JGR Oceans

RESEARCH ARTICLE

10.1029/2020JC017031

Convergence of Daily GRACE Solutions and Models of Submonthly Ocean Bottom Pressure Variability

Michael Schindelegger[1], Alexander A. Harker[1], Rui M. Ponte[2], Henryk Dobslaw[3], and David A. Salstein[2]

[1]Institute of Geodesy and Geoinformation, University of Bonn, Bonn, Germany, [2]Atmospheric and Environmental Research, Inc., Lexington, MA, USA, [3]Department 1: Geodesy, GFZ German Research Centre for Geosciences, Potsdam, Germany

Key Points:

- We rigorously compare daily Gravity Recovery and Climate Experiment (GRACE) gravity solutions with bottom pressure output from five ocean models at periods <60 days
- Southern Ocean mass-field variability in current de-aliasing model is too energetic; dedicated barotropic simulations better match GRACE
- Daily gravity fields have errors of 0.5–1.5 cm (water height) over basin interiors and may guide improvements to existing ocean models

Correspondence to:

M. Schindelegger,
schindelegger@igg.uni-bonn.de

Citation:

Schindelegger, M., Harker, A. A., Ponte, R. M., Dobslaw, H., & Salstein, D. A. (2021). Convergence of daily GRACE solutions and models of submonthly ocean bottom pressure variability. *Journal of Geophysical Research: Oceans*, *126*, e2020JC017031. https://doi.org/10.1029/2020JC017031

Received 30 NOV 2020
Accepted 1 FEB 2021

Abstract Knowledge of submonthly variability in ocean bottom pressure (p_b) is an essential element in space-geodetic analyses and global gravity field research. Estimates of these mass changes are typically drawn from numerical ocean models and, more recently, GRACE (Gravity Recovery and Climate Experiment) series at daily sampling. However, the quality of p_b fields from either source has been difficult to assess and reservations persist as to the dependence of regularized GRACE solutions on their oceanographic priors. Here, we make headway on the subject by comparing two daily satellite gravimetry products (years 2007–2009) both with each other and with p_b output from a diverse mix of ocean models, complemented by insights from bottom pressure gauges. Emphasis is given to large spatial scales and periods <60 days. Satellite-based mass changes are in good agreement over basin interiors and point to excess p_b signals (~2 cm root-mean-square error) over Southern Ocean abyssal plains in the present GRACE de-aliasing model. These and other imperfections in baroclinic models are especially apparent at periods <10 days, although none of the GRACE series presents a realistic ground truth on time scales of a few days. A barotropic model simulation with parameterized topographic wave drag is most commensurate with the GRACE fields over the entire submonthly band, allowing for first-order inferences about error and noise in the gravimetric mass changes. Estimated p_b errors vary with signal magnitude and location but are generally low enough (0.5–1.5 cm) to judge model skill in dynamically active regions.

Plain Language Summary Changes in the pressure at the seafloor tell us how ocean masses move in time and space. These environmental signals are important for understanding variations in Earth's shape, rotation, and gravity field. We assess how well we know the rapid, submonthly portion of bottom pressure changes by analyzing output from oceanographic models and observations from the Gravity Recovery and Climate Experiment (GRACE) dual satellite mission. We show that two different GRACE solutions, sampled daily, are in good agreement with each other over the deep interior of the ocean basins. Moreover, bottom pressure changes simulated with a simple single-layer model are remarkably consistent with GRACE, providing an independent measure of the quality of both products. Based on these grounds, and by aid of an approximate error assessment, we suggest that nonstandard daily GRACE fields are realistic enough to help identifying deficiencies in oceanographic models and guide solutions to these issues. We particularly highlight an overestimation of Southern Ocean bottom pressure variability in two widely used general circulation simulations and speculate on ways how to improve the underlying models.

1. Introduction

The Gravity Recovery and Climate Experiment (GRACE) and its follow-on mission (GRACE-FO; Tapley et al., 2019) have been providing a unique series of measurements of temporal gravity variations at large spatial scales. The technique has rung in a new era in the ability to quantify mass anomalies and fluxes across the Earth system, including changes in terrestrial water storage (Rodell et al., 2018), temporal variations in ocean circulation (Zlotnicki et al., 2007), the mass balance of ice sheets (Velicogna et al., 2020), and manometric contributions to regional sea level (Rietbroek et al., 2016). Although the satellites are orbiting the Earth once in 90 min, it typically takes 1 month for the ground track coverage to become dense enough such that the data can be inverted for a full global gravity field. Mass variations associated with geophysical phenomena on time scales shorter than the nominal GRACE Nyquist frequency (0.5 cyc/

month) are accounted for in the precise orbit determination process by subtracting model-based predictions of their effect on the satellite motion and intersatellite range measurements. However, models of high-frequency mass changes (both tidal and nontidal) are necessarily imperfect, resulting in aliasing errors and unpredictable artifacts in the monthly gravity solutions (e.g., Elsaka et al., 2012; Seo et al., 2016). In fact, full-scale simulations of the GRACE-FO satellite system (Flechtner et al., 2016) indicate that errors in the oceanic de-aliasing models for both tidal and nontidal signals are, along with accelerometer noise, the limiting source of spatially correlated errors in the monthly gravity field time series. Better knowledge of submonthly nontidal mass variations is therefore critical to the success of satellite gravimetry missions on climate time scales.

A precise depiction of nontidal oceanic variability is also important for geometric space-geodetic techniques (e.g., Global Positioning System), since corrections for crustal loading are required to regularize station coordinate time series that contribute to a stable global reference frame (Glomsda et al., 2020; Williams & Penna, 2011). Moreover, mass redistributions in the ocean cause perturbations to the angular momentum of Earth's fluid envelope, which manifests in orientation changes of the solid Earth with respect to inertial space (e.g., Ponte & Ali, 2002). Forecasts of these short-term ocean angular momentum variations are particularly relevant for operational activities like real-time spacecraft navigation or tracking of deep-space objects (Dill et al., 2019). Fluctuations of ocean bottom pressure (p_b), equivalent to the anomalous mass in the column of fluid above, and its spatial gradients are also of interest to physical oceanographers. Specifically, rapid p_b signals can provide insight into barotropic transport variability (Bergmann & Dobslaw, 2012) and effects of bottom topography, dissipation, and atmospheric forcing on ocean circulation and sea-level variability (e.g., Fukumori et al., 2015; Stepanov & Hughes, 2006).

Mapping the accuracy of numerical models involved in these questions has been a notoriously difficult task. Tests of GRACE range-rate residuals aside, assessments of oceanic de-aliasing models have mostly relied on measurements from globally distributed bottom pressure recorders (BPRs, e.g., Dobslaw et al., 2017a; Quinn & Ponte, 2011; Zenner et al., 2014) and comparisons with altimetric sea-level anomalies (e.g., Bonin & Chambers, 2011; Bonin & Save, 2020; Dobslaw et al., 2017a). Neither dataset gives access to the full spectrum of rapid p_b variability relevant to GRACE. Coverage with seafloor pressure measurements is sparse in both space and time, and variance of particular deployments may reflect eddies or other localized effects (Androsov et al., 2020). In contrast, satellite altimetry provides good sampling over much of the ocean, but differences with modeled bottom pressure are not easily interpreted in a quantitative sense where baroclinic signals contribute to observed sea-level changes.

An alternative and potentially useful knowledge reservoir for p_b variations are GRACE gravity fields with up to daily resolution from constrained parameter estimation approaches (Bonin & Save, 2020; Kurtenbach et al., 2009; Kvas et al., 2019). Several lines of evidence suggest these solutions represent true submonthly oceanic mass signals (e.g., Bonin & Save, 2020; Poropat et al., 2019), but an explicit quality assessment in terms of bottom pressure, especially on large scales, has been lacking. Accordingly, the extent to which GRACE-derived p_b retrievals can help improve numerical models down to the 1 cm accuracy level (of equivalent water height) and beyond remains unknown. These vital questions set the tone for the study at hand. We address them through a systematic comparison of two advanced daily GRACE gravity field series with a variety of ocean forward models, each carrying its own strengths and weaknesses. If the two GRACE solutions agree on deficiencies in particular models, that would be a novel result, testifying to the credibility of the daily gravity field estimates over the ocean. Here, we explore such "convergence" in a broadband sense, at periods below 60 and 10 days. Placing the secondary cut-off at 10 days is somewhat arbitrary but emphasizes dynamic signals that are driven by both wind stress and atmospheric pressure (Hirose et al., 2001). These rapid barotropic motions explain 30%–60% of all submonthly p_b variability, depending on the region.

The paper starts with a description of the utilized GRACE series and ocean models (Sections 2 and 3), before all bottom pressure datasets are compared in detail in Section 4. We complement this analysis by mapping errors of the GRACE-derived submonthly bottom pressure change in space (Section 5), and end with a discussion of potential model improvements and further lines of research. Throughout the paper, the term "submonthly" is used as synonym for periods of less than 60 days.

2. Assessment of Daily GRACE Solutions

Considerations of ground track geometry imply that the estimation of daily gravity fields from GRACE observations alone is impractical. Resolving these data into high temporal frequencies and wavelengths comparable to standard monthly solutions thus requires additional information, for example, on how the gravity field is expected to evolve in time. Kurtenbach et al. (2012) proposed to cast statistical information on process dynamics and noise from external geophysical models (atmosphere, ocean, hydrology) into temporal correlations and thereby stabilize each daily solution with information from adjacent days. Stochastic priors and measurements of the process are combined in a recursive Kalman filter—an approach that has matured into a series of daily gravity fields by ITSG (Institute of Geodesy at Graz University of Technology). We use ITSG-Grace2018 (ITSG2018 for short, Kvas et al., 2019; Mayer-Gürr et al., 2018) given in spherical harmonic coefficients for degrees $n = 2...40$ that were estimated as constrained daily parameters relative to a long term static field with annual and secular variations (Kvas, 2020). Background models subtracted during the processing include the Atmosphere and Ocean De-Aliasing Level-1B (AOD1B) product (Release 06 or RL06, Dobslaw et al., 2017a). The covariance model characterizing the expected signal over both continents and oceans is taken from a 12 years-long series of synthetic time-variable gravity fields (Dobslaw et al., 2016) that share common heritage with AOD1B. The degree to which these particular choices of deterministic and stochastic priors bias the daily ITSG2018 solution toward the de-aliasing background model is a subtle issue we seek to address below.

We also use a daily GRACE solution created by the Center for Space Research (CSR) at the University of Texas, Austin (Bonin & Save, 2020; Save, 2019). These time series are based on a "swath" formulation that updates predefined mass concentration (mascon) blocks with exactly one day's worth of GRACE data if the center of the mascon is within 250 km from the satellites' ground track. A particular day's inversion is stabilized by means of a spatio-temporal regularization scheme, in part derived from the GRACE signal content of the respective month (Bonin & Save, 2020). Correlations between land and ocean blocks are penalized to zero to mitigate leakage of continental mass anomalies across coastlines. Release 05 of AOD1B (Dobslaw et al., 2013) was applied as the de-aliasing model during processing and restored in the final product. These swath data are not yet available in their full extent, so we content ourselves with a subset that has several marginal seas (Hudson Bay, Mediterranean, Baltic) and latitudes beyond ±66° masked out. For our analysis, we chose the time span from 2007 to 2009, which contained minimum fluctuations in observation quality and little orbit decay. A 3-year period is practicable in terms of computational demands (Section 3) yet sufficiently long to make robust assertions about submonthly variability.

