Abrykosov, P., Sulzbach, R., Pail, R., Dobslaw, H., Thomas, M. (2022): Treatment of ocean tide background model errors in the context of GRACE/ GRACE-FO data processing. - Geophysical Journal International, 228, 3, 1850-1865.

https://doi.org/10.1093/gji/ggab421

Geophysical Journal International

Geophys. J. Int. (2022) **228,** 1850–1865
Advance Access publication 2021 October 14
Management and Production

https://doi.org/10.1093/gji/ggab421

Treatment of ocean tide background model errors in the context of GRACE/GRACE-FO data processing

Petro Abrykosov,[1] Roman Sulzbach,[2,3] Roland Pail,[1] Henryk Dobslaw[2] and Maik Thomas[2,3]

[1]*Chair of Astronomical and Physical Geodesy, Technical University of Munich (TUM), Arcisstraße 21, 80333 München, Germany. E-mail: petro.abrykosov@tum.de*
[2]*Helmholtz Centre Potsdam, GFZ German Research Centre for Geosciences, Section 1.3, Earth System Modelling, Telegrafenberg, 14473 Potsdam, Germany*
[3]*Institute of Meteorology, Freie Universität Berlin (FUB), Carl-Heinrich-Becker-Weg 6-10, 12165 Berlin, Germany*

Accepted 2021 October 12. Received 2021 October 11; in original form 2021 July 12

SUMMARY

Ocean tide (OT) background models (BMs) used for *a priori* de-aliasing of GRACE/GRACE-FO observations feature distinct spatial uncertainties (primarily in coastal proximity and in latitudes above ±60°), and therefore pose one of the largest contributors to the overall retrieval error. The retrieval performance can be expected to increase if this underlying spatial error distribution is stochastically modelled and incorporated into the data processing chain. In this contribution, we derive realistic error variance-covariance matrices (VCM) based on a set of five state-of-the-art OT models. The additional value of using such VCMs is assessed through numerical closed-loop simulations, where they are rigorously propagated from model to observation level. Further, different approximations of the resulting VCM of observations are assumed, that is full, block-diagonal and diagonal, in order to evaluate the trade-off between computational efficiency and accuracy. It is asserted that correctly weighting the OT BM error can improve the gravity retrieval performance by up to three orders of magnitude, provided no further error contributors are considered. In comparison, the overall gain in retrieval performance is reduced to 75 per cent once instrument noise is taken into account. Here, it is shown that simultaneously modelling the OT BM and the instrument errors is critical, as each effect induces different types of correlations between observations, and exclusively considering covariance information based on the sensor noise may degrade the solution. We further demonstrate that the additional benefit of incorporating OT error VCMs is primarily limited by the de-aliasing performance for non-tidal mass variations of atmosphere (A) and oceans (O). This emphasizes the necessity of best-possible AO-de-aliasing (e.g. through optimized processing techniques and/or improved BMs) in order to optimally exploit the OT BM weighting.

Key words: Satellite gravity; Tides and planetary waves; Time variable gravity.

1. INTRODUCTION

The GRACE twin-satellite mission (Tapley *et al*. 2004), in orbit from 2002 until 2017, has provided the scientific community with unique insights into variations within the Earth's climate. In May 2018, the GRACE Follow-On mission (GRACE-FO; Kornfeld *et al*. 2019) with near-identical payload was launched to extend the climate studies initiated by GRACE. In order to facilitate the retrieval of gravitational variations related to for example hydrology or ice from GRACE/GRACE-FO data, high-frequency high-amplitude signals must first be eliminated in the data processing. This step is imperative because such signal components, typically attributed to tidal mass redistributions as well as short-term non-tidal variations within the atmosphere, cannot be resolved by a single satellite or satellite pair and would therefore induce temporal aliasing. Thus, *a priori* de-aliasing is carried out by means of geophysical background models (BMs). The BM-based de-aliasing, however, inevitably features some level of uncertainty due to shortcomings in the underlying models. For example, since ocean tide (OT) models are based on altimetric measurements, they generally feature high uncertainties in shallow waters (Bosch *et al*. 2009) as well as in latitudes beyond ±60° (e.g. King & Padman 2005 for Antarctica). In case of non-tidal atmospheric and oceanic (AO) mass variations, largest model errors are typically closely related to short-term variations in air and water pressure due to meteorological events (e.g. Bergmann & Dobslaw 2012). Because of these shortcomings, that

is because of the discrepancies between the 'true' observations carried out by the satellite and the 'reference' ones generated on the basis of BMs, the retrieved gravity field will still notably suffer from temporal aliasing. As of now, BM errors constitute one the largest limitations to gravity retrieval performance (Flechtner *et al*. 2015).

In gravity field processing, BMs are usually assumed to be error-free. However, it is reasonable to incorporate uncertainty information into the processing chain (if available), and rigorously propagate it through the adjustment process onto the final temporal gravity solutions in order to increase the de-aliasing performance. Non-tidal AO signals are composed of a multitude of different frequencies, requiring a statistical approach to describe the time behaviour, or the availability of a time-variable variance-covariance matrix (VCM) to appropriately describe the underlying time- and location-dependent uncertainties. Since correlations between each (epoch-wise) set of underlying spherical harmonic coefficients must be considered as well in order to describe the error behaviour as realistically as possible, the task becomes computationally quite demanding. A method to incorporate BM uncertainties for non-tidal atmosphere and ocean signals into the gravity field recovery process was proposed and assessed by Kvas & Mayer-Gürr (2019). They show that either augmenting the covariance matrix of observations, or co-estimating model corrections as stochastic parameters with known prior covariance, is equivalent and leads to identical results. The related background error covariance matrices were derived from the ESA Earth System Model (Dobslaw *et al*. 2015) in the form of an autoregressive model.

In contrast, in case of OTs the time behaviour is defined by specific excitation frequencies and is therefore deterministic. Consequently, the underlying uncertainty information is time-invariant and depends only on the geographic location. It can therefore be described by a single static VCM of the underlying spherical harmonic coefficients. The temporal correlations are then described by the functional model which includes the corresponding tidal excitation frequencies.

Instead of an accurate stochastic modelling in the frame of the gravity retrieval process, other methods to reduce OT model errors have been proposed. Visser (2010) investigated a number of (post-)processing methods, including the temporal filtering of time-series of gravity field solutions, the spatial smoothing of these time-series and evaluation of solutions for different geographical regions for GRACE-type single-pair and Bender-type (Bender *et al*. 2008) double pair missions. An a posteriori method for de-aliasing of OT errors in future double-pair missions was proposed by Liu *et al*. (2016), also giving hints about optimal choices for orbit constellations to mitigate OT aliasing effects. Liu & Sneeuw (2021) describe OT aliasing as a two-step mechanism, with orbit sampling as the first step and gravity recovery as the second step. Another promising approach is the direct estimation of selected OT constituents. However, with single-pair mission this requires time spans of several years, as the major tidal constituents have known alias frequencies with periods between 9 d (for Q1 and N2) and 2725 d (for K1) within the GRACE data (Ray & Luthcke 2006). The extended length of the GRACE data time-series of more than 15 yr, which is now further continued with GRACE-Follow On, provides the possibility to estimate corrections to the amplitudes and phases of the major solar and lunar OT constituents from dedicated OT models (Han *et al*. 2005; Killett *et al*. 2011; Mayer-Gürr *et al*. 2012; Wiese *et al*. 2016), predominantly for latitudes higher than 66°. In Hauk & Pail (2018) the increased feasibility of co-estimating OT parameters from Bender double-pair missions compared to single-pair missions could be demonstrated.

In this paper, we disregard the treatment of AO model errors, and instead focus on the stochastic modelling of OT errors for single-pair missions without explicit co-estimation of OT parameters. The model uncertainty information is introduced in terms of error VCMs, and their additional value for GRACE/GARCE-FO-based gravity products is assessed through numerical closed-loop simulations. The manuscript is structured as follows. In Section 2, the simulation environment is presented alongside the methodology for deriving realistic OT BM error VCMs and propagating them through the entire simulation chain. In Section 3, we present and analyse the retrieval performance of simulation scenarios which incorporate the OT and AO de-aliasing error as well as sensor noise. In Section 4, the key findings are summarized and an outlook to future work is given.

2. DATA AND METHODOLOGY

2.1 Set-up of OT error covariance matrices

Five state-of-the-art OT atlases are used for the derivation of realistic spatial error information: FES2014 (Lyard *et al*. 2021), EOT11a (Savchenko & Bosch 2012), TPX09 (Egbert & Erofeeva 2002, updated), GOT4.10c (Ray 2013) and DTU10 (Cheng & Andersen 2011). For a summarized overview of these models and their respective characteristics we refer the reader to Stammer *et al*. (2014). Each of the models used in this study encompasses at least the four principal diurnal (Q1, O1, P1, K1) and the four semi-diurnal tidal constituents (N2, M2, S2, K2), each represented in terms of global in-phase grids A_ν (tidal elevations at a certain phase $\phi = 0$) and quadrature grids B_ν (tidal elevations at a certain phase $\phi = \pi/2$). The tidal height at a position $P(\theta, \lambda)$ can then be expressed as

$$\zeta(\theta,\lambda) = \sum_\nu f_\nu [A_\nu(\theta,\lambda)\cos(\chi_\nu + u_\nu) + B_\nu(\theta,\lambda)\sin(\chi_\nu + u_\nu)], \quad (1)$$

where f_ν and u_ν are nodal modulation terms (for each constituent ν) stemming from the precession of the lunar node, and χ_ν is the astronomical argument. The corresponding amplitude of a tidal constituent's signal in P can be computed as

$$C_\nu(\theta,\lambda) = f_\nu \sqrt{A_\nu^2(\theta,\lambda) + B_\nu^2(\theta,\lambda)}. \quad (2)$$

As the considered tidal atlases are provided in different formats and with different resolutions (1°/2 to 1°/16), a format unification procedure has to be defined. Regridding transformations must be carried out with caution in order to not engineer additional model differences that would lead to an unrealistic representation of uncertainty in the VCMs

As the tidal amplitude typically increases towards the coasts and then abruptly drops to zero on dry land grid cells, the representation of coastlines is an important issue. As coastlines in non-polar latitudes are well-established at the considered model resolution, they should not be a source of uncertainty of the stochastic model. We unify coastlines in low latitudes ($|\theta| < 66°$) by extending the tidal information onto dry grid cells in terms of nearest-neighbour extrapolation, performing a first order conservative interpolation to a resolution of 1°/8 and afterwards imposing the DTU10 land–sea mask onto all five models. This eliminates inconsistencies that arise from differently represented coastlines at unequal resolutions. The atlases are then conservatively regridded to a Gaussian Grid (resolution 364 × 182), which is used for further processing.

While the differing representations of low-latitude coastlines are mainly a question of the initially provided model resolution, model differences in polar latitudes, especially around Antarctica, are significantly more complicated to treat. The models' individual land–sea masks of the Antarctic ice shelf regions differ notably (cf. Fig. 1), provoking the existence of grid cells holding only $N < 5$ data points. Because missing data (treated as '0') has the same effect as dry land when transforming the tidal solution to Stokes' coefficients, this would falsify the estimation of local mean values and covariances and induce unrealistic error information. To rectify this bias from the quasi-incompleteness of the employed data set, different strategies can be applied (e.g. Schafer 1997). Here, we treat each grid cell for which at least one model provides data as a wet ocean cell and impute values for the remaining models. As tidal solutions are smoothly varying functions in space, missing values are estimated by bilinear extrapolation from the neighbouring cells of the same model. This ultimately results in an overall unified land–sea mask that preserves the information initially provided by the models while enabling the realistic representation of uncertainties. Tidal atlases that do not include ice shelf areas, for example OSU (Fok 2012) or HAMTIDE (Taguchi *et al.* 2014), were not used, since the construction of missing values by large-scale extrapolation would lead to unrealistic values. Fig. 2 shows the mean of the homogenized in-phase and quadrature grids as well as their grid-point-wise standard deviations, which demonstrate the underlying spatial error pattern exemplary for the M2 constituent.

