

🔗 Northern North Atlantic Sea Level in CMIP5 Climate Models: Evaluation of Mean State, Variability, and Trends against Altimetric Observations

KRISTIN RICHTER,^a JAN EVEN ØIE NILSEN,^b ROSHIN P. RAJ,^b INGO BETHKE,^c JOHNNY A. JOHANNESSEN,^{b,d} AIMÉE B. A. SLANGEN,^{e,f} AND BEN MARZEION^g

^a*Institute of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria*

^b*Nansen Environmental and Remote Sensing Center, Bjerknes Centre for Climate Research, Bergen, Norway*

^c*Uni Research Climate, Bjerknes Centre for Climate Research, Bergen, Norway*

^d*Geophysical Institute, Bergen, Norway*

^e*Department of Estuarine and Delta Systems, Royal Netherlands Institute for Sea Research (NIOZ), Yerseke, Netherlands*

^f*Utrecht University, Yerseke, Netherlands*

^g*Institute of Geography, University of Bremen, Bremen, Germany*

(Manuscript received 12 May 2017, in final form 11 August 2017)

ABSTRACT

The northern North Atlantic comprises a dynamically complex area with distinct topographic features, making it challenging to model oceanic features with global climate models. As climate models form the basis for assessment reports of future regional sea level rise, model evaluation is important. In this study, the representation of regional sea level in this area is evaluated in 18 climate models that contributed to phase 5 of the Coupled Model Intercomparison Project. Modeled regional dynamic height is compared to observations from an altimetry-based record over the period 1993–2012 in terms of mean dynamic topography, interannual variability, and linear trend patterns. As models are expected to reproduce the location and magnitude but not the timing of internal variability, the observations are compared to the full 150-yr historical simulations using 20-yr time slices. This approach allows one to examine modeled natural variability versus observed changes and to assess whether a forced signal is detectable over the 20-yr record or whether the observed changes can be explained by internal variability. The models perform well with respect to mean dynamic topography. However, model performances degrade when interannual variability and linear trend patterns are considered. The modeled regionwide average steric and dynamic sea level rise is larger than estimated from observations, and the marked observed increase in the subpolar gyre is not consistent with a forced response but rather a result of internal variability. Using a simple weighting scheme, it is shown that the results can be used to reduce uncertainties in sea level projections.

1. Introduction

Global sea level is rising and will continue to do so in the future. Regional sea level changes have a direct impact on infrastructure, population, and coastal ecosystems. Yet, regional sea level changes deviate largely from the global

mean, and their simulation requires complex and computationally expensive atmosphere–ocean global climate models (AOGCMs) that form the basis for the projection of regional sea level changes (Church et al. 2013).

Quantifying future regional sea level change and its impacts is pivotal for coastal communities in order to adapt to rising sea levels. Because of the presence of unforced internal variability, climate model projections become increasingly more uncertain when going to smaller spatial and temporal scales as the dominating processes become more complex and the need for simplifying parameterizations due to computational constraints impede

🔗 Denotes content that is immediately available upon publication as open access.

📎 Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-17-0310.s1>.

Corresponding author: Kristin Richter, kristin.richter@uibk.ac.at



This article is licensed under a Creative Commons Attribution 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

the accurate representation of local processes. It is therefore important to evaluate the performance of models on a regional scale with available observations.

For the dynamic sea level contribution to sea level changes, [Landerer et al. \(2014\)](#) evaluated the performance of climate model simulations for the global ocean with a focus on the Southern Ocean and equatorial regions where biases of the mean state were largest. In particular, because of large observed rates of sea level rise, the tropical Pacific Ocean has received considerable attention (e.g., [Meysignac et al. 2012](#); [Palanisamy et al. 2015](#)) compared to other regions such as the northern North Atlantic. We plan to direct our attention to the latter region for several reasons: 1) the sea level in the Atlantic subpolar gyre has been rising over the altimetric record (e.g., [Häkkinen et al. 2013](#)), which is associated with a weakening of the gyre circulation; 2) the Nordic seas and subpolar North Atlantic are surrounded by glaciated landmasses and, through gravitational adjustment, regional sea level trends are expected to be affected by changes in land-based ice masses in a way that counteracts the global mean effect; 3) the additional freshwater input to the ocean is suspected to already have an impact on ocean dynamics ([Rahmstorf et al. 2015](#)), although recent research shows that the impact of meltwater runoff is still small ([Böning et al. 2016](#)); and 4) freshwater input is not yet routinely implemented in AOGCMs used for climate projections. Given these potential weaknesses, and the fact that the region is surrounded by densely populated coasts where the models form the basis for numerous national assessment reports (e.g., [Simpson et al. 2015](#); [Grinsted et al. 2015](#)), it is particularly important to evaluate the models here. For the Norwegian coast, AOGCMs have been evaluated earlier ([Simpson et al. 2014](#)); however, for the abovementioned reasons, we wish to focus on the whole northern North Atlantic region.

Recent studies attempted to quantify the magnitude of internal variability in climate model simulations in order to assess the time of emergence of an externally forced signal in regional sea level ([Lyu et al. 2014](#); [Richter and Marzeion 2014](#); [Bilbao et al. 2015](#)). [Richter and Marzeion \(2014\)](#) found that it takes at least 30 years in the area of interest for an externally forced signal to emerge from the noise when considering a period starting in 1990 (the approximate advent of satellite altimetry). Furthermore, from an observational point, [Dangendorf et al. \(2014b\)](#) found that sea level data from tide gauges exhibit significant decadal and multidecadal correlations independent of any systematic rise. It is therefore not clear whether a forced signal in sea level is already detectable in the northern North Atlantic and Nordic seas or whether observed changes over the altimetric period are caused by internal variability.

Global observational coverage is available for the past few decades (1993–2012 is used in this study). As models are expected to reproduce the locations and magnitudes but not the timing of internal variability, the observations have to be compared to the full historical simulations (1850–2012) of sea level. In this way, it can be assessed whether a forced signal is already detectable (i.e., the observed changes are reproduced by all models in the observational period), whether the observed changes show similarity with modeled internal variability (i.e., the observed changes are reproduced in random 20-yr periods), and whether the models are at all able to simulate variability similar to the observed.

We will focus on the northern North Atlantic region and address the following questions: Is the mean dynamic topography correctly simulated? Is the magnitude and location of observed sea level variability reproduced? Are observed sea level trends consistent with modeled trends over the same period (and therefore forced) or more likely due to internal variability? In [section 2](#) we describe the model data as well as the observations and explain how we will compare the two datasets in order to evaluate the simulated sea level. The results are presented in [section 3](#) for the three variables that are being investigated, namely, mean dynamic topography, regional interannual sea level variability, and regional linear sea level trends. The results are discussed in [section 4](#), followed by the conclusions in [section 5](#).

2. Data and methods

a. CMIP5 model output

We use output from 18 climate models (Table S1 in the supplemental material) participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5; [Taylor et al. 2012](#)) to compare modeled sea level with observations with respect to the mean state, interannual variability, and linear trend patterns. The first realization of each model (r1i1p1) is used. We combine the sea level above the geoid (“zos” in CMIP5 terminology) with the global mean (thermo)steric sea level change (“zostoga” or “zossga”). Most but not all of the models conserve volume rather than mass, that is, a net increase in seawater temperature does not necessarily lead to a global mean sea level rise in zos in these models. This effect is instead computed from the density fields and represented by zossga or zostoga. For consistency, we therefore remove the global mean from the spatial fields of zos at every time step prior to adding the global mean steric change. The resulting variable comprises the regional dynamic and steric sea level change as well as the global mean steric change.

The historical simulations used include all known climate forcings (Slangen et al. 2015). With some exceptions, they cover the period 1850–2005. They are extended up to 2012 using the representative concentration pathway 4.5 (RCP4.5) scenario (Van Vuuren et al. 2011). However, the choice of scenario is not critical over this short period as the scenarios only start to diverge in the second half of the twenty-first century. The climate drift (long-term trends in the absence of an external forcing; Sen Gupta et al. 2013) is accounted for by removing the linear trend found in the control simulations for all variables. Since we are eventually interested in multidecadal variability, annual data are computed.

The grid size of the models varies widely in the area (Table S1), from the comparatively dense grid of $0.9^\circ \times 0.5^\circ$ (longitude \times latitude) or finer for the MPI-ESM-LR, CCSM4, and NorESM models to the coarser grid of $2.5^\circ \times 1.2^\circ$ for the CMCC-CMS model. (Expansions of acronyms are available online at <http://www.ametsoc.org/PubsAcronymList>.) For model-observation intercomparison, the processed data are regridded to a regular grid of $1^\circ \times 1^\circ$.

b. Sea level anomalies and mean dynamic topography from observations

Sea surface height (SSH) is defined as the height of the ocean surface above the reference ellipsoid while sea level anomalies are the deviations of the instantaneous SSH from a reference mean sea surface (MSS). In this study, we use sea level anomalies to investigate variability and trends (and refer to it as sea level). Monthly fields of sea level for the period 1993–2012 are obtained from the European Space Agency (ESA) Climate Change Initiative (CCI) project (ESA-CCI; Ablain et al. 2015, 2017). Annual means for the 20-yr time period are then calculated from the monthly data.