The handling of the two GRACE products is as follows. We convert each daily ITSG2018 Kalman estimate from its spherical harmonic representation to mass anomalies on a 1° grid and add back the ocean background model (i.e., the GAD product, see page 61 in Dobslaw et al., 2017b), also provided as daily means up to $n = 40$ by ITSG. We omit, and for CSR swath explicitly remove, the GAD degree-1 term, as GRACE cannot see geocenter variations and our comparison with ocean model output is performed in the center-of-mass frame. No adjustments are made for the zonal degree-2 coefficients (Bonin & Save, 2020; Eicker et al., 2020) and signatures of co- and postseismic deformations, which remain in the data record.

Each daily CSR swath field is resampled to a nominal 1° latitude-longitude grid and subsequently smoothed in space using a 1,000 km full-width half-maximum Gaussian filter on the sphere to suppress noise at the mascon grid-scale and conform length scales to what ITSG2018 can resolve. Unwanted blending of oceanic signal with hydrological variability on land (or the lack of it) is avoided by the use of a coastal buffer. Similar to Eicker et al. (2020), we re-synthesize a binary land-sea mask given in spherical harmonics up to degree $n = 40$ in the grid domain and mask out points below the 0.8 threshold. The buffer is further extended into the ocean where the hydrological background model of the ITSG process dynamics indicates standard deviations larger than 3 mm. Over the 3 years considered, the ITSG2018 series contain 15 days with less than half the average number of GRACE observations and 4 days of no observations at all. As these "gaps" amount to less than 0.4% of the entire time series and the Kalman filter may still carry forward information from adjacent days, we leave the affected epochs in the record. 11 days are missing in the CSR swath series, which we fill in using an iterative EOF (empirical orthogonal function) technique.

Ocean bottom pressure changes p_b are considered to contain contributions from both ocean dynamics and atmospheric surface pressure fluctuations p_a

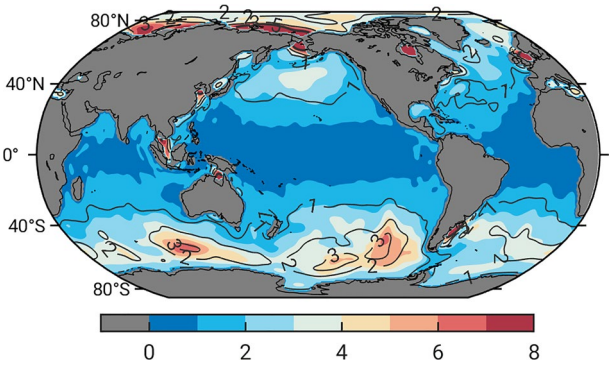

Figure 1. Standard deviation of ocean bottom pressure p_b (in cm of water height) from ITSG2018 over 2007–2009 for periods <60 days (filled contours) and <10 days (isolines at 1, 2, 3, and 5 cm). The background model is AOD1B RL06 without the degree-1 term. Points inside the coastal buffer are masked out; see the main text. AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; ITSG, Institute of Geodesy at Graz University of Technology.

$$p_b = p_a + g\int_{-H}^{0}\left(\rho - \rho_0\right)\mathrm{d}z + g\rho_0\eta \tag{1}$$

where z is the vertical coordinate (pointing upwards), H is the ocean depth, η represents the sea level anomaly with respect to a reference surface $z = 0$, g is the acceleration due to gravity, $\rho = \rho(z)$ is total density, and ρ_0 is a constant reference density. One can generally write η as the sum of a dynamical signal η' and an inverse barometer (IB) component, as defined in Ponte (1993). The first and third terms on the right-hand side of Equation 1 can then be combined to yield $g\rho_0\eta' + \overline{p}_a$, where $\overline{p}_a$ is the spatial mean of p_a over the global ocean (Wunsch & Stammer, 1997). The quantity $\overline{p}_a$ is a function of time (with a standard deviation of ~0.3 hPa on submonthly time scales) but induces no dynamically relevant horizontal gradients. Note that the adopted definition of p_b may be extended to include more subtle surface loading effects associated with, e.g., freshwater flux into the ocean (Ponte, 2006). For clarity, p_b has standard SI units (Pa) in Equation 1 but is quantified as equivalent water thickness in the text and figures.

The signal content of the ITSG2018 series at periods T shorter than 60 and 10 days is illustrated in Figure 1. Largest p_b amplitudes are found on the East Siberian Shelf (up to 14 cm), around the Bering Strait, and in numerous shallow shelf seas (e.g., Hudson Bay, North Sea, Yellow Sea, Gulf of Thailand), which lack vertical stratification and respond barotropically to the imposed atmospheric stresses (Fukumori et al., 2015; Vinogradova et al., 2007). A general intensification of p_b variance with increasing latitude is evident in the deep ocean. At $T < 60$ days, fluctuations of 4–7 cm in bottom pressure coincide with major abyssal plains in the Southern Ocean, such as the Bellingshausen, Australian-Antarctic, and Argentine basins. Confinement of these areas by closed contours of potential vorticity (cf. Figure 7 of Stepanov & Hughes, 2006) and the presence of similar patterns in satellite altimeter measurements (Fu, 2003) have led to the understanding that the enhanced intraseasonal variability in the Southern Ocean reflects resonant excitation of topographically trapped barotropic wave modes by wind stress curl (e.g., Weijer & Gille, 2005). Isolines in Figure 1 reveal an almost identical spatial distribution of p_b variance for periods below 10 days, albeit smaller by a factor of two. Comparison with modeling results (Hirose et al., 2001) suggests that the GRACE observations at these time scales manifest the ocean's dynamic response to both wind and pressure forcing.

A consistency check between ITSG2018 and CSR swath, filtered for spectral content above 0.5 cyc/month, is provided in Figure 2a in terms of a correspondence score r (e.g., Vinogradova et al., 2007), defined as

$$r = \frac{\mathrm{V}\left(p_b^{ITSG} - p_b^{CSR}\right)}{\mathrm{V}\left(p_b^{ITSG}\right)} \tag{2}$$

where p_b denotes filtered bottom pressure anomalies and V is the variance operator. Scores close to zero indicate that the two GRACE time series at a particular grid point are largely equal. Values of $r \leq 0.3$ are evident in both semienclosed and open-ocean regions that have an expected standard deviation in p_b of more than 3 cm on submonthly time scales. The agreement between the solutions is still pronounced ($r \sim 0.1$–0.5) when the signal threshold is lowered to 1 cm (dashed line in Figure 2a), but zones of weak correspondence emerge as one approaches the basins' margins, e.g., in the North Atlantic, the Tasmanian Sea, or over the Humboldt Current. Patches of high values of r also occur outside low-magnitude areas (such as the Bay of Bengal, Labrador Sea, and the seas around Japan), raising the question which GRACE series accounts for the mismatch. To that end, we have validated both ITSG2018 and CSR swath against in situ measurements from a quasiglobal collection of BPRs (Macrander et al., 2010), initially processed by Gebler (2013). The dataset was refined to 40 locations with more than 182 days of observations during 2007–2009, after manually correcting spurious drifts at the start and end of individual time series. We concatenated different deployments from the same BPR site, homogenized temporal sampling, reduced the AOD1B RL05 degree-1 term, and subtracted off estimated tidal variability.

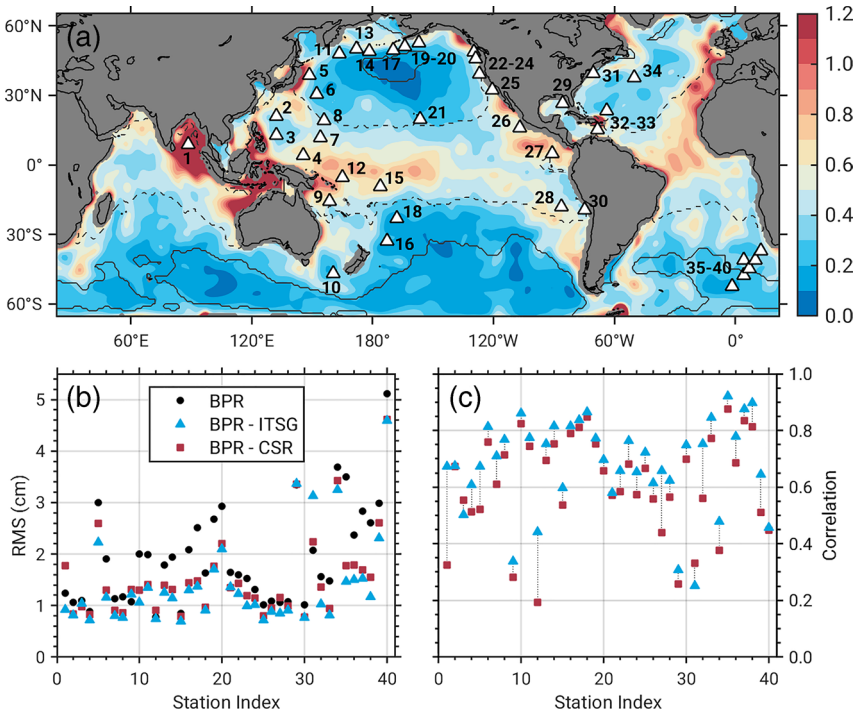

Figure 2. (a) Dimensionless correspondence score between p_b variations from ITSG2018 and CSR swath for periods shorter than 60 days, computed according to Equation 2. Dashed and solid isolines indicate standard deviations of 1 and 3 cm in ocean bottom pressure from AOD1B RL06; cf. Section 3. Panels (b and c) show RMS differences and correlations from the comparison of submonthly ITSG2018 (blue triangles) and CSR swath (red squares) mass anomalies with observations from 40 BPRs, plotted in panel (a) and labeled by increasing longitude. Both GRACE and the BPR series are corrected for variability associated with degree $n = 1$ over the ocean. AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; BPR, bottom pressure recorder; CSR, Center for Space Research; GRACE, Gravity Recovery and Climate Experiment; ITSG, Institute of Geodesy at Graz University of Technology; RMS, root mean square.

Figure 2b shows that the RMS (root mean square, or standard deviation) of the BPR series is in the range of 1–4 cm for periods below 60 days and typically drops by 0.3–1.5 cm if either ITSG2018 or CSR swath is removed from the in situ measurements (station 31 on the Northwest Atlantic Shelf being a notable exception). Largest RMS reductions (1–2 cm) are obtained for stations in the Southern Ocean (indices 10 and 35–40). Importantly, there is a systematic tendency for ITSG2018 to explain more variance in the BPR series than CSR swath, as is the case at 38 out of 40 locations by 0.15 cm water height on average. The median correlation with BPRs (Figure 2c) is 70% for ITSG2018 and 63% for CSR swath, prompting us to conclude that ITSG2018 is a more suitable dataset for validating numerical models. Statistics at BPRs in areas of poor correspondence scores (e.g., indices 1, 5, 12, and 27) suggest that the CSR series still suffer from spurious variability in regions of relative quietness in p_b. While not explicitly shown, we find that measures of similarity are markedly lower (e.g., correlation drops by 5%–20%) when ITSG2018 is replaced by the nontidal AOD1B background model, consistent with results in Poropat et al. (2019). Daily gravity field estimates from the Kalman Smoother approach thus provide genuine information on ocean mass variations not known a priori.

