JGR Oceans

RESEARCH ARTICLE
10.1029/2022JC019342

Key Points:

- Assimilating observations of sea surface height improves initialization of subsurface ocean temperatures in climate forecasting systems
- By including altimetry assimilation, monthly forecasts are improved of upper-ocean heat content and sea level
- Sea surface temperature is only minimally affected by altimetry assimilation, at least in the seasonal forecasts assessed here

Correspondence to:
M. J. Widlansky,
mwidlans@hawaii.edu

Citation:
Widlansky, M. J., Long, X., Balmaseda, M. A., Spillman, C. M., Smith, G., Zuo, H., et al. (2023). Quantifying the benefits of altimetry assimilation in seasonal forecasts of the upper ocean. *Journal of Geophysical Research: Oceans*, *128*, e2022JC019342. https://doi.org/10.1029/2022JC019342

Received 28 SEP 2022
Accepted 18 APR 2023

Quantifying the Benefits of Altimetry Assimilation in Seasonal Forecasts of the Upper Ocean

Matthew J. Widlansky[1], Xiaoyu Long[2,3], Magdalena A. Balmaseda[4], Claire M. Spillman[5], Grant Smith[5], Hao Zuo[4], Yonghong Yin[5], Oscar Alves[5], and Arun Kumar[6]

[1]School of Ocean and Earth Science and Technology, Cooperative Institute for Marine and Atmospheric Research, University of Hawai'i at Mānoa, Honolulu, HI, USA, [2]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, [3]NOAA Physical Sciences Laboratory, Boulder, CO, USA, [4]European Centre for Medium-Range Weather Forecasts, Reading, UK, [5]Bureau of Meteorology, Melbourne, VIC, Australia, [6]Climate Prediction Center, NCEP/NWS/NOAA, College Park, MD, USA

Abstract Satellite altimetry measurements of sea surface height provide near-global ocean state observations on sub-monthly time scales, which are not always utilized by seasonal climate forecasting systems. As early as the mid-1990s, attempts were made to assimilate altimetry observations to initialize climate models. These experiments demonstrated improved ocean forecasting skill, especially compared to experiments that did not assimilate subsurface ocean temperature information. Nowadays, some operational climate forecasting models utilize altimetry in their assimilation systems, whereas others do not. Here, we assess the impact of altimetry assimilation on seasonal prediction skill of ocean variables in two climate forecasting systems that are from the European Centre for Medium-Range Weather Forecasts (SEAS5) and the Australian Bureau of Meteorology (ACCESS-S). We show that assimilating altimetry improves the initialization of subsurface ocean temperatures, as well as seasonal forecasts of monthly variability in upper-ocean heat content and sea level. Skill improvements are largest in the subtropics, where there are typically less subsurface ocean observations available to initialize the forecasts. In the tropics, there are no noticeable improvements in forecast skill. The positive impact of altimetry assimilation on forecast skill related to the subsurface ocean does not seem to affect predictions of sea surface temperature. Whether this is because current forecasting systems are close to the potential predictability limit for the ocean surface, or perhaps altimetry observations are not fully exploited, remains a question. In summary, we find that utilizing altimetry observations improves the overall global ocean forecasting skill, at least for upper-ocean heat content and sea level.

Plain Language Summary Sea surface heights have been nearly continuously and globally measured by satellite-based altimeters for almost 30 years, yet not all climate models utilize these observations during their initialization phases of seasonal forecasting. Since the local sea surface height (SSH), or sea level, is mostly determined by the ocean temperature and salinity-controlled subsurface density structure, the altimetry measurements contain a vast amount of integrated information about the ocean climate state. Climate forecasting systems are usually most skillful when they start from the most accurate initial conditions, including the state of the ocean density structure. Unfortunately, temperature and salinity observations are sparse in many parts of the world's oceans, which degrades the initialization of models and thus possibly diminishes forecast skill. Here, we quantify the benefits of using altimetry observations to initialize two state-of-the-art climate forecasting systems by verifying seasonal forecasts using observational products of upper-ocean temperature, sea surface temperature, and SSH. In the experiments that did not use altimetry data assimilation, we find an overall decrease of seasonal forecasting skill for the subsurface temperature and SSH, which is especially pronounced in the global subtropics, although we find no widespread change in skill for sea surface temperature. Thus, for predicting the subsurface ocean, climate forecasting systems may benefit from incorporating altimetry data assimilation into their ocean initial conditions. More work is needed to assess how much of the subsurface skill improvements can potentially affect the surface ocean.