Through geostrophy the ocean surface circulation relates to the oceans mean dynamic topography (MDT), which yields the long-term average strength of the ocean currents (e.g., Knudsen et al. 2011). The precise knowledge of the geoid height, together with MSS, known within centimeter accuracy (Schaeffer et al. 2012), enables us to compute the MDT of the ocean using (e.g., Raj 2017):

$$\text{MDT} = \text{MSS} - \text{geoid}. \quad (1)$$

The geoid is the equipotential surface of Earth's gravity field. More accurately, it is the sea surface in the absence of winds, tides, and currents, only influenced by gravity (e.g., Raj 2017). The *Gravity Field and Steady-State Ocean Circulation Explorer (GOCE)* High-Level Processing Facility (HPF) models Earth's geopotential as a truncated spherical harmonic expansion in the

spectral domain (Bruinsma et al. 2014). Using the gravity field coefficients from the level-2 global gravity models provided by HPF, the geoid height can be determined as detailed in previous studies (e.g., Johannessen et al. 2003; Jin et al. 2014; Raj 2017). In our study, we use the *GOCE* user toolbox (GUT; Benveniste et al. 2007) to estimate the geoid height from the gravity models and the National Space Institute of the Technical University of Denmark (DTU Space) 2013 (DTU13) MSS model (Andersen et al. 2015) to compute MDT. Even though conceptually simple, the computation of MDT needs to satisfy important a priori conditions (Benveniste et al. 2007): both the geoid height and the MSS used should be referenced to the same reference ellipsoid and estimated on the same tide system. The reference ellipsoid, a sphere flattened at its pole, is an arbitrary reference surface that is a raw approximation of Earth's shape. Here, similar to the DTU13 MSS, the geoid height is estimated in the mean tide system relative to the TOPEX ellipsoid. The TOPEX ellipsoid is the first-order definition of the nonspherical shape of the Earth as an ellipsoid of revolution with an equatorial radius of 6378.1363 km and a flattening coefficient of 1/298.257 (Tapley et al. 1994). In the mean tide system, the effects of the permanent tides are included in the definition of the geoid in contrast to the zero tide system where they are excluded. Further, the noise in the MDT due to inconsistencies in the resolution of MSS and geoid is removed by a Gaussian filter (90 km).

Both maps of sea level and mean dynamic topography, initially on a $0.25^\circ \times 0.25^\circ$ grid, have been remapped to a $1^\circ \times 1^\circ$ grid for comparison with the model data (section 2a). Since MDT is the sea surface relative to the geoid, and the climate models have no geoid, the observed MDT is comparable to the modeled mean sea surface.

c. Contribution from land-based ice and GIA

The research area (Fig. 1) is surrounded by (partly) ice-covered land and therefore located in the near-field of potential land-ice changes. The mass exchange between land and ocean affects the gravitational field, that is, the geoid, and thus the shape of the sea surface. The effect on regional sea level gradients is largest close to changing land ice (e.g., Mitrovica et al. 2001). In addition, the altimetry-based sea level observations include the change of the geoid due to the ongoing response of the viscoelastic Earth to the last deglaciation [glacial isostatic adjustment (GIA); Tamisiea 2011]. The CMIP5 models are neither coupled to land-ice models nor do they have geoids that could change, and sea level changes originating from past as well as present changes in land ice have to be removed from the observations prior to comparing observed and modeled trends.

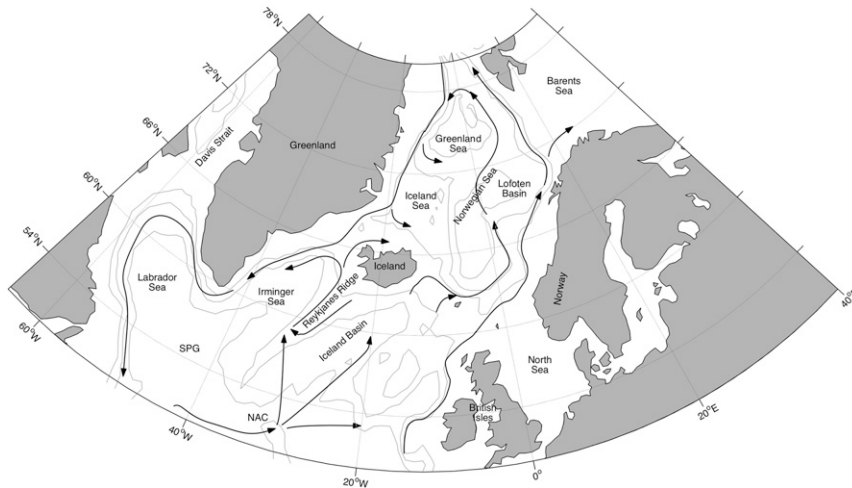


FIG. 1. The northern North Atlantic and the Nordic seas. Schematic surface currents are based on literature (see Furevik and Nilsen 2005). Gray lines represent the 1000-, 2000-, and 3000-m isobaths. (NAC is North Atlantic Current, and SPG is subpolar gyre.)

The glacier mass change over the observational period is obtained by forcing the glacier model of Marzeion et al. (2012) with temperature and precipitation from gridded climate observations [Climatic Research Unit (CRU) time series; New et al. 2002]. Reconstructions of the model over the twentieth century were shown to be consistent with other reconstructive methods based on observations of glacier length change, mass balance observations, and remotely sensed estimates (Marzeion et al. 2015). The contributions from the Greenland and Antarctic ice sheets are taken from Shepherd et al. (2012). The data include both the contribution from surface mass balance changes as well as dynamical changes of the ice.

To translate the land-ice changes into absolute regional sea level changes, the corresponding fingerprints have been computed by assuming the melt of a uniform ice layer over the glaciated area (e.g., Riva et al. 2010; Slangen et al. 2012, 2014; Perrette et al. 2013) and by computing the induced sea level change after solving the sea level equation (Farrell and Clark 1976) using a pseudospectral approach (Mitrovica and Peltier 1991), including the rotational feedback (Milne and Mitrovica 1998), on a compressible elastic Earth (Dziewonski and Anderson 1981).

To account for the geoid-related signal due to GIA from the sea level observation, we use the correction from the ICE-5G GIA model as provided by Peltier (2004).¹ As we are interested in absolute sea level changes, only the rate of change of the geoid is used.

Figure 2a shows the observed 1993–2012 sea level trends in the region. The average trend over the region is 2.35 mm yr^{-1} , less than the global average of $3.16 \pm 0.5 \text{ mm yr}^{-1}$ (Ablain et al. 2015). Sea level change is positive except east and southwest of Svalbard and in a very localized area south of the Iceland Basin. The rise is largest in the subpolar gyre and, to a lesser degree, along the European coast. The contribution from changing glaciers and ice sheets to absolute sea level over the same period (Fig. 2b) is characterized by a sea level decrease around Svalbard and Greenland in accordance with the loss of land ice in these regions. However, compared to observed sea level, changing land ice contributes little to the spatial variability of observed trends whereas the average sea level change over the area due to changes in land ice is significant, contributing 0.49 (glaciers) and 0.43 mm yr^{-1} (ice sheets), respectively, to the observed regionwide average trend (Fig. S2 in the supplemental material). The same is true for the geoid-related GIA contribution: it contains little spatial variability compared to the observed trend pattern while the corresponding regionwide average presents a significant, notably negative, contribution of -0.31 mm yr^{-1} (Fig. 2c).

To assess the impact of land-ice melting on regional sea level trends, we removed the contribution from glaciers and ice sheets as well as from GIA from the observed sea level prior to computing linear trends. The result is a reduced regionwide average trend of 1.75 mm yr^{-1} , with few changes in the regional trend pattern (cf. Figs. 2a,d). The pattern is still dominated by the stronger sea level rise in the subpolar gyre around the southern tip of Greenland. The negative trend

¹Dataset available from http://www.atmosph.physics.utoronto.ca/~peltier/datasets/GRID/dGeoid250.1grid.ICE5Gv1.3_VM2_L90_2012.nc.

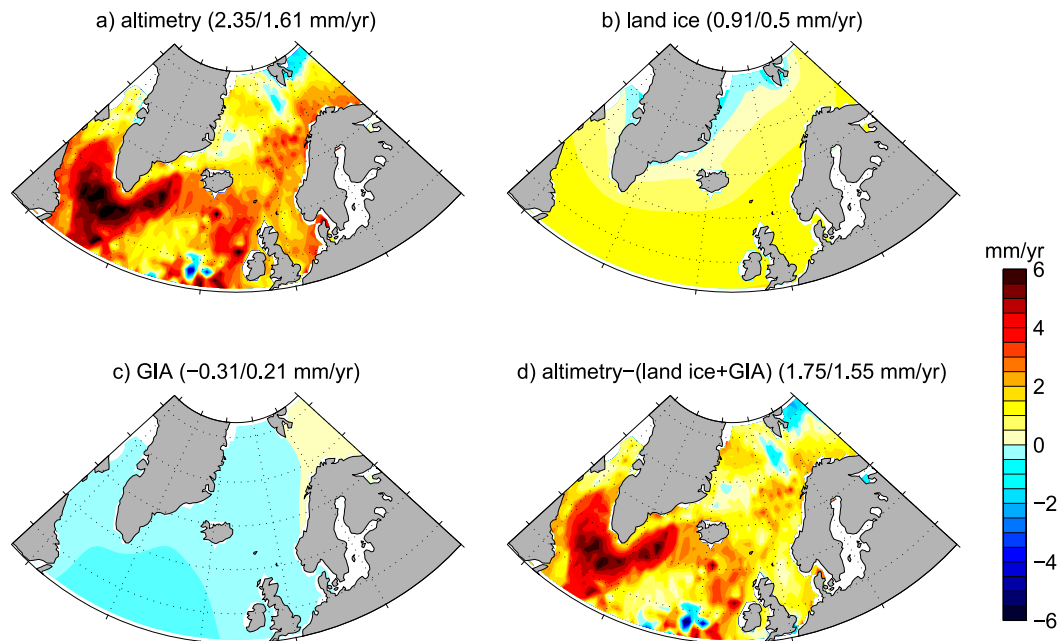


FIG. 2. (a) Observed trends from satellite altimetry during 1993–2012, (b) contribution from glaciers and ice sheets, and (c) contribution from the geoid-related component of GIA. (d) Trends in sea level corrected for land-ice contributions and GIA. The numbers above each map represent the regionwide average and root-mean-square, respectively.

southwest of Svalbard is slightly enhanced while the sea level rise on the Norwegian shelf, and therefore the across-shelf gradient, is reduced. To be consistent with the models, we use these trends (presented in Fig. 2d) in our trend analysis (section 3c).

d. Methods

The sea level observations from satellite altimetry used in this study span the period 1993–2012. From this 20-yr record of annual means, we calculate maps of sea level variability (in terms of temporal standard deviation) and linear sea level trends. These maps, together with maps of MDT, are used for the model–observation analysis. The model simulations cover the period 1850–2012. To account for the random phases of internal variability in the models, we slide a 20-yr window over the annual data and compute maps of MDT, sea level variability, and linear trends for each 20-yr period, thus giving us a total of 143 maps per model per variable, available for comparison with the observations. Linear trends are removed from every 20-yr period prior to calculating variability.