3. Ocean Models

We probe the high-frequency p_b content of three volume-conserving Boussinesq models and two barotropic (constant-density) models, forced with re-analyzed atmospheric data (ERA-Interim, Dee et al., 2011) in all but one case. Specifically, these models are (i) the oceanic component of AOD1B RL06 (Dobslaw et al., 2017a), (ii) Release 4 of the ECCO (Estimating the Circulation and Climate of the Ocean) version 4 global ocean state estimate (ECCO Consortium et al., 2020; Forget et al., 2015), (iii) an eddy-permitting

Table 1
Main Characteristics of the Participating Ocean Models[a]

	AOD1B	ECCOv4	LLC270	DAC	DEBOT
Horizontal resolution	1°	1/3–1°	1/3°	10–100 km	1/3°
Vertical layers	40	50	50	1	1
Atmospheric forcing[b]	ECMWF	ERA[c]	ERA	ERA	ERA
Provenance of wind stress[d]	Bulk	Model	Bulk	Bulk	Model
Parameterized wave drag	–	–	–	Yes	Yes
Sea-ice module	Yes	Yes	Yes	–	–
Cavities under ice shelves	Closed	Closed	Closed	Water	Water

Abbreviations: AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B; DAC, Dynamic Atmospheric Correction; DEBOT, David Einšpigel's Barotropic Ocean Tide model; ECCOv4, Estimating the Circulation and Climate of the Ocean version 4; ECMWF, European Centre for Medium-Range Weather Forecasts; ERA, re-analyzed atmospheric data; LLC, Lat-Lon-Cap.

[a]See the main text for references. [b]ECMWF: operational analysis by ECMWF, ERA: ERA-Interim reanalysis. Forcing frequency is 6-hourly except for AOD1B (3-hourly). All models include the effect of atmospheric pressure loading. [c]ECCOv4 is constrained to oceanographic data and integrated forward using adjusted atmospheric forcing. [d]Wind stress either computed through bulk formulae or adopted directly from atmospheric model output.

experiment with the MITgcm (Massachusetts Institute of Technology general circulation model) on a nominal 1/3° horizontal LLC (Lat-Lon-Cap) grid, (iv) the Dynamic Atmospheric Correction (DAC) derived from a MOG2D-G reanalysis (2D Gravity Waves model, Carrère et al., 2016), and (v) an adaption of DEBOT (David Einšpigel's Barotropic Ocean Tide model, Einšpigel & Martinec, 2017) for broadband atmospheric forcing (Schindelegger et al., 2017). The three barcoclinic models are henceforth abbreviated as AOD1B, ECCOv4, and LLC270, where the latter refers to a decomposition of the globe into 13 square tiles with 270 divisions per side. Both the LLC270 and DEBOT simulations were conducted in the frame of this study. The appendix gives a brief description of all five models, while Table 1 summarizes their general attributes, e.g., horizontal and vertical discretization, or the handling of air-sea momentum flux.

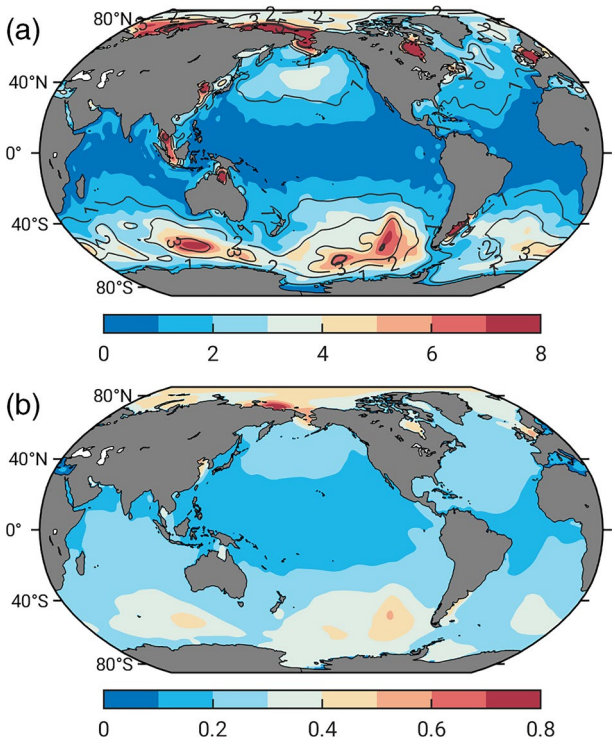

Figure 3. (a) Standard deviation of p_b (cm) from AOD1B RL06 (2007–2009) up to spherical harmonic degree $n = 40$ for periods <60 days (filled contours) and <10 days (isolines at 1, 2, 3, and 5 cm, the latter plotted with greater line width). (b) Standard deviation of the SAL-induced p_b signal (cm) in the 2–60-day band, deduced from two DEBOT simulations with and without the SAL term. AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; DEBOT, David Einšpigel's Barotropic Ocean Tide model; SAL, self-attraction and loading.

Some peculiarities are worth pointing out. None of the models is constrained to oceanographic data except for ECCOv4, which represents an iterative fit of the MITgcm to in situ hydrography and satellite data using weighted least squares minimization. The optimization is realized by variations to uncertain input parameters, including forcing fields. Because the fitting procedure corrects atmospheric forcing at 14-day intervals and the so-derived adjustments are interpolated in between (ECCO Consortium et al., 2020), submonthly p_b variability in ECCOv4 is largely controlled by the forward model itself. Both DAC and DEBOT include the water bodies under Antarctic ice shelves in their respective ocean domains and account for barotropic-baroclinic energy conversion by means of a topographic wave drag scheme (Carrère & Lyard, 2003), see the appendix for details. A unique feature in DEBOT is the dynamic implementation of ocean self-attraction and loading (SAL) effects, achieved by decomposing bottom pressure anomalies at each time step into spherical harmonics and applying a degree-dependent load Love number factor to estimate the additional body force in the momentum equations (Vinogradova et al., 2015). Figure 3b shows the standard deviation of the SAL signal in p_b across the 2–60-day band, calculated from two identical DEBOT integrations with the SAL physics turned off in one of the runs. The extracted variability is consistent with Figure 4 of Vinogradova et al. (2015), though slightly attenuated owing to the temporal filtering applied to the DEBOT

output. For most of the global ocean, the SAL effect is less than 0.3–0.4 cm on average, but increased values of ∼0.5 cm and frequent transient peaks exceeding 1 cm are found in the Bellingshausen Basin and the Arctic Ocean. Omission of these bottom pressure signals due to the lack of a SAL module is a known (albeit small) source of error in AOD1B.

In all model output, periodic signals in bottom pressure driven by diurnal and semidiurnal atmospheric tides were either removed during postprocessing (Dobslaw et al., 2017b) or suppressed by filtering barometric pressure fields prior to the simulations (e.g., Carrère et al., 2016). Daily means of p_b were obtained directly in the case of ECCOv4 but otherwise derived from averaging higher-frequency model output to a center time of 12 UTC. To conform varying spatial resolutions, each model-based daily p_b field was expanded harmonically to a maximum degree consistent with the daily GRACE series (40 in the case of ITSG2018, 180 for CSR swath combined with 1,000 km smoothing) and projected back to a 1° grid. We used SHTOOLS (Wieczorek & Meschede, 2018) to perform all offline spherical harmonic transformations. Figure 3a depicts the standard deviation of submonthly bottom pressure variability in AOD1B, which reveals a distinct similarity with the corresponding ITSG2018 map in Figure 1. Yet, p_b amplitudes in the Southern Ocean are systematically higher in AOD1B, suggesting that the Kalman Smoother acts to dampen energetic, basin-wide features in the background model. We will expand on this early insight through a more explicit analysis below.

4. Comparisons With GRACE

Figure 4 presents, in a very condensed form, the model-to-GRACE comparison at periods below 60 days. In particular, we show RMS differences and percentage variances explained for all ocean models except DAC, which assumes a purely static (IB) ocean at low frequencies (20 days cut-off, A4). A common feature across all models is their high percentage (70%–80%) of ITSG2018 variance explained over the Arctic Ocean and Russian shelf seas, although RMS residuals remain large and exceed 5 cm on the East Siberian Shelf. Lack of a second GRACE series for polar latitudes precludes a more detailed interpretation of these disparities, but errors in bathymetry and atmospheric forcing are likely playing a role. We have also carried out comparisons of the different p_b fields to gridded 3-day sea-level anomalies from ENVISAT north of 52°N (not shown) and found larger RMS differences in AOD1B relative to other models in a few locations (Kara Sea, North Sea, Hudson Bay), consistent with Figure 4. Likewise, the fidelity of AOD1B in the Mediterranean Sea is questionable from both our comparison to ITSG2018 (∼3 cm RMS, 40% explained variance) and the altimetry-based evaluation by Bonin and Save (2020).

In latitudes bounded by ±66° and for long wavelengths, the two GRACE products suggest that AOD1B is the most accurate model in the North Pacific and North Atlantic, while DEBOT produces highest explained variances (close to 90%) and smallest RMS differences in the Southern Ocean. We observe a degradation of DEBOT in terms of explained variances toward lower latitudes, which is an indication of excessive drag due to latitude-dependence in the adopted buoyancy frequencies (Equation A1). ITSG2018 and CSR swath are in accord that ECCOv4 misrepresents (in fact, overestimates) the topographically trapped barotropic variability in the abyssal plains of the Southern Ocean. Similar deficits are present in other models with limited horizontal and vertical resolution (AOD1B) and blocked ice-shelf cavities (e.g., LLC270). These observations imply that a proper representation of bathymetry is important to capture the region's modal variations (cf. Bonin & Save, 2020). Nonetheless, RMS differences in ECCOv4 of up to 4 cm stand out and may also reflect a lack of dissipation. When differences in grid spacings are accounted for, dampening by horizontal eddy viscosity in LLC270 is 50% higher than in ECCOv4 and alleviates some of the latter's energetic features. A spatially more confined signal is bottom pressure variability in the Argentine Basin, dominated by a 20–25 days anticyclonic oscillation of ∼5 cm RMS amplitude around the Zapiola Rise (45°S, 43°W). The mode is present in daily ITSG gravity time series (Yu et al., 2018) but only emerges in numerical models which generate a realistic mesoscale field (i.e., none of the models examined here).

From the analyses shown so far, no ocean model stands out with particularly close correspondence to the GRACE data. Consideration of p_b variability at periods shorter than 10 days is more illuminating. Comparisons with ITSG2018 (Figure 5) and CSR swath (Figure 6) suggest that at such rapid time scales the two barotropic models (DEBOT and DAC) provide a systematically better fit to the daily GRACE solutions