1. Introduction

Satellite-based altimeters have been measuring sea surface height (SSH) with routine near-global coverage since the early 1990s (Chelton et al., 2001). In addition to providing critical information about regional sea level variability on sub-monthly to decadal timescales (Chelton & Schlax, 1996), as well as long-term global sea level rise (Nerem et al., 2018), altimetry observations have been used to describe the ocean circulation (L. Fu & Smith, 1996) and subsurface density structure (Feng & Zhong, 2015; García et al., 2007). Since the local SSH is mostly determined by the integrated effect of underlying temperature anomalies on seawater density in many locations (e.g., Widlansky et al., 2020), satellite-derived observations of SSH provide information about the subsurface ocean that may be otherwise unknown. Use of altimetry information about the subsurface ocean can provide an opportunity to improve the ocean-state description (Webb & Moore, 1986), which may be especially useful for the initialization of seasonal climate forecasts.

Since the operational use of coupled ocean-atmosphere models to predict seasonal changes in the climate (e.g., Anderson, 2012; Stockdale, 1997; Stockdale et al., 1998), their forecast skill has been shown to depend at least in part on accurately initializing the ocean physical state (Palmer & Anderson, 1994). Ocean physics can be fully described by the combined observation of seawater temperature and salinity, as well as the zonal and meridional circulation (respectively, abbreviated T, S, U, and V), and the associated SSH. Observations of T, S, U, and V are extensive in many parts of the world's oceans, especially since the global network of Argo floats and profilers began in 2002 (Feder, 2000). Mooring networks have provided comprehensive oceanic observations for a much longer period in some places, such as the TAO/TRITON array in the equatorial Pacific where the subsurface T data (Hayes et al., 1991) was used to improve understanding of the El Niño-Southern Oscillation (ENSO) and advance forecasting systems (e.g., Stockdale et al., 1998; Vidard et al., 2007). Elsewhere, however, subsurface ocean observations are typically more limited.

To supplement the in-situ observation network (e.g., floats, moorings, and profilers), and provide climate models with additional information about the ocean physical state, altimetry observations of SSH are assimilated into several climate reanalysis products. Examples include the Ocean Reanalysis System 5 (ORAS5; Zuo et al., 2019) from the European Centre for Medium Range Weather Forecasts (ECMWF) as well as each of the other models contributing to the Global Reanalysis Ensemble Product (GREP; Storto, Masina, et al., 2019) from the European Copernicus Marine Environment Monitoring Service (CMEMS). Altimetry assimilation follows different methodologies in the different systems, but most of them try to project the SSH information onto the vertical density structure of the ocean (Alves et al., 2001; Vidard et al., 2009). For instance, in the ORAS5, the relationship between SSH and subsurface density variations (T and S) uses the linearized version of the buoyancy frequency (Balmaseda et al., 2013). Hence, the SSH increment affects each of the ocean state variables in the model, with the subsurface T profile usually influenced the most by the assimilation of altimetry due to the strong dependence of the ocean temperature on the density and, consequently, on sea level variability (e.g., Widlansky et al., 2020).

There are also examples of climate forecasting systems that include altimetry observations in their ocean initializations (e.g., Balmaseda, 2017), which is usually accomplished similarly to the procedure used by the climate reanalysis products. The first successful use of altimetry to improve the initialization of a climate forecasting model was shown by Segschneider et al. (2000). However, it took several years for these early attempts to be transferred to operations (Balmaseda et al., 2008). Later, Balmaseda and Anderson (2009) showed that the assimilation of altimetry improved the seasonal forecasts of sea surface temperature (SST), primarily by correcting the ocean mean state in the initial conditions, which was worse than now. Present-day examples include the operational climate forecast model from ECMWF (i.e., SEAS5; S. J. Johnson et al., 2019) as well as the previous operational forecast system from the Australian Bureau of Meteorology (i.e., EESS-S1; Hudson et al., 2017); both of which assimilate the Global Ocean Along Track SSH product from CMEMS (i.e., altimetry with Level 3 processing). The current operational version of ACCESS-S (i.e., -S2; Wedd et al., 2022) does not include altimetry assimilation, which provides a convenient opportunity for comparing what are otherwise mostly similar forecasting systems.

Despite these earlier studies demonstrating the feasibility of using altimetry assimilation to improve climate forecasting models, and the later incorporation of the procedure into some operational frameworks, the impact of altimetry on seasonal forecasting skill has not been quantified in current-generation systems. Furthermore, many climate models used currently for seasonal forecasting do not include altimetry assimilation, such as NOAA's CFSv2 or any of the other models participating in the North American Multi-Model Ensemble (NMME; Kirtman et al., 2014). Recently, some of the above-mentioned models (i.e., SEAS5, ACCESS-S1, CFSv2, and five other models from the NMME) were assessed together for their seasonal forecasting skill of monthly sea level anomalies (Long et al., 2021). In general, the models that did assimilate altimetry (i.e., SEAS5 and ACCESS-S1) had higher skill compared to the other models. However, in addition to the inclusion of altimetry or not in the forecasting models, there were other important variations among the models (especially their horizontal resolutions of the ocean) that made it impossible for Long et al. (2021) to determine the exact cause of skill differences.