To assess the model performance, we use basic statistical measures: modeled and observed fields are compared by computing the area-weighted pattern correlation coefficients (PCCs) and root-mean-square errors (RMSEs) between modeled and observed maps.

In this case, the PCC is the equivalent to the correlation coefficient between two time series, with the only difference being that the correlation is computed between corresponding locations instead of corresponding points in time (Thomson and Emery 2001). Prior to computing PCCs and RMSEs, the regionwide average is subtracted from all maps. That is, we compute centered statistics to assess the similarity between observed and modeled regional anomalies (e.g., Santer et al. 1995). Note that PCCs are not sensitive to errors in the amplitudes in the modeled patterns. These errors are reflected in the RMSEs. Centered statistics are insensitive to the regionwide average. This is of particular importance for the comparison of observed and modeled trends. In line with the global mean, we expect simulated regionwide average trends to be small in the beginning of the historical record and grow increasingly more positive throughout the modeled period considered here (1850–2012). To take this fact into account, we will also report the total RMSE (RMSE without removing the regionwide average) when presenting results of the trend analysis.

As mentioned, we account for the different phases of internal variability in models and observations by comparing the observations with modeled fields from all 20-yr windows. Since, on a time scale of only two decades, internal variability is expected to be strong in

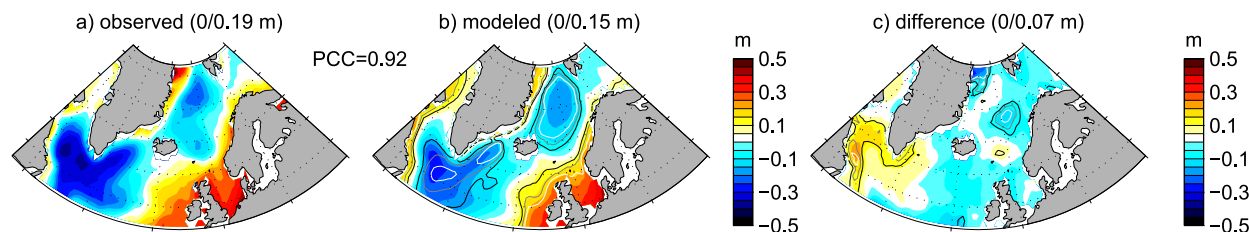


FIG. 3. MDT from (a) observations and (b) multimodel mean for the observational period. The PCC between those fields is also shown. (c) Ensemble mean of differences between models and observations for the observational period. Black, gray, and white contours in (b) and (c) represent signal-to-noise ratio of 1, 1.5, and 2, respectively, defined as the ratio of ensemble mean and ensemble standard deviation of the regional anomalies. The numbers above each map represent the regionwide average and root-mean-square, respectively.

both models and observations, we anticipate some variability in the PCCs and RMSEs as we slide the window across the model record. The approach presents us with several possibilities to compute multimodel means: (i) by taking the mean over the observational period (1993–2012); (ii) by selecting, for each model, the period with maximum PCC with the observed field and taking the mean over these model maps (max corr); or (iii) as in (ii), but for periods of minimum RMSE (min RMSE). In the following, we will present results for ensemble means computed following method (i) and (ii), as results for ensemble (iii) are similar to those of ensemble (ii). Note that multimodel means are formed by first computing maps of MDT, variability, and linear trends for each model, and subsequently averaging over these fields. In the same manner as for individual maps (by computing PCCs and RMSEs), we compare the three different ensemble means with observations. This enables us to assess whether the observed regional trend patterns are due to internal variability or externally forced. In the latter case, we expect a good agreement not only between models and observations but also between the individual models over the observational period (1993–2012).

3. Results

a. Mean dynamic topography

The mean dynamic topography determines the surface circulation of the oceans and is thus crucial not only for heat and salt transport, but also directly for sea level changes, particularly along the margins of the oceans. Observation-based MDT is shown in Fig. 3a with the regionwide average removed for easier comparison with modeled MDT. The cyclonic circulation in the Nordic seas is visible by local minima in the basins of the Nordic seas and elevated MDT along the surrounding coasts of Greenland and northern Europe. The even more pronounced and larger minimum south of Greenland represents the subpolar gyre. The

North Atlantic Current, the inflow to the Nordic seas, and the Norwegian Atlantic Current are represented by the northwest–southeastward sea level gradient in the southeastern half of the map, in particular along the continental slope (see Fig. 1).

Figure 4 shows the RMSE of the MDT fields for each model and the PCC for the observational period together with the corresponding ranges that represent the spread of RMSE and PCC over all 20-yr periods. The PCCs are large (around 0.9) for seven models with corresponding low RMSEs (~ 0.1 m). These models capture the western extent of the subpolar gyre correctly (Fig. S2). Models with $\text{PCC} < 0.6$ (GFDL-ESM2G, MRI-CGCM3, and BCC_CSM1.1) and a mean $\text{RMSE} > 0.15$ m do not simulate the characteristic sea level minimum corresponding to the subpolar gyre at all. The MDT minimum in the Nordic seas is present in all models except for one (INM-CM4.0), although the

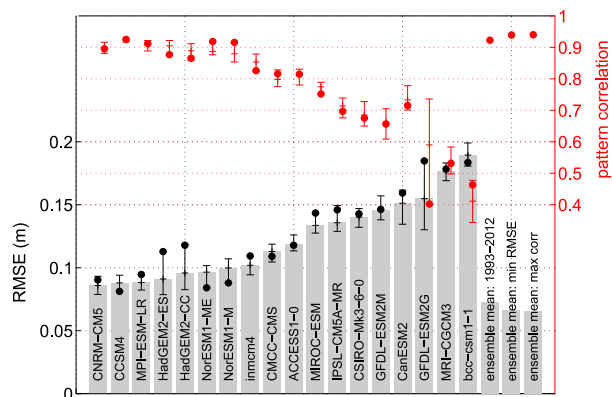


FIG. 4. Results of model–observation comparison of MDT: mean RMSE over all 20-yr periods from 1850 to 2012 (bars) with minimum–maximum range (black error bars) as well as RMSE over observational period (black dots) for each model. The equivalent is shown for the PCCs (red error bars and dots). Also shown are the RMSEs and PCCs for the three different multimodel means over the observational period, over modeled maps of minimum RMSE, and over modeled maps of maximum PCC (section 2d).

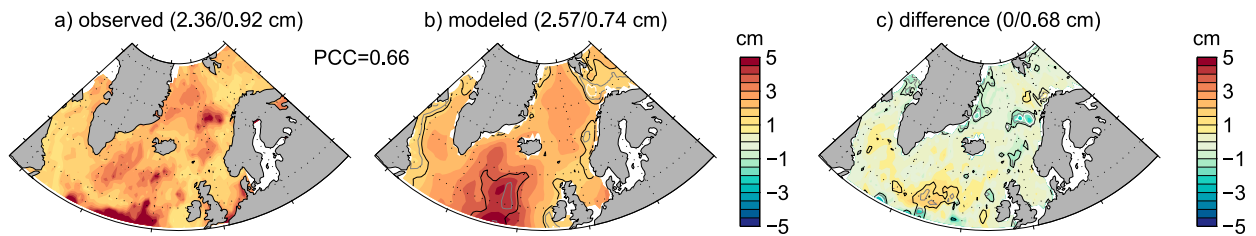


FIG. 5. Sea level variability from (a) observations and (b) multimodel mean over the periods of maximum pattern correlation. The PCC between those fields is also shown. (c) Difference of modeled and observed variability averaged over all models for the period of maximum pattern correlation. The mean difference has been removed and only anomalies are shown. Black, gray, and white contours represent signal-to-noise ratio of 1, 1.5, and 2, respectively, defined as the ratio of ensemble mean and ensemble standard deviation of the regional anomalies. The numbers above each map represent the regionwide average and root-mean-square, respectively.

strength and location vary. The spatial variability of the multimodel mean over the observational period (Fig. 3b) is slightly smaller than the observed variability. The ensemble mean difference between modeled and observed MDT over the observational period is shown in Fig. 3c. The most pronounced feature is a systematic underestimation of the strength of the subpolar gyre in the Labrador Sea. There is also a significant lack of regional details in the North Atlantic Current in the Norwegian Sea.