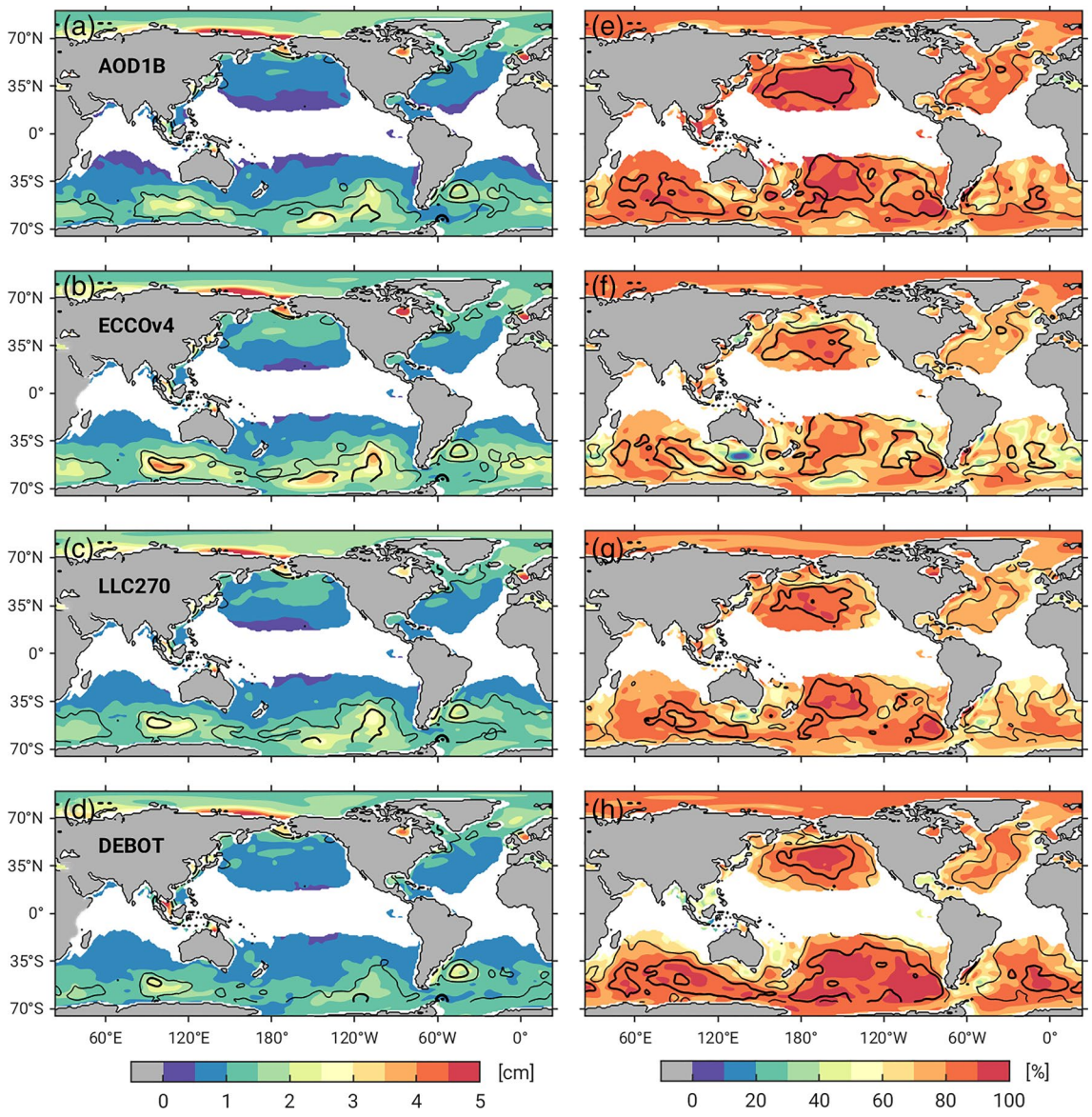

Figure 4. Comparison of four ocean models against daily bottom pressure estimates from ITSG2018 and CSR swath at periods <60 days during 2007–2009. Evaluation metrics are (a–d) RMS differences in cm of water height, and (e–h) percentage of GRACE variance explained by the models. Colored contours show the statistics from the comparison to ITSG2018, whereas the overlain black lines are for CSR swath at levels 1.5 cm (thin lines) and 2.5 cm (thick lines) in (a–d) and levels 60% and 80% in (e–h). Along with points in the coastal buffer (cf. Section 2), regions where AOD1B indicates standard deviations in p_b smaller than 1 cm are masked out. AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; CSR, Center for Space Research; GRACE, Gravity Recovery and Climate Experiment; ITSG, Institute of Geodesy at Graz University of Technology; RMS, root mean square.

than the present de-aliasing standard or LLC270. Focusing on deep areas covered by both GRACE series, we find that DEBOT and DAC have maximum RMS values of 1.3 cm w.r.t. ITSG2018 and 1.6 cm relative to CSR swath. Area-averaged RMS differences are well below 1 cm of water height and explained variances consistently above 70%. Such tight agreement must be close to the noise floor inherent to our analysis, governed by uncertainties in atmospheric forcing data, modeling choices, drag parameterizations, and noise in the GRACE series themselves. More to the point, the consistency is reassuring in that both GRACE solutions converge to numerical models that have not been involved in the daily parameter estimation (To be exact, DAC does have a minor entanglement with ITSG2018 as it is one out of three models used in the computation of the stochastic Kalman Smoother constraints, Dobslaw et al., 2016). In other words, the daily mascon and Kalman Smoother updates impose realistic corrections onto AOD1B RL05/RL06 and enforce a large-scale agreement with models of shallow-water dynamics. These corrections appear as RMS "hot

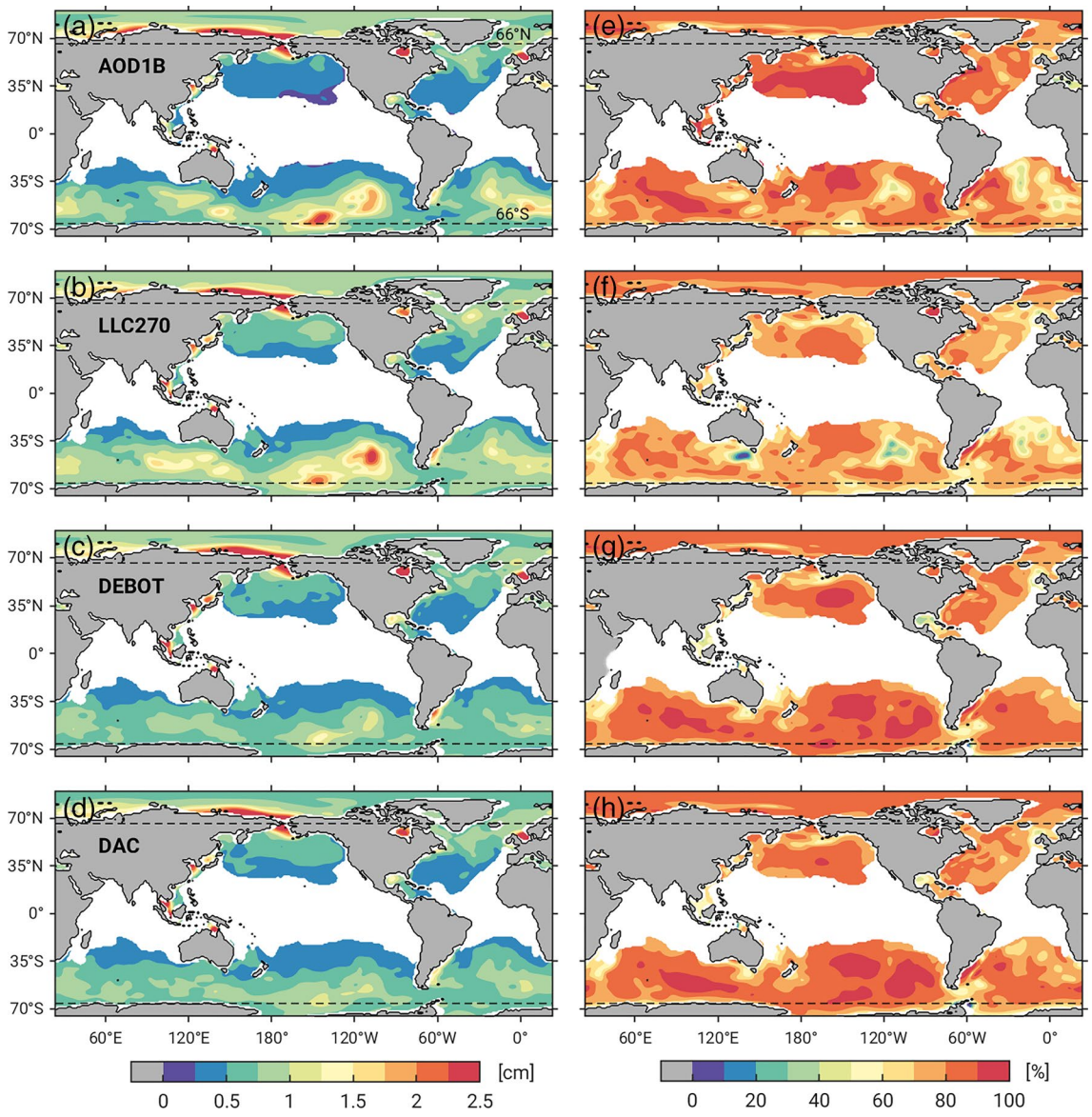

Figure 5. Comparison of numerical models against ITSG2018-based p_b fields as in Figure 4 but for periods <10 days and a different model selection (DAC included, ECCOv4 left out). Corresponding results for CSR swath are shown separately in Figure 6. Left panels depict RMS differences, right panels the percent of variance explained. Regions where AOD1B indicates standard deviations in p_b smaller than 0.8 cm are masked out. AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; CSR, Center for Space Research; DAC, Dynamic Atmospheric Correction; ECCOv4, Estimating the Circulation and Climate of the Ocean version 4; ITSG, Institute of Geodesy at Graz University of Technology; RMS, root mean square.

spots" of ~2 cm for AOD1B (Figures 5a and 6a) in the Bellingshausen Basin (45–55°S, 110°W), the Ross Gyre (65°S, 145°W), the Enderby Abyssal Plain (55°S, 10°E), and to a lesser extent the Australia–Antarctic Basin.

We observe no such pleasing accord in shallow and semienclosed regions, as expected from the low local correspondence scores between ITSG2018 and CSR swath in Figure 2a. A solitary common feature is the relatively poor performance of DEBOT in the Gulf of Mexico and the South China Sea. Elsewhere, the CSR swath series appear to be marred by systematic effects that produce similar patterns of evaluation metrics irrespective of the numerical model, for example, in the Japan Sea, the Gulf of Carpentaria, and the Labrador Sea. Spurious p_b variability is also evident in the Argentine and Bellingshausen Basins (close to 120°W), consistent with Figure 4 of Bonin and Save (2020).

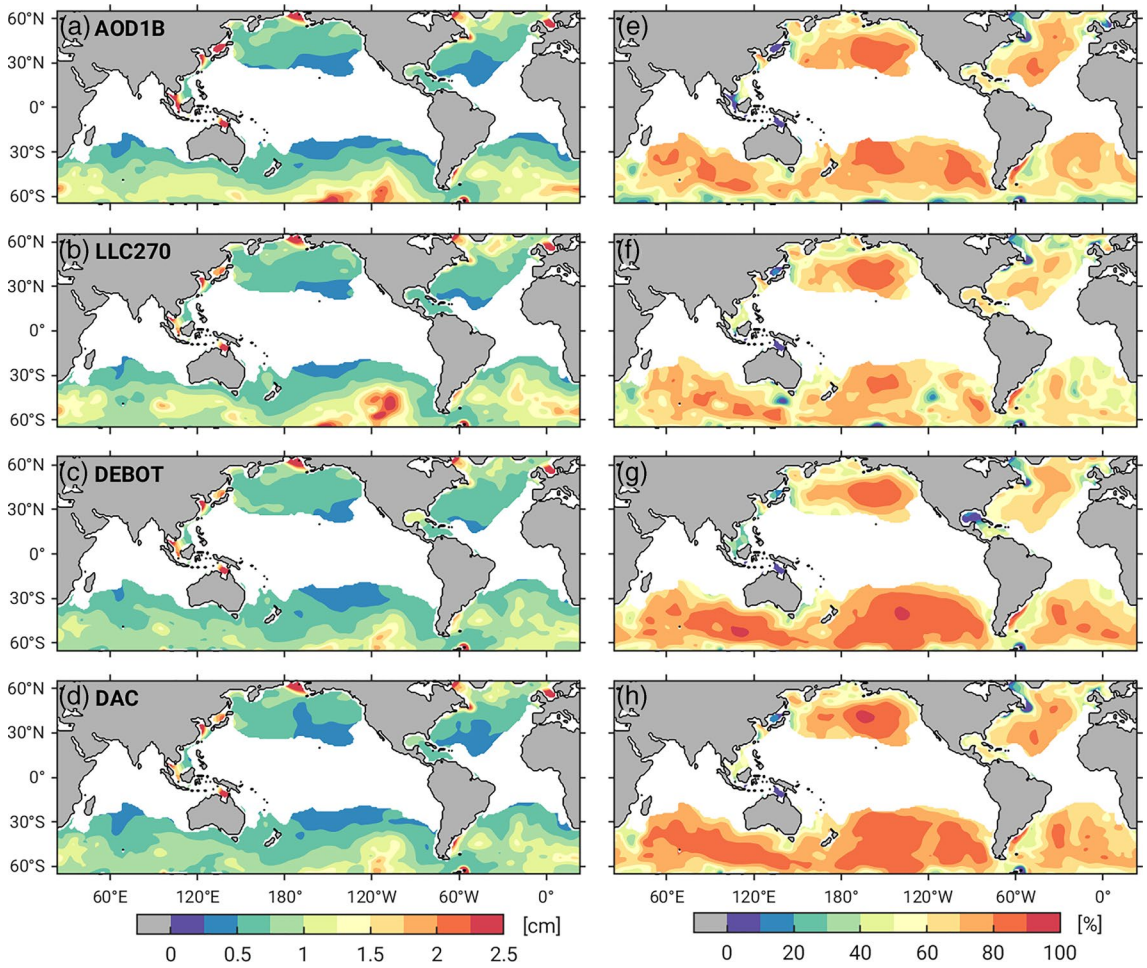

Figure 6. Comparison of four ocean models against daily GRACE series as in Figure 5 ($T < 10$ days) but using the CSR swath bottom pressure fields, smoothed to wavelengths of 1,000 km and longer. CSR, Center for Space Research; GRACE, Gravity Recovery and Climate Experiment.

