Assessment of Ocean Bottom Pressure Variations in CMIP6 HighResMIP Simulations

Le Liu[1], Michael Schindelegger[1], Lara Börger[1], Judith Foth[1], and Junyang Gou[2]

[1]Institute of Geodesy and Geoinformation, University of Bonn, Bonn, Germany
[2]Institute of Geodesy and Photogrammetry, ETH Zurich, Zurich, Switzerland

Correspondence: Le Liu (le.liu@igg.uni-bonn.de)

Abstract. Ocean bottom pressure (p_b) variations from high-resolution climate model simulations under the CMIP6 (Coupled Model Intercomparison Project Phase 6) HighResMIP protocol are potentially useful for oceanographic and space-geodetic research, but the overall signal content and accuracy of these p_b estimates have hitherto not been assessed. Here we compute monthly p_b fields from five CMIP6 HighResMIP models at $1/4°$ grid spacing over both historical and future time spans and compare these data, in terms of temporal variance, against observation-based p_b estimates from a $1/4°$ downscaled GRACE (Gravity Recovery and Climate Experiment) product and 23 bottom pressure recorders, mostly in the Pacific. The model results are qualitatively and quantitatively similar to the GRACE-based p_b variances, featuring—aside from eddy imprints—elevated amplitudes on continental shelves and in major abyssal plains of the Southern Ocean. Modeled p_b variance in these regions is ~10–80% higher and thus overestimated compared to GRACE, whereas underestimation relative to GRACE and the bottom pressure recorders prevails in more quiescent deep-ocean regions. We also form variance ratios of detrended p_b signals over 2030–2049 under a high-emission scenario relative to 1980–1999 for three selected models and find statistically significant increases of future p_b variance by ~30–50% across ~~the Arctic and in eddy-rich regions of the~~ deep Arctic basins and the southern South Atlantic. The strengthening ~~is consistent with~~ appears to be linked to projected changes in high-latitude surface winds and, in the case of the South Atlantic, intensified ~~Agulhas leakage~~ eddy kinetic energy. The study thus points to possibly new pathways for relating observed p_b variability from (future) satellite gravimetry missions to anthropogenic climate change.

1

Ocean bottom pressure (p_b) measures the weight of the total water column exerted on the seafloor, and thus its variations are inherently linked to mass fluxes and redistribution in the Earth system. For example, contemporary increases in ocean mass and thus static p_b result from loss of land ice and continued groundwater depletion (Ludwigsen et al., 2024). Dynamic p_b signals, on the other hand, are largely driven by atmospheric forcing (e.g., Ponte, 1999; Boening et al., 2011; Petrick et al., 2014; Fukumori et al., 2015), but may also arise internally within the ocean through non-linear processes (Zhao et al., 2021). To the extent the ocean is in geostrophic balance, horizontal p_b gradients can be used to infer barotropic transport variability of deep currents (e.g., Makowski et al., 2015) or characteristics of the meridional overturning circulation (Hughes et al., 2018). Physical oceanography aside, knowledge of p_b variability is also essential in geodesy, particularly for de-aliasing purposes in

satellite gravimetry processing (Shihora et al., 2022; Schindelegger et al., 2021). In addition, ocean mass redistributions excite Earth rotation changes (e.g., Harker et al., 2021) and cause crustal loading signals that may mask geophysically more relevant effects in station position time series (van Dam et al., 2012).

Works in the noted disciplines have been served well by p_b estimates on a global scale from time-resolved satellite gravimetry (i.e., the Gravity Recovery and Climate Experiment, GRACE) and numerical ocean models. Pertinent model diagnostics are typically drawn from data-constrained or free simulations of the real ocean (e.g., Piecuch et al., 2015; Hughes et al., 2018; Androsov et al., 2020; Shihora et al., 2022; Ponte et al., 2024). With a few exceptions (Ponte et al., 2002; Bingham and Haines, 2005; Landerer et al., 2007), coupled climate models have so far played no role in studies of p_b variability and its implications~~, mainly because they are arbitrary in time~~. However, such synthetic mass change records are potentially useful as background fields in future gravimetry mission simulations to assess the error characteristics and science return of specific satellite constellations (Daras et al., 2024). Moreover, climate models can be used to delineate likely changes in surface mass variability that may form target signals for future satellite gravity missions (cf. Jensen et al., 2020). Whether or not there are any significant changes in dynamic p_b to expect in a warming world is an interesting question in its own right.

We are thus motivated to take a first look at p_b estimates from the latest realization of the Coupled Model Intercomparison Project (CMIP), i.e., CMIP6. Specifically, our objectives are to (i) validate the p_b diagnostics from CMIP6 experiments by comparing them in, a statistical sense, to observations, (ii) identify source models that could possibly serve as input for end-to-end satellite simulations, and (iii) test for likely future changes in p_b variability under a high-emission scenario. Although the coarse-resolution (∼100-km grid spacing) branch of CMIP6 would suffice most of these purposes, we here analyze the output of eddy-permitting (∼25-km) ocean models under the CMIP6 HighResMIP (CMIP6-HR) protocol. Horizontal model resolution in this range is generally advantageous for representing topographically constrained dynamics (e.g., Harker et al., 2021). It also admits transient eddy features, which have their own unique imprint on p_b (Bingham and Hughes, 2008; Hughes et al., 2018; Zhao et al., 2021). Eddy activity, in turn, is projected to change profoundly in the 21st century (Beech et al., 2022) and therefore may strengthen or weaken p_b variability in certain locations. In total, and for practical reasons, we consider five out of the seven complete institutional contributions to the CMIP6-HR activity (cf. Haarsma et al., 2016; Roberts et al., 2020). The specific time periods examined ~~by us~~ are 1980–2014, representing recent historical climatic conditions, and ~~2030–2949~~ 2030–2049 in the near future. Given the monthly sampling of the data and the ~~fact that model drift (Sen Gupta et al., 2013) precludes the study of trends, we~~ possibility of spurious model drifts (Sen Gupta et al., 2013), we are precluded from studying trends and exclusively analyze p_b signals at intraseasonal to interannual periods.

2 Data and methods

2.1 Climate model data

CMIP6-HR is an effort dedicated to the investigation of the role of horizontal resolution in CMIP-type climate simulations (Haarsma et al., 2016), see, e.g., Docquier et al. (2019), Roberts et al. (2020) or Shi et al. (2021) for ~~valorizations~~ evaluations of these simulations with respect to the ocean. We use the model outputs from the coupled historical runs (labeled as 'hist-

Table 1. Selected specifications of the CMIP6-HR models used in this study[a]

	ECMWF-IFS	HadGEM3-GC31	CNRM-CM6-1	CMCC-CM2	EC-Earth3P
Short name	ECMWF	HadGEM	CNRM	CMCC	Earth3P
Resolution label	HR	HM	HR	VHR4	HR
Atmospheric model[b]	IFS cycle 43r1	UM GA7.0	ARPEGE-Climat6.3	CAM4	IFS cycle 36r4
Atmosphere resolution in km	25	50	35	25	50
Ocean/sea ice					
Ocean model	NEMO3.4	NEMO3.6	NEMO3.6	NEMO3.6	NEMO3.6
Horizontal resolution in ° (km)	1/4° (25 km)	1/4° (25 km)	1/4° (25 km)	1/4° (25 km)	1/4° (25 km)
No. vertical levels	75	75	75	50	75
Variant label	r1i1p1f1	r1i1p1f1	r1i1p1f2	r1i1p1f1	r3i1p2f1
Reference	Roberts et al. (2018)	Roberts et al. (2019)	Voldoire et al. (2019)	Cherchi et al. (2019)	Haarsma et al. (2020)

[a] The analysis window is 1980–2014 (hist-1950 experiment) and, in addition, 2030–2049 for the HadGEM, CNRM, and Earth3P highres-future experiments under SSP5-8.5.

[b] Abbreviations: IFS (Integrated Forecasting System of the ECMWF), UM GA7.0 (Global Atmosphere 7.0 configuration of the Met Office Unified Model), ARPEGE (Action de Recherche Petite Echelle Grande Echelle), CAM4 (Community Atmosphere Model Version 4).

1950' in HighResMIP), complemented by coupled scenario simulations ('highres-future') in several cases. The hist-1950 simulations cover the period 1950–2014 and provide the initial conditions to highres-future, nominally integrated to the year 2050. External forcing fields comprise time-varying solar, volcanic, anthropogenic aerosol, ozone depletion, and greenhouse gas effects (Haarsma et al., 2016). For the highres-future simulations, these forcings follow the Representative Concentration Pathway 8.5 (i.e., high-end emission) scenario with Shared Socioeconomic Pathway 5 (SSP5, O'Neill et al., 2016) conditions, abbreviated as SSP5-8.5.

The five CMIP6-HR models examined in our work are ECMWF-IFS-HR (European Centre for Medium-Range Weather Forecasts Integrated ~~Forecasts~~ Forecasting System), HadGEM3-GC3.1 (Hadley Centre Global Environment Model 3 – Global Coupled vn 3.1), CNRM-CM6-1 (Centre National de Recherches Météorologiques Climate Model version 6), CMCC-CM2 (Euro-Mediterranean Centre on Climate Change Coupled Climate Model version 2), and EC-Earth3P; see Table 1 for an overview. ~~Ad hoc short names~~ Abbreviations, used throughout the paper, are ECMWF, HadGEM, CNRM, CMCC, and Earth3P, respectively. Given the scope of our study, we focus on the high-resolution (label HR, HM, or VHR4) configuration of each model, comprising ~~a horizontal grid spacing~~ horizontal grid spacings of 1/4° (∼25 km) for the ocean and 25 to 50 km for the atmosphere. The selected models differ with regard to the atmospheric component but not for the ocean (Table 1); the ocean component in all five cases is NEMO (Nucleus for European Modelling of the Ocean, either version 3.4 or 3.6, Madec, 2016), a

hydrostatic, primitive equation general circulation model. Processes and properties of the ocean are therefore represented very similarly in all five climate models. Hence, any spread in p_b variability reported below rather bears on parameter uncertainty, internal variability, or structural uncertainty in the atmospheric component. Values for selected parameters of the ocean component (e.g., viscosity and diffusivity) can be found in Docquier et al. (2019) and Roberts et al. (2020).

To deduce each model's bottom pressure, we retrieve monthly mean fields for potential temperature (standard name 'thetao'), salinity ('so'), and sea surface height above geoid ('zos'). ~~For the hist-1950 experiments, we focus on the last 35 simulation years (1980–2014) , giving us a temporal overlap with the gridded validation data from satellite gravimetry (~~Alternatively, we could use the models' diagnostic p_b output (standard name 'pbo') without invoking assumptions in the calculation of p_b; cf. Sect. ~~2.3) of nearly 13 years. Comparisons of signal variance, as attempted in this study, could also be done for non-overlapping time periods if one assumes stationarity of the~~ 2.2. Not all of the tested models provide 'pbo' output, though. In any event, differences between the original and our inferred p_b ~~variability at a given location. We will return to this assumption in the summary section~~values in terms of their temporal RMS (root-mean-square) are very small (Supplementary Fig. S1). For the hist-1950 experiments, we focus on the last 35 simulation years (1980–2014). The time span considered for the highres-future simulations is 2030–2049. These choices were made in view of ~~data availability and~~ manageable data volumes~~.~~, data availability, and the requirement for extended overlap with the p_b time series from satellite gravimetry (Sect. 2.3). For complementary analyses and tentative interpretation of the p_b results, we additionally process monthly eastward and northward seawater velocities (u, v; standard names 'uo', 'vo') and the eastward and northward components of the near-surface atmospheric wind ('uas', 'vas', usually referred to an altitude of 10 m).

2.2 Calculation of ocean bottom pressure anomalies

A mathematical expression for p_b, obtained by integrating the hydrostatic balance over full depth along the vertical coordinate z (positive upward), is (e.g., Gill and Niller, 1973; Ponte, 1999)

$$p_b = \int_{-H}^{0} \rho g \mathrm{d}z + \int_{0}^{\eta} \rho_0 g \mathrm{d}z \underline{\approx} \int_{-H}^{0} \rho g \mathrm{d}z + \rho_0 g \eta \tag{1}$$

where $\rho = \rho(z)$ is the (in situ) density of seawater with surface value ρ_0, η represents the sea level anomaly, H is the local water depth, and g is the acceleration due to gravity. While atmospheric pressure contributes to p_b in nature, it is not part of the exchange variables between atmospheric and oceanic components of CMIP6-HR and therefore neglected in Eq. (1). We perform the vertical integration on the native model grid and specifically derive the value of ρ at a given location from the modeled potential temperature and salinity, using the Gibbs SeaWater Oceanographic Toolbox (McDougall and Barker, 2011). Exact evaluation of Eq. (1) from CMIP6-HR data would also require each model's internal representation of bottom topography, e.g., in the form of a three-dimensional array of wet and dry cell fractions (or partial cell thicknesses; see Adcroft et al., 1997; Bernard et al., 2006). Given that these fractions are not provided in the CMIP6 archive, we calculate approximate versions of them based on given layer thicknesses and the ETOPO1 bathymetry dataset (Amante and Eakins, 2009), interpolated to the horizontal model grid. Although the information on the provenance of model bathymetries in CMIP6-HR is ~~sketchy~~uncertain,

4

105 the HadGEM ocean component description (Storkey et al., 2018) explicitly points to ETOPO1, and that is indeed a standard choice for NEMO simulations on the ORCA025 grid (cf. Börger et al., 2023).

We reduce the full bottom pressure to dynamic anomalies (also denoted p_b) by subtracting the time-mean pressure and trend values at each grid point, along with the monthly-varying spatial average of p_b. Global ocean mass changes, which may have physical and non-physical causes in numerical ocean models (e.g., Sen Gupta et al., 2013), are therefore not part of the 110 analyzed p_b variability. To convey some sense of how this variability is distributed across frequencies while also keeping the description of the results compact, we consider (i) full-time series of detrended dynamic p_b and (ii) non-seasonal anomalies obtained by removing least-squares fits of the annual and semiannual oscillations from the full-time series. Variant (ii) thus emphasizes intraseasonal and interannual p_b signals. The seasonal cycle, which has been the subject of several recent GRACE and ocean model studies (e.g., Cheng et al., 2021; Chen et al., 2023; Ponte et al., 2024), is only treated in passing. Note that p_b 115 has standard SI units (Pa) in our calculations but is expressed as centimeters of equivalent water height (EWH) below, assuming the same references density (1025 kg m^{-3}) as in the GRACE data.