Following the homogenization, we derive the OT parameters, that is the Stokes coefficients which describe the in-phase and quadrature grids of each partial tide given in each of the five tidal atlases according to

$$\begin{Bmatrix} A_v(\theta,\lambda) \\ B_v(\theta,\lambda) \end{Bmatrix} = \frac{R}{3}\frac{\rho_e}{\rho_w}\frac{2n+1}{1+k_n'}\sum_n\sum_m \left(\begin{Bmatrix} a_{nm,v} \\ c_{nm,v} \end{Bmatrix} \cos\, m\lambda + \begin{Bmatrix} b_{nm,v} \\ d_{nm,v} \end{Bmatrix} \sin\, m\lambda \right) \bar{P}_{nm}(\cos\theta) \tag{3}$$

up to d/o 180 by means of numerical quadrature. The advantage of this formulation is the linearity of the observation equation regarding the OT parameters, as $\{a, b, c, d\}_v$ can be transformed to Stokes' coefficients representing the Earth's gravitational potential at each epoch according to

$$\begin{Bmatrix} \Delta\bar{C}_{nm}^{\mathrm{OT}} \\ \Delta\bar{S}_{nm}^{\mathrm{OT}} \end{Bmatrix} = \sum_v f_v \left[\begin{Bmatrix} a_{nm,v} \\ b_{nm,v} \end{Bmatrix} \cos(\chi_v + u_v) + \begin{Bmatrix} c_{nm,v} \\ d_{nm,v} \end{Bmatrix} \sin(\chi_v + u_v) \right]. \tag{4}$$

We then construct $[p, p]$-dimensional error VCMs for tidal constituents' $\{a, b\}$ and $\{c, d\}$, respectively, up to d/o 30 (that is $p = 961$) using their previously derived $N = 5$ realizations. Because the VCM's size scales with a potency of 4 with respect to the maximum harmonic degree, d/o 30 was chosen as a reasonable compromise between computational demand and the precision of approximating the full-resolution spatial errors. However, we acknowledge that increasing the maximum d/o may be beneficial and could be considered for application to real GRACE data. Note that since $N < p$, the empirical covariance matrices

$$\mathrm{cov}(X_v, Y_v) = \frac{1}{N-1}\sum_{k=1}^{N}\left(X_{v,k} - \bar{X}_v\right)\left(Y_{v,k} - \bar{Y}_v\right) \tag{5}$$

with

$$X = \begin{Bmatrix} a, b \\ c, d \end{Bmatrix}_{n_1 m_1}, \quad n_1 = 0, \ldots 30, \quad m_1 = 0, \ldots n_1$$

$$Y = \begin{Bmatrix} a, b \\ c, d \end{Bmatrix}_{n_2 m_2}, \quad n_2 = 0, \ldots 30, \quad m_2 = 0, \ldots n_2$$

$$\begin{Bmatrix} \bar{X}_v \\ \bar{Y}_v \end{Bmatrix} = \frac{1}{N}\sum_{k=1}^{N}\begin{Bmatrix} X_k \\ Y_k \end{Bmatrix}$$

are not positive-semidefinite (e.g. Ledoit & Wolf 2003; Rajaratnam *et al.* 2008; Won *et al.* 2012; Hannart & Naveau 2014) and can consequently not be inverted, which contradicts their usage in the following steps. To resolve this problem, we employ the Graphical Lasso (GLASSO) approach (Friedman *et al.* 2008). Here, the precision matrix Θ (inverse of the VCM) is sconced with a $\|\|_1$-penalty (sum of the absolute values of all matrix elements) that is controlled by a parameter λ while maximizing the log-likelihood with respect to the underlying ensemble. λ increases the sparsity of the precision matrix, thus reducing the number of connections in the underlying Graph, where only the most vital connections remain with an increasing λ. The GLASSO method is thus founded on solid mathematical principles and has been successfully applied to related problems including weather and climate data processing (e.g. Röpnack *et al.* 2011; Keune *et al.* 2014; Zerenner *et al.* 2014). Note that λ can exert a critical influence on the VCM's properties. As we wish to minimize the effect of regularization on the VCM itself, λ is chosen as small as 0.1, which results in a VCM that only shows minor deviations to the initially mentioned empirical VCM (eq. 5). The calculations were performed employing the GLASSO implementation of Laska & Narayan (2018) for the aforementioned eight major partial tides.

2.2. Implementation in simulation environment

For this study, IAPG's closed-loop reduced-scale simulation software is used, which employs various simplifications compared to the real GRACE/GRACE-FO processing scheme in favour for improved computational performance. For a detailed description of the simulation tool the reader is referred to Murböck (2015) and Murböck *et al.* (2014), while here only the central simplification aspect is noted, which is that the observations are given in terms of line-of-sight (LOS) or range-acceleration differences instead of ranges and/or range-rates:

$$\Delta\,\mathrm{acc}_{\mathrm{LOS}} = \langle \Delta a_{\mathrm{grav}}, \Delta r_{\mathrm{LOS}} \rangle + \langle \Delta a_{\text{non-grav}}, \Delta r_{\mathrm{LOS}} \rangle. \tag{6}$$

Here, Δa_{grav} denotes the differences in gravitational acceleration observed by the satellites in along-track, cross-track and radial directions and Δr_{LOS} denotes the LOS vector pointing from one satellite's centre of mass to the other's. $\Delta a_{non-grav}$ describes non-gravitation accelerations which in this study are assumed to arise from the instrument noise. This approach ultimately turns the gravity field adjustment into a linear problem, thereby greatly simplifying the processing. The components of $\Delta\mathrm{acc}_{\mathrm{LOS}}$ stemming from the OT signal can be computed from the used OT model's underlying set of a, b, c, d coefficients as

$$\Delta\,\mathrm{acc}_{\mathrm{LOS}} = F \cdot E \cdot D \cdot \begin{bmatrix} x_{A_1} \\ x_{B_1} \\ \ldots \\ x_{A_8} \\ x_{B_8} \end{bmatrix}, \tag{7}$$

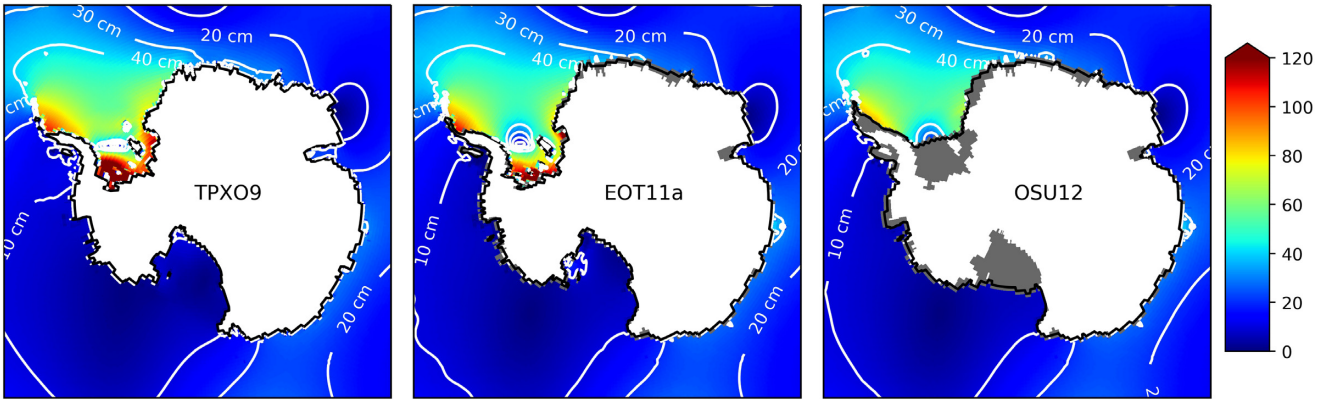

Figure 1. M2 tidal amplitude (cm) around Antarctica for TPXO9 (left), EOT11a (middle) and OSU12 (right) tidal atlas at a resolution of 1°/8. Differences in the land–sea mask with respect to TPXO9 are shown in grey. Models that do not include ice shelf areas (e.g. OSU) were not considered for covariance processing.

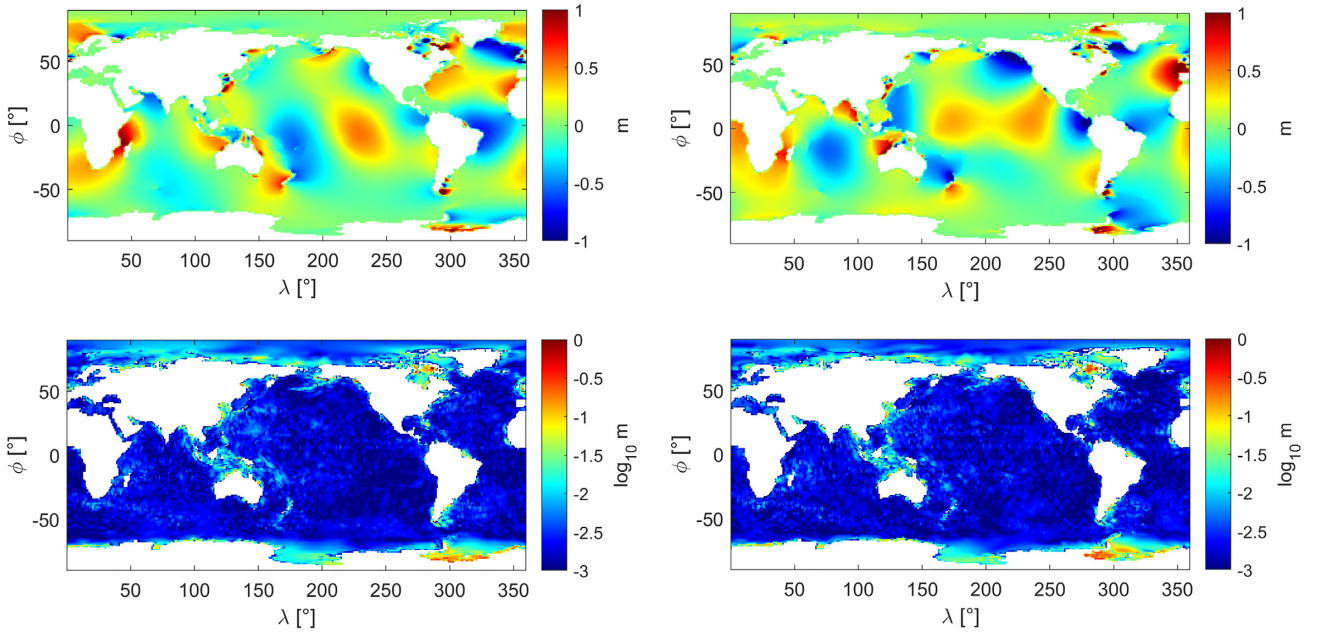

Figure 2. Mean values (top row) and standard deviations (bottom row) for in-phase (left) and quadrature (right) sea level elevation of the main lunar M2-tide. Please note the different scales and increased uncertainty in high latitudes and coastal regions.

