Texas at Austin

Boy JP, Chao B (2005) Precise evaluation of atmospheric loading effects on earth's time-variable gravity field. *J Geophys Res B Solid Earth* 110(8):1–10. <https://doi.org/10.1029/2002JB002333>

Chao B (2005) On inversion for mass distribution from global (time-variable) gravity field. *J Geodyn* 39(3):223–230. <https://doi.org/10.1016/j.jog.2004.11.001>

- Chao B (2016) Caveats on the equivalent water thickness and surface mascon solutions derived from the GRACE satellite-observed time-variable gravity. *J Geodesy* 90(9):807–813. <https://doi.org/10.1007/s00190-016-0912-y>
- Chao B, O'Connor W, Chang A, Hall D, Foster J (1987) Snow load effect on the earth's rotation and gravitational field, 1979–1985. *J Geophys Res* 92(B9):9415–9422
- Ditmar P, da Encarnação JT, Farahani HH (2012) Understanding data noise in gravity field recovery on the basis of inter-satellite ranging measurements acquired by the satellite gravimetry mission GRACE. *J Geodesy* 86(6):441–465. <https://doi.org/10.1007/s00190-011-0531-6>
- Dobslaw H, Bergmann-Wolf I, Dill R, Poropat L, Flechtner F (2016) Product description document for AOD1B product Reslease 06. GFZ German Research Centre for Geosciences, Potsdam
- Flechtner F (2007) AOD1B product description document for product reslease 01 to 04. GeoForschungsZentrum, Potsdam
- Flechtner F, Neumayer KH, Dahle C, Dobslaw H, Fagiolini E, Raimondo JC, Gntner A (2016) What can be expected from the GRACE-FO laser ranging interferometer for Earth science applications? *Surv. Geophys* 37(2):453–470. <https://doi.org/10.1007/s10712-015-9338-y>
- Flury J, Bettadpur S, Tapley BD (2008) Precise accelerometry onboard the grace gravity field satellite mission. *Adv Space Res* 42:1414–1423
- Forootan E, Didova O, Kusche J, Lcher A (2013) Comparisons of atmospheric data and reduction methods for the analysis of satellite gravimetry observations. *J Geophys Res Solid Earth* 118(5):2382–2396. <https://doi.org/10.1002/jgrb.50160>
- GLOBE Task Team, Hastings DA, Dunbar PK, Elphinstone GM, Bootz M, Murakami H, Maruyama H, Masaharu H, Holland P, Payne J, Bryant NA, Logan TL, Muller JP, Schreier G, MacDonald JS (eds) (1999) The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80305-3328, U.S.A., Digital data base on the World Wide Web (URL: <http://www.ngdc.noaa.gov/mgg/topo/globe.html>) and CD-ROMs
- Groh A, Ewert H, Fritsche M, Rülke A, Rosenau R, Scheinert M, Dietrich R (2014) Assessing the current evolution of the Greenland Ice Sheet by means of satellite and ground-based observations. *Surv Geophys* 35(6):1459–1480. <https://doi.org/10.1007/s10712-014-9287-x>
- Heiskanen WA, Moritz H (1967) Physical geodesy. W.H. Freeman, San Francisco
- Mouginot J, Rignot E, Scheuchl B (2014) Sustained increase in ice discharge from the Amundsen Sea Embayment, West Antarctica, from 1973 to 2013. *Geophys. Res. Lett* 41(5):1576–1584. <https://doi.org/10.1002/2013GL059069>
- Siemes C, Ditmar P, Riva REM, Slobbe DC, Liu XL, Farahani HH (2013) Estimation of mass change trends in the Earth's system on the basis of GRACE satellite data, with application to Greenland. *J Geodesy* 87:69–87. <https://doi.org/10.1007/s00190-012-0580-5>
- Swenson S, Wahr J (2002) Estimated effects of the vertical structure of atmospheric mass on the time-variable geoid. *J Geophys Res B Solid Earth* 107(9):2194. <https://doi.org/10.1029/2000JB000024>
- Wahr J, Molenaar M, Bryan F (1998) Time variability of the Earth's gravity field: hydrological and oceanic effects and their possible detection using GRACE. *J Geophys Res* 103(B12):30,205–30,229
- Wouters B, Bonin J, Chambers D, Riva R, Sasgen I, Wahr J (2014) GRACE, time-varying gravity, Earth system dynamics and climate change. *Rep Prog Phys* 77(11):116801. <https://doi.org/10.1088/0034-4885/77/11/116801>