Here, we will use a dedicated experiment with altimetry assimilation turned off in the SEAS5 initialization to quantify the impact of that modification on ocean seasonal forecasting skill. We will also utilize the recent change from ACCESS-S1 to ACCESS-S2 to make a similar comparison, since only the former version includes altimetry assimilation. However, we note that the latter comparison is imperfect because a new data assimilation system was implemented for ACCESS-S2, which was developed in-house at the Bureau of Meteorology (Wedd et al., 2022), in addition to other mostly minor changes. Using the first comparison, we will isolate the role of altimetry assimilation on ocean forecasting skill in the SEAS5 model and, likewise, the comparison of ACCESS-S models will provide further validation in an additional operational forecasting system. All of the forecasting systems are fairly comparable in modeling design (Hudson et al., 2017; S. J. Johnson et al., 2019; Wedd et al., 2022), with each using the same ocean model (Nucleus for European Modeling of the Ocean; NEMO; Gurvan et al., 2022) ran at a nominal 0.25° eddy-permitting horizontal resolution that is assimilated using mostly similar ocean observations (except of course for whether or not altimetry is included), which makes feasible the comparisons presented here.

Our hypothesis consists of two parts, which are primarily based on the results of the Segschneider et al. (2000) as well as Palmer and Anderson (1994) studies mentioned above. First, we expect that the altimetry assimilation included in SEAS5 (i.e., the Control setup), as well as ACCESS-S1, improves the initialization of the ocean subsurface density structure. Hence, we expect these assimilation systems to have a more accurate SSH state estimation (mostly via a thermosteric sea level response), compared to the systems with no altimetry assimilation (i.e., the SEAS5-Experiment or ACCESS-S2). Second, we anticipate that models with improved ocean state estimation (i.e., the SEAS5-Control or ACCESS-S1) will have higher forecast skill at the monthly-to-seasonal leads, compared to the systems lacking altimetry assimilation. However, if the altimetry assimilation does not in fact improve the initial ocean density structure (e.g., if temperatures are nudged in the models at the wrong vertical levels), then we hypothesize that ocean forecasts using such methods are either not improved or possibly made worse. We also do not expect altimetry assimilation to have a large effect on SST for several reasons: (a) SST is already well observed (primarily by other satellite instruments) and assimilated into the models with fidelity; (b) there is a multitude of processes affecting SST forecasts (e.g., air-sea interactions) beyond the quality of the ocean initial conditions; and, (c) the mean state of the upper-ocean temperature is now better constrained even in the absence of altimetry observations thanks to improvements in the in-situ observing system and data assimilation algorithms (M. A. Balmaseda & Anderson, 2009), especially in the equatorial Pacific where SST seasonal variability is large. The latter reasoning also suggests a geographical dependence (i.e., outside of the equatorial Pacific) on the usefulness of altimetry assimilation with regards to improving forecasts of other variables related to upper-ocean temperatures.

By assessing the retrospective forecasts from the Control and Experiment setups of the SEAS5 model, and likewise comparing ACCESS-S1 with ACCESS-S2, we will determine whether the assimilation of altimetry-measured SSH improves the forecasts of SSH and upper-300 m ocean heat content (OHC-300), since these variables capture the potential impact of ocean subsurface physics (e.g., dynamical and thermal memory; Bulgin et al., 2020; Shi et al., 2022). Furthermore, both SSH and OHC-300 are of interest on their own for forecasting applications, such as concerning sea level coastal impacts (Dusek et al., 2022; Stephens et al., 2014) and the effects of marine heatwaves on ecosystems (Behrens et al., 2019; C. M. Spillman & Smith, 2021). We will also analyze the forecasts of SST, which we will use to compare the effects of altimetry assimilation at the ocean surface versus the subsurface characteristics represented by OHC-300 and SSH.

The remainder of the paper is organized as follows. In the next section, the data and methods are described with attention to the choice of verification data as well as the experimental setup. Forecast assessment results are then presented in Section 3. A summary of the study is provided in the final section, followed by discussion to interpret the results and consider opportunities for improving future seasonal climate forecasting systems.