where x_{A_ν}, x_{B_ν}, $\nu = 1, \ldots 8$ are the OT parameters of eight main tidal constituents, D are the functionals (i.e. the design matrix) denoting the transition from these OT parameters to $\Delta\bar{C}_{nm}^{\mathrm{OT}}$, $\Delta\bar{S}_{nm}^{\mathrm{OT}}$ according to eq. (4), E are the functionals describing the transition from $\Delta\bar{C}_{nm}^{\mathrm{OT}}$, $\Delta\bar{S}_{nm}^{\mathrm{OT}}$ to the satellites' gravitational acceleration measurements in radial, along-track and cross-track directions, and F represents the functionals for the LOS projection of the observed accelerations and their consecutive subtraction at each epoch. According to propagation of uncertainty, the full VCM of observations then reads

$$\Sigma_{\Delta\mathrm{acc}_{\mathrm{LOS}}} = (FED)\,\Sigma_{xx}(FED)^T \tag{8}$$

with

$$\Sigma_{xx} = \begin{bmatrix} \Sigma_{x_{A_1}x_{A_1}} & & & 0 \\ & \Sigma_{x_{B_1}x_{B_1}} & & \\ & & \cdots & \\ & & & \Sigma_{x_{A_8}x_{A_8}} \\ 0 & & & \Sigma_{x_{B_8}x_{B_8}} \end{bmatrix}. \tag{9}$$

Here, $\Sigma_{x_{A_\nu}x_{A_\nu}}$ and $\Sigma_{x_{B_\nu}x_{B_\nu}}$ denote the error VCMs of the underlying OT parameters (cf. Section 2.1). Note that $\Sigma_{\Delta\mathrm{acc}_{\mathrm{LOS}}}$ encompasses both spatial and temporal correlations between observations, the former having been introduced through Σ_{xx} and the latter through E. The mean gravity field can then be derived by means of a regular least-squares adjustment:

Table 1. Orbit parameters.

	Altitude (km)	Inclination (°)	Rev/nodal day	Ascending node (°)	Mean anomaly difference (°)
GRACE-A	475.14	89	457/30	0	–
GRACE-B					1.674

$$\hat{x} = \left(G^T \Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}^{-1} G\right)^{-1} G^T \Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}^{-1} \Delta \mathrm{acc}_{\mathrm{LOS}}. \quad (10)$$

Note that $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$ may become singular due to extremely high correlations between observations of certain epochs. This would restrict its use in eq. 10, so a Tikhonov regularization is applied according to

$$\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}} = \Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}} + \alpha I \quad (11)$$

where the regularization factor α is chosen as

$$\alpha = 10^{-7} \cdot \min\left(\mathrm{diag}\left(\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}\right)\right) \quad (12)$$

This choice is made empirically in such a way that the regularization's impact is minimal while also yielding a regular $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$.

The simulation results are evaluated in terms of residuals, that is the retrieval error, which allows one to assess the additional benefit of the BM error weighting. Since the processing scheme at hand is linear, the OT de-aliasing error can be simulated by directly basing $\Delta \mathrm{acc}_{\mathrm{LOS}}$ on $\Delta\{a, b, c, d\}$, that is the coefficient differences of two models, one of which is assumed to represent the true observable, while the other serves as basis for the reference observations.

3. SIMULATION PARAMETERS AND RESULTS

3.1. Simulation environment and parameters

The additional benefit of OT BM weighting is evaluated within a GRACE-type mission scenario which is assumed to be flying on a repeat orbit with a repeat cycle of 30 d (cf. Table 1). Mean gravity solutions are derived for a retrieval period of 5 and 30 d. In a first set of simulations, we assume the imperfect OT de-aliasing, which we base on the difference of TPX09 and DTU10, to be the sole contributor to the total error budget. Later, we subsequently add the instrument noise as well as non-tidal variations of the gravity field into the simulation environment. Stochastic modelling is applied for the OT BM error and the sensor noise. For the underlying OT VCMs, we distinguish the following weighting scenarios:

(i) *W5*:

Σ_{xx} is based on all five OT models (cf. Section 2.1). This scenario can be regarded as realistic, because while the model ensemble features highly similar large-scale error patterns, small-scale differences originating from, for example the choice and the processing of the underlying data may be unique to a pair of models. Such an error VCM therefore features a small degree of imperfection with regard to the two models used in the simulation.

(i) *W2*:

Σ_{xx} is based on exclusively TPX09 and DTU10. This approach can be regarded as best-case, since the error information represented by the VCM is now perfectly tailored to the OT data used in the simulation.

(i) *WI*:

Σ_{xx} is taken as a unit matrix. From a spatial point of view this implies a homogenously (i.e. randomly) distributed error pattern. Propagating such matrices onto the level of observations in accordance with eq. (8) will still yield a full $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$, but it will now exclusively encompass temporal correlations stemming from the OT signal's periodicities.

The retrieval performance yielded by these weighting approaches is compared against a reference scenario denoted as *W0*, where no weighting is applied at all.

Since covariance information is limited to d/o 30, we maintain consistency in the processing by limiting gravity-related signal components to d/o 30 as well as executing the retrieval up to d/o 30 in accordance with eq. (10). Note that while employing a full $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$ is expected to yield the best result, its size increases quadratically with the number of observation epochs, and the computation of its inverse may quickly become unfeasible. In order to evaluate the trade-off between computational performance and efficiency, arc-wise processing can be applied (Mayer-Gürr 2006). In this approach, normal equations are set up separately for consecutive blocks (arcs) of observations and then stacked. Correlations are in this case solely preserved for observations within each individual arc, which corresponds to a block-diagonal (arc-wise) $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$. Here, we consider a 1-d arc approximation for the 5-d retrieval as well as 1- and 5-d arc approximations for the 30-d retrieval period. Further, the impact of using a diagonal $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$, which essentially implies single-epoch arcs, is studied.

3.2. OT de-aliasing as the only error contributor

First, the OT BM weighting performance is evaluated separately for each partial tide, that is the simulated tidal signal is assumed to stem from just one partial tide, in an otherwise error-free simulation environment. The simulation results are presented in Fig. 3 in terms of degree RMS of the retrieval error. Degree RMS of the dimensionless Stokes' coefficients are computed as

$$c_{n,\mathrm{RMS}} = \sqrt{\sum_{m=0}^{n} \Delta \bar{C}_{nm}^2 + \Delta \bar{S}_{nm}^2} \quad (13)$$

where $\Delta \bar{C}_{nm}$ and $\Delta \bar{S}_{nm}$ are the Stokes' coefficients describing the retrieval error. This expression can further be represented in terms of equivalent water heights (EWH) according to

$$c_{n,\mathrm{RMS\ EWH}} = \frac{R}{3} \frac{\rho_{av}}{\rho_w} \frac{2n+1}{1+k_n'} c_{n,\mathrm{RMS}} \ [m] \quad (14)$$

where R is the Earth radius, ρ_w and ρ_{av} are the average water density and the average density of the Earth, respectively, and k_n' are the degree-dependent 2nd load love numbers.

When working with a diagonal $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$ based on *WI*, next to no change in retrieval performance can be seen in comparison to *W0*. On the other hand, if the diagonal $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$ is based on *W5* or *W2*, some minor improvement can be established in case of the 5-d solutions in the diurnal tides as well as N2 and M2, while the solutions for S2 and K2 seem to remain widely unaffected by the weighting. A similar behaviour can be seen in the 30-d solutions

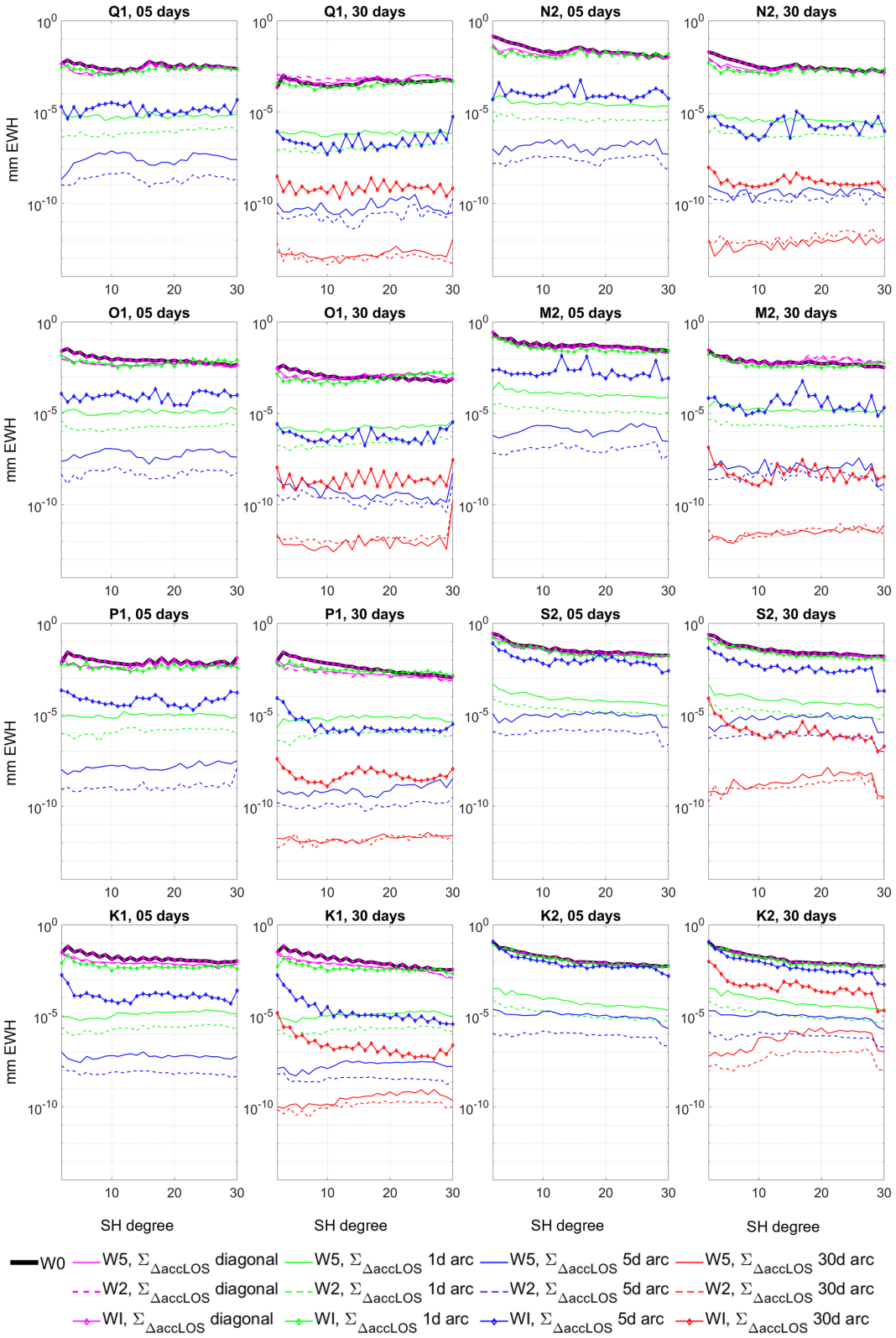

Figure 3. Retrieval error degree RMS of simulation scenarios based on the error signal of single tidal constituents for a retrieval period of 5 (first and third columns) and 30 d (second and fourth columns).