All comparisons to observations are conducted based on measures of signal variance, i.e., the temporal RMS or quantities derived thereof. As will be shown below (Sect. 3.5), the p_b RMS at a given location is generally a non-stationary statistic that fluctuates under both constant and evolving climatic conditions. These fluctuations may render modeled and observed 120 p_b variability statistically indistinguishable from each other. We indicate such cases in our presentation by testing whether the GRACE-based p_b RMS is inside the 10th–90th percentile range of an approximate distribution of model RMS values at the considered grid point. For each model, the distribution is constructed by computing RMS values in rolling 15-year windows, progressively shifted forward by one year from 1980–1994 to 1990–2014. We thus have 21 RMS estimates, drawn from windows that are similar in length (although somewhat shorter) than the period of the adopted GRACE data. In figures 125 pertaining to the historical simulations, we always plot the median RMS of each model's distribution, which closely agrees with the RMS estimate over the full 1980–2014 period (typically within 0.2 cm).

2.3 ~~Downscaled~~ GRACE ~~product~~data

The primary dataset for validating the CMIP6-HR bottom pressures is a high-resolution satellite-based p_b product (Gou et al., 2025), obtained by downscaling available monthly p_b anomalies from GRACE/-Follow On (Tapley et al., 2019) beyond their 130 typical effective resolution of ∼3°. This product, called GRACE-DS hereinafter, offers ~~gridded~~ p_b information from April 2002 to December 2020 on a $1/4° \times 1/4°$ grid~~, commensurate with the resolution of the CMIP6-HR models~~. The downscaling was realized using a self-supervised deep-learning algorithm, guided by p_b diagnostics from two eddy-permitting ocean reanalyses ~~(Lellouche et al., 2013; Zuo et al., 2017). In particular, spatial~~ based on the NEMO model (Lellouche et al., 2013; Zuo et al., 2017). Spatial detail on sub-GRACE grid scales was obtained by maximizing similarities between the downscaled product and 135 the reanalysis output, while constraining the solution to the original GRACE values at larger scales (Gou et al., 2025). The data fusion algorithm was further enhanced by supervision signals based on input features, comprising, e.g., wind stress and bathymetry. From the three downscaled solutions provided by Gou et al. (2024), we use the one based on GRACE mascons from the Center for Space Research ~~(Save et al., 2016)~~(CSR, RL06.2, Save et al., 2016).
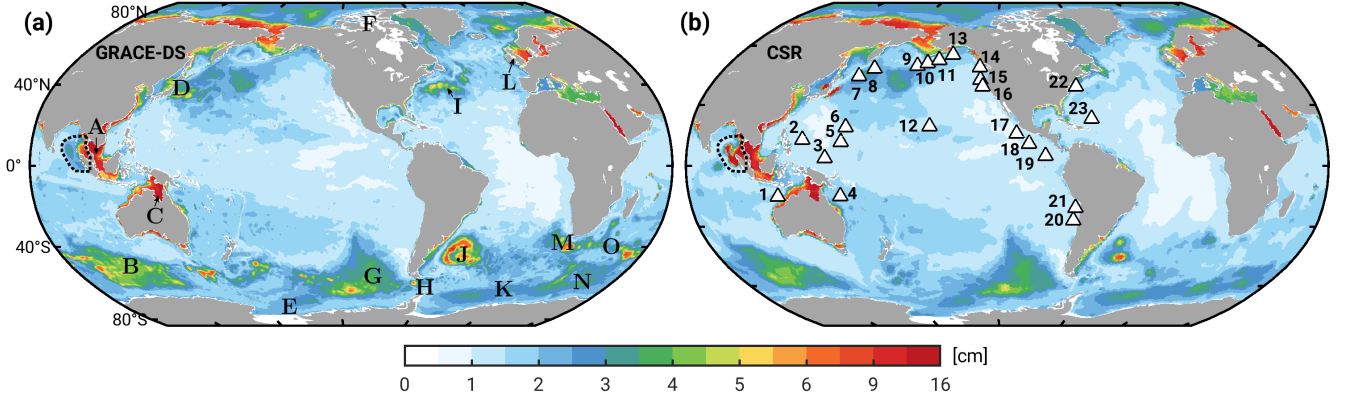
Figure 1. RMS of monthly p_b anomalies (in cm of EWH) from (a) GRACE-DS, obtained by downscaling 1° ~~GRACE~~ CSR mascons (Save et al., 2016) to a 1/4° × 1/4° grid. ~~Time~~ The RMS of the CSR solution is illustrated in panel (b). Analyzed time series are from April 2002 to December 2020, with trends and spatial mean p_b signals removed. White triangles in (b) indicate the 23 BPR stations used in this study, labeled by increasing longitude. The dashed black polygon in the Andaman Sea marks the region excluded from the comparison in Sect. 3 due to the presence of seismic deformation signals in the GRACE fields. Lettered keys in panel (a) denote the (**A**) Gulf of Thailand, (**B**) Australian Antarctic Basin, (**C**) Gulf of Carpentaria, (**D**) Kuroshio Current, (**E**) Ross Sea, (**F**) Canadian Arctic Archipelago, (**G**) Bellingshausen Basin, (~~**GH**~~) Drake Passage, (~~**HI**~~) Gulf Stream Extension, (~~**IJ**~~) Argentine Basin/Zapiola Gyre, (~~**JK**~~) Weddell Basin, (~~**KL**~~) ~~Enderby Basin~~Celtic Sea, (~~**LM**~~) Cape Basin/Agulhas leakage, (~~**MN**~~) Enderby Basin, (**O**) Agulhas Return Current.

~~The temporal RMS (root-mean-square) of the monthly~~ A visual comparison of the GRACE-DS ~~fields at each grid point is depicted~~ and CSR fields in terms of the RMS of dynamic, detrended p_b anomalies is provided in Fig. 1~~and discussed in relation to the CMIP6-HR p_b signal content in Sect. 3.2. As with the models, the~~. As per design, the deep-learning algorithm retains the signal content of the CSR parent solution at wavelengths of a few hundred kilometers and longer. At the same time, short-scale features are added throughout the ocean, including marked p_b increases in eddy-active regions (e.g., Southern Ocean, western boundary current regions). The downscaling approach also enhances near-shore amplitudes on several continental shelves, usually by 20–30% (e.g., around Australia, China, the Atlantic coast of South America, or on the Siberian Shelf). As these spatial details are mainly derived from the NEMO-based ocean reanalyses, the model-data comparison in eddying regions and near continental boundaries should to be treated with caution. In the supplementary to this article, we thus repeat parts of our model assessment using (i) the CSR RMS map in Fig. 1b, and (ii) another p_b mascon solution from Jet Propulsion Laboratory (JPL, RL06.3M, Watkins et al., 2015; Wiese et al., 2023). Whereas CSR employs the same 1/4° × 1/4° grid as GRACE-DS~~time series represents dynamic, detrended~~, the JPL mascons are discretized in the form of 3° × 3° equal-area caps. Prior to comparing to JPL data, we average the model p_b ~~anomalies.~~fields to the exact geometry of these caps (cf. Ponte et al., 2024). Such coarsening suppresses most of the mesoscale imprints on p_b, which are, in any case, difficult to assess. For the presentation in the main text, we focus on GRACE-DS, as its characteristics are most commensurate with the CMIP6-HR models across scales and from deep to shallow regions.

155   Sources of inconsistency between ~~the two datasets are~~CMIP6-HR and the GRACE(-DS) data, among other effects, the absence of barometric pressure-driven p_b variability in the models (cf. Piecuch et al., 2022, $\lesssim 0.5$ cm in places) and seismic gravity/deformation signals associated with major earthquakes in ~~GRACE-DS~~ the gravity fields solutions (e.g., Zhang et al., 2020). The largest such spurious contribution to p_b, apparent in Fig. 1 around the Malay Peninsula, stems from the 2004 Sumatra-Andaman and 2012 Indian Ocean earthquakes. We approximately delimit the affected region based on insight from

160   literature (De Linage et al., 2009) and enhanced (≥ 2 cm) open-ocean p_b values, and exclude it from all direct comparisons to the model fields~~below~~.

2.4   In situ observations

We also compare the CMIP6-HR p_b diagnostics with direct seafloor pressure measurements from the DART® (Deep-ocean Assessment and Reporting of Tsunamis) program (Bernard and Meinig, 2011). DART is a US-led tsunami early warning

165   system comprising more than 50 active deep-ocean bottom pressure recorders (BPRs) located mainly along the Circum-Pacific belt. The DART data are disseminated in meters of EWH with a nominal sampling of 15 min. From the full network, we extract 23 stations with nearly continuous observations over an 11-year time span from 2013 to 2023. The year 2013 is a reasonable start date, as it marks the completion of a major upgrade to the DART technology (Bernard and Meinig, 2011). Each of the 23 time series is a concatenation of recordings from successive deployments in slightly different locations, typically differing by

170   O(10 m) in the overlying pressure. We process each of these segments individually, removing (i) 3 days of observations at the start and end of the time series, (ii) a fitted second-order polynomial (to account for sensor drift, cf. Poropat et al., 2018), (iii) occasional spikes, and (iv) signals of 16 largest tidal constituents (Hart-Davis et al., 2021, including fortnightly, monthly, and semiannual constituents). Steps (i)–(iv) are applied to hourly-sampled versions of the segments, which we subsequently splice together and average into a monthly p_b series per station.

175   For consistency with our treatment of the CMIP6-HR fields (Sect. 2.1), the DART data should also be cleaned from global mean p_b signals, reflecting both net atmospheric pressure variations over the ocean and true ocean mass changes (Ponte, 1993; Johnson and Chambers, 2013). While the former is readily corrected for using monthly sea level pressures from a modern atmospheric reanalysis (Hersbach et al., 2020), precise estimates for the mean ocean mass change over 2013–2023 are harder to come by. However, apart from a trend, these changes mainly consist of an annual oscillation, reaching $\sim$0.5 cm in amplitude

180   (Johnson and Chambers, 2013). We thus remove from the detrended DART data a synthetic time series of the seasonal cycle in ocean mass, constructed from p_b output (1992–2017) of an ocean state estimate (the Estimating the Circulation and Climate of the Ocean, Version 4 Release 4b, ECCOv4rb, Forget et al., 2015). ECCOv4rb is deemed suitable for this purpose, as it is constrained to most available oceanographic data, including GRACE-based p_b and observed global ocean mass changes.

    Our final network, illustrated in Fig. 1b, comprises 23 BPRs featuring a mean cumulative record length of 10.8 years, with

185   monthly averages typically built from 260 to 350 days of valid observations per year. The in situ time is, therefore, much shorter than the CMIP6-HR records but sufficiently long to constrain variability at the annual and higher frequencies (cf. Ray et al., 2021, for the case of sea level). Interannual p_b variability, which is less well represented by these $\sim$10-year records, accounts for about 30% of the variance in the full monthly DART time series. The network also heavily emphasizes the Pacific,

with only two stations in the western North Atlantic and one sensor located in the Indian Ocean (North Australian Basin). We nevertheless prefer to perform our analysis based on a homogeneous, consistently processed in situ dataset and refrain from adding other BPR time series, which are usually short in duration ($\lesssim$ 1–4 years, Poropat et al., 2018; Androsov et al., 2020; Schindelegger et al., 2021) or confined to special locations (e.g., the Arctic Beaufort Gyre, Kemp et al., 2005).

3 Results

3.1 Overview of modeled p_b variability

Figure 2 (left column) illustrates the RMS of monthly dynamic p_b anomalies, calculated for each of the five CMIP6-HR models over the period 1980–2014 (as median of 21 RMS values in rolling 15-year windows). The model results are qualitatively and quantitatively very similar to the p_b variability in the GRACE-DS product (Fig. 1a), ocean state reconstructions (Piecuch et al., 2015; Gou et al., 2025), and ocean forward models forced by atmospheric analysis fields (e.g., Hughes et al., 2018; Poropat et al., 2018; Androsov et al., 2020). Among the defining features are a general increase in p_b variability from temperate to polar latitudes, elevated amplitudes in several marginal seas (e.g., Hudson Bay, Mediterranean Sea, Red Sea), and largest signals near continental boundaries. Shallow depths facilitate a vigorous barotropic response to the imposed wind stress (Vinogradova et al., 2007), leading to peak values of 10–15 cm in Fig. 2 over, e.g., the Siberian Shelf, Chukchi Sea, Gulf of Thailand, Gulf of Carpentaria, and Baltic Sea. In some of these places, coastally trapped waves contribute to the spreading of locally generated mass anomalies (e.g., Oliver and Thompson, 2011; Fukumori et al., 2015).

In the deep ocean, known regions of relatively high barotropic variability are the Arctic Ocean, the subpolar Northwest Pacific (Petrick et al., 2014), and the three major abyssal plains in the Southern Ocean (Bellingshausen, Australian-Antarctic, and Weddell-Enderby basins). These plains are encased by nearly closed contours of planetary potential vorticity, thus trapping and amplifying forced barotropic circulations (e.g., Fukumori et al., 1998; Weijer and Gille, 2005; Ponte and Piecuch, 2014; Weijer, 2015). The Australian-Antarctic basin is the most energetic of the three regions, with RMS values of ∼4–6 cm in all five models. Bottom topography also plays a key role in generating the characteristic, near-uniform fluctuation in p_b (up to 4 cm, Fig. 2) in the interconnected deep basins of the Arctic Ocean and Nordic Seas. In detail, wind stress along the continental slope results in cross-slope Ekman transport, which creates p_b anomalies of opposite sign between the shallow shelf and the adjacent deep ocean (Fukumori et al., 2015). The deep-ocean mass anomaly propagates away from its source regions as trapped Kelvin wave, separates from its shallow counterpart at straits and sills, and eventually equilibrates across the deep Arctic basins that are bounded by gradients of potential vorticity.