2. Data and Methods

We describe the seasonal variability of upper-ocean physics using monthly anomalies of three variables: OHC-300, SST, and SSH. For each variable, we choose an observational product and then compare it to the equivalent forecast model output. We quantify the observation-to-model comparison using temporal anomaly correlation coefficients (ACC) and root-mean square error (RMSE), which we calculate by first removing the monthly climatology from the respective data sets and then determining either the correlations or errors between the remaining anomalies. The assessment epoch was chosen to maximize the amount of retrospective data available since the altimetry era began (1993–2014 for SEAS5 and 1993–2012 for ACCESS-S), and we consider together only the May and November forecast starts because those are the only forecasts available for the SEAS5-Experiment (no altimetry assimilated). Finally, we assess the effect of altimetry assimilation on the forecast skill (i.e., ACC and RMSE at a particular lead time from 0 to 5 months) by subtracting results for the forecasts with no altimetry assimilation (i.e., the experiment) from the control results. We perform all of the assessments for the global oceans that are generally ice free (60°S–60°N), using data regridded to a uniform 1° × 1° latitude-longitude resolution (i.e., 43,200 grid points). For each ACC and RMSE map throughout, the global-average of values between 60°S and 60°N is indicated in the respective panel title (spatial data is weighted by grid-cell area prior to averaging; i.e., by multiplying the gridded data by the cosine of its central latitude). Thus, we quantify both where and by how much the use of altimetry assimilation affects the seasonal forecast skill of the upper-ocean monthly variability.

We first show that OHC-300 and SSH are expected to be strongly correlated nearly everywhere. We see this strong correlation in the NEMO ocean model forced by the atmospheric fluxes from ERA-Interim (Dee et al., 2011) but with no data assimilation in the ocean except for strong nudging to the observed SST (Figure 1a). For reference, the NEMO configuration used in both SEAS5 and ACCESS-S has been shown to produce a realistic representation of observed ocean physics (e.g., Zuo et al., 2017).

The correlation between OHC-300 and SSH in the NEMO model (Figure 1a) is not surprising since large-scale sea level variability has been shown to be linked to the thermocline variability and thus the amount of heat and relative buoyancy of the upper ocean, especially in the tropical Pacific (Long et al., 2020; Timmermann et al., 2010). Besides some regions poleward of 60°N/S with weak or negative correlations, which are outside of the ice-free domain that we will assess forecast skill, the only areas of negative ACC between OHC-300 and SSH are confined to parts of the mid-latitudes (e.g., in parts of the North Atlantic Ocean, Sea of Okhotsk in the northwestern Pacific Ocean, and the Southern Ocean near Africa) as well as in the equatorial Atlantic. Sea level variability is relatively small in the latter region, especially compared to most of the equatorial Pacific and Indian Oceans (Long et al., 2021). The equatorial Atlantic SSH also may be more influenced by salinity than temperature, which is unusual compared to most of the global oceans (Widlansky et al., 2020). There are multiple reports of wind-forced barotropic processes affecting regional sea levels as well (L. L. Fu & Davidson, 1995; Fukumori et al., 1998; Hughes et al., 2018; Wunsch, 1991), which may explain other areas equatorward of 60°N/S with relatively weak correlations between OHC-300 and SSH (e.g., along the U.S. East Coast and in parts of the tropical Indian Ocean; see respectively, Piecuch et al., 2016; Rohith et al., 2019).

Global OHC-300 data must be either interpolated from heterogeneous subsurface observations or inferred by other means. Thus, the question of how to verify OHC-300 forecasts is not trivial since the choice of observational product is likely to affect the assessment. Examples of interpolated data sets are in-situ based objective analyses using primarily Argo data such as the Met Office Hadley Centre EN4 and Institute of Atmospheric Physics (IAP) products, which are respectively described by Good et al. (2013) and Cheng et al. (2017). An alternative approach uses the dynamics inherent with climate reanalysis products (such as GREP) to simulate the OHC-300 variability. This is in contrast with the SST and SSH variables, which are readily studied using globally-gridded data sets that are produced almost entirely from satellite measurements. Here, we use the NOAA Optimum Interpolated (OI) SST V2 and the SSALTO/DUACS multi-mission altimetry data set distributed by the CMEMS (i.e., the Level 4 processing of altimetry), which are respectively described by Reynolds et al. (2002) and Long et al. (2021). Since none of the models include atmospheric pressure forcing on the sea level, we use the standard SSH product from CMEMS that includes a dynamic atmospheric correction to remove the inverse-barometer effect from the altimetry measurements.

We considered that the SEAS5 forecasts verified using the ORAS5 reanalysis might appear to have higher skill, compared to if a more independent observational product was used for the verification. To assist selection of the OHC-300 verification product, we compared the ACC between observed SSH and the OHC-300 from either a multi-model reanalysis (GREP; Figure 1b), single-model reanalysis (ORAS5; Figure 1c), or two different objective analyses (IAP and EN4; Figures 1e and 1g). We note that GREP and ORAS5 are both forced by ERA-Interim atmospheric fluxes (similar to the NEMO experiment; Figure 1a), although different bulk formulas and data