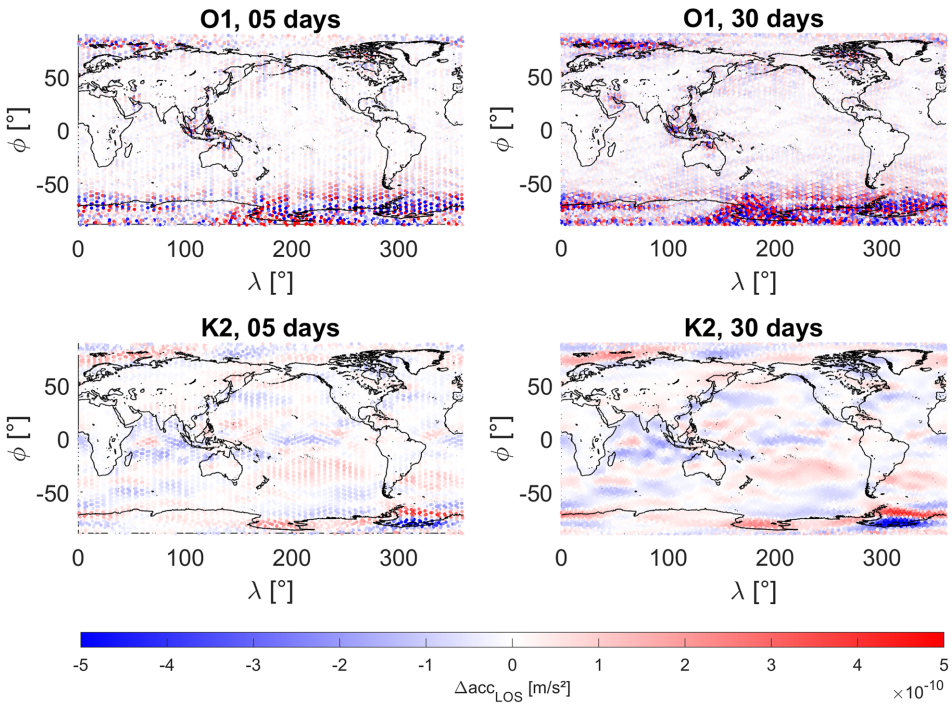

Figure 4. LOS acceleration differences from observation of the O1 and K2 error signal.

with the exception of Q1 and M2, where a slight degradation of the retrieval quality occurs.

The retrieval error decreases drastically once covariances among observations are considered. If daily arc-wise weighting based on *W5* is used, the improvement constitutes up to four orders of magnitude (depending on the tidal constituent) in terms of degree RMS, and an additional factor of 5 in case of *W2*. For *WI*, however, the gain is far smaller, as the retrieval performance is only marginally better than that of the solution obtained with a diagonal $\Sigma_{\Delta acc_{LOS}}$ based on *W5* and *W2*.

The solution quality further improves with increasing arc length. In general, the best results can be achieved with a full $\Sigma_{\Delta acc_{LOS}}$. In this case, up to five orders of magnitude in terms of degree RMS (O1) can be gained even with *WI*. However, although the gain in retrieval performance is immense in either case, it is notable that the overall retrieval performance of individual tidal constituents strongly varies. Up to nine orders of magnitude can be gained by applying *W5*- or *W2*-based weighting for the Q1 and M2 tides (30 d, full $\Sigma_{\Delta acc_{LOS}}$), while in the same manner only around four orders of magnitude can be gained for K2. It is also notable that the *W0*-based 5- and 30-d solutions for the tides with the comparatively smallest BM weighting gain barely change at all.

This behaviour can be explained with the underlying aliasing periods (e.g. Liu 2019), the effect of which is exemplary demonstrated in Fig. 4. for O1 and K2. In this example, the observations based on O1, which has a comparatively short aliasing period, are rather heterogeneous at any given geolocation, that is the signal is observed at various phases in a certain region. This can be seen especially well in the Antarctic region, where the observation error values feature a near-white-noise behaviour. This behaviour is further amplified with an increasing observation period, and since these heterogeneous observations are adjusted to zero in the parameter estimation, a 30-d retrieval yields a more favourable solution in comparison to the 5-d one even in case of *W0*. The solution is adjusted towards zero even further due to the highly stochastic nature of the observation error when BM weighting is introduced. The aliasing period of K2 is, on the other hand, significantly longer, and results in the observation of near-identical phases of the tidal signal at any given location throughout the retrieval period. Hence, we observe a nearly static, that is non-stochastic, signal. The contribution of BM weighting is therefore greatly reduced. In order to optimally compensate for this effect, significantly longer retrieval periods, preferably exceeding the alias period, would be required.

In the second set of simulations the full tidal signal consisting of all eight principal tides is employed and stochastically modelled. Analogously to the previously presented single-tide simulations, the results are shown in Fig. 5 in terms of degree RMS of the retrieval error. The respective retrieval performances are also summarized in Table 2 in terms of (latitude-weighted) RMS of global EWH grids. Once again, *W5* and *W2* allow for a significant increase in retrieval performance if non-diagonal VCMs of observations are used. The overall gain is now, however, decreased by around two orders of magnitude. In case of a diagonal $\Sigma_{\Delta acc_{LOS}}$ based on *W5* and *W2* minor gains can be asserted with respect to *W0* (up to 20 per cent primarily in the spectrum between degrees $n = 5$ and 15). With *WI*, on the other hand, no additional benefit can be achieved with a diagonal $\Sigma_{\Delta acc_{LOS}}$, as has already been shown in the single-tide simulation scenarios. Nevertheless, *WI* still allows for an improved retrieval performance once (at least) an arc-wise VCM of observations is implemented. In comparison to *W5* and *W2*, however, longer arcs are required in order to establish a notable gain in retrieval performance. In case of the 5-d solution, a full *WI*-based $\Sigma_{\Delta acc_{LOS}}$ results in a retrieval error decrease of up to 28 per cent in terms of degree RMS in the degrees below n = 20, while a full $\Sigma_{\Delta acc_{LOS}}$ in case of a 30-d solution decreases the retrieval error by over 60 per cent throughout the entire spectrum.

While the results presented in this section vividly demonstrate the value for OT BM weighting, they nevertheless somewhat fail to prove the original claim of OT BM weighting being able to reduce temporal aliasing effects, that is the GRACE-typical striping pattern.

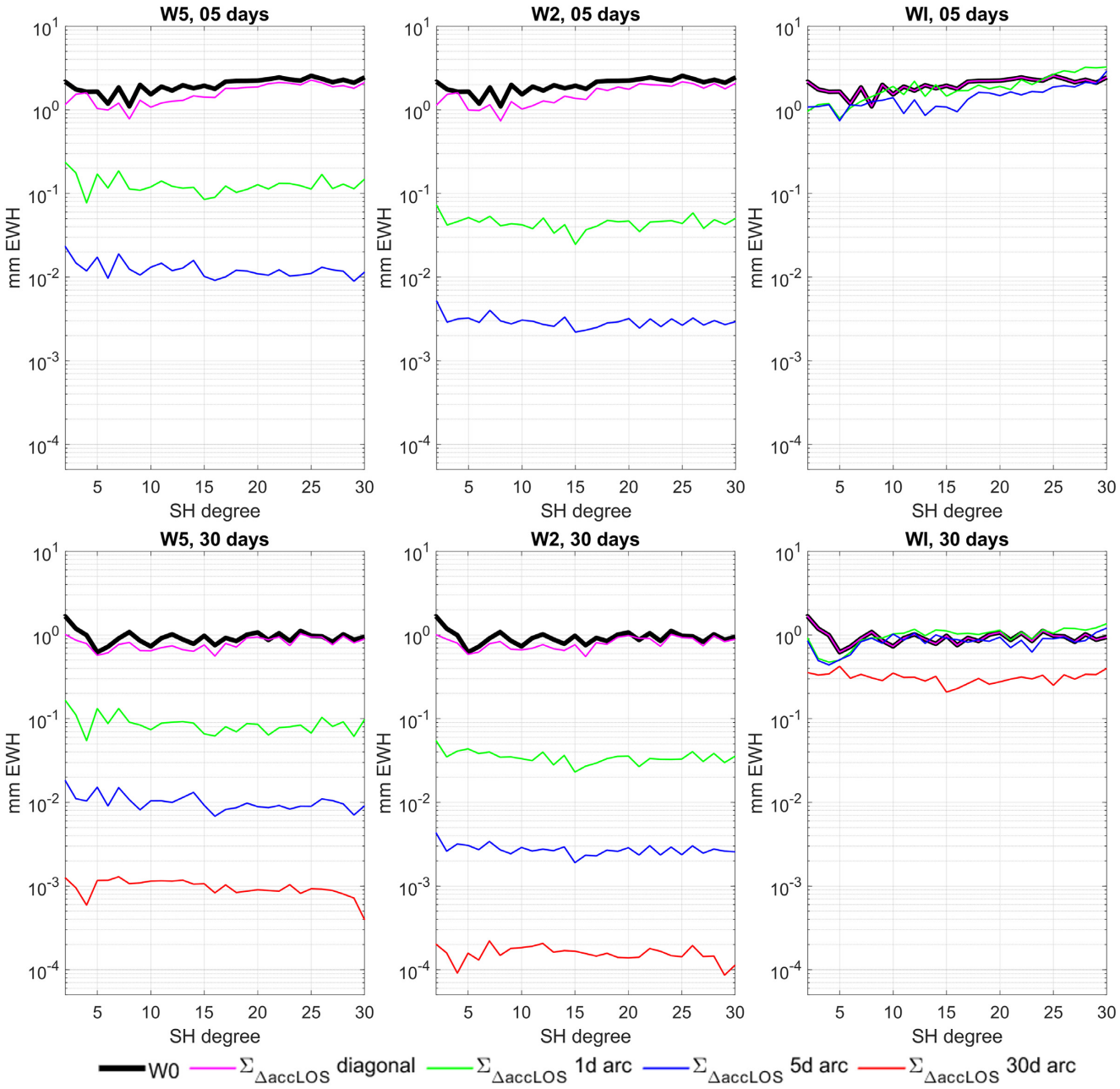

Figure 5. Retrieval error degree RMS of simulation scenarios based on the combined error signal of eight principal tides for a retrieval period of 5 (top row) and 30 d (bottom row) and W5- (left column), W2- (middle column) and WI-based (right column) Σ_{xx}.

This can be attributed to the simple fact that the spatial resolution of d/o 30 is simply too low to produce the typical high-frequency stripes. Therefore, we additionally conduct simulations based on the full OT signal up to d/o 60 as an explicit demonstration for this matter. Two weighting scenarios are considered here. In the first approach, in the following denoted as *W5_true60*, we derive full error covariance matrices up to d/o 60 from all five available OT models applying the *W5*-approach explained in Section 3.1. In the second approach which we denote as *W5_exp60*, we retain *W5*-based full covariance information up to d/o 30 and extend it with variances for d/o 31–60. These variances are set to an empirically chosen constant value of $10^3 \cdot \min(\mathrm{diag}(\Sigma_{\mathrm{xx,\ W5}}))$ for the individual in-phase and quadrature fields of each partial tide, respectively. This choice can be justified with the fact that the underlying spatial error patterns are of higher-frequency nature (high values in coastal regions, abrupt decrease to zero on continents), and thus, the error level in the higher-frequency spectrum can be assumed—without prior knowledge—to be quite substantial.