Generally, the lower the PCC is, the higher the RMSE is (Fig. 4). The RMSE is of the same order of magnitude as found in Landerer et al. (2014) for the entire world ocean over the period 1993–2002. The spread of RMSEs and PCCs derived from all available 20-yr periods is small for most models, indicating that internal variability has little effect on modeled MDT patterns. The exception is GFDL-ESM2G, where the RMSE increases gradually and the PCC decreases gradually throughout the historical simulation (not shown). PCCs are highest (>0.9) with corresponding low RMSEs for the multimodel mean. This applies to any 20-yr period over which MDT is calculated: the observational period, the period of minimum RMSE, and the period of maximum pattern correlation (Fig. 4).

b. Variability

The observed interannual sea level variability (Fig. 5a) exhibits local maxima in the Lofoten Basin, as well as in the Iceland Basin and Irminger Sea at the eastern rim of the subpolar gyre. In contrast, observed interannual variability is weak in the western branch of the subpolar gyre and on the shallow European shelf from the British Isles along the Norwegian coast up to the entrance to the Barents Sea. In the coastal areas of the North Sea, however, variability is stronger.

Models are designed to reproduce the main climate modes in terms of amplitude and location. However, the phase and timing of simulated variability is not necessarily

the same as that of the observed variability. Therefore, we do not expect the models to reproduce the observed variability over the observational period, and the RMSEs and the PCCs are expected to vary when comparing all available modeled 20-yr periods with the observations. This is indeed the case (Fig. 6). The range of PCCs is much larger than the equivalent range for the MDT (Fig. 4). RMSEs are on the order of 1–2 cm, with some models showing a spread as large as 1 cm. The PCCs are low (between -0.1 and 0.5), and some models display both negative and positive correlation coefficients. The PCCs for the observational period are not systematically in the higher end of the range, indicating that the observed variability is not in phase with the variability in the forced simulations and thus indeed unforced. In the same way as for MDT, the ensemble means of variability for the observational period and periods of minimum RMSE and maximum PCC, respectively, show smaller RMSEs and larger PCCs than any of the individual models.

Figure 5b shows the modeled variability as the multimodel mean over the periods of maximum pattern correlation of the individual models. These periods are

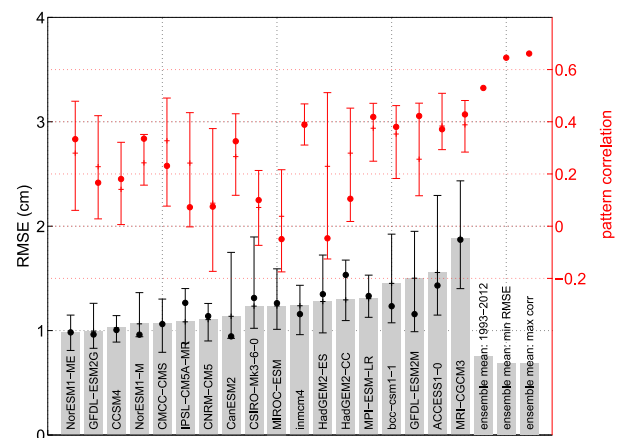


FIG. 6. As in Fig. 4, but for results of model–observation comparison of sea level variability.

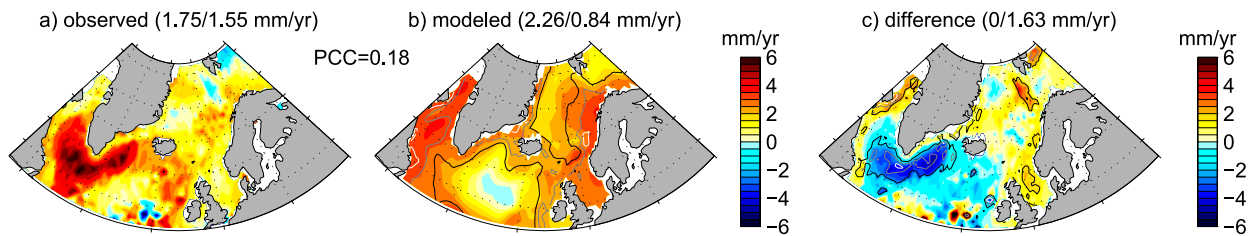


FIG. 7. Linear sea level trends over the observational period from (a) observations (as in Fig. 2d) and (b) multimodel mean of modeled trends. The PCC between those fields is also shown. (c) Multimodel mean of the differences between modeled and observed trend anomalies for the observational period. The regionwide average difference that has been removed (as we look at anomalies) is 0.82 mm yr^{-1} . Black, gray, and white contours represent signal-to-noise ratio of 1, 1.5, and 2, respectively, defined as the ratio of ensemble mean and ensemble standard deviation of the regional maps. The numbers above each map represent the regionwide average and root-mean-square, respectively.

different for each model as the internal variability is random and not in phase between the individual models. The mean difference averaged over all models is 0.21 cm (not included in Fig. 5c), indicating that models overestimate the variability in this region. The multimodel mean variability shows enhanced variability away from the coasts in accordance with the observations (Figs. 5a,b). However, the pattern is very smooth and does not reproduce the finer spatial features. This is partly due to the ensemble averaging: when looking at individual models (Fig. S3 in the supplemental material) local details like the enhanced variability in the central Nordic seas are evident in most models (except IPSL-CM5A-MR). The strength and location of these details, however, differ from model to model.

The relatively good statistical performance of the ensemble mean as shown in Fig. 6 is therefore due to smoothing of local features, and the performance comes from good reproduction of large-scale features. However, a systematic underestimation of variability in the Lofoten Basin is identified, as well as an overestimation of the strength of the variability in the southern Iceland Basin (Fig. 5c, at lower right). In the remaining areas, no systematic differences are found.

c. Linear trends

As mentioned, we use the altimetric sea level corrected for contributions from land ice (Figs. 2d and 7a) in the following trend analysis. Linear trends in sea level over the observational period may be unforced, that is, due to internal variability only; externally forced (natural or anthropogenic); or a combination of both. To account for the presence of unforced trends, we apply the same procedure as for the investigation of variability and compare the observed trends with trends from all 20-yr periods available in the models.

Statistical parameters from the observation–model comparison of trend anomalies are shown in Fig. 8. The PCCs for trends span an even wider range (from -0.45 to

0.45) than for variability (cf. Fig. 6), as is to be expected. For most models, PCCs are symmetric around zero. Notable exceptions are GFDL-ESM2G and IPSL-CM5A-MR, with maximum positive PCCs much larger than absolute negative PCCs. The RMSEs are on the order of the observed mean sea level rise itself or higher. For the ensemble means, RMSEs are smaller than for the individual models. For the ensemble mean over the periods of minimum RMSE and maximum pattern correlation, the PCCs exceed 0.6 (0.18 for observational period), indicating that the observed trend pattern is at least partly simulated at some point in time by some of the models. It is also noteworthy that PCCs are not largest and RMSEs not smallest over the observational period (red and black dots in Fig. 8, respectively), hinting toward the importance of internal variability. Note that the RMSE is not the total RMSE but the RMSE of the fields with the regionwide average trend removed. That is, we compare observed and modeled trend anomalies. This way, we account only for spatial variability and not for biases in the regionwide average trend. This is important as the modeled trends, and therefore the total

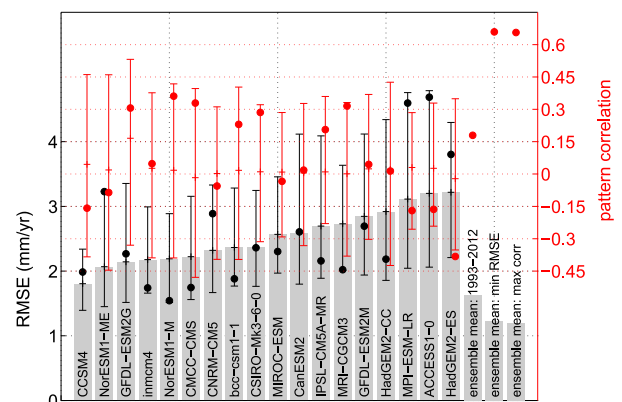


FIG. 8. As in Fig. 4, but for results of model–observation comparison of trends.

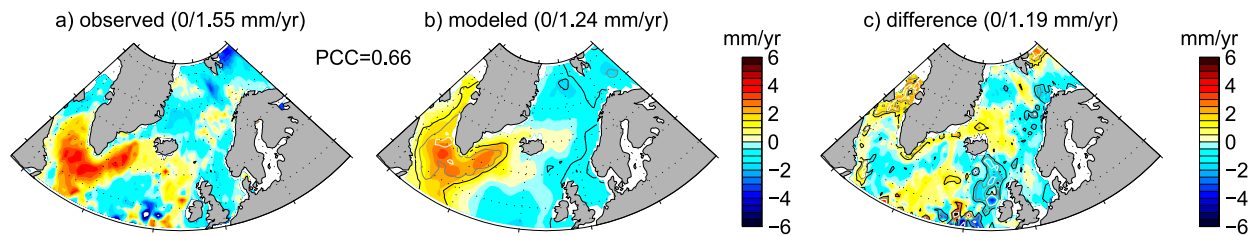


FIG. 9. Linear sea level trend anomalies (relative to regionwide average) from (a) observations and (b) ensemble mean over periods of maximum pattern correlation. The PCC between those fields is also shown. (c) The multimodel mean of the differences between modeled and observations. Black, gray, and white contours represent signal-to-noise ratio of 1, 1.5, and 2, respectively, defined as the ratio of ensemble mean and ensemble standard deviation of the regional maps. The numbers above each map represent the regionwide average and root-mean-square, respectively.

RMSE, are strongly affected by the global average sea level rise. As we expect the observed trends to be at least partly affected by internal variability and therefore correspond to any modeled time period with potentially lower global sea level rise, it is important to take this into account.