10.1029/2022JC019342

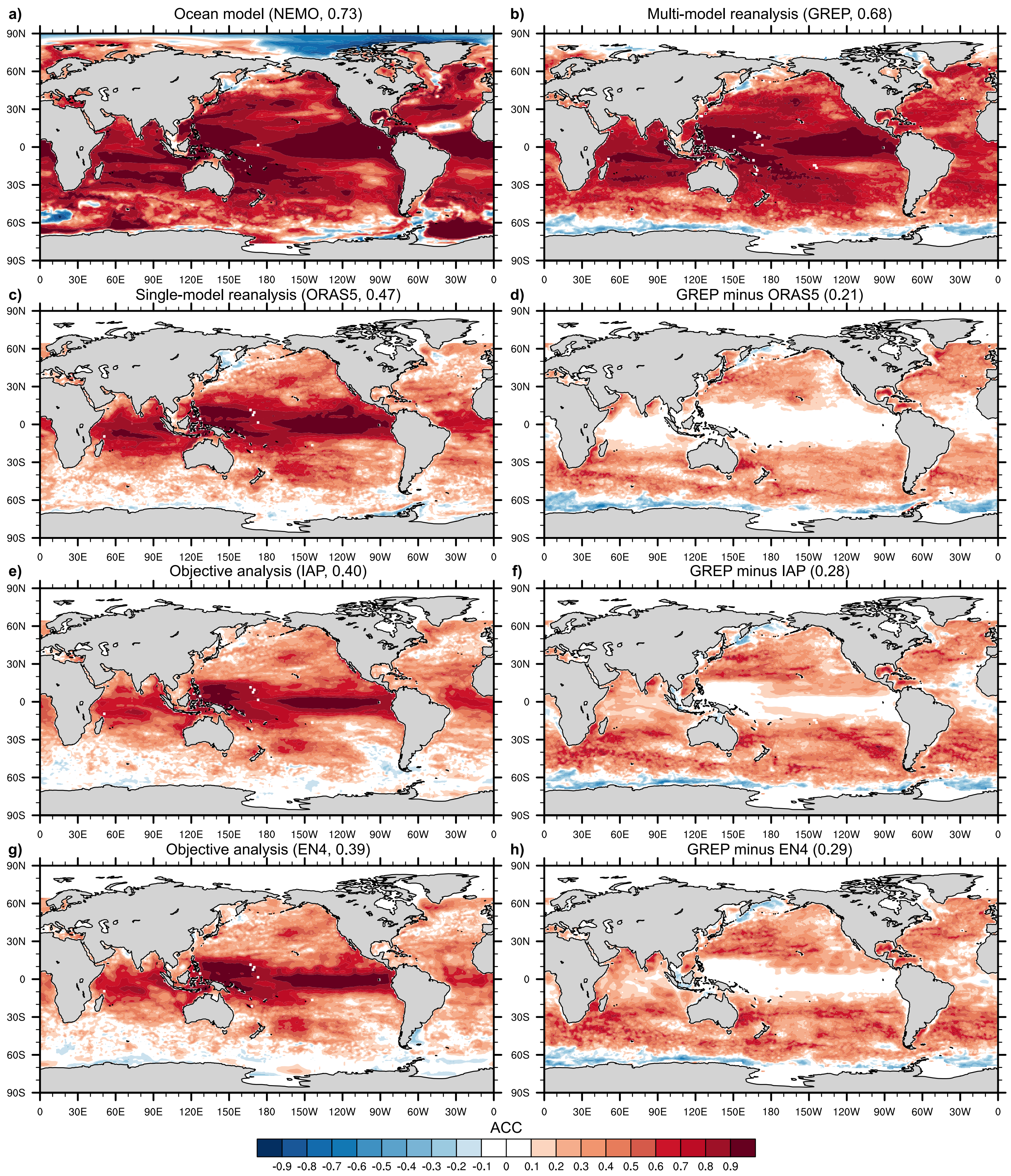

Figure 1. Comparison of the ACC between SSH and OHC-300 in a climate model, reanalysis products, and objective analysis products. The OHC-300 is from either the (a) ocean model (NEMO), (b) multi-model mean reanalysis (GREP), (c) single-model reanalysis (ORAS5) or, (e and g) objective analyses of in-situ ocean temperature observations (IAP and EN4, respectively). The SSH is from either the (a) climate model or, otherwise altimetry observations. (d, f, and h) Differences in ACC between products compared to the multi-model reanalysis (GREP minus ORAS5, GREP minus IAP, and GREP minus EN4, respectively). Global averages of the ACC maps (60°N/S domain) are indicated in parentheses throughout.