The results are presented in terms of degree RMS of the retrieval error in Fig. 6 (cf. also Table 2). It is evident that in case of *W5_true60*-based weighting the results perform close to identically to the previously shown *W5*-based scenarios up to d/o 30. More notable, however, is the fact that the reference solution (*W0*) now severely suffers from temporal aliasing (significantly increased degree RMS values from $n = 40$ upwards), and that applying *W5_true60*-based weighting mostly rectifies this behaviour (the degree RMS curve remains rather flat). The spatial error distribution, previously dominated by the aliasing-induced striping pattern, thus

Table 2. Retrieval performance in terms of (latitude-weighted) RMS of global EWH grids synthesized up to SH degree 30 (and 60, if available) for various scenarios presented in this paper. The values in the table are given in [mm EWH].

			$\Sigma_{\Delta acc_{LOS}}$ diag		$\Sigma_{\Delta acc_{LOS}}$ 1d arc		$\Sigma_{\Delta acc_{LOS}}$ 5d arc		$\Sigma_{\Delta acc_{LOS}}$ 30d arc	
Signal	Retrieval period (d)	Weighting strategy	$n = 30$	$n = 60$	$n = 30$	$n = 60$	$n = 30$	$n = 60$	$n = 30$	$n = 60$
Full OT (max. d/o 30)	5	W0 (ref.)	6.61		6.61		6.61			
		W5	5.47		0.53		0.05			
		W2	5.49		0.17		0.01			
		WI	6.61		6.33		5.07			
	30	W0 (ref.)	3.13		3.13		3.13		3.13	
		W5	2.65		0.36		0.04		0.004	
		W2	2.73		0.13		0.01		0.001	
		WI	3.13		2.81		2.55		1.11	
Full OT (max. d/o 60)	30	W0 (ref.)	3.14	44.30	3.14	44.30	3.14	44.30	3.14	44.30
		W5_exp60	2.91	40.42	1.10	6.41	0.61	3.88	0.14	2.04
		W5_true60	3.18	32.12	0.75	1.57	0.18	0.34	0.01	0.02
Full OT (max. d/o 30), NS1	30	W0 (ref.)			18.62		18.62		18.62	
		W5 + instr.			13.90		9.30		9.40	
		instr. only			19.88		11.83		10.52	
Full OT (max. d/o 30), NS2	30	W0 (ref.)			4.82		4.82		4.82	
		W5 + instr.			2.70		2.09		1.93	
		instr. only			12.36		7.63		5.09	
Full OT, HIS (max. d/o 30), NS1	30	W0 (ref.)			18.86		18.86		18.86	
		W5 + instr.			14.00		9.41		9.60	
Full OT, 10% AO, HIS (max. d/o 30), NS1	30	W0 (ref.)			19.11		19.11		19.11	
		W5 + instr.			14.40		9.73		9.99	
Full OT, AOHIS (max. d/o 30), NS1	30	W0 (ref.)			32.95		32.95		32.95	
		W5 + instr.			30.73		27.99		28.36	
Full OT, HIS (max. d/o 30), NS2	30	W0 (ref.)			5.46		5.46		5.46	
		W5 + instr.			3.50		3.06		2.99	
Full OT, 10% AO, HIS (max. d/o 30), NS2	30	W0 (ref.)			6.11		6.11		6.11	
		W5 + instr.			4.37		3.98		4.01	
Full OT, AOHIS (max. d/o 30), NS2	30	W0 (ref.)			27.19		27.19		27.19	
		W5 + instr.			26.73		26.62		26.73	

becomes highly isotropic, as can be clearly seen in the global EWH grids presented in Fig. 7. Overall, it can be stated that the performance of OT BM weighting does not depend on the maximum considered degree as long as it is done correctly and consistently. This further underlines the overall validity of the simulations considering gravity signal up to d/o 30.

It is further remarkable that notable improvements can also be achieved when applying *W5_exp60*-based weighting. While the major gains are unsurprisingly located in the frequency bands below $n = 30$ (nearly two orders of magnitude when using a full $\Sigma_{\Delta acc_{LOS}}$), the upper part of the spectrum is also retrievable with higher accuracy. In this case, processing observations based on daily arcs yields an improvement of more than one order of magnitude at $n = 60$, while employing a full $\Sigma_{\Delta acc_{LOS}}$ results in an improvement of two orders of magnitude in the same frequency band. Moreover, in the latter case the error level from $n = 31$ upwards remains rather flat, indicating a rather homogenous spatial error. This behaviour can also be verified with the corresponding global EWH grid in Fig. 7.

The findings regarding the performance of *W5_exp60*-based weighting can be considered of great significance for application of OT BM error modelling in real data processing. Because there the OT signal is (theoretically) present up to d/o infinity, a cut-off degree up until which full error covariance information is provided must be chosen in order to retain a reasonable computation time. Also, especially with regard to the inclusion of additional error sources such as instrument noise or aliasing of non-tidal signals which may become the dominant error contributors (cf. Sections 3.3 and 3.4) a rather low cut-off degree might be well sufficient. The choice of an optimal cut-off degree, however, exceeds the scope of this work and might be subject of a future manuscript.

3.3 OT de-aliasing errors and instrument noise

In the following set of simulations, instrument noise is introduced to the simulation environment in addition to the OT de-aliasing error. Here, we consider a 3-axis accelerometer in accordance with the SuperSTAR specifications (Touboul *et al*. 2016) and the KBR inter-satellite link following the specification provided in Iran Pour *et al*. (2015). In addition to this realistic sensor noise (in the following denoted as *NS1*) we also consider an optimistic noise scenario, where the previously denoted instrument performances are collectively downscaled by a factor of 5 (in the following referred to as *NS2*). These specifications are graphically presented in Fig. 8. In the simulations, the instrument noise components are projected onto

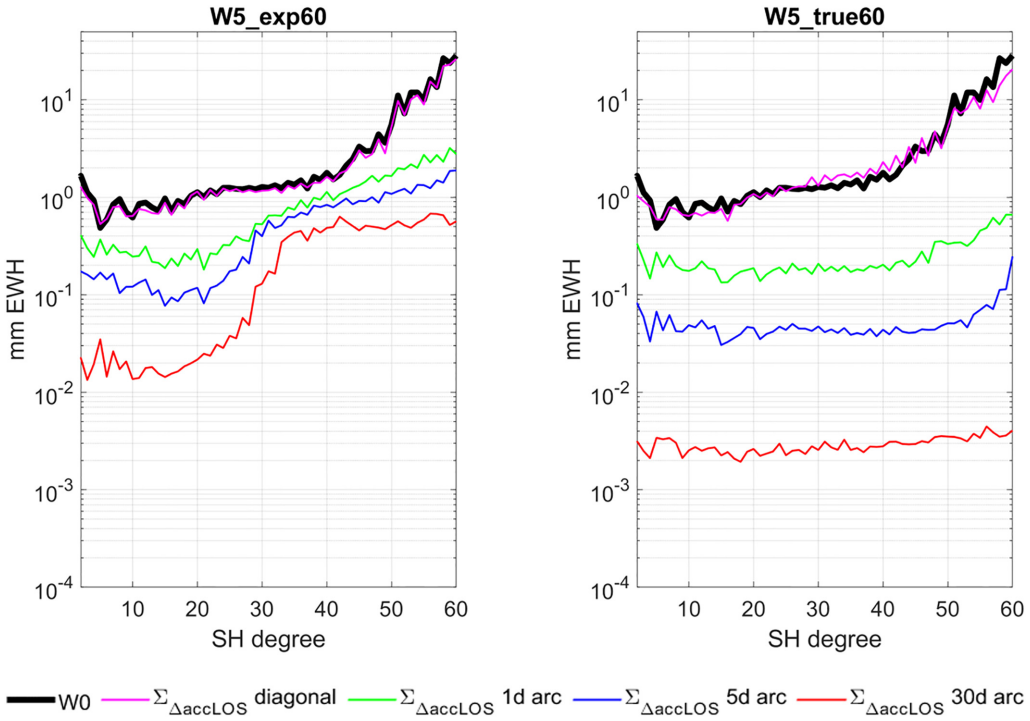

Figure 6. Retrieval error degree RMS of simulation scenarios based on the combined error signal of eight principal tides taken up to d/o 60 for a retrieval period of 30 d. For Σ_{xx} we use W5_exp60 (left), that is the full covariance information up to d/o 30 expanded with variances up for d/o 31 to 60, as well as W5_true60 (right), that is the full covariance information up to d/o 60.

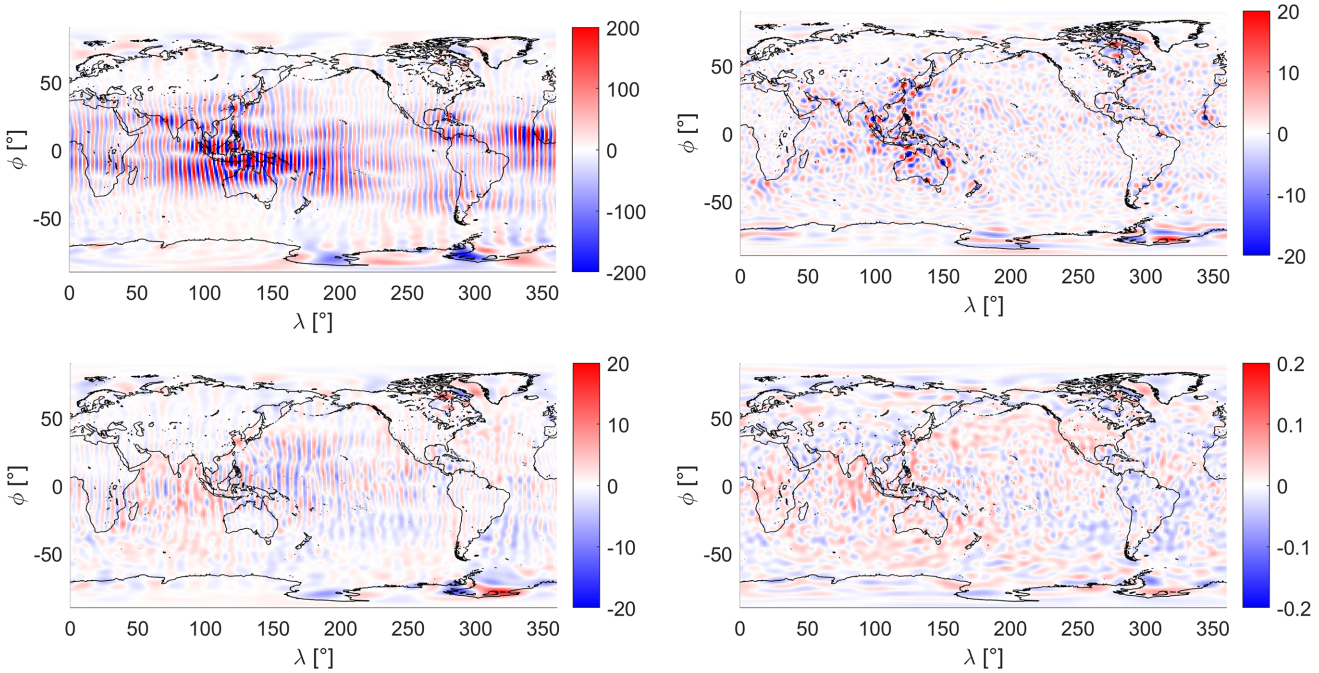

Figure 7. Global retrieval error in terms of [mm EWH] of simulation scenario based on the combined error signal for eight principal tides taken up to d/o 60 for a retrieval period of 30 d. Top left: no weighting is applied (reference scenario); top right: W5_exp60-based Σ_{xx}, full $\Sigma_{\Delta acc_{LOS}}$; bottom left: W5_true60-based Σ_{xx}, 1d-arc-structured $\Sigma_{\Delta acc_{LOS}}$; bottom right: W5_true60-based Σ_{xx}, full $\Sigma_{\Delta acc_{LOS}}$. Please note the different colour bar ranges.

the LOS and the resulting time-series is added to the observations. An error VCM representing the stochastic properties of the noise is then set up based on the LOS-projected time-series' empirical auto-covariance function. The weighting scheme for the OT BM errors remains unchanged, and thus, the total VCM of observations reads

$$\Sigma_{\Delta acc_{LOS}} = \Sigma_{\Delta acc_{LOS},\ \mathrm{OT}} + \Sigma_{\Delta acc_{LOS},\ \mathrm{sensors}} \tag{15}$$

implying that the OT and instrument errors are mutually uncorrelated.