Forced barotropic dynamics aside, the CMIP6-HR models also contain an active mesoscale field, which can modulate the p_b diagnostics on local to basin-wide scales at intraseasonal and longer periods (Hughes et al., 2018; Zhao et al., 2021, 2023). Localized imprints of eddies and baroclinic instabilities are most obvious in western boundary current regions (Agulhas Retroreflection, Kuroshio, Gulf Stream) and broadly along the path of the Antarctic Circumpolar Current (Androsov et al., 2020). A stand-out feature in all models, perhaps with the exception of CMCC (Fig. 2d), is a ∼8-cm G-shaped p_b structure in the Argentine Basin. The pattern is the signature of the Zapiola Anticyclone, a well-known recirculation supported by high eddy

8
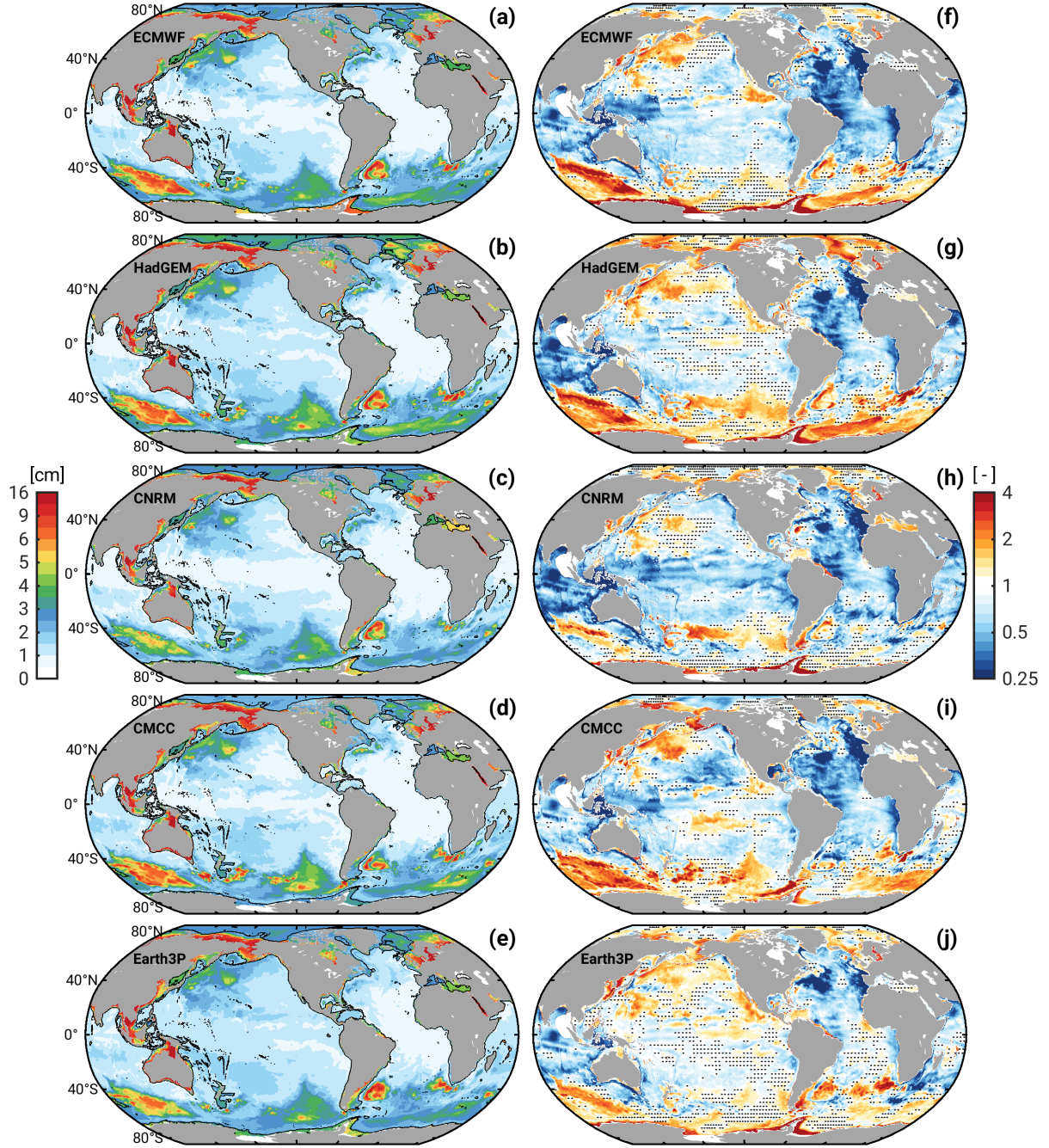
Figure 2. Plots on the left show the RMS (cm) of detrended p_b time series ~~, 1980–2014,~~ from five ~~CMIP-HR~~ CMIP6-HR models~~(left column) and~~, computed as median of the RMS distribution over 1980–2014, see Sect. 2.2 for details. Plots on the right show each model's variance ratio, R, relative to the p_b variance from GRACE-DS (~~right column). Values of~~ i.e., $R > 1$ indicates ~~an overestimation of~~ higher bottom pressure variability in the model~~, while $R < 1$ suggests underestimation~~). Stippling marks $3° \times 3°$ cells where the RMS of GRACE-DS (Fig. 1a) is within the 10th–90th percentile range of the model RMS for at least a third of the contained $1/4°$ grid points. The black solid line ~~marks~~ is the 1000-m isobath, and the region of the Andaman Sea anomaly seen in GRACE-DS is masked out~~; cf.Fig. 1.~~

9

activity in the western parts of the basin and interactions of that eddy field with bottom topography (de Miranda et al., 1999; Hughes et al., 2007). The GRACE-DS product suggests a similarly G-shaped structure around the Zapiola Drift (Fig. 1a), albeit ~~largely~~ inherited from the utilized ocean reanalyses ~~rather than~~ and not the GRACE gravity fields themselves ~~(Gou et al., 2025)~~ ~~.~~(Fig. 1b).

3.2 Comparison with ~~GRACE-DS~~ GRACE – broadband variability

For a more quantitative analysis of the agreement between modeled and observation-based p_b fluctuations, we compute variance ratios, R, defined as

$$R = \frac{\sigma^2_{mod}}{\sigma^2_{obs}} \qquad (2)$$

where σ^2_{mod} represents the p_b variance of a chosen model and σ^2_{obs} is the observed variance at a given location, calculated from GRACE-DS, ~~or from BPRsbelow.~~ CSR or JPL mascons, or the BPRs.

Figure 2 (right column) illustrates ~~global~~ estimates of R for the full (i.e., unfiltered) p_b model time series relative to GRACE-DS, deduced after interpolating σ^2_{mod} from the native model grid to the GRACE-DS grid. Despite differences in detail, the five maps have several features in common. In particular, they suggest a tendency for excess model variance ($R > 1$) in areas of enhanced p_b amplitudes (Fig. 2, left column), while underestimation by the models ($R < 1$) prevails in less energetic regions. One may argue that the latter is simply due to noise in the chosen GRACE product, dominating the low and mid-latitude RMS in Fig. 1a. However, random errors in a single monthly gravity field solution are typically estimated to be 3–4 cm (e.g., Kvas et al., 2019), which limits the noise component ~~on~~ in the temporal RMS in Fig. 1 to 0.3 cm (inferred from a pessimistic variance propagation over 192 monthly p_b fields).

~~Less easy to rebut is the existence of systematic effects in the~~ The prevalence of low variance ratios outside energetic regions is therefore more likely the result of systematic effects in the GRACE-DS ~~product, comprising, e.g., imperfect corrections for low-degree zonal spherical harmonics and geocenter motion. In addition, GRACE-derived p_b estimates contain the effects of gravitational attraction and loading (GAL). The physics of GAL are not encoded in CMIP6-HR but can induce ocean mass changes in the order of 0.5 cm, particularly at the annual frequency and near continental boundaries (Ponte et al., 2024). Parts of the higher variance in GRACE-DS in the Northern Indian Ocean and off the Amazon Shelf, which adjoin regions of large terrestrial water storage variations, are therefore likely due to GAL. We also note that the underestimation of variance in the deep-water tropics and some eastern boundary regions in CMIP6-HR (Fig. 2f–j) coincides with areas where p_b variability is influenced by baroclinic dynamics and coupling between baroclinic and barotropic modes; cf. Hughes et al. (2018) but see also Piecuch et al. (2015) and Zhao et al. (2021). Proper depiction of these processes in the models may pose a challenge, especially in non-eddy-resolving configurations. Overall, we conclude that the dominance of $R < 1$ in quiescent ocean regions in Fig. 2 is due to a mixture of data noise, remaining systematic effects in the GRACE-DS fields, and mismodeled or entirely absent p_b signals in the simulations~~product (e.g., imperfect corrections for low-degree zonal spherical harmonics, Gałdyn and Sośnica, 2024) and structural errors in the CMIP6-HR models. Thoughts on possible model limitations are offered in Sect. 3.4. Out of the five cases considered, HadGEM and Earth3P ~~appear to be least affected by modeling issues~~compare most favorably to GRACE-DS;

Table 2. Global median variance ratios, $\overline{R}$, of the CMIP6-HR models relative to GRACE-DS and BPRs[a]

Model Name	GRACE-DS, deep		GRACE-DS, shelf		BPRs[b]
	RMS > 3 cm	RMS ≤ 3 cm	$|\phi| > 60°$	$|\phi| \leq 60°$	
ECMWF	~~1.54~~1.58 (1.40)	~~0.72~~0.71 (0.66)	~~1.39~~ ~~(1.50~~1.37 (1.41)	~~1.11~~ ~~(1.36~~1.10 (1.03)	~~1.05~~ ~~(0.63~~1.02 (0.59)
HadGEM	~~1.52~~ ~~(1.49~~1.50 (1.47)	~~0.82~~0.81 (0.74)	~~1.86~~ ~~(1.71~~1.95 (1.66)	1.30 (~~1.64~~1.34)	~~1.04~~ ~~(0.69~~1.01 (0.68)
CNRM	~~1.16~~ ~~(1.11~~1.15 (1.09)	~~0.73~~ ~~(0.62~~0.72 (0.61)	~~1.50~~ ~~(1.53~~1.43 (1.36)	~~0.94~~ ~~(1.20~~0.91 (1.00)	~~0.84~~ ~~(0.60~~0.81 (0.59)
CMCC	~~1.78~~ ~~(1.57~~1.75 (1.58)	~~0.74~~ ~~(0.59~~0.73 (0.57)	1.38 (~~1.36~~1.30)	~~1.19~~ ~~(1.32~~1.15 (0.95)	~~0.76~~ ~~(0.53~~0.77 (0.47)
Earth3P	~~1.32~~ ~~(1.24~~1.29 (1.23)	~~0.91~~ ~~(0.74~~0.90 (0.73)	~~1.49~~ ~~(1.53~~1.44 (1.40)	~~1.25~~ ~~(1.59~~1.26 (1.30)	~~1.26~~ ~~(0.72~~1.28 (0.74)

[a] For each model, we list the global median value of R, relative to GRACE-DS, in four regions, two regions each in deep water (depths > 1000 m) and shallow water (depths < 1000 m). The partitioning in the deep ocean is by model RMS (> 3 cm or ≤ 3 cm), while shelf regions are separated by latitude ϕ (higher or lower than 60°). Values in brackets are the same estimates for the non-seasonal signal content. The last column shows the median variance ratios relative to BPRs.

[b] When compared to BPRs, GRACE-DS has a median variance ratio of 1.30 for the full time-series, and 0.92 for non-seasonal p_b changes.

global median variance ratios, $\overline{R}$, deduced from all grid points with a model RMS of ≤ 3 cm and in depths greater than 1000 m, are ~~0.82~~0.81 (HadGEM) and ~~0.91~~0.90 (Earth3P). These numbers are ~~considerably~~ somewhat closer to 1 than for the other three models ($\overline{R} \approx$ ~~0.73~~0.72, Table 2).

Shifting the focus to the red (i.e., $R > 1$) patches in Fig. 2, we find that all models ~~overestimate the~~ have higher p_b variance ~~relative to~~ than GRACE-DS in the Northwest Pacific, the Australian-Antarctic Basin, and (to some extent) the Weddell-Enderby abyssal plain. Excess values are least pronounced in CNRM and largest in CMCC, evident also from the global median variance ratios taken over deep energetic regions ($\overline{R} =$ ~~1.16~~1.15 vs. $\overline{R} =$ ~~1.78~~1.75, see Table 2). ~~Given the importance of wind stress curl and bottom topography in shaping the~~ Excessive p_b ~~variability in these deep basins (cf. Sect. 3.1), misrepresentation of either of these factors could reduce the fidelity of the modeled barotropic dynamics. For example, both the horizontal and vertical model resolution may be too coarse to produce the necessary topographic steering and dissipation by form stress (Weijer and Gille, 2005). Other suspects are imperfect wind stress parametrizations or overly strong surface winds as simulated~~

11

265 ~~by each model's atmosphere. However, supplementary checks of the 10-m wind speeds from several of the tested models revealed only small deviations from atmospheric reanalysis fields.~~

~~A propensity for excess p_b~~ variance in the CMIP6-HR models is also seen for most continental shelf regions—in subpolar and polar latitudes more so than in the tropics (cf., e.g., the Gulf of Thailand and Indonesian Seas). This geographical dependence is further brought out in Table 2, where we ~~compute~~ list global median variance ratios in shallow water ($H < 1000$ m)
270 both for latitudes higher and lower than $60°$ (a somewhat arbitrarily chosen cut-off). The overestimation of shallow p_b variations within $60°$ from the equator is rather subtle, bounded by $\overline{R} = 1.30$ for HadGEM. The HadGEM value increases to $\overline{R} = $ ~~1.86~~ 1.95 in polar latitudes, followed by ~~$\overline{R} = 1.38$–1.50~~ $\overline{R} \approx 1.40$ in the other models. Common regions of ~~systematic overestimation~~ overly strong p_b variations are the East Siberian Shelf, the Bering Sea, and the Antarctic continental shelf. ~~While the red patches in the Arctic may again point to issues regarding wind stress forcing and dissipation mechanisms, the overestimated~~
275 ~~p_b variability in the Weddell and Ross Seas bears on the blocking of Antarctic ice shelf cavities in CMIP6-HR. This modeling choice constricts the solution to a narrower body of water than in the real ocean and necessarily distorts the dynamics of the region.~~

As for the p_b imprints of mesoscale eddies, areas of ~~over- and underestimation mostly vary~~ high and low variance ratios tend to change from model to model, see, e.g., the Drake Passage or the Agulhas Retroreflection and its leakage to the west.
280 However, ~~at these small scales, the signal content~~ most eddy-related variability in the GRACE-DS product is ~~the result of trading off reanalyses output with GRACE fields that have an effective resolution of ~300 km (Gou et al., 2025). This type of blend certainly~~ propagated through from ocean reanalysis output (Sect. 2.3, Gou et al., 2025). This limits the validity of the comparison in Fig. 2 at the mesoscale. We nevertheless find a common patch of low variances ratios ($R \approx$ ~~0.5~~ 0.3) in the Northwest Atlantic south of Grand Banks, where the p_b signature of eddies in the GRACE-DS product spread considerably
285 further into the Atlantic interior (Fig. 1a) than in the CMIP6-HR models. The situation bears close resemblance to Figure 8 in Bernard et al. (2006), which suggests that data-free NEMO simulations on the ORCA025 grid typically underestimate the magnitude and spatial spread of eddy kinetic energy in the Gulf Stream Extension.