Figure 7 provides a compact spectral view of the above findings. We computed domain-averaged RMS differences between daily GRACE fields (global for ITSG2018, ±66° for CSR swath) and each ocean model from the respective Fourier coefficients $\hat{p}_b = \hat{p}_b(\sigma)$, as standard in the tidal literature,

$$\overline{\Delta p_b}(\sigma) = \sqrt{\frac{\iint \left| \hat{p}_b^{GRACE} - \hat{p}_b^{model} \right|_\sigma^2 \, \mathrm{d}A}{2 \iint \mathrm{d}A}} \tag{3}$$

where σ is frequency and dA is the area element. Points inside the coastal buffer and of little signal content ($\sqrt{\mathrm{V}(p_b)} < 1$ cm at $T < 60$ days) were excluded from the spatial average. The number of daily available samples (1,096) implies that the spectral resolution is $\sim 10^{-3}$ cyc/day. For visualization purposes, we grouped the RMS error estimates from these original frequency bands to a coarser sampling of 10^{-2} cyc/day, using summation under variance propagation rules. The resulting RMS curve over periods is raised relative to its baseline version at higher spectral resolution, as exemplified for AOD1B in Figure 7 (thin and thick black lines). When condensed to a single value for all periods below 60 days, the global RMS difference of AOD1B with ITSG2018 (CSR swath) is 1.22 cm (1.35 cm), identical to what area-averaging results in Figure 4a would give.

The presence of systematic errors in CSR swath and its nonglobal extent make it difficult to fully reconcile results from Figures 7a and 7b, see, for example, the difference in RMS curves for ECCOv4 on weekly time scales. Nonetheless, the analysis confirms that p_b fields from DEBOT and DAC are closest to the daily GRACE

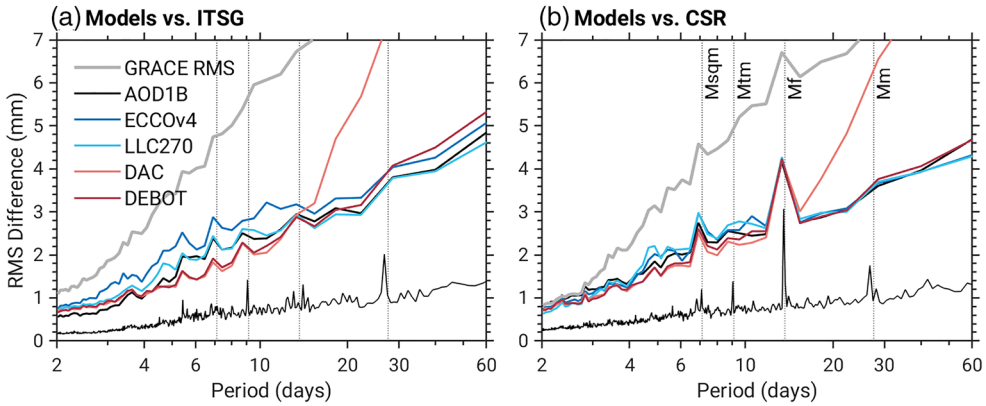

Figure 7. Globally averaged RMS differences in bottom pressure (mm) between five ocean models and (a) ITSG2018 and (b) CSR swath, respectively. Gray thick lines denote RMS amplitudes of the GRACE series before subtracting $\hat{p}_b^{model}$ (Equation 3). Results are plotted over period in intervals corresponding to a frequency resolution of 10^{-2} cyc/day, but the thin solid black line shows the RMS differences for AOD1B at the original spectral resolution of $\sim 10^{-3}$ cyc/day. Dashed vertical lines indicate submonthly tidal constituents with periods 27.55 (Mm), 13.66 (Mf), 9.13 (Mtm), and 7.10 days (Msqm), taken from Hartmann and Wenzel (1995). AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; CSR, Center for Space Research; GRACE, Gravity Recovery and Climate Experiment; ITSG, Institute of Geodesy at Graz University of Technology; RMS, root mean square.

solutions at shortest periods up to 12 days. DAC starts to diverge from all other models at $T \sim 16$ days, as dynamic p_b variability is gradually suppressed (Appendix A4). Statistics relative to ITSG2018 tend to favor baroclinic modeling frameworks with increasing period, as expected from assertions in Vinogradova et al. (2007). A salient feature in the CSR swath spectrum (Figure 7b) are sharp excess RMS values at, or very close to, major long-period tidal constituents, in particular at the fort-nightly frequency Mf (13.66 days). We have harmonically analyzed the swath series and traced back the 3 mm global RMS peak at Mf to anomalous tidal signals in the Gulf of Carpentaria (3–4 cm), the Gulf of Thailand (2 cm) and the Bellingshausen Basin (1 cm)—all of which are regions where Mf departs from its prevalent zonal structure (e.g., Ray & Erofeeva, 2014). The ITSG2018 processing allows for GRACE-driven corrections of selected constituents from tide model priors in the parameter estimation (Kvas et al., 2019). Indeed, no gross artifact is present at the Mf frequency in Figure 7a, but the enhanced RMS values at Mm (27.55 days) and Mtm (9.13 days) are comparable to those in CSR swath. Mitigating residual mass signals associated with long-period tidal constituents may therefore contribute to improved daily gravity field solutions.

Table 2
Comparison of p_b Signals From Various Sources at 26 BPR Locations in Two High-Frequency Bands[a,b]

	2–4 days		4–12 days	
	RMS	PVE	RMS	PVE
ITSG2018	3.6 (–)	45 (–)	6.5 (–)	62 (–)
AOD1B	4.2 (2.4)	27 (82)	7.2 (7.2)	54 (81)
DEBOT	3.4 (3.3)	50 (62)	6.5 (4.9)	62 (78)
DAC	3.7 (3.3)	41 (61)	6.3 (4.7)	64 (80)
Signal[c]	5.1 (5.6)		10.9 (10.8)	

Abbreviations: AOD1B RL06, Atmosphere and Ocean De-Aliasing Level-1B Release 06; BPR, bottom pressure recorder; DAC, Dynamic Atmospheric Correction; DEBOT, David Einšpigel's Barotropic Ocean Tide model; PVE, percentage variance explained; RMS, root mean square.
[a]Tabulated are RMS differences (mm) between BPR signals and gridded datasets of p_b (models and ITSG2018), and the associated PVE. Values in brackets show the statistics when the BPR measurements are replaced by a local evaluation of ITSG2018.
[b]BPRs considered are a subset of those shown in Figure 2a, with site 31 (anomalous) and all locations in low-magnitude areas ($\sqrt{\mathrm{V}(p_b)} < 1$ cm) excluded. [c]Average of the RMS values computed from the filtered BPR time series.

At periods shorter than 4 days, RMS differences between ITSG2018 and AOD1B become suspiciously low, implicating a bias of the daily Kalman Smoother solutions toward their de-aliasing prior. When the evaluation metrics are mapped in space, akin to Figure 5, AOD1B produces explained variances of ~90% even in middle latitudes of the main ocean basins, where p_b drops to standard deviations of 3 mm at $T < 4$ days. We deem such consistency for low-magnitude signals implausible. In fact, no other model provides explained variances in excess of 50% in these areas. Supplemental checks against most of the BPRs already used in Section 2 corroborate our conjecture; cf. Table 2. In the 2–4 days band, RMS differences between AOD1B and BPR observations are 4.2 mm on average, higher than the corresponding statistics for DEBOT or DAC (and ITSG2018). However, when the ITSG2018 series are adopted as "observations" at the BPR sites, the RMS difference with AOD1B drops to 2.4 mm and explained variances increase from 27% to 82%. Results for all other ocean models are far less sensitive to this change of ground truth. Hence, although our BPR database is sparse, we conclude that at the shortest periods ITSG2018 tends to fall back to AOD1B and the Kalman

Smoother correction imposes little realistic oceanic mass variability. Updates to the gravity field from actual GRACE observations may only be expected once a full sampling of the globe is completed (every 4–5 days); cf. Eicker et al. (2020). A repeat of the assessment with BPRs for the 4–12-day band in Table 2 suggests that (i) the ITSG2018 bias becomes indeed less severe with decreasing frequency and (ii) the skill of the barotropic models in this frequency band (Figure 7a) is also borne out by in situ bottom pressure observations.

5. Quantifying GRACE Errors

An important question left out so far is whether the various p_b datasets available to us allow for a spatial mapping of the uncertainties in the GRACE-derived submonthly bottom pressure changes. Such error quantification would provide context to the model-data differences seen in the previous section and contribute toward formulating realistic weights for the use of daily GRACE fields in ocean state estimation (ECCO Consortium et al., 2020) or particular types of sequential data assimilation (e.g., Saynisch et al., 2015). In seeking approximate, first-order estimates of these errors, we follow the recipe laid out by Quinn and Ponte (2008) and compare either GRACE product (ITSG2018, CSR swath) to p_b output from an independent ocean model. Ideally, the GRACE (or "data") error would be made up by instrument noise and deficiencies in the background models and regularization constraints. However, as we compare with simulation results, it will also contain the ocean model's representation error—i.e., the part of the observed signal not captured by the model due to physical simplifications, unresolved processes, or imperfect boundary conditions. We specifically employ DEBOT, as the above analysis suggests that its representation error is smaller than those of other models.

Let d and m be the GRACE and ocean model estimates of the true p_b signal at a grid point. Under the restrictive assumption that correlations among the signal, model errors (m'), and data errors (d') are absent (Quinn & Ponte, 2008), the variance of d' may be computed from

$$\mathrm{V}\left(d'\right) \approx \mathrm{V}\left(d\right) - \mathrm{W}\left(d, m\right) \tag{4}$$

where W is the covariance operator. Negative values of $\mathrm{V}\left(d'\right)$ are rare and eventually avoided by setting the data error variance to 10% of $\mathrm{V}\left(d - m\right)$ where it is smaller than that threshold. An estimate for the noise-to-signal ratio (NSR) is $\mathrm{V}\left(d'\right) / \mathrm{V}\left(d\right)$, the square root of which gives the data error relative to the signal RMS (e.g., Figure 1).