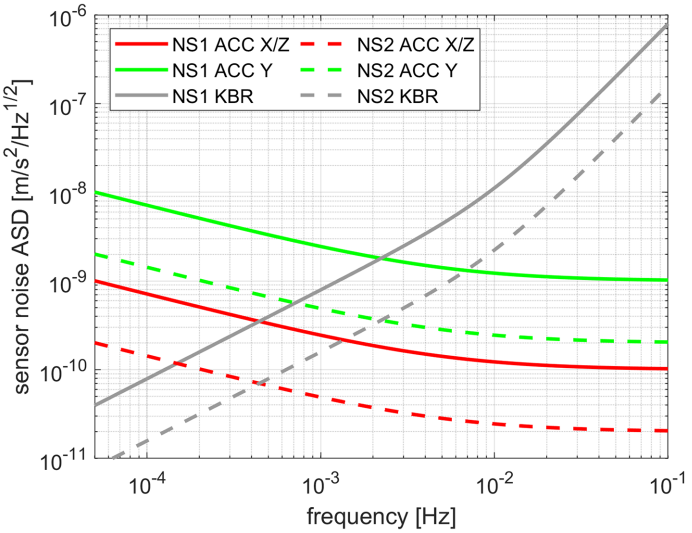

Figure 8. Instrument noise specifications for the three-axis accelerometer and the KBR link according to the realistic performance scenario (NS1, solid lines) and the improved performance scenario (NS2, dashed lines).

In a first step, we investigate the impact of *W5*-based OT BM weighting, and evaluations are carried out for a 30-d retrieval period using a full $\Sigma_{\Delta acc_{LOS}}$ as well as for its arc-wise approximations. The scenario employing a diagonal $\Sigma_{\Delta acc_{LOS}}$ is now disregarded, since the sensor noise is an exclusively temporal effect, and therefore cannot be stochastically modelled without considering correlations (or covariances, respectively) between observations. A scenario where no stochastic modelling is applied for either of the two considered error contributors is taken as reference (once again denoted as *W0*). The corresponding results are shown in Fig. 9.

It is evident that the *NS1* sensor noise is the dominant contributor to the total error budget from degree $n = 5$ onwards, while *NS2* noise remains subordinate to the OT de-aliasing error up to n = 22. On average, an improved retrieval quality of 50 per cent can now be asserted with respect to the reference scenario for *NS1* with a full $\Sigma_{\Delta acc_{LOS}}$. A quite similar retrieval performance can be achieved with a $\Sigma_{\Delta acc_{LOS}}$ based on 5-d arcs, while a computation with daily arcs seems to partially degrade the solutions in the harmonic degrees 15 to 25 compared to the 5-d- and the 30-d-arc approach. For *NS2* an even larger overall gain can be seen in case of a full (on average 75 per cent) and a 5-d-arc-based $\Sigma_{\Delta acc_{LOS}}$ (on average 60 per cent), while the daily-arc solution once again shows a slightly worse performance in the spectrum between harmonic degrees 10 and 25.

It is further clear that the maximum increase in retrieval performance can only be achieved when stochastic modelling is applied both to the OT BMs and the instrument noise. If weighting is applied exclusively with respect to instrument noise (which can be considered as the state of the art for the current data processing methodology), the solution quality remains unchanged at best in the spectral bands where the OT de-aliasing error is dominant, and only improves in the spectral range dominated by the sensor noise (as stated earlier, longer arcs are preferrable). In case of *NS2* even a deterioration of the solution in the degrees above 12 can be observed.

It can overall be asserted that the gain in the current set of simulations is, although still significant, distinctly smaller than in the OT-only simulations shown in Section 3.2. The logical explanation for this behaviour is the fact that OT signal and instrument noise introduce different types of correlations between observations (mainly spatial correlations in case of OT, exclusively temporal correlations in case of sensor noise). Therefore, when the partial VCMs are combined, the influence of each contributor to an observation cannot be distinguished.

Now, the impact of *WI*-based OT BM weighting is studied. While in Section 3.2 a unit matrix was taken for Σ_{xx}, such a VCM would now fully cover the covariance information representing the instruments' stochastic properties, and must therefore be rescaled accordingly. Thus, Σ_{xx} now reads

$$\Sigma_{xx} = \beta I. \tag{16}$$

The scale factor β is chosen empirically to 1.4×10^{-23} in case of *NS1* and to 5.6×10^{-25} in case of *NS2* in order for the resulting $\Sigma_{\Delta acc_{LOS}}$ to feature a similar relative impact from $\Sigma_{\Delta acc_{LOS,\,sensors}}$ and $\Sigma_{\Delta acc_{LOS,\,OT}}$ as in case of *W5*-based OT weighting. In accord with the findings obtained in Section 3.2, the corresponding simulations are carried out only for a 30-d retrieval using a full $\Sigma_{\Delta acc_{LOS}}$. The simulation results shown in Fig. 10 indicate that the retrieval performance can still be improved to some extent when weighting is applied both to OT BM errors and the sensor noise.

However, the overall gain is now significantly smaller than in case of a realistic Σ_{xx} (*W5*), and it is now mostly limited to the degrees below $n = 5$ and above $n = 20$. Both for *NS1* and *NS2* the solution quality can be improved by up to 30 per cent (below $n = 5$) or 15 per cent (above $n = 20$), respectively. Also, for *NS1* one can achieve the best overall solution in the spectrum above $n = 5$ when applying observation weighting based exclusively on the noise component. In contrast, for *NS2* the combined weighting of OT error and instrument noise still yields the best overall solution. We conclude that using *WI* to stochastically model the OT BM errors is only advisable in case no realistic, that is *W5*-based, Σ_{xx} are available. However, we emphasize that this approach can only be expected to yield an improved retrieval performance when the error contribution of imperfect OT de-aliasing supersedes that of the instruments.

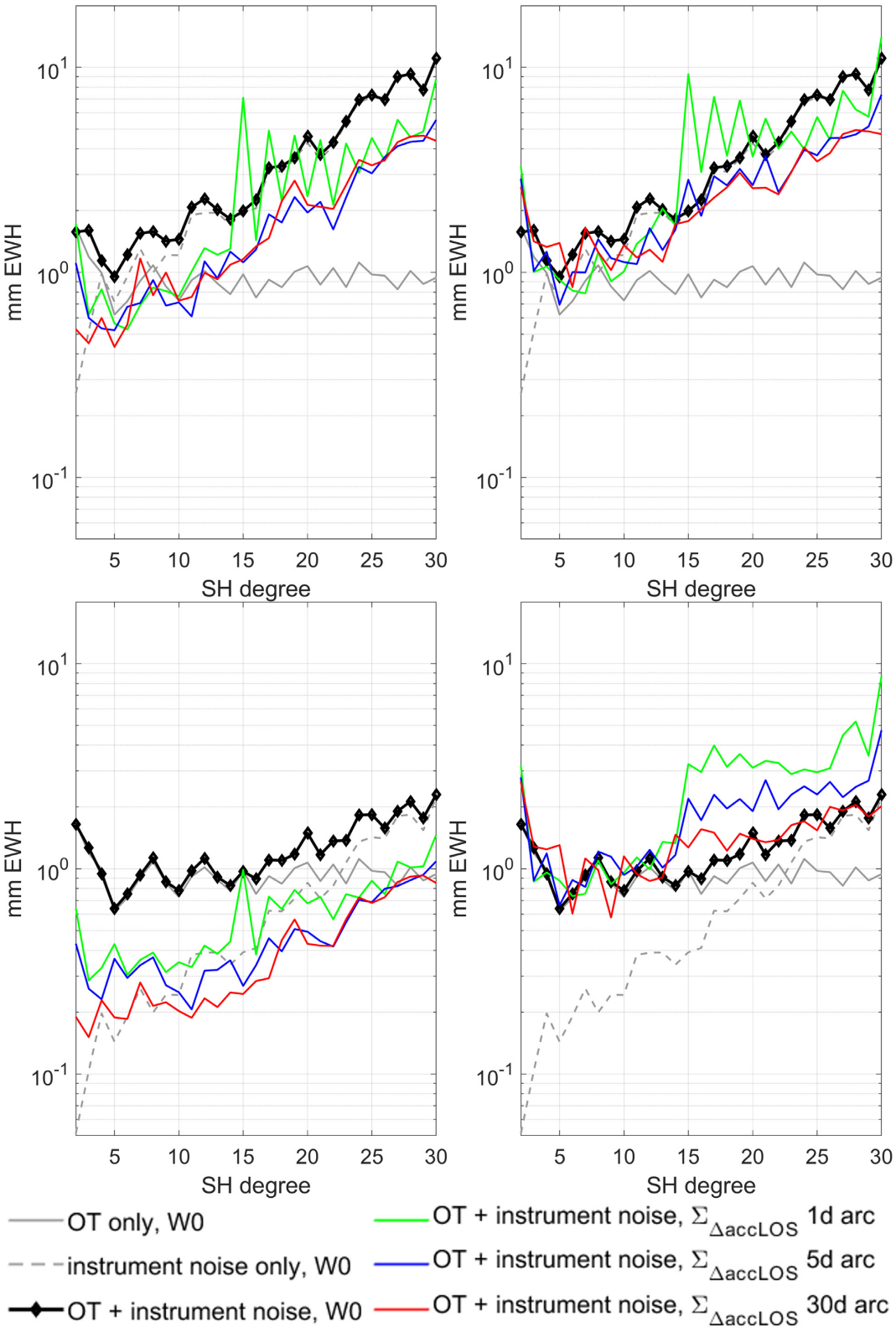

Figure 9. Retrieval error degree RMS of simulations based on the combined error signal of eight principal tides as well as NS1 (top row) or NS2 (bottom row) sensor noise for a retrieval period of 30 d. Observation weighting is applied for the OT signal (Σ_{xx} is based on W5) and the sensor noise (left column) as well as exclusively for the sensor noise (right column).