The quantitative results above can be tested for robustness by replacing GRACE-DS with either the CSR or JPL mascons as reference data (Supplementary Figs. S2–S3, Supplementary Tables S1–S2). These analyses suggest a spatial partitioning of
290 variance ratios that is very similar to Fig. 2. We again observe a tendency for $R < 1$ in many quiet deep-ocean regions and Southeast Asian seas, whereas values of $R > 1$ are common in the Southern Ocean, Northwest Pacific, and on continental shelves in polar latitudes. Among the more prominent deviations from Fig. 2 (GRACE-DS) and Supplementary Fig. S2 (CSR) are increased variance ratios with JPL throughout most of the Atlantic. These differences hint at uncertainties in the gravity field recovery in regions of small signal levels. In addition, the patches of red color are more saturated in the evaluation against CSR
295 mascons, which lack signal content on sub-grid GRACE scales (Fig. 1). Hence, the estimates of $\overline{R}$ from CSR in energetic deep regions in Supplementary Table S1 cannot be taken at face value. Variance ratios of the coarsened CMIP6-HR fields relative to the $3° \times 3°$ JPL data offer a cleaner form of comparison, and that analysis does indeed confirm many of the numerical values and model rankings noted in relation to Table 2 (column 'deep'). On shelf regions, the situation is mostly reversed, that is, the finer-resolved CSR mascons are better suited as alternative reference data than JPL. Here, Supplementary Table S1
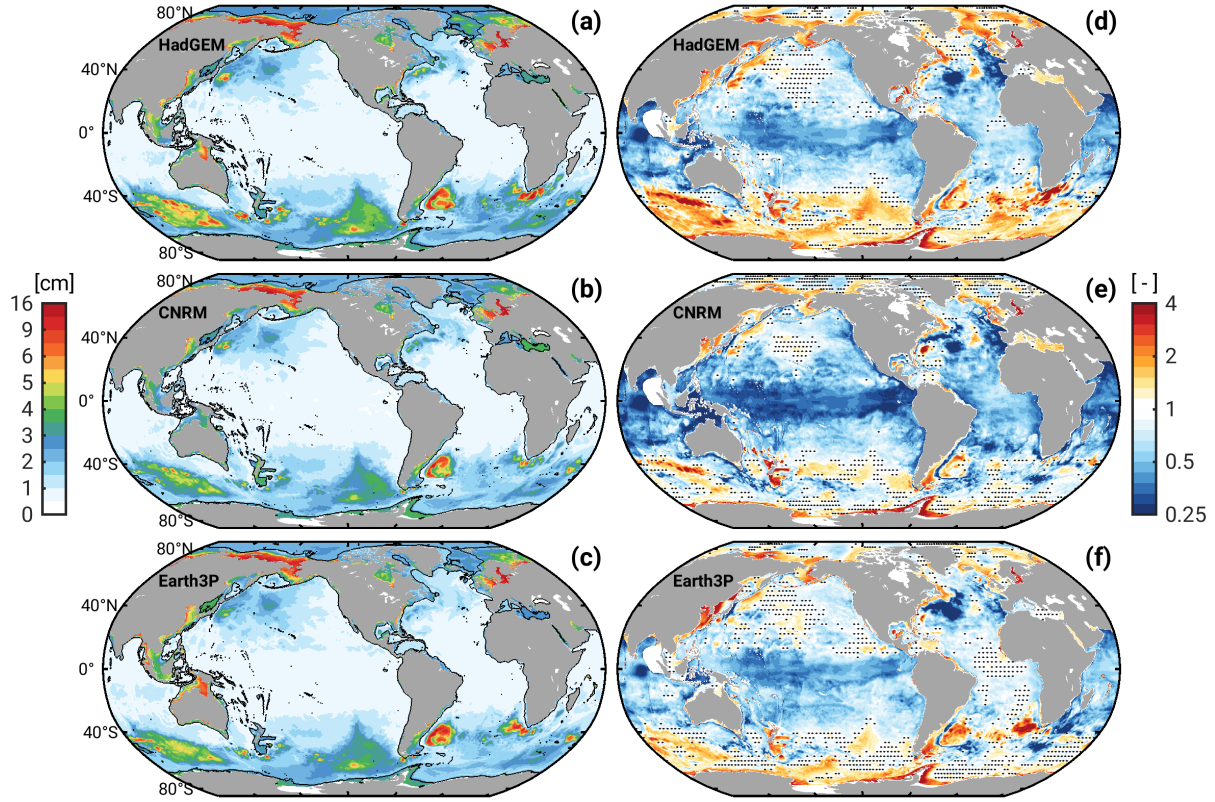
Figure 3. As in Fig. 2 but restricted to three models (HadGEM, CNRM, Earth3P) and non-seasonal signals.

300 (column 'shelf') again reveals no substantial changes from the corresponding $\overline{R}$ values with GRACE-DS in Table 2. Given these successful cross-checks, we only employ the downscaled product for the remaining model-data comparisons.

3.3 Non-seasonal p_b variability

Modeled p_b variability with annual and semiannual oscillations subtracted is briefly assessed for three of the five available models. We select CNRM and Earth3P, which compare favorably to ~~GRACE~~ GRACE-DS for energetic deep-ocean regions
305 in Fig. 2, and ~~HadGEM—essentially to have somewhat of~~ HadGEM—a somewhat arbitrary choice but still a counterpoint to the other two models. Global median variance ratios at non-seasonal periods for all five models are compiled in Table 2. As evident from Fig. 3, removal of the seasonal cycle reduces RMS values in deep waters of temperate latitudes to ~1 cm. At these magnitudes, comparisons with GRACE-DS are likely limited by data noise and subtle systematic errors (cf. Sect. 3.2)~~,~~ ~~inhibiting~~ . These issues inhibit more solid inferences as to why variance ratios in quiet deep-ocean regions decrease relative
310 to the full time series (Table 2, Fig. 2f–j vs. Fig. 3d–f).

13

Non-seasonal signal levels in many of the more energetic areas in Fig. 3 remain comparable to the full variability~~(Fig. 2)~~, and so do the variance ratios over GRACE-DS. Obvious examples are the Arctic Ocean and Nordic Seas, the Bellingshausen Basin, and several eddy-active regions (e.g., Zapiola Drift, Agulhas leakage, Gulf Stream Extension).

~~Wherever p_b has an appreciable seasonal cycle (cf. Ponte et al., 2024)~~By contrast, RMS values ~~in Fig. 3 are lower than in~~ for non-seasonal variations are clearly smaller than their broadband counterparts (Fig. 2~~. Aside from numerous shelf areas, such as~~) in the Gulf of Thailand, ~~non-seasonal amplitudes are also weaker than seasonal amplitudes in the~~Northwest Pacific, ~~the~~Australian-Antarctic Basin, and ~~the~~Weddell-Enderby Abyssal Plain. These changes go along with an improvement of the variance ratios relative to GRACE-DS toward $R \approx 1$ (even for HadGEM), suggesting that wind-driven geostrophic ~~modes variability~~ in abyssal-plain regions ~~are~~is indeed too energetic at seasonal periods in CMIP6-HR. ~~Another noteworthy change~~ Over mid- and low-latitude continental shelves, estimates of R ~~over~~ are largely unaltered as we switch from full to non-seasonal ~~variability is the increase in global median variance ratios over mid- and low-latitudinal continental shelves by about 0.2 for each of the five models~~ time series, and that behavior is also reflected in the global median statistics (Table 2). ~~The increase is mainly caused by an excess of~~

3.4 Intermediate discussion

What are potential reasons for the documented differences between modeled and satellite-based p_b variability? As alluded to in Sect. 3.2, the dominance of $R < 1$ in quiescent ocean regions may be attributable to a mixture of data noise, residual systematic effects in the GRACE-DS fields, and mismodeled or entirely absent p_b signals in the simulations. One specific process disregarded in CMIP6-HR are the effects of gravitational attraction and loading (GAL, Vinogradova et al., 2011). GAL induces ocean mass changes in the order of 0.5 cm, particularly at the annual frequency and near continental boundaries (Ponte et al., 2024). Parts of the higher variance in GRACE-DS in the Northern Indian Ocean and off the Amazon Shelf, which adjoin regions of large terrestrial water storage variations, may be due to GAL. We also note that the low model p_b amplitudes relative to GRACE-DS in the ~~shallow waters along the China and North American east coasts, as well as around New Zealand. For eastern continental boundaries, this overestimation potentially points to modeling problems regarding the propagation and decay of coastally trapped waves (?).~~deep-water tropics and some eastern boundary regions (Fig. 2f–j) coincide with areas where p_b variability is influenced by non-linearities and coupling between baroclinic and barotropic modes (Piecuch et al., 2015; Hughes et al., 2018; Zhao et al., 2021). Proper depiction of these processes in the models may pose a challenge and may depend on spatial resolution or certain parameter settings concerning, e.g., the viscosity of the flow.

Deep energetic regions, on the other hand, show a propensity for too vigorous p_b signals ($R > 1$). Given the importance of wind stress curl and bottom topography in shaping the dynamics of these deep basins (cf. Sect. 3.1), misrepresentation of either of these factors could reduce the fidelity of the modeled barotropic p_b variability. For example, both the horizontal and vertical model resolution may be too coarse to produce the necessary topographic steering and dissipation by form stress (Weijer and Gille, 2005). Other suspects are imperfect wind stress parametrizations or overly strong surface winds as simulated by each model's atmosphere. Similarly, the enhanced p_b variations on several continental shelves (e.g., in the Arctic) at both seasonal and non-seasonal periods point to issues regarding model topography and dissipation mechanisms. In particular, as the

14
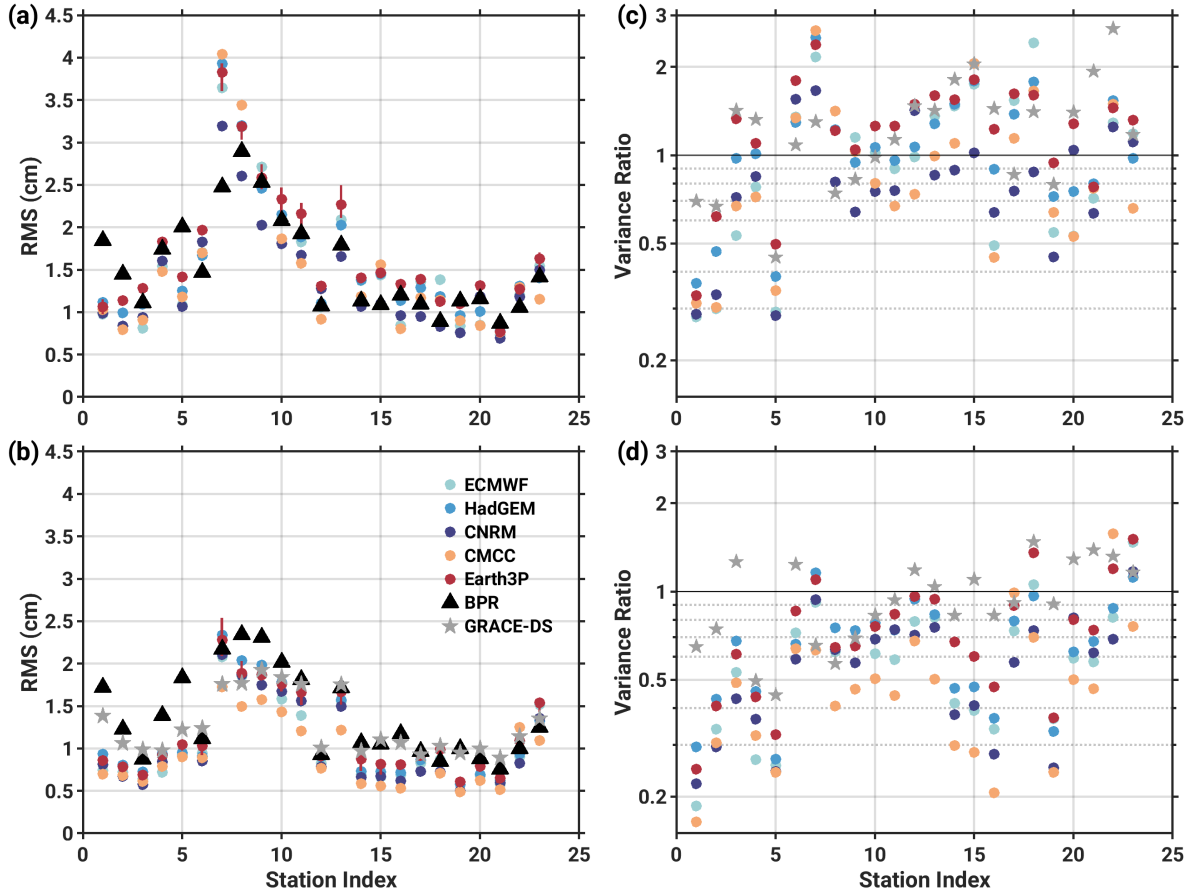
Figure 4. Comparison of the CMIP6-HR and GRACE-DS p_b variability against BPR observations at 23 locations (see Fig. 1b) in terms of RMS (left panels) and variance ratios (right panels, $R = 1$ indicated by horizontal gray line). For context, the left panels also include the 10th–90th percentile interval of the RMS from one model (Earth3P, red vertical bars). The top row shows the statistics for the full broadband p_b variations, while the bottom row is for non-seasonal signals.

barotropic ocean response to wind stress forcing scales with the inverse of both bed friction and water depth (Csanady, 1982), large p_b variations in shelf regions can occur if the models are too shallow or not frictional enough. These processes may also contribute to the inflated modeled p_b variability in the shallow parts of the Weddell and Ross Seas (Figs. 2–3). An arguably more critical factor, though, is the blocking of Antarctic ice shelf cavities in CMIP6-HR. This modeling choice constricts the solution to a narrower body of water than in the real ocean and necessarily distorts the dynamics of the region.