The modeled average sea level trends for the observational period range from 0.12 (MRI-CGCM3) to 4.97 mm yr^{-1} (BCC_CSM1.1; Fig. S4 in the supplemental material) with a multimodel mean of $2.26 \pm 1.23 \text{ mm yr}^{-1}$ (one standard deviation), which is larger than the observed average trend of 1.75 mm yr^{-1} . The ensemble mean of modeled regional trends over the observational period is shown in Fig. 7b. Compared to the observations (Fig. 7a), the simulated spatial variability is small (0.84 versus 1.48 mm yr^{-1}). The marked sea level rise in the subpolar gyre that dominates the observed trends is not simulated at all. This indicates that it is a manifestation of internal variability, the response to an external forcing that is not included in the models, or processes that are not properly represented or resolved. Instead, a negative trend in the central North Atlantic is simulated that is also present in the observations but has a signal-to-noise ratio of less than 1. The sea level rise along the shallow shelf areas and west of Greenland is systematic across the models (high signal-to-noise ratio), indicating that it may be a forced signal. However, it should be noted that the relatively high signal-to-noise ratio shown in Fig. 7b is due to the regionwide average sea level trend. Once this trend is removed and only anomalies are considered, the signal-to-noise ratio is less than 1 in the entire area (not shown).

To assess whether the models systematically over- or underestimate observed trends, we compute the multimodel mean of the differences between modeled and observed trend anomalies for the observational period (Fig. 7c). The regionwide average difference is 0.51 mm yr^{-1} , reflecting the overestimation of the sea level rise in the

region by the models. The differences shown in Fig. 7c resemble the spatial pattern of the observed trend anomalies (Fig. 7a), and indeed the PCC between the two fields is -0.86 .

Compared to the observed trends, the modeled trend field is rather smooth. While a large grid size contributes to a smoother result, this also indicates that the modeled trend anomalies during this period are offsetting each other partially in the ensemble mean and are therefore induced by internal variability rather than forced. This is confirmed when looking at individual models (Fig. S4). While all models simulate a regionwide average sea level rise, the spatial patterns are very different. From this, we conclude that the observed regional sea level trend pattern results from internal variability while the regionwide average positive sea level change is externally forced.

To investigate whether internal variability as simulated by the models reproduces the observed spatial trend pattern, we turn again to the pattern correlation analysis and select, for each model, the 20-yr period with the maximum pattern correlation. The results are shown in Fig. S5 of the supplemental material for each model, and the multimodel mean is presented in Fig. 9b. Indeed, the multimodel mean now displays positive trends in the region of the subpolar gyre with a relatively high signal-to-noise ratio. The PCC of the ensemble mean with the observed anomalies is 0.66. The regionwide average of the multimodel mean (subtracted in Fig. 9b) is 1.32 mm yr^{-1} , indicating that the periods of maximum pattern correlation occur at times of lower regionwide sea level rise in the majority of the models. The regionwide average sea level trend for individual models ranges from -0.68 mm yr^{-1} for HadGEM2-ES (period of maximum pattern correlation: 1954–73) to 4.45 mm yr^{-1} for GFDL-ESM2M (1981–2000).

For the period of maximum PCC, the multimodel mean difference (Fig. 9c) between modeled and observed anomalies is less systematic than the differences

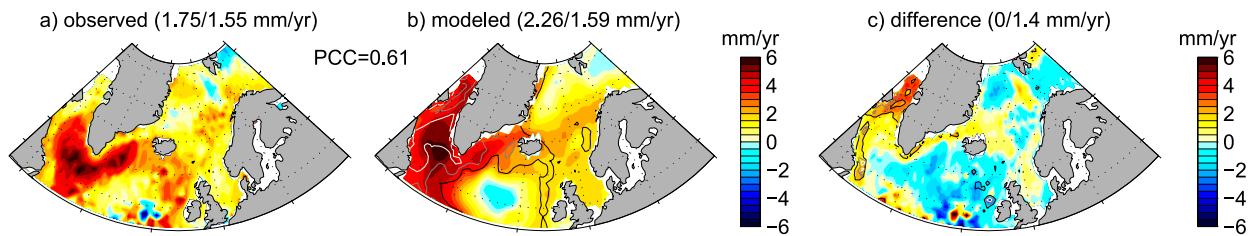


FIG. 10. Linear trends in sea level from (a) observations (as in Fig. 2d) and (b) combined modeled signal from the observational period (Fig. 7b) and periods of maximum pattern correlation (Fig. 9b). The PCC between those fields is also shown. (c) The multimodel mean differences with observations. Black, gray, and white contours represent signal-to-noise ratio of 1, 1.5, and 2, respectively, defined as the ratio of ensemble mean and ensemble standard deviation of regional anomalies. The numbers above each map represent the regionwide average and root-mean-square, respectively.

for the observed period (Fig. 7c) with a lower corresponding PCC of -0.62 . It appears that a part of the observed signal can be identified as an internal signal.

In the next step, we combine the forced signal that appears to consist of a regionwide sea level rise (Fig. 7b) with the signal that originates from internal variability (Fig. 9b) to obtain the full signal (Fig. 10b). The PCC with the observations is 0.61. The approach implicitly assumes the perfect cancellation of internal variability during the observational period, which cannot be expected to be entirely true due to the different representation of internal variability in the models in terms of strength, location, and frequency. The anomalously high sea level rise in the subpolar gyre is well captured, as is the sea level rise along the east coast of Greenland. The regional pattern of the differences between modeled and observed trend anomalies has a very low signal-to-noise ratio (PCC of -0.42 with observations), except in the Davis Strait, where the models appear to systematically overestimate the change in sea level (Fig. 10c).

4. Discussion

We have assessed the representation of simulated sea level in terms of its mean state (MDT), interannual variability, and linear trends in 18 CMIP5 climate models by comparing the model output with observations from satellite altimetry. As the altimetric record

covers only 20 years, a time span on which internal variability dominates over externally forced signals (Richter and Marzeion 2014), the observations were compared to the entire model record (1850–2012) using 20-yr time slices.

A summary of the statistical results with respect to multimodel means is shown in Table 1 and for individual models in Fig. 11. While the multimodel means represent the observed MDT rather well (PCC > 0.9), the multimodel mean performance degrades with respect to regional interannual variability and trends. This is also true for individual models: the PCCs for MDT can exceed 0.9 for some models, while the maximum PCC is below 0.6 for linear trends and variability.

With the exception of GFDL-ESM2G, the model performance with respect to MDT depends only a little on the chosen period (small spread in RMSE and correlations in Fig. 4), indicating that internal variability has little influence on MDT in a 20-yr period and is of the same order of magnitude as intermodel differences. The most striking intermodel differences are the simulation of the location, strength, and shape of the sea level minimum associated with the subpolar gyre and to some degree the Nordic seas minimum (Fig. S2). Shortcomings in the representation of open ocean circulation can be expected to have consequences for the ability to simulate relevant regional coastal sea level, since it is the transport of steric anomalies toward the shelves and the dynamics of the slope currents that constitute the

TABLE 1. PCCs and RMSEs between observations and different ensemble means: observational period 1993–2012, periods of maximum pattern correlation (maxcorr), and for trends in periods of maximum pattern correlation plus observational period (maxcorr2012). For variability and trends, the total RMSE (regionwide average included) is shown in the bottom row.

	MDT (m)		Variability (cm)		Linear trends (mm yr^{-1})		
	1993–2012	Maxcorr	1993–2012	Maxcorr	1993–2012	Maxcorr	Maxcorr2012
PCC	0.92	0.95	0.53	0.66	0.18	0.66	0.61
RMSE	0.073	0.065	0.76	0.68	1.63	1.19	1.40
Total	—	—	0.78	0.69	1.69	2.14	1.48

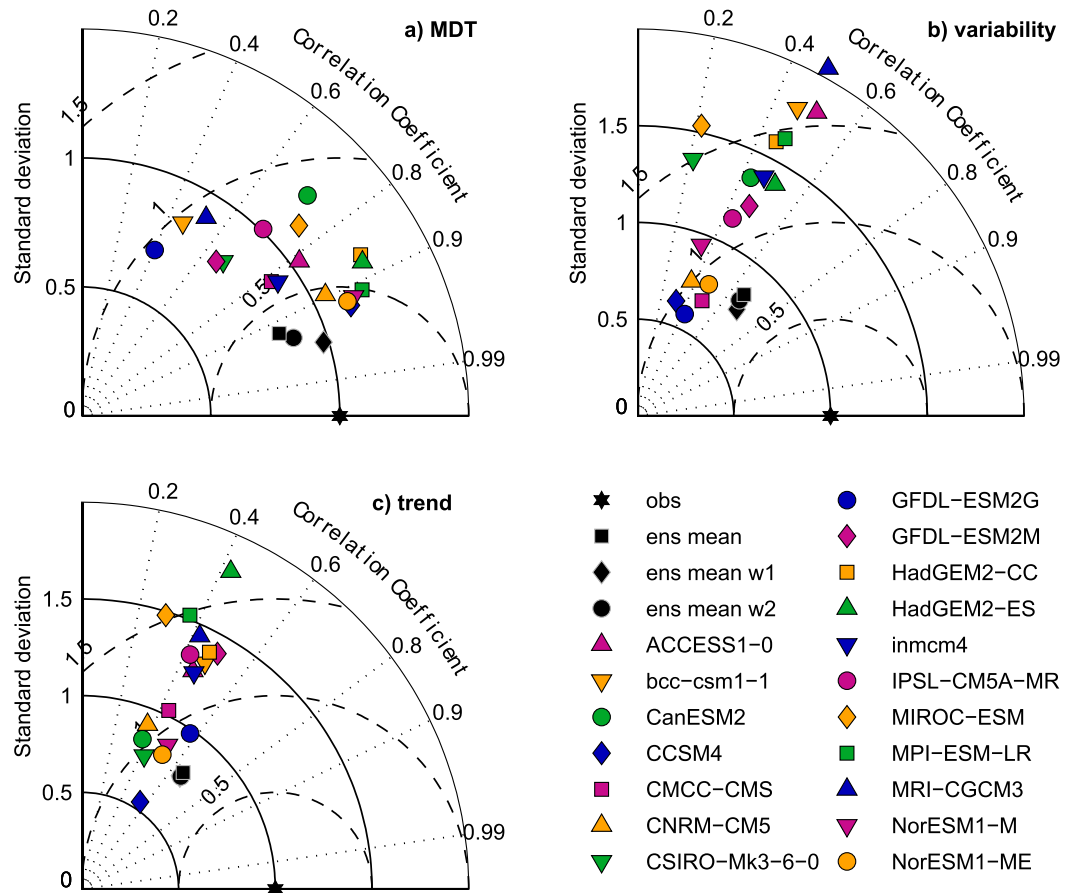


FIG. 11. Taylor diagrams to summarize the performance of individual models with respect to (a) MDT (observational period), (b) variability (maximum PCC), and (c) linear trends (maximum PCC). The dashed arcs represent the RMSE. Standard deviation and RMSE are standardized with the respective observed spatial standard deviation. The respective multimodel mean and the observations as well as the multimodel mean based on selection (w1) and weighting (w2) are also shown.

regional steric/dynamic contribution to coastal sea level change and variability.