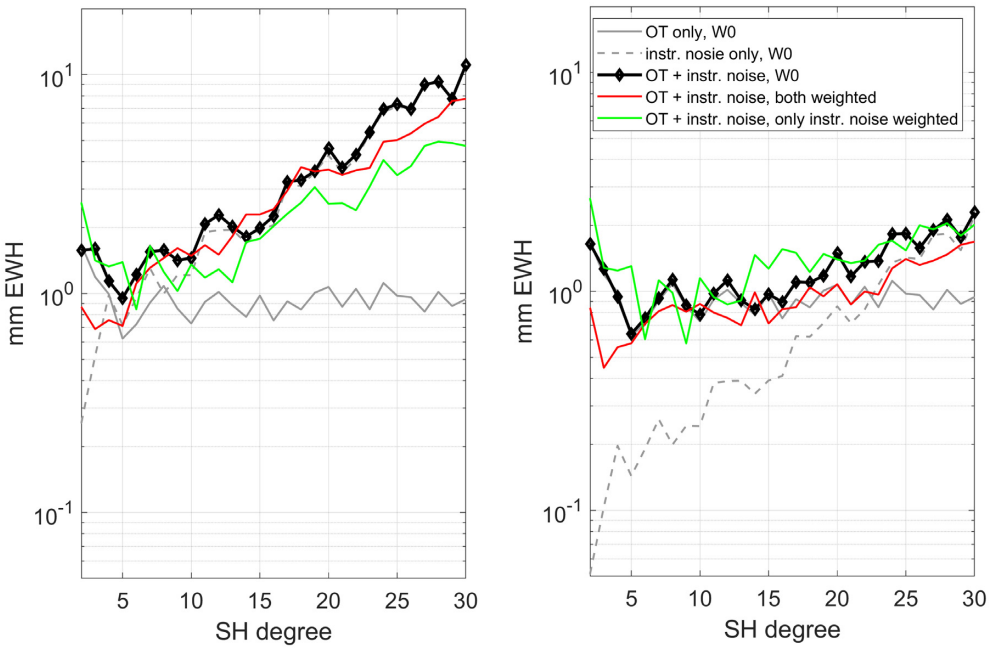

Figure 10. Retrieval error degree RMS of simulations based on the combined error signal of eight principal tides as well as NS1 (left) or NS2 (right) sensor noise, respectively, for a retrieval period of 30 d. Observation weighting is applied for the OT signal (Σ_{xx} is based on WI) and sensor noise as well as exclusively for the sensor noise.

3.4. OT de-aliasing errors, instrument noise and AO de-aliasing error

In this final set of simulations, we now consider the non-tidal temporal variations taken from the updated ESA Earth System Model in addition to the OT BM and instrument errors in accord with *NS1* and *NS2*. Here, we distinguish between three scenarios. First, we consider only the signal components stemming from hydrology, ice and solid Earth (HIS) within the simulations, thereby implicitly assuming that the AO components were fully removed with the help of geophysical BMs. We then consider a more realistic scenario where we assume that 10 per cent of the AO signal could not be compensated within the *a priori* de-aliasing. Finally, the full AOHIS signal is taken, which implies that no *a priori* AO de-aliasing was carried out, and therefore represents the worst-case scenario. Analogously to the previous section, stochastic modelling is carried out for the OT BMs (this time, only *W5* is applied) and the sensor noise.

Fig. 11 shows that the contribution of non-tidal HIS variations to the overall retrieval error is fully superimposed by the erroneous OT de-aliasing and the sensor noise. Therefore, the weighting performance remains almost unchanged in comparison to the results shown in Fig. 9. For *NS2* the retrieval performance slightly decreases around $n = 15$ in comparison to the corresponding OT and sensor noise simulation, since the HIS error now constitutes around 50 per cent of the total retrieval error.

The effect of the realistic AO de-aliasing error manifests itself primarily in the long-wavelength spectrum of the gravity solution. Consequently, the retrieval performance is notably degraded in the low degrees, although an improvement can still be seen in comparison to the reference scenario. For *NS1* the combined OT and noise weighting now yields only up to 15 per cent improvement below $n = 6$ (in comparison: over 30 per cent improvement in case of perfect AO de-aliasing). For *NS2*, the decrease in retrieval quality can be observed up to $n = 15$, where now only an average gain of around 35 per cent can be asserted (in comparison: over 65 per cent in case of perfect AO de-aliasing). Further, the performance in these spectral bands seems to be independent of the approximation chosen for $\Sigma_{\Delta \mathrm{acc}_{\mathrm{LOS}}}$.

In case that no AO de-aliasing is carried out at all, the hereby induced temporal aliasing becomes the most dominant error contributor, and the applied stochastic modelling therefore loses nearly all of its additional benefit. Only for *NS1*, where the instrument noise once again dominates the retrieval performance above $n = 23$, one can see some minor improvements due to the observation weighting. For *NS2*, on the other hand, no gain can be established with respect to *W0*.

This finding strongly underlines the necessity of improved AO de-aliasing, be it through improved geophysical BMs or through novel processing strategies, for example sophisticated parametrization or stochastic modelling analogous to the approach presented in this paper, in order to fully exploit the benefit of OT BM weighting.

4. SUMMARY AND OUTLOOK

In this study, the additional benefit of stochastically modelling OT BM errors is investigated in the context of a GRACE-type satellite gravimetry mission. For this purpose, realistic error VCMs depicting a time-invariant spatial error pattern are derived up to d/o 30 for each of the eight major tidal constituents based on a set of five modern-day OT models. As an additional scenario, these VCMs are taken as unit matrices in order to exclusively use temporal variations stemming from the tidal signals' periodicity. The error VCMs are then incorporated into numerical closed-loop simulations, where they are rigorously propagated onto the level of observations. Instrument noise as well as non-tidal temporal gravity field variations are also introduced to simulate a realistic gravity retrieval performance. Different approximations of the resulting VCM of observations are then used to evaluate the trade-off between computational efficiency and accuracy.

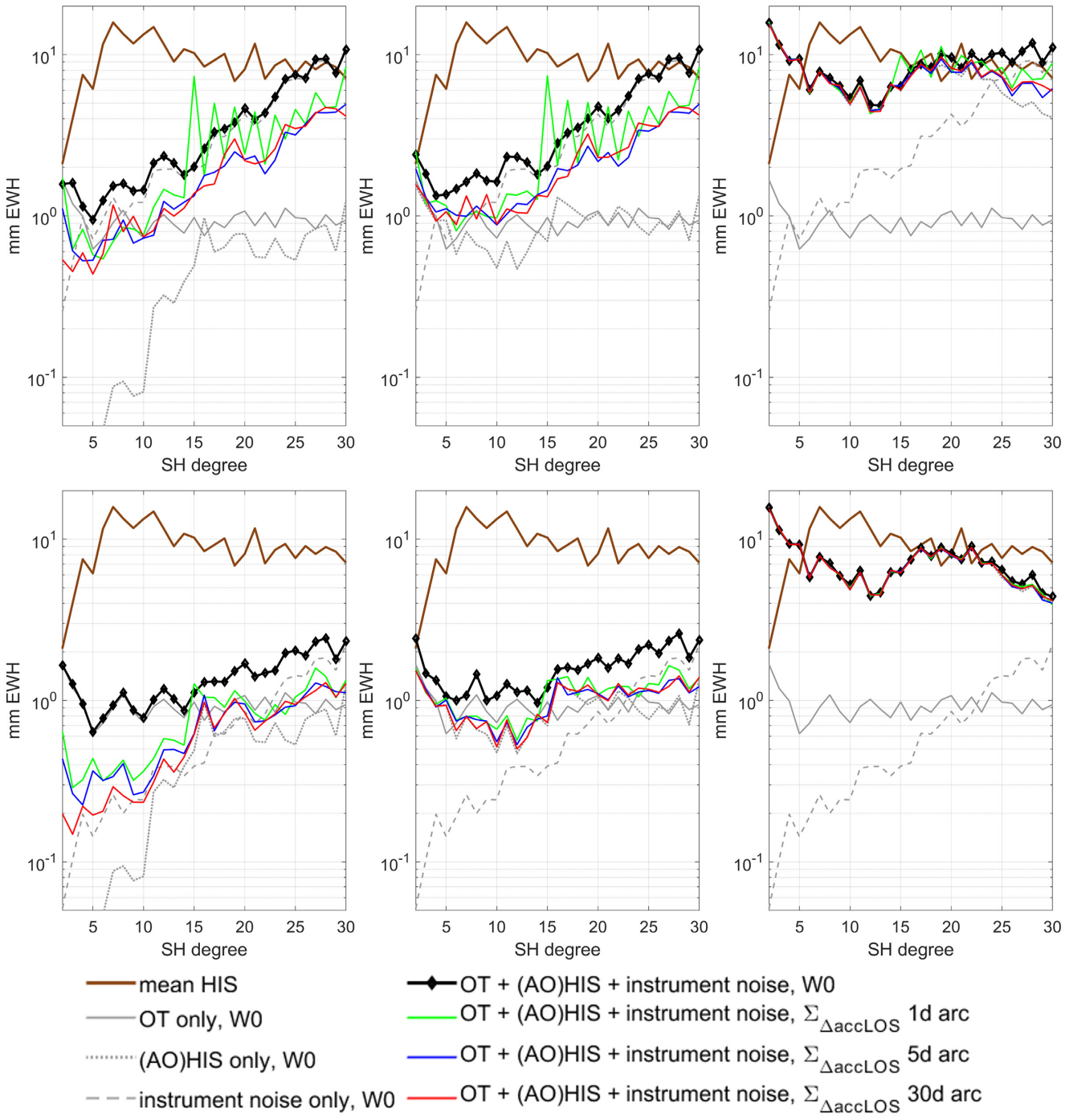

Figure 11. Retrieval error degree RMS of simulations based on the combined error signal of eight principal tides and NS1 (top row) or NS2 (bottom row) sensor noise, respectively, as well as non-tidal temporal signal, respectively assuming perfect AO de-aliasing (left column), an AO de-aliasing error of 10 per cent (middle column) or no AO de-aliasing (right column) for a retrieval period of 30 d. Observation weighting is applied for the OT signal (Σ_{xx} is based on W5) and the sensor noise.

It is shown that the weighting performance varies between tidal constituents due to the different underlying aliasing periods. Thus, the improvement in retrieval performance is smaller for partial tides with long aliasing periods, as the resulting error signal is of nearly deterministic nature. Longer retrieval periods, ideally constituting at least the aliasing period, are therefore recommended.

In case the imperfect OT de-aliasing constitutes the sole contribution to the retrieval error, we show that significant performance gains can be asserted when realistic OT BM error weighting is applied. Here, the retrieval error can be decreased by four orders of magnitudes if a full VCM of observations is used, and up to 1.5 orders of magnitude in terms of degree RMS in case of a VCM of observations based on daily arcs. Comparatively minor gains are obtained when correlations between observations are entirely neglected. In this set of simulations, it is further established that even if exclusively temporal correlations between observations are

considered, the solution can already be improved by 60 per cent in terms of degree RMS with a full VCM of observations. It is additionally established that the overall retrieval performance remains widely invariant if OT signal is considered up to a higher d/o, as long as the weighting is done correctly and consistently. Using full error covariance information only up to a certain SH degree lower than that of the considered signal has also been proven to yield a significant gain in retrieval performance, which can be considered of great interest for application of OT BM weighting in real data processing.

When instrument noise is applied in addition to the OT de-aliasing error, notable improvements can still be asserted if stochastic modelling is applied for both effects. In this manner, a retrieval performance gain of up to 50 per cent in terms of degree RMS can be established when using a full VCM of observations based on realistic spatial OT error information. The additional benefit can further be increased by improving the sensor performance. Here an additional gain of over 20 per cent in solution quality can be reached if the instrument noise is decreased by half an order of magnitude. Generally, in this set of simulations the overall solution quality does not depend as strongly on the VCM's approximation level as in case of OT-only simulations (although the full VCM still allows for the overall lowest retrieval error). It is further noted that neglecting the spatial correlations in favour of temporal ones when stochastically modelling OT BM errors can also lead to an improved gravity retrieval. However, this approach yields significantly smaller improvements as when realistic error VCMs are used. Another disadvantage here is that the underlying error VCMs must be empirically rescaled for it to fit the instruments' error level. Therefore, it is not recommended to use this weighting approach unless no realistic error VCMs are available.