3.5 Comparison with in situ observations

Comparisons against BPRs in terms of signal levels (RMS), variance ratios, and median variance ratios over all 23 sites are illustrated in Fig. 4 and Table 2. The analysis complements the previous validation against GRACE-DS in relatively quiet (RMS

15

≤ 3 cm) deep-ocean regions, emphasizing in particular the Pacific; cf. the BPR locations in Fig. 1b. For additional context, we also include (but do no extensively discuss) point-wise statistics from GRACE-DS in Fig. 4. A first general observation is

355 that the BPRs have signal levels very similar to those present in the model-based and GRACE-DS p_b estimates at the tested locations. Exceptions ~~from~~ to this behavior are rather nuanced and also depend on whether unfiltered or non-seasonal signals (Fig. 4a–b) are analyzed. Specifically, we can discern a tendency for higher in situ than model ~~p_b~~ p_b variance for several sites near the equator (IDs 1–5 and 19), i.e., latitudes where baroclinic and partly non-linear dynamics govern bottom pressure changes (Piecuch, 2015; Hughes et al., 2018). The phenomenon appears to be common to both the full and filtered time series

360 at most of the noted stations.

Somewhat more puzzling, at least at first glance, is that for all sites in the North and Northeast Pacific (IDs 7–16, and partly elsewhere), the model variance ratios hover around 1–1.5 with the seasonal cycle included (Fig. 4c) but drop well below $R = 1$ once it is eliminated (Fig. 4d). This contrast points to systematic differences between BPRs and models in their representation of seasonal p_b signals. At the points analyzed here, seasonal oscillations evidently make up a larger fraction of the overall p_b

365 variance in the models than they do in the DART data. We suspect that at a resolution of 1/4°, the model bathymetry is too coarse (or too smooth) to capture the damping effects of small-scale topography on the p_b response to seasonal wind stress forcing; cf. Chen et al. (2023) for an in-depth study of the problem in the North Pacific Ocean. Accordingly, the apparent agreement between CMIP6-HR and the BPRs in the broadband case (Fig. 4a) may be partly coincidental; that is, anomalous seasonal variability in the models matches the variance due to non-seasonal local p_b fluctuations in the in situ series. ~~On the~~

370 ~~subject of topography, a final interesting case is station 22 (DART ID 44402, Southeast Block Canyon), which sits at the lower New England continental slope at a depth of about 2600 m. This setting likely suppresses mesoscale signals in bottom pressure from the nearby Gulf Stream (Hughes et al., 2018), hence the diminished RMS in Fig. 4a. However, there is little hope that such effects are represented accurately with ocean models run at ~25-km grid spacings.~~ Twin experiments with NEMO in the ORCA025 configuration, employing either a standard 1/4° batyhmetry or smoothed versions of it, could be used to test this

375 hypothesis.

Adopting a more quantitative view, median variance ratios $\overline{R}$ over all BPRs range from ~~0.76~~ 0.77 (CMCC) to ~~1.25~~ 1.28 (Earth3P) and decrease to $\overline{R} =$ ~~0.53~~ 0.47–~~0.73~~ .74 after removal of the seasonal cycle (Table 2). The comparison of CMIP6-HR with GRACE-DS on a global scale suggests, in fact, very similar ranges and model rankings. This is a reassuring result that alleviates some of our concerns as to the quality of the satellite-based p_b solutions at low signal levels (Sect. 3.3). Consistency in

380 detail is particularly seen along the North and East Pacific margins (IDs 9–20), where both GRACE-DS and the BPRs attribute lowest variance ratios to either CNRM and CMCC and highest variance ratios to Earth3P (Figs. 2–4). Thus, the DART series offer a useful point of comparison, despite their limited spatial coverage and the possible presence of short-scale signals in the records. Joint consideration of the lumped statistics in Table 2 and the GRACE-DS variance ratio maps (Figs. 2–3) nevertheless suggest that Earth3P outperforms the other models in most dynamical regimes. CNRM is least affected by the common issue

385 of excess p_b variance in deep energetic regions, whereas HadGEM comes close to Earth3P in quiescent areas. We proceed with these three models for the remainder of the analysis and note that while other choices would have been possible, the overall p_b characteristics in all five models are still very similar.
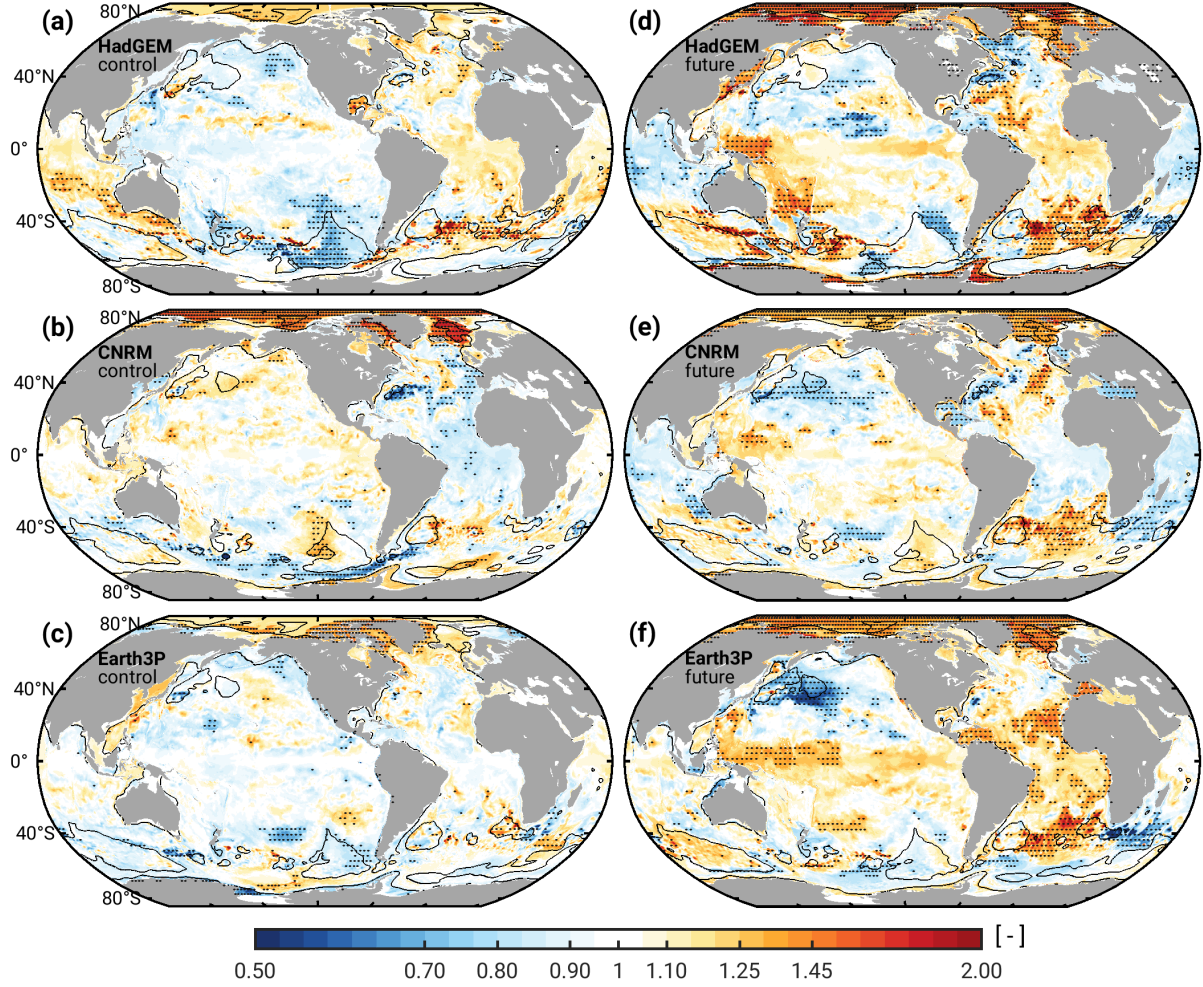
Figure 5. ~~Dimensionless variance ratios~~ Ratio of ~~unfiltered~~ p_b ~~time series~~ variance over the ~~2030–2049~~ time period 2030–2049 relative to 1980–1999 for HadGEM (top row), CNRM (middle row), and Earth3P (bottom row). Left panels (a–c) present the variance ratios ~~for p_b~~ from each model's control simulation, whereas the right panels (d–f) show the ratios formed from the ~~combined~~ scenario and historical simulations. Stipples indicate $3° \times 3°$ cells where the variance ratio is significantly different from zero at 90% confidence for at least a third of the contained $1/4°$ grid points~~(~~. To determine these confidence levels~~were determined from the proper~~, we used an F-test ~~considering~~ based on the effective degrees of freedom of the two p_b series compared at each grid point~~)~~. Black contours in all panels encase areas where the model RMS in Fig. 2, smoothed to 200 km length scales, exceeds 3 cm. Note that the color scale is logarithmic from -0.3 to 0.3, discretized at steps of 0.02.

3.6 Projected changes in p_b variability

We now examine whether the CMIP6-HR models suggest any notable changes in p_b variability under anthropogenic climate change (Fig. 5). To that end, we compute variance ratios, R, of the detrended p_b series from the three previously selected models over the time span 2030–2049 (highres-future) with respect to 1980–1999 (hist-1950 simulations). The mid-points of the two windows are therefore 50 years apart, representing the maximum separation possible with the data at hand. We additionally consider 2030–2049 vs. 1980–1999 variance ratios from the control run (Fig. 5, left column), which was integrated for 150 years with fixed forcing (Haarsma et al., 2016). These estimates give an approximate idea of the magnitude of regional p_b changes that emerge internally within the simulations in the absence of time-evolving climate conditions. Evidently, the control run variance ratios from the three models have very little in common and are mostly restricted to the range $R = 0.7$–1.3, with only a few broader statistically significant changes (e.g., in the Bellingshausen Basin for HadGEM). These findings provide useful context for to the discussion of climate-driven changes in the simulated p_b variance (see below). A region that needs to be interpreted with caution, though, is the deep Arctic Ocean, which sees strengthened p_b variability in the control run of all three models. However, that increase is statistically significant only in CNRM.

The maps of R formed using the CMIP6-HR historical and scenario runs (Fig. 5, right column) paint a relatively complex picture of future changes in p_b variance. Amplitude decreases typically alternate with increases on sub-basin scales, involving changes in regions with distinct dynamical regimes. Although the three maps show many differences in detail, we can still discern a few common, statistically significant features that point to a robust response of dynamic p_b to climate change. One such region appears to be . One region of interest is the connected deep basin of the Arctic Ocean and Nordic Seas, with widespread variance increases of 40–60%, partly also extending to adjacent continental shelves. The increase in CNRM, though, is smaller than in the control run (Figs. 5b and 5e) and therefore questions the results from HadGEM and Earth3P. To examine the matter further—and given the tight connection of p_b fluctuations in the Arctic region to wind forcing (Fukumori et al., 2015)—we plot projected changes in the zonally averaged 10-m compute variance ratios R_V for the near-surface wind speed from the three CMIP6-HR simulations ($R_V = \sigma^2_{2030-2049}/\sigma^2_{1980-1999}$). These estimates are shown for the global ocean in Fig. ??, 2030–2049 relative to 1980–1999 as before).All three models suggest enhanced surface wind speeds poleward of 60°N by 2030–2049, but in CNRM the climate-induced signal is indiscriminate from internal variability (as quantified by means of the control run; blacked dashed line in 6 and in a more focused manner for the Arctic Ocean and its neighboring seas in Supplementary Fig. S4.