10.1029/2022JC019342

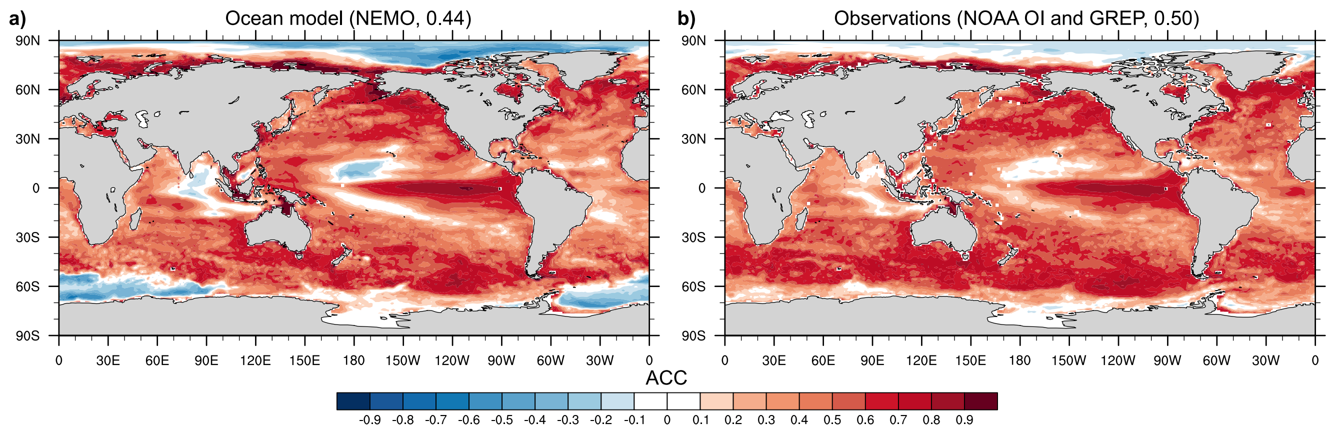

Figure 2. The ACC between SST and OHC-300. (a) Data is from the ocean model (NEMO). (b) Data is from observations (NOAA OI) or the multi-model reanalysis (GREP) for SST or OHC-300, respectively.

assimilation schemes are used across the systems. The multi-model reanalysis result most closely resembles the expected strong correlation between SSH and OHC-300, as inferred from the independent model result (Figure 1a). Furthermore, the ACC using GREP was higher than any of the other products (Figures 1c, 1e and 1f), especially outside of the tropics. Interestingly, there are almost no differences in the equatorial Pacific. Considering the overall stronger ACC between observed SSH and the OHC-300 from GREP, we chose this product for verifying the OHC-300 forecasts.

OHC-300 is also positively correlated with SST in many places, both in the NEMO model and observations (Figure 2). We remind that the SST in this particular NEMO simulation is strongly nudged to observations, which is applied using a surface non-solar heat flux computed with a constant restoration factor (-200 W/m^2/K^{-1}) times the difference between the model background SST and observed SST (Zuo et al., 2019), so the model is not completely representative of independent ocean physics. The relationship between the upper-ocean and surface temperatures is often reported in climate assessments (e.g., G. C. Johnson et al., 2021), although the OHC-300 correlation with SST is not as strong as we saw with SSH. Together, the correlations between OHC-300 and either SSH (Figure 1) or SST (Figure 2) demonstrate the strong relationship between the three ocean variables considered here. Furthermore, McAdam et al. (2022) showed that in SEAS5 the seasonal forecasting skill for OHC-300 is similar to SST in most places, which they interpreted as suggesting that dynamical climate forecasting systems benefit from the thermo-dynamical memory in the ocean initial conditions.

We assess the monthly anomalies of OHC-300, SST, and SSH from climate forecasting systems, and focus on the forecasts for lead-0 and lead-3 months. The former lead time is representative of the initial conditions after the data assimilation, and the latter represents the outlook at the beginning of the next season. Differences of the results between each of the leads that are available for all of the forecasting systems (i.e., out to lead-5 months) are discussed in the final section.

We calculate the forecast monthly anomalies following the methodology in Long et al. (2021), which is namely to remove both the lead-time dependent climatology as well as any remaining long-term trends in each of the variables. Removal of the trend from the forecasts and observations is especially important for consideration of SSH because global sea level rise is treated differently among the models (see Long et al., 2021). Also, substantial trend errors exist in the models for SST (e.g., in the tropical Pacific; L'Heureux et al., 2022). Furthermore, removing the trends is necessary to assess correlations between variables, which we do for OHC-300 and SSH (Figures 1 and 3) as well as SST (Figure 2).

Here, we measure forecast skill using the ACC and RMSE metrics of how well the predictions compare to the respective verification products (i.e., OHC-300 from the GREP multi-model reanalysis, as well as satellite-observed SST and SSH). ACC measures the association between the model forecast and observed variability, whereas RMSE measures the accuracy or amplitude errors of the forecast (Wilks, 2011). Our methodology for quantifying the role of altimetry assimilation in forecast skill consists of simply comparing the ACC and RMSE from retrospective forecasts (i.e., hindcasts) that were made either with or without altimetry assimilation.