The larger range in PCCs for variability (Fig. 5) as compared to MDT reflects the importance of internal variability in the area. The finer details (e.g., in the Lofoten Basin and the areas east and west of Reykjanes Ridge) are not captured by the models. When looking at individual models, variability is strikingly overestimated by some models in large parts of the region. A more detailed analysis is necessary to reveal the reasons for the under- or overestimation of interannual variability (i.e., steric vs dynamic changes).

Unsurprisingly, we found that linear 20-yr trends are heavily impacted by internal variability and no forced pattern of regional sea level trend could be detected over the observational period, supporting previous studies (Richter et al. 2017). According to our results, the observed increase in sea level in the subpolar gyre

appears to be induced by internal variability. This is in line with the results of Häkkinen et al. (2013), who found that the increase is part of Atlantic multidecadal variability mainly forced by wind stress changes. Climate modes involving atmospheric circulation (such as the North Atlantic meridional dipole) have been shown to govern variability in both on-shelf sea level (Richter et al. 2012a; Calafat et al. 2012; Dangendorf et al. 2014a; Chafik et al. 2017) and ocean circulation (e.g., Nilsen et al. 2003; Sandø et al. 2012; Richter et al. 2012b) in the North Atlantic, on time scales from months to decades. Although considered part of natural variability, spatial shifts and changes in these modes (Ulbrich and Christoph 1999) may lead to significant sea level changes, especially along the coasts (Chafik et al. 2017).

A recent study by Becker et al. (2016) compared long-term correlation in sea level as observed by century-long tide gauge records with simulated nearby sea level from

climate models and found that the models overestimated long-term trends at the North Atlantic coasts as induced by internal variability. Here, we found no evidence of an overestimation of 20-yr trends, but our approach is not perfectly suited to detect such a misfit. Note also that we are focusing on larger spatial scales and the open oceans to shed light on the model performance in the regions from where long-term sea level changes are likely to originate. For a more thorough assessment of both variability and trends, and to disentangle the steric and dynamic contribution, we would need to produce a steric estimate from both models and observations. This is, however, out of the scope of this study.

The performance of the individual models is summarized in Fig. 11. In the Taylor diagrams, the distance between the points representing the simulated fields and the point representing the observed field is inversely proportional to the overall similarity of the two fields. For variability and trends (Figs. 11c,d), models are grouped along a line as we display their performance during periods of maximum PCC, which is similar for most models (Figs. 6 and 8). From Fig. 11 outliers can be identified visually, that is, MRI-CGCM3 in Fig. 11b and HadGEM2-CC in Fig. 11c. It is also clear that the ensemble mean delivers the best performance. However, because of the ensemble averaging, finer details are lost and the spatial standard deviation of the ensemble mean is always smaller than the observed one (Fig. 11), except when weighting is applied (see below).

From Fig. 11 it can be seen that there is no clear relationship between the model performances regarding MDT and their performances regarding variability and linear trends. Also, there is no consistent relationship between model grid spacing and their performance. While models with comparatively small grid spacing like CCSM4 and NorESM are consistently in the upper half of the ranking, so is CMCC-CMS—a model with relatively large grid spacing in the area. In contrast, MPI-ESM-LR does perform well with respect to MDT, but performance degrades considerably regarding variability and trends. A good performance with respect to the mean state is therefore no guarantee for an equally good performance with respect to decadal trends and interannual variability.

One might ask whether the results of our analysis can be used to improve projections and/or reduce uncertainties of regional sea level and risk assessments that are based on the climate simulations we investigated here. Weighting model output has been used widely with respect to, for example, mean surface temperature (Räisänen et al. 2010), the meridional overturning circulation (Schneider et al. 2007), and sea ice extent

(Knutti et al. 2017). Various approaches are conceivable, such as selecting the model that performs best, excluding models that perform inadequately, or assigning different weights to models based on their performance.

Prior to computing weights, it needs to be assessed whether our diagnostics (mean state, interannual variability, or short-term trends) are relevant for projections on long time scales. The regional trend patterns are mostly due to internal variability. It is therefore questionable to draw conclusions about sea level projections (i.e., forced changes) based on the models' performance with respect to linear trends over the observational period. The accurate representation of the mean state is more relevant to projections (and has also been shown to be time independent), as is the simulation of the regionwide sea level rise that has been shown to be at least partly forced. If weights are to be computed, these two diagnostics should be used.

We emphasize that our study focuses on evaluating the large-scale features of CMIP5 models in contrast to details on shelves and near coasts, where we do not expect these coarse grid models to perform well (e.g., Griffies et al. 2014). The overestimation of near-coast sea level changes pointed out by Becker et al. (2016) may indicate that CMIP5 models may be lacking in representation of relevant processes. For example, the decoupling between on-shelf and off-shelf sea level changes on short–interannual time scales (e.g., Hughes and Williams 2010; Bingham and Hughes 2012) may play a role on longer time scales, but it is not necessarily modeled correctly, neither are transfer mechanisms involving coastally trapped waves (Calafat et al. 2012; Dangendorf et al. 2014a; Frederikse et al. 2016). The ability to model exchanges with the Mediterranean Sea has also been proposed to be important for the representation of sea level changes on the European shelves (Hughes et al. 2015). Such small-scale processes should be taken into consideration for improvement and evaluation of future models but are outside the scope of this study. Our motivation is that the circulation and variability in the open oceans need to be realistic to model the off-shelf, long-term steric and dynamic changes that will, through the abovementioned or other mechanisms, affect long-term sea level changes on shelves and coasts.

Here, we provide a test of our hypothesis that there is potential benefit in basing model selection or weighting on the models' performance with respect to the large-scale mean state. Comparisons were done between the full ensemble (equal weights), an ensemble of models with MDT RMSE below median (zero weights for deselected models), and an ensemble weighted by the

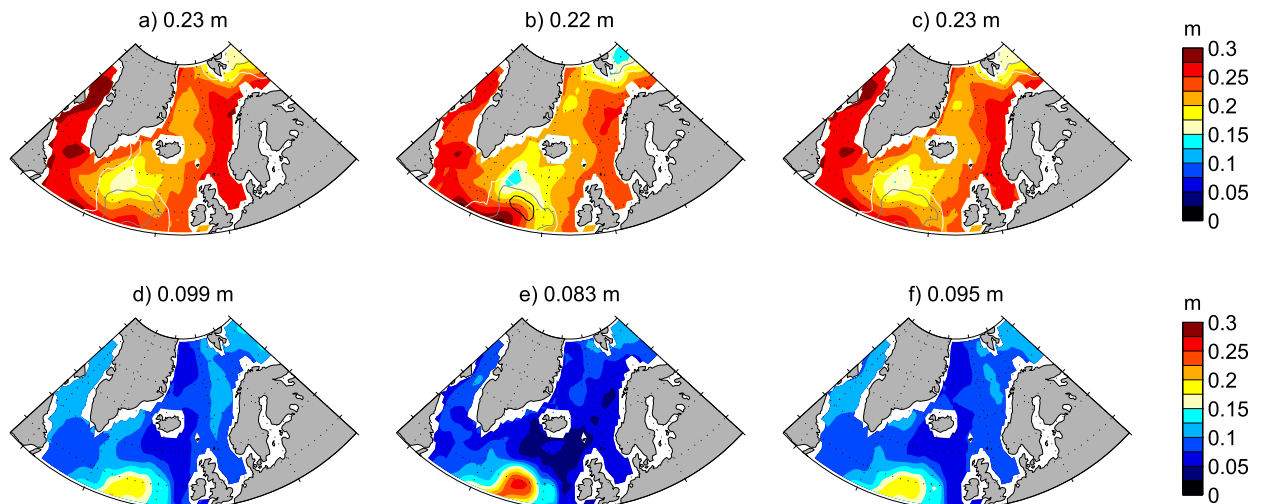


FIG. 12. Multimodel mean of sea level change from 1986–2005 to 2081–2100 (RCP4.5) with (a) no weighting (i.e., equal weights), (b) a subselection of models with MDT RMSE smaller than median, and (c) models weighted with inverse of MDT RMSE. (d)–(f) The respective ensemble standard deviation (also weighted when appropriate). The numbers on top of each map indicate the regionwide average.

inverse of the MDT RMSE. The according weights are presented in Table S2, and the skill of the modified multimodel means is included in Fig. 11. The performance of the ensemble mean MDT improves only marginally with the selection and weighting (Fig. 11a and Fig. S6 in the supplemental material), as the full ensemble already performed reasonably. However, the strength and shape of the subpolar gyre becomes more realistic with selection, but not with weighting. As expected, the performance with respect to trends and variability does not improve with weighting or selecting, but it does not deteriorate either (see also Figs. S7 and S8 in the supplemental material).