A first pass of the R_V maps indicates that while the three models attribute variance gains and reductions to somewhat different geographic locations, the Arctic Ocean and northern North Atlantic are, in general, leading regions for future changes in near-surface wind speed variability. To link these changes to the modeled increase in Arctic p_b variance, a look at results in Fukumori et al. (2015) is enlightening. Their Figures 9 and 10 reveal the continental slope circling around Northeast Greenland/Canada all the way to the East Siberian Shelf as spatially connected forcing region for the near-uniform Arctic p_b fluctuation. Similarly, the fluctuation can be driven by along-slope winds south of Iceland and the Celtic Sea, but such remote contributions counteract the basin-wide bottom pressure signals excited inside the Arctic domain (Fukumori et al., 2015). This type of cancellation may
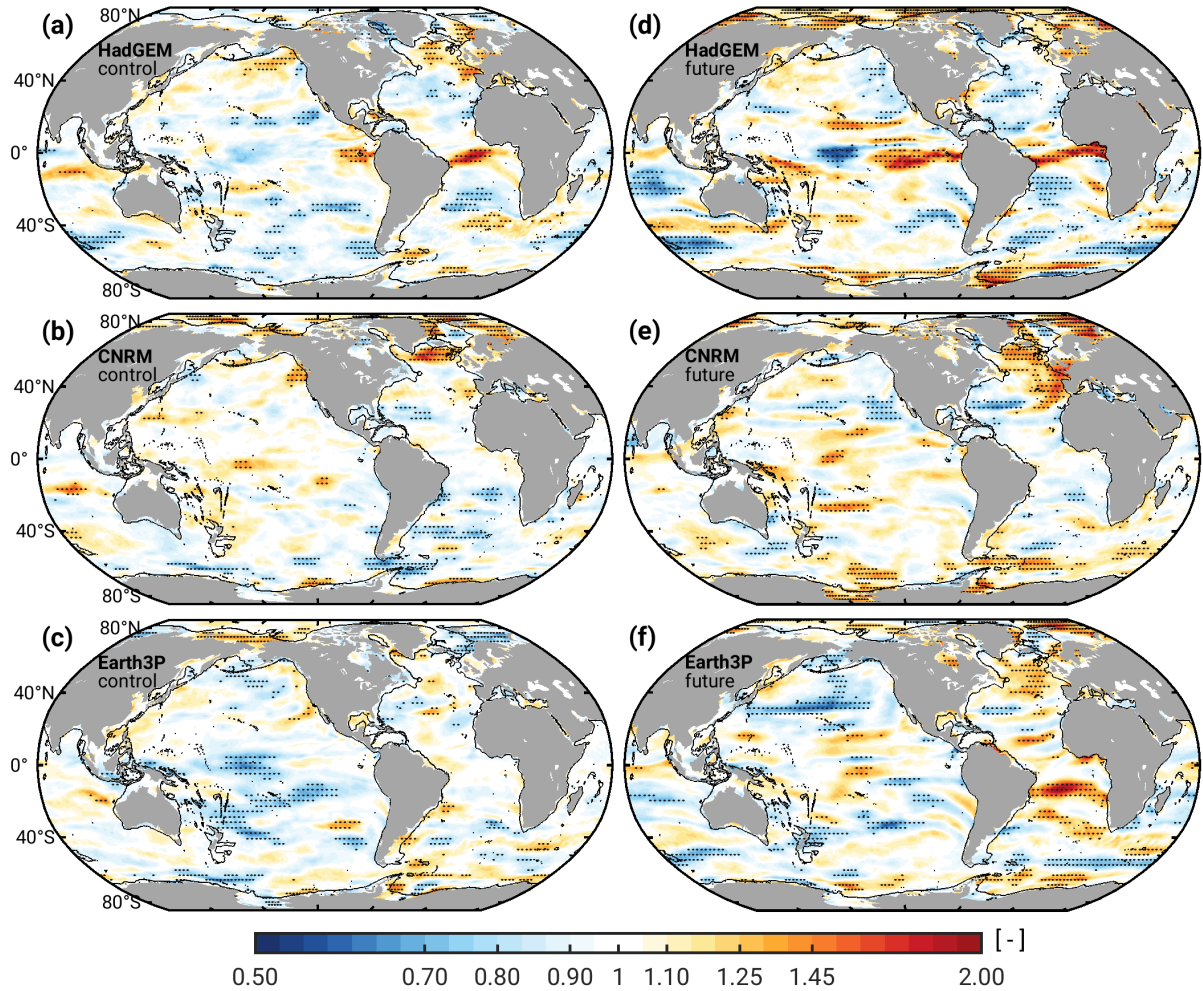
Figure 6. As in Fig. 5 but for the variance ratio R_V of the near-surface wind speed (i.e., the magnitude of the horizontal wind vector). Ratios are again formed for the period 2030–2049 relative to 1980–1999. Time-mean wind speeds do not contribute to the picture. To help pinpoint wind speed changes near the continental slope, the 1000-m isobath is included.

indeed occur in the CNRM scenario simulation (Fig. ~~??b). However, similar analyses of high-latitude surface wind speeds from most available 1° CMIP6 models (??)~~ rather favor the strengthening evident in HadGEM and ~~5e),~~ which suggests increases in wind speed variance both in the Celtic Sea and over the continental slope off Northeast Greenland (Fig. 6e, Supplementary

425 Fig. S4e). In contrast, HadGEM, Earth3P ~~We therefore conclude that the increase in simulated Arctic p_b variability by the mid of the 21st century is indeed attributable to anthropogenic forcing and specifically caused by accelerated surface winds.~~, and the CNRM control run see little change in near-surface wind speed variability over remote forcing regions. Rather, patches of $R_V > 1$ sit over relevant continental slope sections inside the Arctic region, i.e., northeast of Greenland in Earth3P and CNRM-control, and north of the Canadian Arctic Archipelago in HadGEM. Enhanced variability of the associated wind stress

19

430   could trigger more frequent or more energetic trapped Kelvin waves (Sect. 3.1), thus amplifying the basin-wide Arctic p_b fluctuation in the three models (Fig. 5, panels b, d, f).

~~Change in zonally averaged global surface wind speed by 2030–2049 relative to 1980–1999 from the (a) HadGEM, (b) CNRM, and (c) Earth3P simulations (unit is m s^{-1}, land areas are included). Colored meridional profiles show the results from the scenario and historical runs (cf. Fig. 5), while the black dotted line indicates the corresponding profile from the control run.~~

435   Along with some interesting cases of diminishing amplitudes (e.g., Northwest Pacific, Gulf Stream region), we also find an intensification of p_b variability under SSP5-8.5 in the western tropical Pacific. North of New Guinea, values of R peak at $R \approx 1.3$–1.5. The area is characterized by low-magnitude p_b signals (Figs. 1 and 2), ~~induced by~~ which are forced by winds, involve a mix of ~~both~~ barotropic and baroclinic processes~~that tend to be~~, and are most active at annual and semiannual frequencies (e.g., Qu et al., 2008; Piecuch et al., 2015; Cheng et al., 2021). Indeed, the patch of intensified p_b amplitudes in the

440 western tropical Pacific vanishes when variance ratios are formed based on non-seasonal model time series (Supplementary Fig. ~~S1). Although more definitive diagnosis of the processes at work is beyond our scope, we note that~~ S5). We also note that in Fig. 5, both HadGEM and Earth3P show increases in R (non-significant) over large parts of the tropical Pacific Ocean~~have positive values of R in both HadGEM and Earth3P. This in turn hints at changes in local Ekman pumping or~~. The patterns could hint at broader changes in the wind-driven ~~Rossby waves, possibly related to increases in stratification or migration of~~

445 ~~the~~ component of the seasonal cycle around the Inter-Tropical Convergence Zone~~(cf. Qu et al., 2008; Piecuch et al., 2015).~~. A more definite diagnosis of the processes at work is beyond our scope, though.

  The third salient feature in Fig. 5, seen across all three models, is a mid-21st century intensification of $\sim$30–60% in p_b variance in the southern South Atlantic. The patch of red color sits between the Zapiola Gyre and Cape Basin (cf. Fig. 1a) and is accompanied in eastward direction by a decrease in p_b variability over the Agulhas Return Current. ~~These structures are~~

450 ~~consistent with the patterns of~~ Closer inspection of the corresponding R_V maps (Fig. 6) fails to reveal a common, larger-scale change in wind speed variance that would hint at an origin of the p_b modifications in forced barotropic dynamics. In particular, it is unclear how the east-west dipole in p_b changes centered around the dip of South Africa could arise from variations in surface winds and associated quantities (e.g., wind stress curl).

  Instead, we suspect that the p_b structures in Fig. 5 around the junction of the Atlantic and Indian oceans largely bear

455 on the projected changes in the regions~~'~~'s eddy kinetic energy (EKE, Beech et al., 2022; Wang et al., 2024). This is not an unreasonable connection to make, as previous model analyses (Hughes et al., 2018; Zhao et al., 2021) suggest a clear imprint of mesoscale variability on p_b in the areas in question. In terms of processes, ~~as implied by future climate simulations (Beech et al., 2022; Wang et al., 2024). In particular,~~ global warming is expected to strengthen Southern Hemisphere westerlies and progressively shift them to the south ~~(Deng et al., 2022, see also Fig. ??), thus weakening westerly winds around the~~

460 ~~southern coast of Africa ($\sim$35°S). Decelerated winds~~ (Deng et al., 2022, see also Supplementary Fig. S6). This change in the large-scale wind forcing will in turn reduce volume transport and ~~eddy kinetic energy~~ EKE in the Agulhas Current, while also leading to increases in Agulhas leakage and mesoscale activity in the southern South Atlantic (Biastoch et al., 2009; van Sebille et al., 2009; Beech et al., 2022). ~~That these dynamics imprint on p_b is clear from Fig. 2 and previous model analyses~~
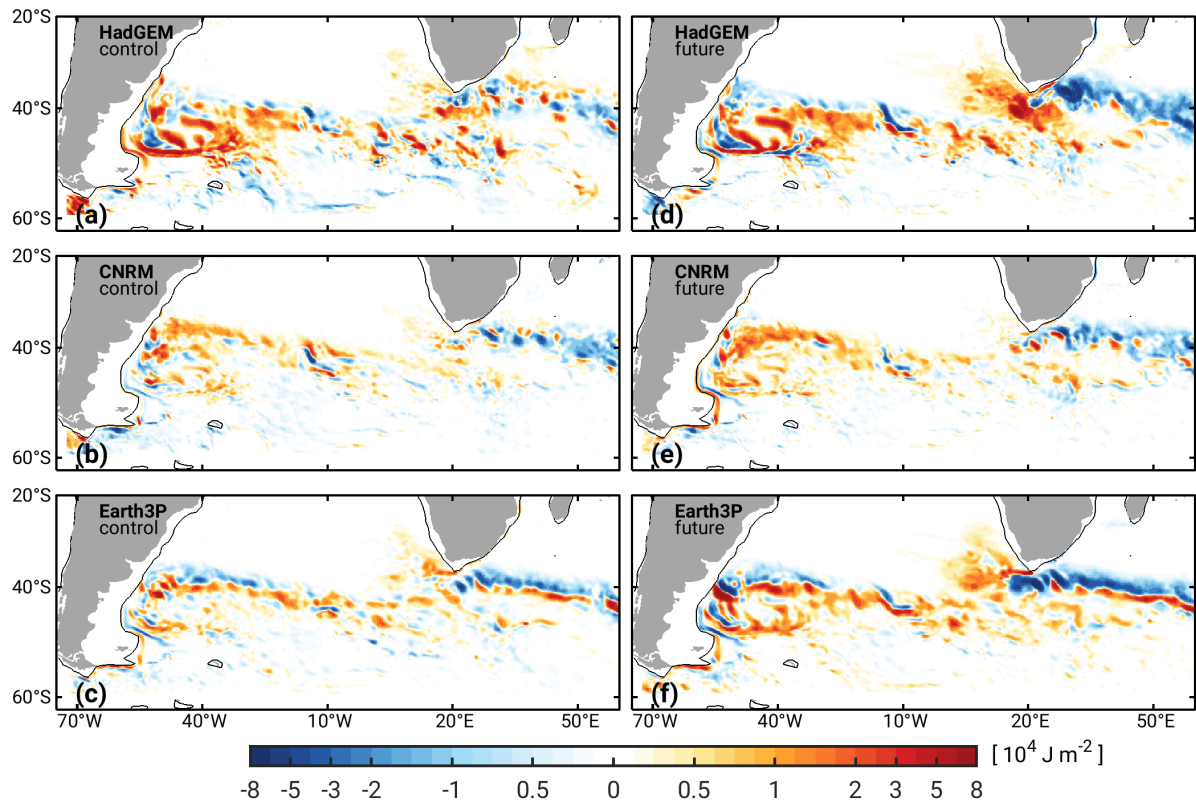
Figure 7. Changes in depth-integrated time-mean EKE (J m^{-2}), 2030–2049 relative to 1980–1999, around the southern South Atlantic. As in Fig. 5, we present results from HadGEM (top row), CNRM (middle row), and Earth3P (bottom row). Left panels (a–c) show the variance ratios from each model's control simulation, whereas the right panels (d–f) show the ratios formed from the scenario and historical simulations. The black solid line is the 1000-m isobath. Note the non-linear color scale.

(Hughes et al., 2018; Zhao et al., 2021). In addition, climate change invigorates eddy activity in the western part of the basin
465 (Beech et al., 2022), and that too could contribute to the enhanced future p_b variance seen in Fig. 5. Thus, the

We have examined the three CMIP6-HR models for some of the described processes and specifically deduced estimates for the depth-integrated EKE (Ni et al., 2023, Sect. S1 in the Supplementary) in the region. Figure 7 illustrates the resultant changes in EKE by 2030–2049 relative to 1980–1999. As evident from the control run (left column), the full-depth EKE is subjected to considerable changes in distribution and magnitude even under constant climatic conditions. The response
470 under SSP5-8.5 (right column) is nevertheless more pronounced, exhibiting also a good degree of consistency among the three models. Shared characteristics are an increase in eddy activity in and west of the Argentine Basin (with signs of a southward shift in HadGEM and Earth3P) and reduced full-depth EKE along the path of the Agulhas Return Current. Agulhas leakage clearly intensifies by the middle of the 21st century in HadGEM and Earth3P, but has not yet emerged in CNRM. Overall, the EKE changes produced by CNRM are rather subdued, and so are the changes in the model's p_b variance in the

analyzed region (Fig. 5e). In comparing the variations in oceanic EKE and p_b, note that there is no prerequisite for an exact spatial correspondence between the two variables, as the mesoscale imprint on p_b is generally spread out wider than eddies themselves (Zhao et al., 2021). We leave the analysis with these pointers and conclude that the southern South Atlantic and Agulhas Return Current region ~~appear to be~~ are an interesting laboratory for linking long-term changes in p_b variability with anthropogenic impacts on the ocean circulation.