Once non-tidal gravity field variations are introduced into the simulation, the performance differences between different approximations of the observations' VCM become marginal, meaning that computational efficiency can be greatly improved by processing and weighting observations based on, for example daily arcs without notable drawbacks in solution accuracy. It also becomes clear that the gain of OT BM error weighting depends on the quality of the *a priori* AO de-aliasing. Full advantage of stochastically modelling the OT model errors can be taken in case the AO-based signal components are fully removed, as they otherwise become the dominant error contributor and therefore greatly diminish the effect of OT BM error weighting. Here it is shown that the retrieval performance can be increased by over 30 per cent in the low-degree spectrum in case of perfect AO de-aliasing, while only up to 15 per cent gain can be asserted in this part of the spectrum when 10 per cent of the AO signal remain uncompensated.

Based on the findings presented in this study we conclude that stochastic modelling of OT BM errors offers great benefit for GRACE/GRACE-FO data processing, as it allows to mitigate the impact of imperfect OT de-aliasing. Optimal results can be achieved by precisely modelling the underlying spatial error patterns and utilizing a maximum of the observations' covariance information. However, since the overall performance depends to a high degree on the interaction with other contributors to the total error budget, stochastic modelling of sensor noise is recommended, and general reduction of the sensors' noise levels can be regarded as highly beneficial. On the other hand, improved AO de-aliasing can be considered as imperative for the future. This can potentially be achieved through general improvement of the corresponding BMs and/or novel processing techniques, for example a BM weighting procedure similar to the strategy presented in this paper. The findings presented here should further be validated on real GRACE data. Also, the benefit of combining OT BM weighting with currently used parametrization approaches, for example explicit OT co-estimation, should be studied.

ACKNOWLEDGEMENTS

The work presented in this paper was performed within the framework of the project 'New Refined Observations of Climate Change from Spaceborne Gravity Missions (NEROGRAV)' funded by the German Research Foundation (Research Unit FOR 2736/1). We also acknowledge the provision of supercomputing resources by the Leibniz Supercomputing Centre (LRZ; Address: Boltzmannstraße 1, 85748 Garching bei München, Germany).

DATA AVAILABILITY

The data underlying this paper will be shared on reasonable request to the corresponding author.

REFERENCES

Bender, P. L., Wiese, D. N. & Nerem, R. S., 2008. A possible dual-grace mission with 90 degree and 63 degree inclination orbits, in *Proceedings of the 3rd International Symposium on Formation Flying, Missions and Technologies,* ESA/ESTEC, Noordwijk, pp. 1–6.

Bergmann, I. & Dobslaw, H., 2012. Short-term transport variability of the Antarctic Circumpolar Current from satellite gravity observations, *J. geophys. Res.,* **117,** C05044, doi:10.1029/2012JC007872.

Bosch, W., Savchenko, R., Flechtner, F., Dahle, C., Mayer-Gürr, T., Stammer, D., Taguchi, E. & Ilk, K. H., 2009). Residual ocean tide signals from satellite altimetry, GRACE gravity fields, and hydrodynamic modelling, *Geophys. J. Int.,* **178**(3), 1185–1192.

Cheng, Y. & Andersen, O. B., 2011). Multimission empirical ocean tide modelling for shallow waters and polar seas, *J. geophys. Res.,* **116**(C11). doi: https://doi.org/10.1029/2011JC007172.

Dobslaw, H., Bergmann-Wolf, I., Dill, R., Forootan, E., Klemann, V., Kusche, J. & Sasgen, I., 2015. The updated ESA Earth System Model for future gravity mission simulation studies, *J. Geod.,* **89**(5), 505–513.

Egbert, G D. & Erofeeva, S Y., 2002. Efficient inverse modeling of barotropic ocean tides, *J. Atmos. Ocean. Technol.,* **19.2,** 183–204.

Flechtner, F., Neumayer, K. H., Dahle, C., Dobslaw, H., Fagiolini, E., Raimondo, J. C. & Güntner, A., 2015. What can be expected from the GRACE-FO laser ranging interferometer for Earth science applications?, *Surv. Geophys.,* **37,** 453–470.

Fok, H. S., 2012. *Ocean Tides Modeling Using Satellite Altimetry*. Ohio State University.

Friedman, J., Hastie, T. & Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso, *Biostatistics,* **9**(3), 432–441.

Han, S., Shum, C. & Matsumoto, K., 2005. GRACE observations of M2 and S2 ocean tides underneath the Filchner-Ronne and Larsen ice shelves, Antarctica, *Geophys. Res. Lett.,* **32,** L20311, doi:10.1029/2005GL024296.

Hannart, A. & Naveau, P., 2014. Estimating high dimensional covariance matrices, *J. Multivariate Anal.,* **131,** 149–162.

Hauk, M. & Pail, R., 2018. Treatment of ocean tide aliasing in the context of a next generation gravity field mission, *Geophys. J. Int.,* **214**(1), 345–365.

Iran Pour, S. *et al.*, 2015. Assessment of satellite constellations for monitoring the variations in earth gravity field, SC4MGV, ESA/ESTEC Contract No. AO/1-7317/12/NL/AF, Final Report.

Keune, J., Ohlwein, C. & Hense, A., 2014 Multivariate probabilistic analysis and predictability of medium-range ensemble weather forecasts, *Mon. Weather Rev.,* **142**(11), 4074–4090.

Killett, B., Wahr, J., Desai, S., Yuan, D. & Watkins, M., 2011. Arctic Ocean tides from GRACE satellite accelerations, *J. geophys. Res.,* **116,** C11005, doi:10.1029/2011JC007111.

King, M. & Padman, L., 2005. Accuracy assessment of ocean tide models around Antarctica, *Geophys. Res. Let.,* **32,** L23608, doi:10.1029/2005GL023901.

Kornfeld, R. P., Arnold, B. W., Gross, M. A., Dahya, N. T., Klipstein, W. M., Gath, P. F. & Bettadpur, S., 2019. GRACE-FO: the gravity recovery and climate experiment follow-on mission, *J. Spacecr. Rockets,* **56**(3), 931–951.

Kvas, A. & Mayer-Gürr, T., 2019. GRACE gravity field recovery with background model uncertainties, *J. Geod.,* **93,** 2543–2552.

Laska, J. & Narayan, M., 2018. skggm 0.2.8: a scikit-learn compatible package for general graphical models (Version 0.2.8), Zenodo, http://doi.org/10.5281/zenodo.1413742.

Ledoit, O. & Wolf, M., 2003. Improved estimation of covariance matrix of stock returns with application to portfolio selection, *J. Empir. Finance,* **10**(5), 603–621.

Liu, W., 2019. Understanding ocean tide aliasing in satellite gravimetry, Universität Stuttgart, OPUS – Online Publikationen der Universität Stuttgart. doi:10.18419/opus-10485.

Liu, W. & Sneeuw, N., 2021. Aliasing of ocean tides in satellite gravimetry: a two-step mechanism, *J. Geod.,* in review.

Liu, W., Sneeuw, N., Iran Pour, S., Tourian, M. J. & Reubelt, T., 2016. A posteriori de-aliasing of ocean tide error in future double-pair satellite gravity missions, in *International Symposium on Earth and Environmental Sciences for Future Generations. International Association of Geodesy Symposia,* Vol. 147, pp. 103–109, eds Freymueller, J.T. & Sanchez, L., Springer, Cham.

Lyard, F. H., Allain, D. J., Cancet, M., Carrère, L. & Picot, N., 2021. FES2014 global ocean tide atlas: design and performance, *Ocean Sci.,* **17,** 615–649.

Mayer-Gürr, T., 2006. Gravitationsfeldbestimmung aus der Analyse kurzer Bahnbögen am Beispiel der Satellitenmissionen CHAMP und GRACE, in *Schriftenreihe /Institut für Geodäsie und Geoinformation,* **9**. Online-Ausgabe in bonndoc: https://hdl.handle.net/20.500.11811/1391.

Mayer-Gürr, T., Savcenko, R., Bosch, W., Daras, I., Flechtner, F. & Dahle, C., 2012. Ocean tides from satellite altimetry and GRACE, *J. Geodyn.,* **59-60,** 28–38.

Murböck, M., 2015. Virtual Constellations of Next Generation Gravity Missions. Deutsche Geodästische Komission der Bayerischen Akademie der Wissenschaften, Reihe C, Dissertationen, Heft 750, Verlag der Bayerischen Akademie der Wissenschaften, ISBN(Print) 978-3-7696-5162-1, ISSN 0065-5325, 2016.

Murböck, M., Pail, R., Daras, I. & Gruber, T., 2014. Optimal orbits fpr temporal gravity recovery regarding temporal aliasing, *J. Geod.,* **88**(2), 113–126.

Rajaratnam, B., Massam, H. & Carvalho, C. M., 2008. Flexible covariance estimation in graphical Gaussian models, *Ann. Stat.,* **36**(6), 2818–2849.

Ray, R., 2013. Precise comparisons of bottom-pressure and altimetric ocean tides, *J. geophys. Res.,* **118**(9), 4570–4584.

Ray, R. & Luthcke, S., 2006. Tide model errors and GRACE gravimetry: towards a more realistic assessment, *Geophys. J. Int.,* **167**(3), 1055–1059.

Röpnack, A., Hense, A., Gebhardt, C. & Majevski, D., 2011. Bayesian model verification of NWP ensemble forecasts, *Mon. Weather Rev.,* **141,** 375–387.

Savchenko, R. & Bosch, W., 2012. EOT11a – Global Empirical Ocean Tide model from multi-mission satellite altimetry, DGFI Report No. 89, München: Deutsches Geodätisches Forschungsinstitut.

Schafer, J. L., 1997. *Analysis of Incomplete Multivariate Data,* Chapmann and Hall.

Stammer, D. *et al.*, 2014. Accuracy assessment of global barotropic ocean tide models, *Rev. Geophys.,* **52**(3), 243–282.

Taguchi, E., Stammer, D. & Zahel, W., 2014. Inferring deep ocean tidal energy dissipation from the global high-resolution data-assimilative HAMTIDE model, *J. geophys. Res.,* **119,** 4573–4592.

Tapley, B., Bettadpur, S., Watkins, M. & Reigber, C., 2004. The gravity recovery and climate experiment: mission overview and early results, *Geophys. Res. Lett.,* **31**(9), L09607, doi: 10.1029/2004GL019779.

Touboul, P. *et al.*, 2016. Gravitation and geodesy with inertial sensors, from ground to space, *Testing in Aerospace Research,* doi:10.12762/2016.AL12-11.

Visser, P., 2010. Designing Earth gravity field missions for the future: a case study., in *Gravity, Geoid and Earth Observation. International Association of Geodesy Symposia,* Vol. 135, pp. 131–138, ed. Mertikas, S., Springer-Verlag.

Wiese, D. N., Killett, B., Watkins, M. & Yuan, D., 2016. Antarctic tides from GRACE satellite accelerations, *J. geophys. Res.,* **121**(5), 2874–2886.

Won, J. H., Lim, J., Kim, S. J. & Rajaratnam, B., 2012. Condition-number-regularized covariance estimation, *J. R. Stat. Soc. B,* **75**(3), 427–450.

Zerenner, T., Friedrichs, P., Lehnertz, K. & Hense, A., 2014. A Gaussian graphical model approach to climate networks, *Chaos,* **24,** 023103, https://doi.org/10.1063/1.4870402.