Figures 8 and 9 depict the stochastic fields from this processing scheme for ITSG2018 and CSR swath at periods below 60 and 10 days. Standard deviations at $T < 60$ days (panels a) are generally 0.5–1.5 cm over energetic basin interiors, implying a relative error of 30%–45% (NSR ~ 0.10–0.20). A tendency toward larger absolute errors (~3 cm) but low NSR is seen for shallow areas which accommodate high p_b variability, for example, the Arctic Ocean, the North Sea, and Hudson Bay. The spatial structure of the CSR swath error in the Southern Ocean (Figure 9a) is more heterogeneous than for ITSG2018, possibly reflecting the local nature of the mascon approach compared to globally fitted spherical harmonics. High noise levels in the CSR swath solution in most coastal regions and away from areas of enhanced p_b variability endorse our assessment of these data in the foregoing sections.

Data errors at $T < 10$ days (panels c) are ~40% smaller than those across the entire submonthly band and amount to less than 1.2 cm for much of the deep ocean. No major qualitative difference is found when comparing the NSR grids between the two spectral windows (panels b vs. d), indicating that the noise in the two GRACE series is largely proportional to the signal across the tested frequency bands. We have repeated our stochastic assessment of the GRACE-derived p_b fields using the second independent ocean model on hand (LLC270) and obtained very similar results to what is shown in Figures 8 and 9. Deviations from the DEBOT-based maps (e.g., an increase of the suggested data error by 1 cm in the Arctic and 0.4 cm in the Southern Ocean) are repercussions of LLC270's larger representation error. By and large, the plotted regional error estimates are comparable in size to the model-data differences shown for DEBOT in Figures 4–6, suggesting that the latter can be generally interpreted as data error. Importantly, the standard deviations derived for the two GRACE products at submonthly periods are sufficiently small to lend credence to our finding of anomalous p_b signals in baroclinic ocean models over the abyssal plains of the Southern Ocean (Section 4).

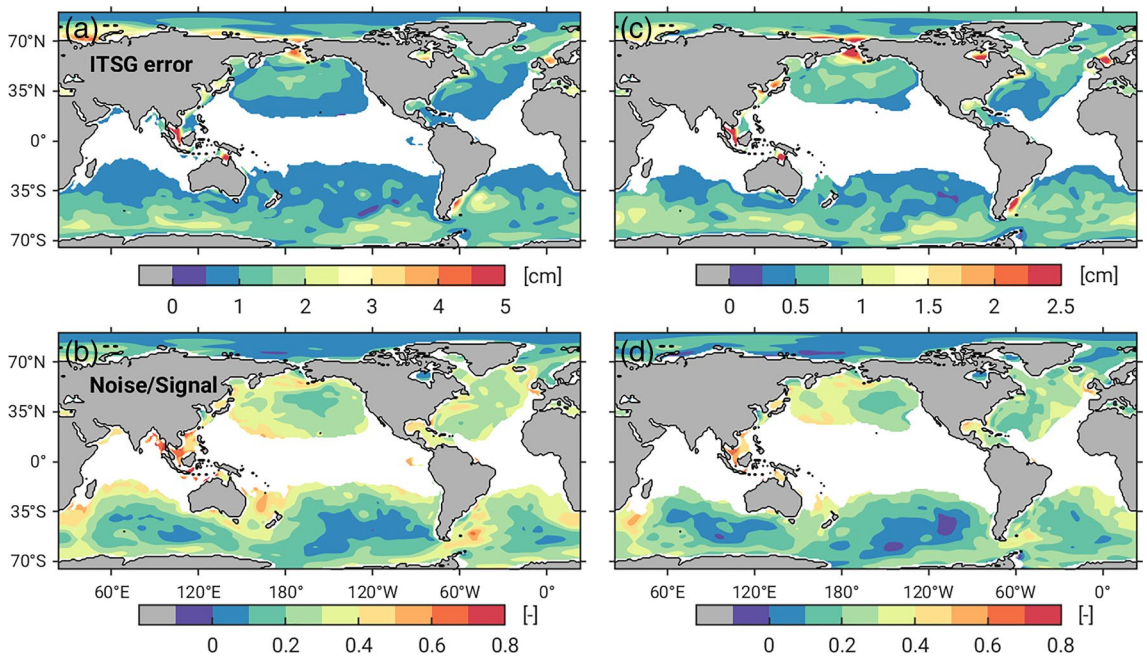

Figure 8. (a and c) Standard deviation of ITSG2018 data errors (in cm) and (b and d) ITSG2018 noise-to-signal ratios $\mathrm{V}(d')$ / $\mathrm{V}(d)$, as deduced from comparisons with p_b fields from DEBOT during 2007–2009. Results are shown for periods <60 days (left column, masking at $\sqrt{\mathrm{V}(p_b)}$ < 1 cm) and <10 days (right column, $\sqrt{\mathrm{V}(p_b)}$ < 0.8 cm). DEBOT, David Einšpigel's Barotropic Ocean Tide model; ITSG, Institute of Geodesy at Graz University of Technology.

6. Summary Remarks and Outlook

We have carried out an extended assessment of high-frequency, nontidal mass changes in a collection of ocean models, ranging from de-aliasing standards (AOD1B, DAC) to forward models involved in ocean state estimation efforts (ECCOv4, LLC270). None of the tested models is free from error, and questionable p_b structures persist in community standards such as AOD1B and ECCOv4, albeit with noticeably smaller magnitudes than in their predecessors (cf. Bonin & Chambers, 2011; Quinn & Ponte, 2011). Preparations

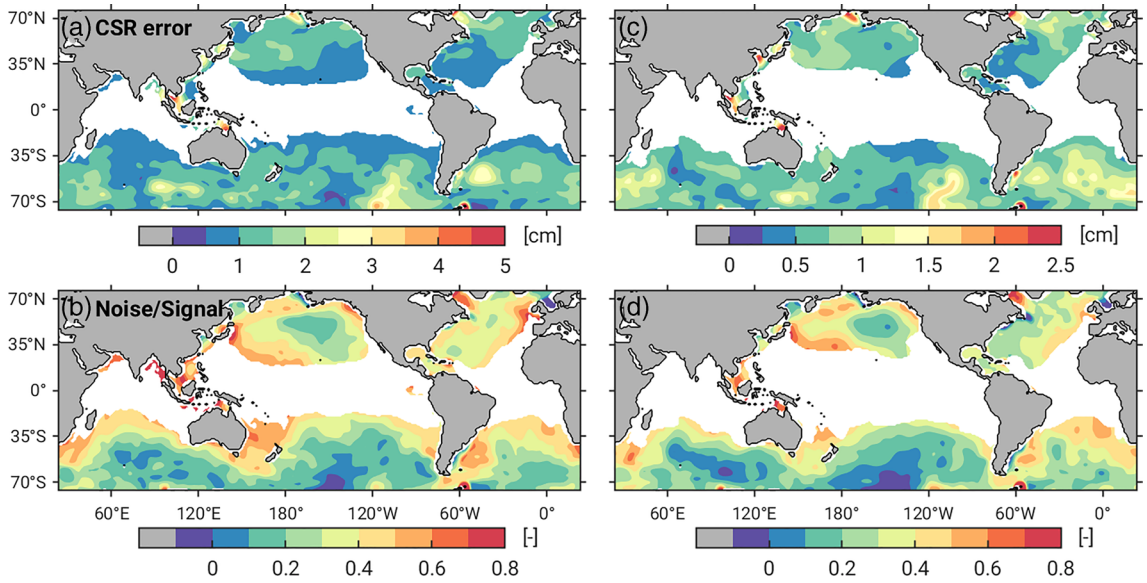

Figure 9. As Figure 8 but for CSR swath. CSR, Center for Space Research.

for new releases of these models, particularly AOD1B, are well along and will mitigate some of the better known model deficiencies related, for example, to the representation of bottom topography and Antarctic ice-shelf cavities. Judging the gains from these and other developments will require external validations, which in the past have relied on BPRs, sea-level observations, and GRACE *K*-band residuals sensitive to all (i.e., also atmospheric) surface mass changes (e.g., Dobslaw et al., 2017a). The key message of our study is that daily GRACE solutions, while not fully independent of their de-aliasing priors, can guide improvements in modeling bottom pressure at large spatial scales, down to RMS differences of 1–2 cm at periods below 60 days.

Most of our analysis has been diagnostic in character, leaving inferences about the origin of particular model errors for further research. Investigations along these lines will be required to disentangle sensitivities of rapid p_b signals to water depths, frictional dampening, and details in the atmospheric forcing (Hirose et al., 2001; Weijer et al., 2009). A relatively clean test case we have started to address, using both numerical models and GRACE, is the 4–6 days mode in basin-wide bottom pressure that manifests a nonequilibrium ocean response to a Rossby-Haurwitz wave in barometric pressure (Ponte, 1997; Stepanov & Hughes, 2006). Because the forcing of this oscillation is well known, mapping its spatial structure can shed light on the relative importance of bottom topography and dissipation for p_b fluctuations at subweekly periods in general. To what extent errors in modeled p_b fields relate to uncertainties in atmospheric forcing (primarily wind stress but also pressure, e.g., Salstein et al., 2008) is a question that pertains to the entire submonthly band. In upcoming work, we plan to address this issue by repeated numerical simulations, using an invariant ocean model configuration and atmospheric forcing taken from raw members of an ensemble variational data assimilation system (Hersbach et al., 2020).

The fidelity with which DEBOT represents submonthly p_b variability over most of the ocean is a pleasing and useful result. In a GRACE context, it suggests that computationally cheap barotropic models are credible tools for assessing sensor performance and de-aliasing procedures (Flechtner et al., 2016), and for deriving process statistics that stabilize the estimation of daily gravity fields. That said, we expect AOD1B endeavors to proceed with a baroclinic framework and gradually refine the representation of rapid p_b signals in support of GRACE. Central targets for such improvements are dynamically active marginal seas, which would benefit from increases in spatial resolution through the use of unstructured model meshes (e.g., Wang et al., 2014). Considerable scope remains for a more elaborate treatment of dissipation effects. Dampening in coarse-resolution baroclinic models is usually conveyed by constant (or grid-scale dependent) parameters for quadratic bottom friction and horizontal eddy viscosity. However, optimal levels of dissipation are known to vary with location and bottom topography (e.g., Hirose et al., 2001). In fact, centering drag over seafloor gradients has proven expedient in barotropic models and similar choices may be pursued in 3D. Forward modeling of rapid oceanic mass variations will ultimately profit from extensions of ongoing ocean synthesis efforts to higher frequencies. ECCO-type parameter estimation techniques in particular may help to improve dissipation schemes and provide realistic adjustments to atmospheric forcing fields.

Appendix A: Description of Ocean Models

A1 AOD1B: The oceanic component of AOD1B RL06 is based on an unconstrained forward simulation with the Max-Planck-Institute for Meteorology Ocean Model (MPIOM; Jungclaus et al., 2013), discretized on a ∼1° horizontal tripolar grid and 40 vertical layers. The model is forced with 3-hourly atmospheric data from ECMWF operational analysis starting in 2007 and ERA-interim reanalysis data prior to that. The oceanic response to atmospheric tides at 12 partial lines has been empirically estimated and removed from the simulated bottom pressure prior to the calculation of Stokes coefficients. In keeping with Equation 1, we use the GAD (OBA) product, which includes the static atmospheric contribution $\overline{p}_a$ to ocean bottom pressure and is given in spherical harmonics up to degree and order 180. We resynthesize model mass anomalies on a 1° grid, upon multiplying the GAD Stokes coefficients with the proper combination of Load Love numbers (Dobslaw et al., 2017b).