More importantly, compared to using the full ensemble, selecting models leads to a reduction in spread of regional sea level ensemble projections along the coastlines. A similar spread reduction is not apparent in our weighted results (Figs. 12d–f). These results indicate that climate model performance with respect to simulating coastal sea level change should be assessed by their oceanwide circulation and that selection should be considered a liable option. This also underlines that climate models need to be investigated further regarding their steric and dynamic features and performance, preferably on longer time scales as well.

5. Conclusions

We decomposed annually averaged sea level in the northern North Atlantic and Nordic seas over a 20-yr period into a mean state, linear trends, and residual interannual variability to evaluate the performance of

climate models that are routinely used to project the steric/dynamic component of future sea level changes against altimetric observations.

We demonstrated that models are in general capable of capturing the main features of observed sea level changes if internal variability is taken into account. Our study shows that the observed linear trend pattern over the period 1993–2012 is likely dominated by internal variability, whereas the mean state pattern is rather insensitive to which 20-yr period is chosen. While no forced signal could be detected in the regional trend patterns, the observed regionwide average sea level rise during the observational period (1993–2012) appears to be partly forced.

According to our results, the ensemble mean outperforms every single model, in line with what has been shown in other studies (e.g., Yin et al. 2010; Simpson et al. 2014). Risk assessments are, however, not only interested in the most likely projection, but also the uncertainty range around it. Our results also highlight that the multimodel projection spread of regional coastal sea level change can be reduced by selecting models based on their ability to simulate the mean dynamic topography over the entire ocean region in question. This highlights the general importance of a realistic mean state representation as the most important feature of ocean models in order to be suitable for sea level projections.

Our selection–weighting approach is rather ad hoc and simple and, as such, is subject to certain caveats (e.g., we are not considering intermodel dependencies). Sea level changes as investigated here are the result of

several processes (air–sea heat exchange, redistribution of heat and freshwater in the ocean, and dynamic changes) and, in line with Knutti et al. (2017), several variables (e.g., ocean temperature and salinity) and diagnostics (e.g., stratification) should be analyzed in conjunction in order to assign sensible weights to each model.

Acknowledgments. This work was funded by the Austrian Science Fund (FP253620) and supported by the Austrian Ministry of Science BWF as part of the UniInfrastrukturprogramm of the Focal Point Scientific Computing at the University of Innsbruck. Support was also provided by the Centre for Climate Dynamics at the Bjerknes Centre, through the project iNCREASE. The work is also partly supported by the ESA-CCI Sea Level project funded under the ESRIN Contract 4000109872/13/1-NB and by the Copernicus Arctic MFC project. This paper has further profited from the outcome of work with the ISSI international team on Contemporary Regional and Global Sea-Level Rise: Assessment of Satellite and In-Situ Observations and Climate Models (www.issibern.ch). We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP5, and we thank the climate modeling groups (listed in Table S1 of this paper) for producing and making available their model output.

REFERENCES

- Ablain, M., and Coauthors, 2015: Improved sea level record over the satellite altimetry era (1993–2010) from the Climate Change Initiative project. *Ocean Sci.*, **11**, 67–82, doi:10.5194/os-11-67-2015.
- , J. Legeais, P. Prandi, M. Marcos, L. Fenoglio-Marc, H. Dieng, J. Benveniste, and A. Cazenave, 2017: Satellite altimetry-based sea level at global and regional scales. *Surv. Geophys.*, **38**, 7–31, doi:10.1007/s10712-016-9389-8.
- Andersen, O., P. Knudsen, and L. Stenseng, 2015: The DTU13 MSS (mean sea surface) and MDT (mean dynamic topography) from 20 years of satellite altimetry. *IGFS 2014: Proceedings of the 3rd International Gravity Field Service*, S. Jin and R. Barzaghi, Eds., IAG Symposia, Vol. 144, Springer, 111–121, doi:10.1007/1345_2015_182.
- Becker, M., M. Karpytchev, M. Marcos, S. Jevrejeva, and S. Lennartz-Sassinek, 2016: Do climate models reproduce complexity of observed sea level changes? *Geophys. Res. Lett.*, **43**, 5176–5184, doi:10.1002/2016GL068971.
- Benveniste, J., and P. Knutsen, and the GUTS Team, 2007: The GOCE User Toolbox. *Proc. 3rd International GOCE User Workshop*, SP-627, Frascati, Italy, European Space Agency, 265–268.
- Bilbao, R. A., J. M. Gregory, and N. Bouttes, 2015: Analysis of the regional pattern of sea level change due to ocean dynamics and density change for 1993–2009 in observations and CMIP5 AOGCMs. *Climate Dyn.*, **45**, 2647–2666, doi:10.1007/s00382-015-2499-z.
- Bingham, R., and C. Hughes, 2012: Local diagnostics to estimate density-induced sea level variations over topography and along coastlines. *J. Geophys. Res.*, **117**, C01013, doi:10.1029/2011JC007276.
- Böning, C. W., E. Behrens, A. Biastoch, K. Getzlaff, and J. L. Bamber, 2016: Emerging impact of Greenland meltwater on deepwater formation in the North Atlantic Ocean. *Nat. Geosci.*, **9**, 523–527, doi:10.1038/ngeo2740.
- Bruinsma, S. L., C. Förste, O. Abrikosov, J.-M. Lemoine, J.-C. Marty, S. Mulet, M.-H. Rio, and S. Bonvalot, 2014: ESA's satellite-only gravity field model via the direct approach based on all GOCE data. *Geophys. Res. Lett.*, **41**, 7508–7514, doi:10.1002/2014GL062045.
- Calafat, F., D. Chambers, and M. Tsimplis, 2012: Mechanisms of decadal sea level variability in the eastern North Atlantic and the Mediterranean Sea. *J. Geophys. Res.*, **117**, C09022, doi:10.1029/2012JC008285.
- Chafik, L., J. E. Ø. Nilsen, and S. Dangendorf, 2017: Impact of North Atlantic teleconnection patterns on Northern European sea level. *J. Mar. Sci. Eng.*, **5**, 43, doi:10.3390/jmse5030043.
- Church, J. A., and Coauthors, 2013: Sea level change. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1137–1216.
- Dangendorf, S., F. M. Calafat, A. Arns, T. Wahl, I. D. Haigh, and J. Jensen, 2014a: Mean sea level variability in the North Sea: Processes and implications. *J. Geophys. Res. Oceans*, **119**, 6820–6841, doi:10.1002/2014JC009901.
- , D. Rybski, C. Muddersbach, A. Müller, E. Kaufmann, E. Zorita, and J. Jensen, 2014b: Evidence for long-term memory in sea level. *Geophys. Res. Lett.*, **41**, 5530–5537, doi:10.1002/2014GL060538.
- Dziewonski, A., and D. Anderson, 1981: Preliminary reference Earth model. *Phys. Earth Planet. Inter.*, **25**, 297–356, doi:10.1016/0031-9201(81)90046-7.
- Farrell, W., and J. Clark, 1976: On postglacial sea level. *Geophys. J. Int.*, **46**, 647–667, doi:10.1111/j.1365-246X.1976.tb01252.x.
- Frederikse, T., R. Riva, M. Kleinherenbrink, Y. Wada, M. Broeke, and B. Marzeion, 2016: Closing the sea level budget on a regional scale: Trends and variability on the Northwestern European continental shelf. *Geophys. Res. Lett.*, **43**, 10864–10872, doi:10.1002/2016GL070750.
- Furevik, T., and J. E. O. Nilsen, 2005: Large-scale atmospheric circulation variability and its impacts on the Nordic seas ocean climate—A review. *The Nordic seas: An Integrated Perspective*, *Geophys. Monogr.*, Vol. 158, Amer. Geophys. Union, 105–136, doi:10.1029/158GM09.
- Griffies, S. M., and Coauthors, 2014: An assessment of global and regional sea level for years 1993–2007 in a suite of interannual CORE-II simulations. *Ocean Modell.*, **78**, 35–89, doi:10.1016/j.oceomod.2014.03.004.
- Grinsted, A., S. Jevrejeva, R. E. Riva, and D. Dahl-Jensen, 2015: Sea level rise projections for northern Europe under RCP8.5. *Climate Res.*, **64**, 15–23, doi:10.3354/cr01309.
- Häkkinen, S., P. B. Rhines, and D. L. Worthen, 2013: Northern North Atlantic sea surface height and ocean heat content variability. *J. Geophys. Res. Oceans*, **118**, 3670–3678, doi:10.1002/jgrc.20268.
- Hughes, C. W., and S. D. Williams, 2010: The color of sea level: Importance of spatial variations in spectral shape for assessing the significance of trends. *J. Geophys. Res.*, **115**, C10048, doi:10.1029/2010JC006102.