4   Summary and conclusions

We have run basic and mostly global assessments of bottom pressure signals in CMIP6-HR simulations, motivated in large part by the potential use of these fields in satellite gravimetry simulation studies. The climate models capture all relevant p_b phenomena that can be expected to emerge at the given resolution and under the encoded physics, e.g., eddy structures and wind-driven variability on continental shelves and in the deep ocean. Model accuracy, as evaluated through comparisons to a downscaled GRACE product and BPR measurements, varies with latitude, depth, frequency, and more generally the specific dynamical regime (Figs. 2 and 3, Table 2). In particular, the CMIP6-HR models tested here are susceptible to overestimation of topographically constrained p_b variability, most notably in the Australian-Antarctic and Weddell-Enderby Abyssal plains and at seasonal frequencies. Similar issues are apparent in NEMO-based ocean reanalyses on the ORCA025 grid (cf. Figure 3 in Gou et al., 2025), pointing to structural errors in that particular ocean model setup. As suggested in Sect. ~~3.2~~3.4, the excess variance may have diverse causes, including imperfect momentum transfer schemes, coarse vertical spacing of deep layers, and the artificial blocking of ice shelf cavities. In this light, analysis of p_b variability may be useful to guide model improvements or at least complement the standard tests of near-surface variables and meridional overturning (e.g., Zuo et al., 2017; Roberts et al., 2019, 2020).

We have exercised a degree of caution in our comparisons of the CMIP6-HR p_b fields against GRACE-DS and BPRs. Neither reference dataset is free of error, whereas the models only approximate reality and thus disregard several legitimate contributions to the observed p_b changes (e.g., GAL, barometric pressure-driven dynamics, sub-grid scale variability). These issues can compromise the model-data comparison, particularly in very quiet deep-ocean regions (Fig. 3) and at individual BPRs (Fig. 4). The spatial median statistics in Table 2 are nonetheless conclusive and internally consistent enough to discriminate more accurate models from less realistic p_b representations. Specifically, in satellite gravimetry simulations, where one would want to portray oceanic mass changes across all spatial and temporal scales, the HighResMIP contribution of Earth3P appears to be a very sensible choice. One minor caveat is that compared to the models (RMS distribution analyzed over 1980–2014), the BPR and GRACE-DS variance estimates are taken over different ~~and much shorter~~ time periods, i.e., 2013–2023 and $\sim$2002–2020. However, when narrowing the model time series to 2002–2014 (Supplementary Table ~~S1~~S3), our basic conclusions drawn from Table 2 ~~remained~~ remain unchanged, except for low-latitude shelf regions. The agreement is not unexpected, because the climatic conditions over the CMIP6-HR late historical period ~~do not evolve as drastically as~~ evolve less drastically than they do in the scenario extension to 2049, implying temporal changes in p_b variance that are well below those shown in ~~Fig~~Figs. 5d–f.

Our model-based diagnosis of future climate impacts on the strength of p_b fluctuations has added a new facet to the existing bottom pressure literature. While these impacts are spread over regions characterized by very different dynamics (i.e., Arctic, tropical Pacific, Agulhas leakage), an origin in altered wind forcing ~~(Fig. ??)~~ is likely common to all three cases. Dedicated modeling work, such as stand-alone ocean simulations under distinct forcing assumptions, will be required to shed light on the exact processes at work. More generally, the reported strengthening of regional p_b variability awaits confirmation through analysis of output from other climate models~~and perhaps, at some point , also through~~ . Consideration of a larger ensemble of simulations would also help better distinguish climate-induced p_b changes from model internal variability (cf. the case of CNRM in Sect. 3.6). At some point it might even be possible to detect the projected p_b changes in actual observations. Here we have made our inferences based on variance metrics in bi-decadal windows 50 years apart, but the noted p_b amplitude increases in the Arctic, tropical Pacific, and parts of the South Atlantic also emerge when using a reference periods shifted forward by 15 years (1995–2014, Supplementary Fig. ~~S2~~S7). Thus, given the planned continuity of space-based mass change monitoring (Daras et al., 2024), climate-driven changes in large-scale, dynamic p_b variability may be soon detectable with satellite gravimetry.

. **L. L.**: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **M. S.**: Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Writing - original draft, Writing - review & editing. **L. B.**: Software, Writing - review & editing. **J. F.**: Formal analysis, Software, Writing - review & editing. **J. G.**: Methodology, Writing - review & editing.

References

535 Adcroft, A., Hughes, C., and Marshall, J.: Representation of topography by shaved cells in height coordinate ocean model, Mon. Weather Rev., 125, 2293–2315, https://doi.org/10.1175/1520-0493(1997)125<2293:ROTBSC>2.0.CO;2, 1997.

Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24, https://doi.org/10.7289/V5C8276M, 2009.

Androsov, A., Boebel, O., Schröter, J., Danilov, S., Macrander, A., and Ivanciu, I.: Ocean bottom pressure variability: Can it be reliably
540 modeled?, J. Geophys. Res. Oceans, 125, e2019JC015 469, https://doi.org/10.1029/2019JC015469, 2020.

Beech, N., Rackow, T., Semmler, T., Danilov, S., Wang, Q., and Jung, T.: Long-term evolution of ocean eddy activity in a warming world, Nat. Clim. Change, 12, 910–917, https://doi.org/10.1038/s41558-022-01478-3, 2022.

Bernard, B., Madec, G., Penduff, T., Molines, J.-M., Treguier, A.-M., Le Sommer, J., Beckmann, A., Biastoch, A., Böning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theeten, M., McClean, J., and De Cuevas, B.: Impact of partial steps
545 and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution, Ocean Dynam., 56, 543–567, https://doi.org/doi.org/10.1007/s10236-006-0082-1, 2006.

Bernard, E. N. and Meinig, C.: History and future of deep-ocean tsunami measurements, in: Proceedings of Oceans'11 MTS/IEEE Kona, 6106894, pp. 1–7, IEEE, 2011.

Biastoch, A., Böning, C. W., Schwarzkopf, F. U., and Lutjeharms, J. R. E.: Increase in Agulhas leakage due to poleward shift of Southern
550 Hemisphere westerlies, Nature, 462, 495–498, https://doi.org/10.1038/nature08519, 2009.

Bingham, R. and Hughes, C.: The relationship between sea-level and bottom pressure variability in an eddy-permitting ocean model, Geophys. Res. Lett., 35, L03 602, https://doi.org/10.1029/2007GL032662, 2008.

Bingham, R. J. and Haines, K.: Decadal ocean bottom pressure variability and its associated gravitational effects in a coupled ocean-atmosphere model, in: Gravity, Geoid and Space Missions: GGSM 2004 IAG International Symposium Porto, pp. 298–303, Springer,
555 https://doi.org/10.1007/3-540-26932-0_52, 2005.

Boening, C., Lee, T., and Zlotnicki, V.: A record-high ocean bottom pressure in the South Pacific observed by GRACE, Geophys. Res. Lett., 38, L04 602, https://doi.org/10.1029/2010GL046013, 2011.

Börger, L., Schindelegger, M., Dobslaw, H., and Salstein, D.: Are ocean reanalyses useful for Earth rotation research?, Earth Space Sci., 10, e2022EA002 700, https://doi.org/10.1029/2022EA002700, 2023.

560 Chen, L., Yang, J. Y., and Wu, L. X.: Topography effects on the seasonal variability of ocean bottom pressure in the North Pacific Ocean, J. Phys. Oceanogr., 53, 929–941, https://doi.org/10.1175/JPO-D-22-0140.1, 2023.

Cheng, X., Ou, N. S., Chen, J. J., and Huang, R. X.: On the seasonal variations of ocean bottom pressure in the world oceans, Geosci. Lett., 8, 29, https://doi.org/10.1186/s40562-021-00199-3, 2021.

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Scoccimarro, E., Materia, S., Bellucci, A., and Navarra,
565 A.: Global mean climate and main patterns of variability in the CMCC-CM2 coupled model, J. Adv. Model. Earth Sy., 11, 185–209, https://doi.org/10.1029/2018MS001369, 2019.

Csanady, G. T.: Circulation in the Coastal Ocean, Springer, https://doi.org/10.1007/978-94-017-1041-1, 1982.

Daras, I., March, G., Pail, R., Hughes, C. W., Braitenberg, C., Güntner, A., Eicker, A., Wouters, B., Heller-Kaikov, B., Pivetta, T., and Pastorutti, A.: Mass-change And Geosciences International Constellation (MAGIC) expected impact on science and applications,
570 Geophys. J. Int., 236, 1288–1308, https://doi.org/10.1093/gji/ggad472, 2024.

De Linage, C., Rivera, L., Hinderer, J., Boy, J.-P., Rogister, Y., Lambotte, S., and Biancale, R.: Separation of coseismic and postseismic gravity changes for the 2004 Sumatra–Andaman earthquake from 4.6 yr of GRACE observations and modelling of the coseismic change by normal-modes summation, Geophys. J. Int., 176, 695–714, https://doi.org/10.1111/j.1365-246X.2008.04025.x, 2009.

de Miranda, A. P., Barnier, B., and Dewar, W. K.: On the dynamics of the Zapiola Anticyclone, J. Geophys. Res. Oceans, 104, 21 137–21 149, https://doi.org/10.1029/1999JC900042, 1999.

Deng, K., Azorin-Molina, C., Yang, S., Hu, C., Zhang, G., Minola, L., and Chen, D.: Changes of Southern Hemisphere westerlies in the future warming climate, Atmos. Res., 270, 106 040, https://doi.org/10.1016/j.atmosres.2022.106040, 2022.

Docquier, D., Grist, J. P., Roberts, M. J., Roberts, C. D., Semmler, T., Ponsoni, L., Massonnet, F., Sidorenko, D., Sein, D. V., Iovino, D., Bellucci, A., and Fichefet, T.: Impact of model resolution on Arctic sea ice and North Atlantic Ocean heat transport, Clim. Dynam., 53, 4989–5017, https://doi.org/10.1007/s00382-019-04840-y, 2019.

Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C.: ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation, Geosci. Model Dev., 8, https://doi.org/10.5194/gmd-8-3071-2015, 2015.

Fukumori, I., Raghunath, R., and Fu, L. L.: Nature of global large-scale sea level variability in relation to atmospheric forcing: A modeling study, J. Geophys. Res. Oceans, 103, 5493–5512, https://doi.org/10.1029/97JC02907, 1998.

Fukumori, I., Wang, O., Llovel, W., Fenty, I., and Forget, G.: A near-uniform fluctuation of ocean bottom pressure and sea level across the deep ocean basins of the Arctic Ocean and the Nordic Seas, Prog. Oceanogr., 134, 152–172, https://doi.org/10.1016/j.pocean.2015.01.0135, 2015.

Gałdyn, F. and Sośnica, K.: Impact of the combination and replacement of SLR-based low-degree gravity field coefficients in GRACE solutions, Prog. Earth Planet. Sci., 11, 7, https://doi.org/10.1186/s40645-024-00608-z, 2024.

Gill, A. E. and Niller, P. P.: The theory of the seasonal variability in the ocean, Deep Sea Res. Oceanogr. Abstr., 20, 141–177, https://doi.org/10.1016/0011-7471(73)90049-1, 1973.

Gou, J. Y., Börger, L., Schindelegger, M., and Soja, B.: SeDA-OBP: Global ocean bottom pressure anomaly fields with a spatial resolution of 0.25 degrees from self-supervised data assimilation, ETH Research Collection, https://doi.org/10.3929/ethz-b-000686843, 2024.

Gou, J. Y., Börger, L., Schindelegger, M., and Soja, B.: Downscaling GRACE-derived ocean bottom pressure anomalies using self-supervised data fusion, J. Geodesy, 99, 19, https://doi.org/10.1007/s00190-025-01943-9, 2025.

Haarsma, R., Acosta, M., Bakhshi, R., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Corti, S., Davini, P., Exarchou, E., Fabiano, F., Fladrich, U., Fuentes Franco, R., García-Serrano, J., von Hardenberg, J., Koenigk, T., Levine, X., Meccia, V. L., van Noije, T., van den Oord, G., Palmeiro, F. M., Rodrigo, M., Ruprich-Robert, Y., Le Sager, P., Tourigny, E., Wang, S., van Weele, M., and Wyser, K.: HighResMIP versions of EC-Earth: EC-Earth3P andEC-Earth3P-HR – description, model computational performance and basic validation, Geosci. Model Dev., 13, 3507–3527, https://doi.org/10.5194/gmd-13-3507-2020, 2020.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Harker, A. A., Schindelegger, M., Ponte, R. M., and Salstein, D. A.: Modeling ocean-induced rapid Earth rotation variations: An update, J. Geodesy, 95, 110, https://doi.org/10.1007/s00190-021-01555-z, 2021.

Hart-Davis, M. G., Piccioni, G., Dettmering, D., Schwatke, C., Passaro, M., and Seitz, F.: EOT20: A global ocean tide model from multi-mission satellite altimetry, Earth Syst. Sci. Data, 13, 3869–3884, https://doi.org/10.5194/essd-13-3869-2021, 2021.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Astron. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hughes, C. W., Stepanov, V. N., Fu, L. L., Barnier, B., and Hargreaves, G. W.: Three forms of variability in Argentine Basin ocean bottom pressure, J. Geophys. Res. Oceans, 112, https://doi.org/10.1029/2006JC003679, 2007.

Hughes, C. W., Williams, J., Blaker, A., Coward, A., and Stepanov, V.: A window on the deep ocean: The special value of ocean bottom pressure for monitoring the large-scale, deep-ocean circulation, Prog. Oceanogr., 161, 19–46, https://doi.org/10.1016/j.pocean.2018.01.011, 2018.

Jensen, L., Eicker, A., Dobslaw, H., and Pail, R.: Emerging changes in terrestrial water storage variability as a target for future satellite gravity missions, Remote Sens., 12, 3898, https://doi.org/10.3390/rs12233898, 2020.

Johnson, G. C. and Chambers, D. P.: Ocean bottom pressure seasonal cycles and decadal trends from GRACE Release-05: Ocean circulation implications, J. Geophys. Res. Oceans, 118, 4228–4240, https://doi.org/10.1002/jgrc.20307, 2013.

Kemp, J., Newhall, K., Ostrom, W., Krishfield, R., and Proshutinsky, A.: The Beaufort Gyre observing system 2004: Mooring recovery and deployment operations in pack ice, Woods Hole Oceanographic Institution, 2005.