A2 ECCOv4: Global ocean state estimates from the ECCO Consortium (Wunsch et al., 2009) are physically consistent, iterated least squares fits of a general circulation model to a large volume of oceanographic data

(Argo, altimetry, monthly GRACE gravity fields, etc. in the case of version 4). The solution minimizes a cost function that penalizes the time-evolving model-data misfits and a set of control variables comprising initial conditions, mixing coefficients, and atmospheric forcing data. The adjusted parameters and forcing fields are used to drive the final estimate, which is a free forward integration of the MITgcm (Marshall et al., 1997). As such, ECCOv4 conserves ocean properties and mass (volume) to machine precision. The model's native grid is referred to as LLC90 (Forget et al., 2015), featuring a regular longitude-latitude grid south of 57°N plus a Cartesian cap in the Arctic. It provides a nominal resolution of 1° in the horizontal, telescoping to 1/3° in the meridional direction near the equator. Here we extract daily averaged p_b fields (variable OBP) over 2007–2009 from the full bidecadal state estimate.

A3 MITgcm LLC270 experiment: The setup for our eddy-permitting simulation is that described by Zhang et al. (2018) as part of the ongoing effort to transition the ECCO ocean and sea-ice estimates to the LLC270 grid (270 grid points along one-quarter of the Earth's circumference at the equator). For our purposes, all ECCO capabilities were switched off and an unconstrained forward integration of the MITgcm was performed using unadjusted atmospheric forcing (including atmospheric pressure). The simulation was started from a state of rest at July 1, 2006, with initial hydrography set to an annual mean climatology of potential temperature and salinity constructed from an early ECCO LLC270 iteration. A short spin up of only half a year is sufficient, given our interest in high-frequency, mostly barotropic variability. Time step (20 min), bathymetry, eddy viscosities, and other parameter choices of the simulation are identical to the Zhang et al. (2018) setup. As in the bottom pressure products furnished by AOD1B and ECCOv4, the MITgcm's diagnostic p_b fields must be corrected for spurious mass fluxes associated with the Boussinesq approximation in the governing equations. Accordingly, we subtract off the global area-averaged pressure at each 6-hourly output interval to enforce mass conservation (Greatbatch, 1994) and add back the static atmospheric contribution.

A4 DAC is the primary de-aliasing product for sea-level measurements from reference altimeter missions. It is released as a combination of an IB response to atmospheric pressure for periods longer than 20 days and a modeled barotropic response to sea-level pressure and 10 m winds at higher frequencies. The associated nonlinear shallow-water model (MOG2D-G, Carrère & Lyard, 2003) uses a finite element approach to increase resolution toward the coast and over major bathymetric features. Energy losses to topographically generated internal waves are considered in parameterized form

$$F_w = -CL\bar{N}(\nabla H)^2 \boldsymbol{u} \tag{A1}$$

where $\bar{N}$ is the depth-averaged buoyancy frequency, H denotes water depth, $\boldsymbol{u}$ represents the barotropic velocity vector, and L is a length scale estimate of $\mathcal{O}(10\ \mathrm{km})$ that can be tuned with a multiplicative factor. We use the global DAC sea-level grids from a dedicated MOG2D-G simulation based on ERA-Interim forcing (Carrère et al., 2016) and extract ocean bottom pressure fields by adding $p_a/\rho g$ at each ocean grid point.

A5 DEBOT: Dedicated barotropic model simulations were performed with a DEBOT configuration on a 1/3° latitude-longitude grid extending from 86°S to 87°N. We use the MOG2D-G topographic wave drag scheme (Equation A1), set the length scale to $L = 10$ km and apply a scaling factor of $C = 7.0$ based on experience with the model in varying discretizations. For output purposes and inside the SAL part of the code, the artificial landmass north of 87°N is filled by linear interpolation in terms bottom pressure. Although we are actively working on a more sophisticated way to address the singularity at the North pole, the present fix may still lead to a reasonably good representation of the intraseasonal Arctic bottom pressure variability given the latter's spatially coherent character and excitation by meridional winds in lower latitudes (Fukumori et al., 2015). A field approach is used to incorporate the ocean's response to surface loading by dynamic bottom pressure. At each time step, anomalies of p_b, after subtraction of a spatial mean, are expanded into spherical harmonics to degree 179 and multiplied with the wavenumber representation of the proper SAL kernel (Stepanov & Hughes, 2004; Vinogradova et al., 2015). The resulting perturbation force in the grid space is then added to the pressure gradient term in the forward step of the momentum equations. Additional SAL effects on ocean masses associated with, e.g., atmospheric pressure variations over land are currently not taken into account.

Data Availability Statement

Daily bottom pressure anomalies (1°, 2007–2009) from the DEBOT and LLC270 simulations have been placed at https://doi.org/10.5281/zenodo.4293426. All other datasets used in this study are available from the following links: ITSG-Grace2018 (ifg.tugraz.at/ITSG-Grace2018), CSR swath (https://doi.org/10.18738/T8/95ITIK), ADO1B releases (ftp://isdcftp.gfz-potsdam.de/grace/Level-1B/GFZ/AOD/), ECCOv4 release 4 (https://ecco-group.org/products-ECCO-V4r4.htm), LLC270 setup and unadjusted forcing (https://ecco.jpl.nasa.gov/drive/files/Version5/Alpha), DAC (by request from info-sealevel@esa-sealevel-cci.org), ERA-Interim (https://apps.ecmwf.int/datasets/data/interim-full-daily/). The first author is happy to share the BPR data and any codes used in this study.

Acknowledgments
The authors are grateful to two anonymous reviewers for their constructive suggestions. Financial support for this work was made available by the Austrian Science Fund (FWF, through project P30097-N29 of first author M.S.), the National Aeronautics and Space Administration (GRACE Follow-On Science Team grant 80NSSC20K0728, R.M.P.), and the German Research Foundation (DFG, DO1311/4-1, H.D.). All numerical simulations were performed using the Vienna Scientific Cluster (VSC).

References

Androsov, A., Boebel, O., Schröter, J., Danilov, S., Macrander, A., & Ivanciu, I. (2020). Ocean bottom pressure variability: Can it be reliably modeled? *Journal of Geophysical Research: Oceans*, *125*(3), e2019JC015469. https://doi.org/10.1029/2019JC015469

Bergmann, I., & Dobslaw, H. (2012). Short-term transport variability of the Antarctic Circumpolar Current from satellite gravity observations. *Journal of Geophysical Research*, *117*, C05044. https://doi.org/10.1029/2012JC007872

Bonin, J. A., & Chambers, D. P. (2011). Evaluation of high-frequency oceanographic signal in GRACE data: Implications for de-aliasing. *Geophysical Research Letters*, *38*(17), L17608. https://doi.org/10.1029/2011GL048881

Bonin, J. A., & Save, H. (2020). Evaluation of sub-monthly oceanographic signal in GRACE "daily" swath series using altimetry. *Ocean Science*, *16*(2), 423–434. https://doi.org/10.5194/os-16-423-2020

Carrère, L., Faugère, Y., & Ablain, M. (2016). Major improvement of altimetry sea level estimations using pressure-derived corrections based on ERA-Interim atmospheric reanalysis. *Ocean Science*, *12*(3), 825–842. https://doi.org/10.5194/os-12-825-2016

Carrère, L., & Lyard, F. (2003). Modeling the barotropic response of the global ocean to atmospheric wind and pressure forcing - comparisons with observations. *Geophysical Research Letters*, *30*(6), 1275. https://doi.org/10.1029/2002GL016473

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Dill, R., Dobslaw, H., & Thomas, M. (2019). Improved 90-day Earth orientation predictions from angular momentum forecasts of atmosphere, ocean, and terrestrial hydrosphere. *Journal of Geodesy*, *93*, 287–295. https://doi.org/10.1007/s00190-018-1158-7

Dobslaw, H., Bergmann-Wolf, I., Dill, R., Poropat, L., & Flechtner, F. (2017). *Product description document for AOD1B release 06, rev. 6.1*. GFZ Potsdam, Potsdam, Germany. Retrieved from ftp://isdcftp.gfz-potsdam.de/grace/DOCUMENTS/Level-1/GRACE_AOD1B_Product_Description_Document_for_RL06.pdf

Dobslaw, H., Bergmann-Wolf, I., Dill, R., Poropat, L., Thomas, M., Dahle, C., et al. (2017). A new high-resolution model of non-tidal atmosphere and ocean mass variability for de-aliasing of satellite gravity observations: AOD1B RL06. *Geophysical Journal International*, *211*(1), 263–269. https://doi.org/10.1093/gji/ggx302

Dobslaw, H., Bergmann-Wolf, I., Forootan, E., Dahle, C., Mayer-Gürr, T., Kusche, J., et al. (2016). Modeling of present-day atmosphere and ocean non-tidal de-aliasing errors for future gravity mission simulations. *Journal of Geodesy*, *90*, 423–436. https://doi.org/10.1007/s00190-015-0884-3

Dobslaw, H., Flechtner, F., Bergmann-Wolf, I., Dahle, C., Dill, R., Esselborn, S., et al. (2013). Simulating high-frequency atmosphere-ocean mass variability for dealiasing of satellite gravity observations: AOD1B RL05. *Journal of Geophysical Research: Oceans*, *118*(7), 3704–3711. https://doi.org/10.1002/jgrc.20271

ECCO Consortium, Fukumori, I., Wang, O., Fenty, I., Forget, G., Heimbach, P., & Ponte, R. M. (2020). *Synopsis of the ECCO central production global ocean and sea-ice state estimate (version 4 release 4)*. Retrieved from https://ecco-group.org/docs/v4r4_synopsis.pdf Accessed 7 August 2020.

Eicker, A., Jensen, L., Wöhnke, V., Dobslaw, H., Kvas, A., Mayer-Gürr, T., et al. (2020). Daily GRACE satellite data evaluate short-term hydro-meteorological fluxes from global atmospheric reanalyses. *Scientific Reports*, *10*, 4504. https://doi.org/10.1038/s41598-020-61166-0

Einšpigel, D., & Martinec, Z. (2017). Time-domain modeling of global ocean tides generated by the full lunisolar potential. *Ocean Dynamics*, *67*, 165–189. https://doi.org/10.1007/s10236-016-1016-1

Elsaka, B., Kusche, J., & Ilk, K.-H. (2012). Recovery of the Earth's gravity field from formation-flying satellites: Temporal aliasing issues. *Advances in Space Research*, *50*(11), 1534–1552. https://doi.org/10.1016/j.asr.2012.07.016

Flechtner, F., Neumayer, K.-H., Dahle, C., Dobslaw, H., Fagiolini, E., Raimondo, J.-C., et al. (2016). What can be expected from the GRACE-FO laser ranging interferometer for Earth science applications? *Surveys in Geophysics*, *37*, 453–470. https://doi.org/10.1007/s10712-015-9338-y

Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, *8*(10), 3071–3104. https://doi.org/10.5194/gmd-8-3071-2015

Fu, L.-L. (2003). Wind-forced intraseasonal sea level variability of the extratropical oceans. *Journal of Physical Oceanography*, *33*, 436–449. https://doi.org/10.1175/1520-0485(2003)033<0436:WFISLV>2.0.CO;2

Fukumori, I., Wang, O., Llovel, W., Fenty, I., & Forget, G. (2015). A near-uniform fluctuation of ocean bottom pressure and sea level across the deep ocean basins of the Arctic Ocean and the Nordic Seas. *Progress in Oceanography*, *134*, 152–172. https://doi.org/10.1016/j.pocean.2015.01.013

Gebler, M. (2013). *Detection and analysis of oceanic motion using Pressure Inverted Echo Sounders (PIES). PhD thesis*. Universität Oldenburg.