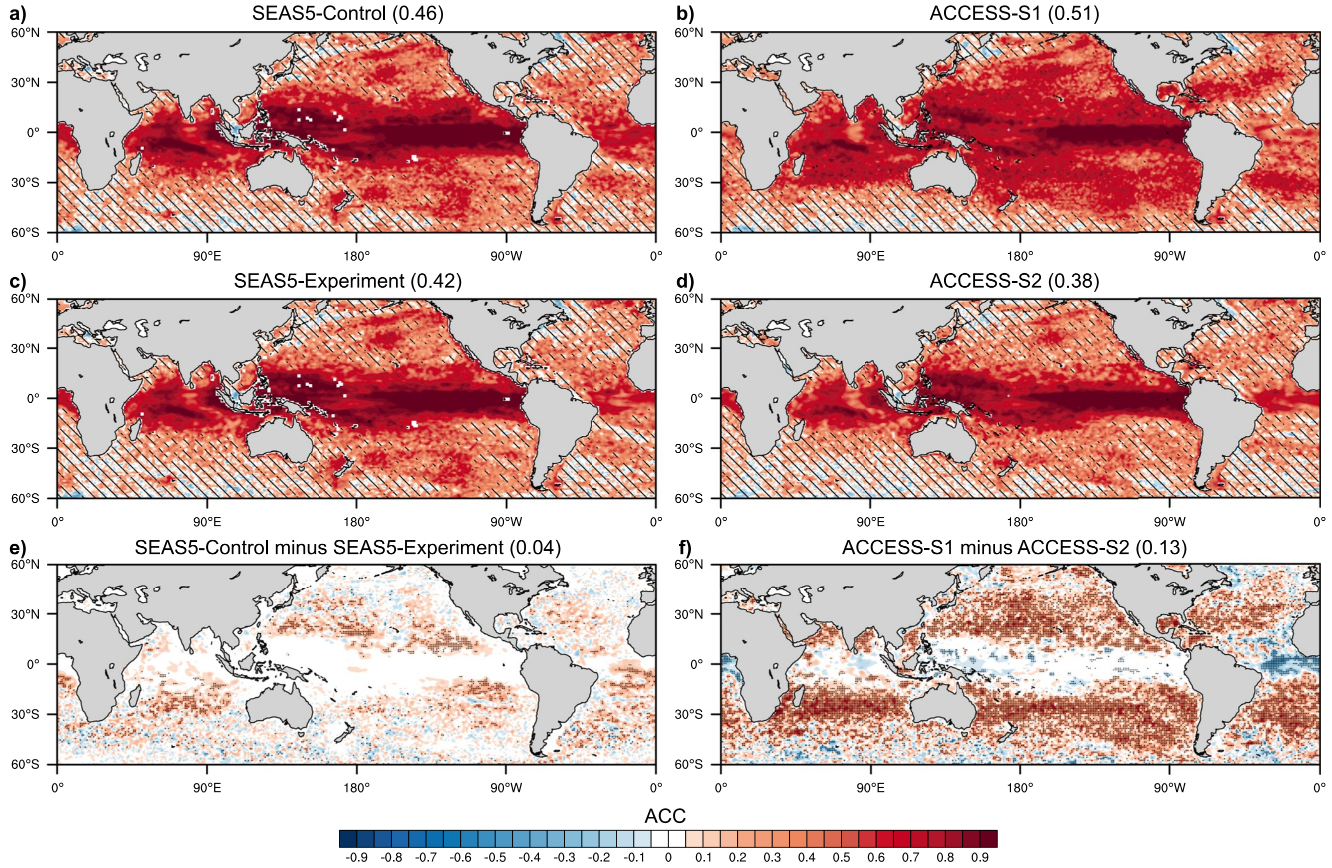

Figure 3. The ACC between observations of SSH (from altimetry) and forecasts of OHC-300 (from the lead-0 month). (a and b) Forecasts with altimetry assimilation (SEAS5-Control and ACCESS-S1, respectively). (c) and (d) Forecasts with no altimetry assimilation (SEAS5-Experiment and ACCESS-S2, respectively). Hatching indicates non-significant ACC values. (e and f) Differences between the forecasts with and without altimetry assimilation. Stippling indicates significant ACC differences.

As mentioned in the introduction, the control forecasts are from SEAS5 and ACCESS-S1, as altimetry assimilation is included in both systems. The experiment forecasts (i.e., without altimetry assimilation) are from a dedicated set of retrospective forecasts (i.e., SEAS5-Experiment) and ACCESS-S2. For each of the models, we use the average of available ensemble members (14 for both configurations of SEAS5, 11 for ACCESS-S1, and 12 for ACCESS-S2). Throughout, we only directly calculate the forecast skill differences between similar models (i.e., SEAS5-Control vs. SEAS5-Experiment and ACCESS-S1 vs. ACCESS-S2).