Kvas, A., Behzadpour, S., Ellmer, M., Klinger, B., Strasser, S., Zehentner, N., and Mayer-Gürr, T.: ITSG-Grace2018: Overview and evaluation of a new GRACE-only gravity field time series, J. Geophys. Res. Sol. Ea., 124, 9332–9344, https://doi.org/10.1029/2019JB017415, 2019.

Landerer, F. W., Jungclaus, J. H., and Marotzke, J.: Ocean bottom pressure changes lead to a decreasing length-of-day in a warming climate, Geophys. Res. Lett., 34, L06 307, https://doi.org/10.1029/2006GL029106, 2007.

Lellouche, J.-M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., Ferry, N., Desportes, C., Testut, C.-E., Bricaud, C., Bourdallé-Badie, R., Tranchant, B., Benkiran, M., Drillet, Y., Daudin, A., and De Nicola, C.: Evaluation of global monitoring and forecasting systems at Mercator Océan, Ocean Sci., 9, 57–81, https://doi.org/10.5194/os-9-57-2013, 2013.

Liu, L., Schindelegger, M., and Börger, L.: Monthly ocean bottom pressure anomalies 1980-2014 and 2030-2049 from CMIP6 HighResMIP models, Zenodo, https://doi.org/10.5281/zenodo.14886818, 2025.

Ludwigsen, C. B., Andersen, O. B., Marzeion, B., Malles, J.-H., Müller Schmied, H., Döll, P., Watson, C., and King, M. A.: Global and regional ocean mass budget closure since 2003, Nat. Commun., 15, 1416, https://doi.org/10.1038/s41467-024-45726-w, 2024.

Madec, G.: NEMO reference manual 3.6_STABLE: NEMO ocean engine, Institut Pierre-Simon Laplace (IPSL), France, 2016.

Makowski, J. K., Chambers, D. P., and Bonin, J. A.: Using ocean bottom pressure from the gravity recovery and climate experiment (GRACE) to estimate transport variability in the southern Indian Ocean, J. Geophys. Res. Oceans, 120, 4245–4259, https://doi.org/10.1002/2014JC010575, 2015.

McDougall, T. J. and Barker, P. M.: Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox, Scor/Iapso WG, 127, 1–28, 2011.

Ni, Q., Zhai, X., LaCasce, J. H., Chen, D., and Marshall, D. P.: Full-depth eddy kinetic energy in the global ocean estimated from altimeter and Argo observations, Geophys. Res. Lett., 50, e2023GL103 114, https://doi.org/10.1029/2023GL103114, 2023.

Oliver, E. C. J. and Thompson, K. R.: Sea level and circulation variability of the Gulf of Carpentaria: Influence of the Madden-Julian Oscillation and the adjacent deep ocean, J. Geophys. Res. Oceans, 116, C02 019, https://doi.org/10.1029/2010JC006596, 2011.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M.: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6, Geosci. Model Dev., 9, 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016, 2016.

650 Petrick, C., Dobslaw, H., Bergmann-Wolf, I., Schön, N., Matthes, K., and Thomas, M.: Low-frequency ocean bottom pressure variations in the North Pacific in response to time-variable surface winds, J. Geophys. Res. Oceans, 119, 5190–5202, https://doi.org/10.1002/2013JC009635, 2014.

Piecuch, C. G.: Bottom-pressure signature of annual baroclinic Rossby waves in the northeast tropical Pacific Ocean, J. Geophys. Res. Oceans, 120, 2449–2459, https://doi.org/10.1002/2014JC010667, 2015.

Piecuch, C. G., Fukumori, I., Ponte, R. M., and Wang, O.: Vertical structure of ocean pressure variations with application to satellite-
655 gravimetric observations, J. Atmos. Ocean. Tech., 32, 603–613, https://doi.org/10.1175/JTECH-D-14-00156.1, 2015.

Piecuch, C. G., Fukumori, I., Ponte, R. M., Schindelegger, M., Wang, O., and Zhao, M.: Low-frequency dynamic ocean response to barometric-pressure loading, J. Phys. Oceanogr., 52, 2627–2641, https://doi.org/10.1175/JPO-D-22-0090.1, 2022.

Ponte, R. M.: Variability in a homogeneous global ocean forced by barometric pressure, Dynam. Atmos. Ocean., 18, 209–234, https://doi.org/10.1016/0377-0265(93)90010-5, 1993.

660 Ponte, R. M.: A preliminary model study of the large-scale seasonal cycle in bottom pressure over the global ocean, J. Geophys. Res. Oceans, 104, 1289–1300, https://doi.org/10.1029/1998JC900028, 1999.

Ponte, R. M. and Piecuch, C. G.: Interannual bottom pressure signals in the Australian–Antarctic and Bellingshausen basins, J. Phys. Oceanogr., 44, 1456–1465, https://doi.org/10.1175/JPO-D-13-0223.1, 2014.

Ponte, R. M., Rajamony, J., and Gregory, J. M.: Ocean angular momentum signals in a climate model and implications for Earth rotation,
665 Clim. Dynam., 19, 181–190, https://doi.org/10.1007/s00382-001-0216-6, 2002.

Ponte, R. M., Zhao, M., and Schindelegger, M.: How well do we know the seasonal cycle in ocean bottom pressure?, Earth Space Sci., 11, e2024EA003 661, https://doi.org/10.1029/2024EA003661, 2024.

Poropat, L., Dobslaw, H., Zhang, L., Macrander, A., Boebel, O., and Thomas, M.: Time variations in ocean bottom pressure from a few hours to many years: In situ data, numerical models, and GRACE satellite gravimetry, J. Geophys. Res. Oceans, 123, 5612–5623,
670 https://doi.org/10.1029/2018JC014108, 2018.

Qu, T., Gan, J., Ishida, A., Kashino, Y., and Tozuka, T.: Semiannual variation in the western tropical Pacific Ocean, Geophys. Res. Lett., 35, L16 602, https://doi.org/10.1029/2008GL035058, 2008.

Ray, R. D., Loomis, B. D., and Zlotnicki, V.: The mean seasonal cycle in relative sea level from satellite altimetry and gravimetry, J. Geodesy, 95, 80, https://doi.org/10.1007/s00190-021-01529-1, 2021.

675 Roberts, C., Senan, R., Molteni, F., Boussetta, S., Mayer, M., and Keeley, S.: Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP, Geosci. Model Dev., 11, 3681–3712, https://doi.org/10.5194/gmd-11-3681-2018, 2018.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., Jackson, L. C., Kuhlbrodt, T., Mathiot, P., Roberts, C. D., Schiemann, R., Seddon, J., Vannière, B., and Vidale, P. L.: Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1
680 model as used in CMIP6 HighResMIP experiments, Geosci. Model Dev., 12, 4999–5028, https://doi.org/10.5194/gmd-12-4999-2019, 2019.

Roberts, M. J., Jackson, L. C., Roberts, C. D., Meccia, V., Docquier, D., Koenigk, T., Ortega, P., Moreno-Chamarro, E., Bellucci, A., Coward, A., Drijfhout, S., Exarchou, E., Gutjahr, O., Hewitt, H., Iovino, D., Lohmann, K., Putrasahan, D., Schiemann, R., Seddon, J., Terray, L.,

27

Xu, X., Zhang, Q., Chang, P., Yeager, S. G., Castruccio, F. S., Zhang, S., and Wu, L.: Sensitivity of the Atlantic Meridional Overturning

685   Circulation to model resolution in CMIP6 HighResMIP simulations and implications for future changes, J. Adv. Model. Earth Sy., 12,
      e2019MS002 014, https://doi.org/10.1029/2019MS002014, 2020.

Save, H., Bettadpur, S., and Tapley, B. D.: High-resolution CSR GRACE RL05 mascons, J. Geophys. Res. Sol. Ea., 121, 7547–7569,
      https://doi.org/10.1002/2016JB013007, 2016.

Schindelegger, M., Harker, A. A., Ponte, R. M., Dobslaw, H., and Salstein, D. A.: Convergence of daily GRACE solutions and models of

690   submonthly ocean bottom pressure variability, J. Geophys. Res. Oceans, 126, e2020JC017 031, https://doi.org/10.1029/2020JC017031,
      2021.

Sen Gupta, A., Jourdain, N. C., Brown, J. N., and Monselesan, D.: Climate drift in the CMIP5 models, J. Climate, 26, 8597–8615,
      https://doi.org/10.1175/JCLI-D-12-00521.1, 2013.

Shi, J. R., Talley, L. D., Xie, S. P., Peng, Q. H., and Liu, W.: Ocean warming and accelerating Southern Ocean zonal flow, Nat. Clim. Change,

695   11, 1090–1097, https://doi.org/10.1038/s41558-021-01212-5, 2021.

Shihora, L., Balidakis, K., Dill, R., Dahle, C., Ghobadi-Far, K., Bonin, J., and Dobslaw, H.: Non-tidal background modeling for satellite
      gravimetry based on operational ECWMF and ERA5 reanalysis data: AOD1B RL07, J. Geophys. Res. Sol. Ea., 127, e2022JB024 360,
      https://doi.org/10.1029/2022JB024360, 2022.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., Calvert, D., Graham, T., Hewitt, H. T., Hyder, P., Kuhlbrodt,

700   T., Rae, J. G. L., and Sinha, B.: UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions, Geosci. Model Dev., 11,
      3187–3213, https://doi.org/10.5194/gmd-11-3187-2018, 2018.

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J. S., Landerer, F. W., Chambers,
      D. P., Reager, J. T., Gardner, A. S., Himanshu, S., Ivins, E. R., Swenson, S. C., Boening, C., Dahle, C., Wiese, D. N., Dobslaw,
      H., Tamisiea, M. E., and Velicogna, I.: Contributions of GRACE to understanding climate change, Nat. Clim. Change, 9, 358–369,

705   https://doi.org/10.1038/s41558-019-0456-2, 2019.

van Dam, T., Collilieux, X., Wuite, J., Altamimi, Z., and Ray, J.: Nontidal ocean loading: amplitudes and potential effects in GPS height time
      series, J. Geodesy, 86, 1043–1057, https://doi.org/10.1007/s00190-012-0564-5, 2012.

van Sebille, E., Biastoch, A., Van Leeuwen, P., and De Ruijter, W.: A weaker Agulhas Current leads to more Agulhas leakage, Geophys. Res.
      Lett., 36, L03 601, https://doi.org/10.1029/2008GL036614, 2009.

710   Vinogradova, N. T., Ponte, R, M., and Stammer, D.: Relation between sea level and bottom pressure and the vertical dependence of oceanic
      variability, Geophys. Res. Lett., 34, L03 608, https://doi.org/10.1029/2006GL028588, 2007.

Vinogradova, N. T., Ponte, R. M., Tamisiea, M. E., Quinn, K. J., Hill, E. M., and Davis, J. L.: Self-attraction and loading effects on ocean mass
      redistribution at monthly and longer time scales, J. Geophys. Res. Oceans, 116, C08 041, https://doi.org/doi.org/10.1029/2011JC007037,
      2011.

715   Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., Colin, J., Guérémy, J.-F., Michou, M., Moine, M.-P.,
      Nabat, P., Roehrig, R., Salas y Mélia, D., Séférian, R., Valcke, S., Beau, I., Belamari, S., Berthet, S., Cassou, C., Cattiaux, J., Deshayes, J.,
      Douville, H., Ethé, C., Franchistéguy, L., Geoffroy, O., Lévy, C., Madec, G., Meurdesoif, Y., Msadek, R., Ribes, A., Sanchez-Gomez, E.,
      Terray, L., and Waldman, R.: Evaluation of CMIP6 DECK experiments with CNRM-CM6-1, J. Adv. Model. Earth Sy., 11, 2177–2213,
      https://doi.org/10.1029/2019MS001683, 2019.

720   Wang, S., Jing, Z., Wu, L., Sun, S., Chen, Z., Ma, X., and Gan, B.: A more quiescent deep ocean under global warming, Nat. Clim. Change,
      14, 961–967, https://doi.org/10.1038/s41558-024-02075-2, 2024.

Watkins, M. M., Wiese, D. N., Yuan, D.-N., Boening, C., and Landerer, F. W.: Improved methods for observing Earth's time variable mass distribution with GRACE using spherical cap mascons, J. Geophys. Res. Sol. Ea., 120, 2648–2671, https://doi.org/10.1002/2014JB011547, 2015.

725 Weijer, W.: Modal variability in the Southeast Pacific Basin: Energetics of the 2009 event, Deep Sea Res. Pt. II, 114, 3–11, https://doi.org/10.1016/j.dsr2.2012.10.002, 2015.

Weijer, W. and Gille, S. T.: Adjustment of the Southern Ocean to wind forcing on synoptic time scales, J. Phys. Oceanogr., 35, 2076–2089, https://doi.org/10.1175/JPO2801.1, 2005.

Wiese, D. N., Yuan, D.-N., Boening, C., Landerer, F. W., and Watkins, M. M.: JPL GRACE and GRACE-FO Mascon
730 Ocean, Ice, and Hydrology Equivalent Water Height Coastal Resolution Improvement (CRI) Filtered Release 06.1 Version 03, https://doi.org/10.5067/TEMSC-3JC63, 2023.

Zhang, L., Tang, H., Chang, L., and Sun, W.: Performance of GRACE mascon solutions in studying seismic deformations, J. Geophys. Res. Sol. Ea., 125, e2020JB019 510, https://doi.org/10.1029/2020JB019510, 2020.

Zhao, M. N., Ponte, R. M., Penduff, T., Close, S., Llovel, W., and Molines, J. M.: Imprints of ocean chaotic intrinsic variability on bottom
735 pressure and implications for data and model analyses, Geophys. Res. Lett., 48, e2021GL096 341, https://doi.org/10.1029/2021GL096341, 2021.

Zhao, M. N., Ponte, R. M., and Penduff, T.: Global-scale random bottom pressure fluctuations from oceanic intrinsic variability, Sci. Adv., 9, eadg0278, https://doi.org/10.1126/sciadv.adg0278, 2023.

Zuo, H., Balmaseda, M. A., and Mogensen, K.: The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties
740 in climate signals, Clim. Dynam., 49, 791–811, https://doi.org/10.1007/s00382-015-2675-1, 2017.