- , R. J. Bingham, V. Roussenov, J. Williams, and P. L. Woodworth, 2015: The effect of Mediterranean exchange flow on European time mean sea level. *Geophys. Res. Lett.*, **42**, 466–474, doi:10.1002/2014GL062654.
- Jin, S., G. Feng, and O. Andersen, 2014: Errors of mean dynamic topography and geostrophic current estimates in China's marginal seas from *GOCE* and satellite altimetry. *J. Atmos. Oceanic Technol.*, **31**, 2544–2555, doi:10.1175/JTECH-D-13-00243.1.
- Johannessen, J. A., and Coauthors, 2003: The European Gravity Field and Steady-State Ocean Circulation Explorer satellite mission: Its impact on geophysics. *Surv. Geophys.*, **24**, 339–386, doi:10.1023/B:GEOP.0000004264.04667.5e.
- Knudsen, P., R. Bingham, O. Andersen, and M.-H. Rio, 2011: A global mean dynamic topography and ocean circulation estimation using a preliminary *GOCE* gravity model. *J. Geod.*, **85**, 861–879, doi:10.1007/s00190-011-0485-8.
- Knutti, R., J. Sedláček, B. M. Sanderson, R. Lorenz, E. M. Fischer, and V. Eyring, 2017: A climate model projection weighting scheme accounting for performance and interdependence. *Geophys. Res. Lett.*, **44**, 1909–1918, doi:10.1002/2016GL072012.
- Landerer, F. W., P. J. Gleckler, and T. Lee, 2014: Evaluation of CMIP5 dynamic sea surface height multi-model simulations against satellite observations. *Climate Dyn.*, **43**, 1271–1283, doi:10.1007/s00382-013-1939-x.
- Lyu, K., X. Zhang, J. A. Church, A. B. Slangen, and J. Hu, 2014: Time of emergence for regional sea-level change. *Nat. Climate Change*, **4**, 1006–1010, doi:10.1038/nclimate2397.
- Marzeion, B., A. H. Jarosch, and M. Hofer, 2012: Past and future sea-level change from the surface mass balance of glaciers. *Cryosphere*, **6**, 1295–1322, doi:10.5194/tc-6-1295-2012.
- , P. W. Leclercq, J. G. Cogley, and A. H. Jarosch, 2015: Brief communication: Global reconstructions of glacier mass change during the 20th century are consistent. *Cryosphere*, **9**, 2399–2404, doi:10.5194/tc-9-2399-2015.
- Meyssignac, B., D. Salas y Melia, M. Becker, W. Llovel, and A. Cazenave, 2012: Tropical Pacific spatial trend patterns in observed sea level: Internal variability and/or anthropogenic signature? *Climate Past*, **8**, 787–802, doi:10.5194/cp-8-787-2012.
- Milne, G., and J. Mitrovica, 1998: Postglacial sea-level change on a rotating Earth. *Geophys. J. Int.*, **133**, 1–19, doi:10.1046/j.1365-246X.1998.1331455.x.
- Mitrovica, J., and W. Peltier, 1991: On postglacial geoid subsidence over the equatorial oceans. *J. Geophys. Res.*, **96**, 20 053–20 071, doi:10.1029/91JB01284.
- Mitrovica, J. X., M. E. Tamisiea, J. L. Davis, and G. A. Milne, 2001: Recent mass balance of polar ice sheets inferred from patterns of global sea-level change. *Nature*, **409**, 1026–1029, doi:10.1038/35059054.
- New, M., D. Lister, M. Hulme, and I. Makin, 2002: A high-resolution data set of surface climate over global land areas. *Climate Res.*, **21**, 1–25, doi:10.3354/cr021001.
- Nilsen, J. E. O., Y. Gao, H. Drange, T. Furevik, and M. Bentsen, 2003: Simulated North Atlantic-Nordic seas water mass exchanges in an isopycnic coordinate OGCM. *Geophys. Res. Lett.*, **30**, 1536, doi:10.1029/2002GL016597.
- Palanisamy, H., B. Meyssignac, A. Cazenave, and T. Delcroix, 2015: Is anthropogenic sea level fingerprint already detectable in the Pacific Ocean? *Environ. Res. Lett.*, **10**, 084024, doi:10.1088/1748-9326/10/8/084024.
- Peltier, W. R., 2004: Global glacial isostasy and the surface of the ice-age Earth: The ICE-5G (VM2) model and GRACE. *Annu. Rev. Earth Planet. Sci.*, **32**, 111–149, doi:10.1146/annurev.earth.32.082503.144359.
- Perrette, M., F. Landerer, R. Riva, K. Frieler, and M. Meinshausen, 2013: A scaling approach to project regional sea level rise and its uncertainties. *Earth Syst. Dyn.*, **4**, 11–29, doi:10.5194/esd-4-11-2013.
- Rahmstorf, S., G. Feulner, M. E. Mann, A. Robinson, S. Rutherford, and E. J. Schaffernicht, 2015: Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nat. Climate Change*, **5**, 475–480, doi:10.1038/nclimate2554.
- Räsänen, J., L. Ruokolainen, and J. Ylhäisi, 2010: Weighting of model results for improving best estimates of climate change. *Climate Dyn.*, **35**, 407–422, doi:10.1007/s00382-009-0659-8.
- Raj, R. P., 2017: Surface velocity estimates of the north Indian Ocean from satellite gravity and altimeter missions. *Int. J. Remote Sens.*, **38**, 296–313, doi:10.1080/01431161.2016.1266106.
- Richter, K., and B. Marzeion, 2014: Earliest local emergence of forced dynamic and steric sea-level trends in climate models. *Environ. Res. Lett.*, **9**, 114009, doi:10.1088/1748-9326/9/11/114009.
- , J. E. Ø. Nilsen, and H. Drange, 2012a: Contributions to sea level variability along the Norwegian coast for 1960–2010. *J. Geophys. Res.*, **117**, C05038, doi:10.1029/2011JC007826.
- , O. Segtman, and T. Furevik, 2012b: Variability of the Atlantic inflow to the Nordic seas and its causes inferred from observations of sea surface height. *J. Geophys. Res.*, **117**, C04004, doi:10.1029/2011JC007719.
- , B. Marzeion, and R. Riva, 2017: The effect of spatial averaging and glacier melt on detecting a forced signal in regional sea level. *Environ. Res. Lett.*, **12**, 034004, doi:10.1088/1748-9326/aa5967.
- Riva, R. E. M., J. L. Bamber, D. A. Lavallée, and B. Wouters, 2010: Sea-level fingerprint of continental water and ice mass change from GRACE. *Geophys. Res. Lett.*, **37**, L19605, doi:10.1029/2010GL044770.
- Sandø, A., J. Nilsen, T. Eldevik, and M. Bentsen, 2012: Mechanisms for variable North Atlantic-Nordic seas exchanges. *J. Geophys. Res.*, **117**, C12006, doi:10.1029/2012JC008177.
- Santer, B. D., K. E. Taylor, T. M. L. Wigley, J. E. Penner, P. D. Jones, and U. Cubasch, 1995: Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dyn.*, **12**, 77–100, doi:10.1007/BF00223722.
- Schaeffer, P., Y. Faugère, J. F. Legeais, A. Ollivier, T. Guinle, and N. Picot, 2012: The CNES_CLS11 global mean sea surface computed from 16 years of satellite altimeter data. *Mar. Geod.*, **35**, 3–19, doi:10.1080/01490419.2012.718231.
- Schneider, B., M. Latif, and A. Schmittner, 2007: Evaluation of different methods to assess model projections of the future evolution of the Atlantic meridional overturning circulation. *J. Climate*, **20**, 2121–2132, doi:10.1175/JCLI4128.1.
- Sen Gupta, A., N. C. Jourdain, J. N. Brown, and D. Monselesan, 2013: Climate drift in the CMIP5 models. *J. Climate*, **26**, 8597–8615, doi:10.1175/JCLI-D-12-00521.1.
- Shepherd, A., and Coauthors, 2012: A reconciled estimate of ice-sheet mass balance. *Science*, **338**, 1183–1189, doi:10.1126/science.1228102.
- Simpson, M. J. R., K. Breili, and H. P. Kierulf, 2014: Estimates of twenty-first century sea-level changes for Norway. *Climate Dyn.*, **42**, 1405–1424, doi:10.1007/s00382-013-1900-z.

- , and Coauthors, 2015: Sea level change for Norway: Past and present observations and projections to 2100. NCCS Rep. 1/2015, Norwegian Centre for Climate Services, 156 pp., <http://www.miljodirektoratet.no/Documents/publikasjoner/M405/M405.pdf>.
- Slangen, A., C. Katsman, R. van de Wal, L. Vermeersen, and R. Riva, 2012: Towards regional projections of twenty-first century sea-level change based on IPCC-SRES scenarios. *Climate Dyn.*, **38**, 1191–1209, doi:10.1007/s00382-011-1057-6.
- , M. Carson, C. Katsman, R. van de Wal, A. Köhl, L. Vermeersen, and D. Stammer, 2014: Projecting twenty-first century regional sea-level changes. *Climatic Change*, **124**, 317–332, doi:10.1007/s10584-014-1080-9.
- , J. A. Church, X. Zhang, and D. P. Monselesan, 2015: The sea level response to external forcings in historical simulations of CMIP5 climate models. *J. Climate*, **28**, 8521–8539, doi:10.1175/JCLI-D-15-0376.1.
- Tamisiea, M. E., 2011: Ongoing glacial isostatic contributions to observations of sea level change. *Geophys. J. Int.*, **186**, 1036–1044, doi:10.1111/j.1365-246X.2011.05116.x.
- Tapley, B. D., and Coauthors, 1994: Precision orbit determination for TOPEX/POSEIDON. *J. Geophys. Res. Oceans*, **99**, 24 383–24 404, doi:10.1029/94JC01645.
- Taylor, K., R. Stouffer, and G. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Thomson, R. E., and W. J. Emery, 2001: *Data Analysis Methods in Physical Oceanography*. Elsevier, 654 pp.
- Ulbrich, U., and M. Christoph, 1999: A shift of the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing. *Climate Dyn.*, **15**, 551–559, doi:10.1007/s003820050299.
- Van Vuuren, D., and Coauthors, 2011: The representative concentration pathways: An overview. *Climatic Change*, **109**, 5–31, doi:10.1007/s10584-011-0148-z.
- Yin, J. J., S. M. Griffies, and R. J. Stouffer, 2010: Spatial variability of sea level rise in twenty-first century projections. *J. Climate*, **23**, 4585–4607, doi:10.1175/2010JCLI3533.1.