Glomsda, M., Bloßfeld, M., Seitz, M., & Seitz, F. (2020). Benefits of non-tidal loading applied at distinct levels in VLBI analysis. *Journal of Geodesy*, *94*, 90. https://doi.org/10.1007/s00190-020-01418-z

Greatbatch, R. J. (1994). A note on the representation of steric sea level in models that conserve volume rather than mass. *Journal of Geophysical Research*, *99*(C6), 12767–12771. https://doi.org/10.1029/94JC00847

Hartmann, T., & Wenzel, H.-G. (1995). The HW95 tidal potential catalogue. *Geophysical Research Letters*, *22*, 3553–3556. https://doi.org/10.1029/95GL03324

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horanyi, A., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*, 1999–2049. https://doi.org/10.1002/qj.3803

Hirose, N., Fukumori, I., Zlotnicki, V., & Ponte, R. M. (2001). Modeling the high-frequency barotropic response of the ocean to atmospheric disturbances: Sensitivity to forcing, topography, and friction. *Journal of Geophysical Research*, *106*(C12), 30987–30995. https://doi.org/10.1029/2000JC000763

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., et al. (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*(2), 422–446. https://doi.org/10.1002/jame.20023

Kurtenbach, E., Eicker, A., Mayer-Gürr, T., Holschneider, M., Hayn, M., Fuhrmann, M., et al. (2012). Improved daily GRACE gravity field solutions using a Kalman smoother. *Journal of Geodynamics*, *59*(60), 39–48. https://doi.org/10.1016/j.jog.2012.02.006

Kurtenbach, E., Mayer-Gürr, T., & Eicker, A. (2009). Deriving daily snapshots of the Earth's gravity field from GRACE L1B data using Kalman filtering. *Geophysical Research Letters*, *36*, L17102. https://doi.org/10.1029/2009GL039564

Kvas, A. (2020). *Estimation of high-frequency mass variations from satellite data in near real-time: Implementation of a technology demonstrator for near real-time GRACE/GRACE-FO gravity field solutions*. TU Graz University Press. https://doi.org/10.3217/978-3-85125-771-7

Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., et al. (2019). ITSG-Grace2018: Overview and evaluation of a new GRACE-only gravity field time series. *Journal of Geophysical Research: Solid Earth*, *124*(8), 9332–9344. https://doi.org/10.1029/2019JB017415

Macrander, A., Böning, C., Boebel, O., & Schröter, J. (2010). Validation of GRACE gravity fields by in situ data of ocean bottom pressure. In F. Fletchner, et al. (Eds.), *System Earth via geodetic geophysical space techniques*. (pp. 169–185). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-10228-8_14

Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research*, *102*(C3), 5753–5766. https://doi.org/10.1029/96JC02775

Mayer-Gürr, T., Behzadpur, S., Ellmer, M., Kvas, A., Klinger, B., Strasser, S., et al. (2018). *ITSG-Grace2018 - Monthly, daily and static gravity field solutions from GRACE*. GFZ Data Services. http://doi.org/10.5880/ICGEM.2018.003

Ponte, R. M. (1993). Variability in a homogeneous global ocean forced by barometric pressure. *Dynamics of Atmospheres and Oceans*, *18*(3), 209–234. https://doi.org/10.1016/0377-0265(93)90010-5

Ponte, R. M. (1997). Nonequilibrium response of the global ocean to the 5-day Rossby–Haurwitz wave in atmospheric surface pressure. *Journal of Physical Oceanography*, *27*(10), 2158–2168. https://doi.org/10.1175/1520-0485(0)027<2158:NROTGO>2.0.CO;2

Ponte, R. M. (2006). Oceanic response to surface loading effects neglected in volume-conserving models. *Journal of Physical Oceanography*, *36*, 426–434. https://doi.org/10.1175/JPO2843.1

Ponte, R. M., & Ali, A. H. (2002). Rapid ocean signals in polar motion and length of day. *Geophysical Research Letters*, *29*(15). https://doi.org/10.1029/2002GL015312

Poropat, L., Kvas, A., Mayer-Gürr, T., & Dobslaw, H. (2019). Mitigating temporal aliasing effects of high-frequency geophysical fluid dynamics in satellite gravimetry. *Geophysical Journal International*, *220*(1), 257–266. https://doi.org/10.1093/gji/ggz439

Quinn, K. J., & Ponte, R. M. (2008). Estimating weights for the use of time-dependent gravity recovery and climate experiment data in constraining ocean models. *Journal of Geophysical Research*, *113*, C12013. https://doi.org/10.1029/2008JC004903

Quinn, K. J., & Ponte, R. M. (2011). Estimating high frequency ocean bottom pressure variability. *Geophysical Research Letters*, *38*, L08611. https://doi.org/10.1029/2010GL046537

Ray, R. D., & Erofeeva, S. Y. (2014). Long-period tidal variations in the length of day. *Journal of Geophysical Research: Solid Earth*, *119*, 1498–1509. https://doi.org/10.1002/2013JB010830

Rietbroek, R., Brunnabend, S.-E., Kusche, J., Schröter, J., & Dahle, C. (2016). Revisiting the contemporary sea-level budget on global and regional scales. *Proceedings of the National Academy of Sciences*, *113*(6), 1504–1509. https://doi.org/10.1073/pnas.1519132113

Rodell, M., Famiglietti, J. S., Wiese, D. N., Reager, J. T., Beaudoing, H. K., Landerer, F. W., et al. (2018). Emerging trends in global freshwater availability. *Nature*, *557*, 651–659. https://doi.org/10.1038/s41586-018-0123-1

Salstein, D. A., Ponte, R. M., & Cady-Pereira, K. (2008). Uncertainties in atmospheric surface pressure fields from global analyses. *Journal of Geophysical Research*, *113*, D14107. https://doi.org/10.1029/2007JD009531

Save, H. (2019). *CSR RL05 GRACE daily swath mass anomaly estimates over the ocean*. Texas Data Repository Dataverse. Retrieved from https://doi.org/10.18738/T8/95ITIK

Saynisch, J., Bergmann-Wolf, I., & Thomas, M. (2015). Assimilation of GRACE-derived oceanic mass distributions with a global ocean circulation model. *Journal of Geodesy*, *89*, 121–139. https://doi.org/10.1007/s00190-014-0766-0

Schindelegger, M., Salstein, D., Einšpigel, D., & Mayerhofer, C. (2017). Diurnal atmosphere-ocean signals in Earth's rotation rate and a possible modulation through ENSO. *Geophysical Research Letters*, *44*(6), 2755–2762. https://doi.org/10.1002/2017GL072633

Seo, K. W., Wilson, C. R., Han, S. C., & Waliser, D. E. (2016). Gravity Recovery and Climate Experiment (GRACE) alias error from ocean tides. *Journal of Geophysical Research: Solid Earth*, *113*, B03405. https://doi.org/10.1029/2006JB004747

Stepanov, V. N., & Hughes, C. W. (2004). Parameterization of ocean self-attraction and loading in numerical models of the ocean circulation. *Journal of Geophysical Research*, *109*, C03037. https://doi.org/10.1029/2003JC002034

Stepanov, V. N., & Hughes, C. W. (2006). Propagation of signals in basin-scale ocean bottom pressure from a barotropic model. *Journal of Geophysical Research*, *111*, C12002. https://doi.org/10.1029/2005JC003450

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., et al. (2019). Contributions of GRACE to understanding climate change. *Nature Climate Change*, *9*, 358–369. https://doi.org/10.1073/pnas.1519132113

Velicogna, I., Mohajerani, Y., A, G., Landerer, F., Mouginot, J., Noel, B., et al. (2020). Continuity of ice sheet mass loss in Greenland and Antarctica from the GRACE and GRACE Follow-On missions. *Geophysical Research Letters*, *47*(8), e2020GL087291. https://doi.org/10.1029/2020GL087291

Vinogradova, N. T., Ponte, R. M., Quinn, K. J., Tamisiea, M. E., Campin, J.-M., & Davis, J. L. (2015). Dynamic adjustment of the ocean circulation to self-attraction and loading effects. *Journal of Physical Oceanography*, *45*(3), 678–689. https://doi.org/10.1175/JPO-D-14-0150.1

Vinogradova, N. T., Ponte, R. M., & Stammer, D. (2007). Relation between sea level and bottom pressure and the vertical dependence of oceanic variability. *Geophysical Research Letters*, *34*, L03608. https://doi.org/10.1029/2006GL028588

Wang, Q., Danilov, S., Sidorenko, D., Timmermann, R., Wekerle, C., Wang, X., et al. (2014). The Finite Element Sea Ice-Ocean Model (FESOM) v.1.4: Formulation of an ocean general circulation model. *Geoscientific Model Development*, *7*, 663–693. https://doi.org/10.5194/gmd-7-663-2014

Weijer, W., & Gille, S. T. (2005). Adjustment of the Southern Ocean to wind forcing on synoptic time scales. *Journal of Physical Oceanography*, *35*, 2076–2089. https://doi.org/10.1175/JPO2801.1

Weijer, W., Gille, S. T., & Vivier, F. (2009). Modal decay in the Australia-Antarctic Basin. *Journal of Physical Oceanography*, *39*, 2893–2909. https://doi.org/10.1175/2009JPO4209.1

Wieczorek, M. A., & Meschede, M. (2018). SHTools: Tools for working with spherical harmonics. *Geochemistry, Geophysics, Geosystems*, *19*(8), 2574–2592. https://doi.org/10.1029/2018GC007529

Williams, S. D. P., & Penna, N. T. (2011). Non-tidal ocean loading effects on geodetic GPS heights. *Geophysical Research Letters*, *38*, L09314. https://doi.org/10.1029/2011GL046940

Wunsch, C., Heimbach, P., Ponte, R. M., Fukumori, I., & The ECCO-GODAE Consortium Members. (2009). The global general circulation of the ocean estimated by the ECCO-Consortium. *Oceanography*, *22*, 88–103. https://doi.org/10.5670/oceanog.2009.41

Wunsch, C., & Stammer, D. (1997). Atmospheric loading and the oceanic "inverted barometer" effect. *Reviews of Geophysics*, *35*, 79–107. https://doi.org/10.1029/96RG03037

Yu, Y., Chao, B. F., García-García, D., & Luo, Z. (2018). Variations of the Argentine Gyre observed in the GRACE time-variable gravity and ocean altimetry measurements. *Journal of Geophysical Research: Oceans*, *123*, 5375–5387. https://doi.org/10.1029/2018JC014189

Zenner, L., Bergmann-Wolf, I., Dobslaw, H., Gruber, T., Günter, A., Wattenbach, M., et al. (2014). Comparison of daily GRACE gravity field and numerical water storage models for de-aliasing of satellite gravimetry observations. *Surveys in Geophysics*, *35*, 1251–1266. https://doi.org/10.1007/s10712-014-9295-x

Zhang, H., Menemenlis, D., & Fenty, I. (2018). *ECCO LLC270 ocean-ice state estimate*. California Institute of Technology. http://hdl.handle.net/1721.1/119821 Accessed 10 August 2020.

Zlotnicki, V., Wahr, J., Fukumori, I., & Song, Y. T. (2007). Antarctic Circumpolar Current transport variability during 2003–2005 from GRACE. *Journal of Physical Oceanography*, *37*(2), 230–244. https://doi.org/10.1175/JPO3009.1