We calculate statistical significance of the individual ACC values using a two-tailed *t*-test (0.05 confidence level) against a null hypothesis of zero correlation, and with the degrees of freedom determined by the observed autocorrelation decay timescale at each location (see Long et al., 2021). For the combined two forecast starts per year, most locations have about 40 degrees of freedom in both the control and experiment forecasts, with the SEAS5 forecasts having slightly more degrees of freedom overall compared to ACCESS-S because of the longer epoch of retrospective forecasts (i.e., 22 vs. 20 years). Despite the rather limited sample sizes, we attempt to test the significance of the forecast skill differences (i.e., the control minus experiment), which was not done in Long et al. (2021). We use a similar two-tailed *t*-test applied to the difference of Fisher Z-transforms of the ACC values, and that accounts for correlation typically existing between the forecasts being compared (Meng et al., 1992).

3. Results

3.1. Effect of Altimetry Assimilation on the Forecast Initialization

Assessment of the two forecast systems (i.e., SEAS5 and ACCESS-S) during the lead-0 month shows mostly positive ACC values between observed SSH (i.e., from satellite altimetry) and predicted OHC-300 (Figure 3), regardless of whether altimetry assimilation is included in the models. All of the forecasts at lead-0 month have ACC characteristics that resemble the single-model reanalysis result (ORAS5; Figure 1c), which suggests that this lead time is within the influence of skill due to the initialization. We consider the lead-0 month forecasts to be mostly representative of each model's conditions around the time of initialization (i.e., the lead-0 month is used as a proxy reanalysis). Interestingly, neither the control forecasts (i.e., with altimetry assimilation; Figures 3a and 3b) or experiment forecasts (i.e., without altimetry assimilation; Figures 3c and 3d) have as strong correlations as we saw in either the uninitialized ocean model (NEMO; Figure 1a) or the multi-model reanalysis (GREP; Figure 1b), although ACCESS-S1 clearly has the strongest ACC of the forecast models. In general, we see that all of the forecast models at least begin with an expected pattern of correlation between observed SSH and simulated OHC-300 (i.e., higher ACC in the tropics compared to the mid-latitudes; see Figure 1a).

In the forecast models with no altimetry assimilation (Figures 3c and 3d), the ACC between observed SSH and OHC-300 at the lead-0 month is somewhat reduced compared to the control forecasts (Figures 3a and 3b). In both SEAS5 and ACCESS-S, altimetry assimilation has a much larger effect on the lead-0 month forecast outside of the tropics (i.e., typically poleward of about 15°N/S). In the subtropics, especially, the ACC differences associated with using altimetry assimilation are generally positive and are likely to be significant at many specific locations (Figures 3e and 3f; stippling). Conversely, in much of the tropics and especially the equatorial Pacific, altimetry assimilation has almost no effect (as assessed by the ACC metric).

Comparing the reduction of ACC values in the SEAS5-Experiment and ACCESS-S2 (i.e., the forecast models with no altimetry assimilation; Figures 3e and 3f), the differences are usually larger in the latter model. For almost everywhere that altimetry assimilation seems to impact the ACC metric in the SEAS5-Experiment, ACCESS-S2 shows a larger sensitivity (i.e., reduced correlations in the lead-0 month forecasts compared to ACCESS-S1). However, there is an exception near the equator (especially in the Atlantic Ocean) where ACC values are mostly larger in ACCESS-S2, as indicated by the blue shading in Figure 3f. We briefly note here that some of the differences comparing ACCESS-S1 and ACCESS-S2 may be related to other changes in the forecast system besides whether or not altimetry assimilation was included (Wedd et al., 2022), which applies to all such assessments (see also the discussion in Section 4).

3.2. Forecast Skill Assessment

We now consider the effect of altimetry assimilation on forecast skill as quantified by the ACC and RMSE between the predicted and observed ocean variables (i.e., OHC-300, SST, and SSH). For each variable, we will assess the forecast at lead-0 and lead-3 months. As done previously, we will present the ACC (as well as RMSE) values for both the models that either used altimetry assimilation or did not (i.e., the control and experiment forecasts, respectively), followed by the respective differences (i.e., the control minus experiment).

3.2.1. OHC-300

Beginning with the assessment of OHC-300, which is an indicator of a model's ability to predict the upper-ocean thermodynamic conditions as well as dynamical processes such as thermocline variability, we see evidence of skillful seasonal forecasts in much of the global ocean for all of the models (Figures 4 and 5). For the lead-0 month forecast, regardless of whether or not the models used altimetry assimilation, ACC values in most places are significantly different from random correlations (i.e., not hatched in Figures 4a–4d) and the RMSE values are mostly small compared to the observed standard deviations (Figures 5a–5d). The skillful forecasts for most places continue to the lead-3 month (Figures 4g–4j and 5g–j), although the ACC values are clearly diminished nearly everywhere and the RMSE values are higher compared to at the lead-0 month.

The forecast skill at lead-0 and 3 months tends to be better for the SEAS5 models (i.e., higher ACC and lower RMSE) compared to the ACCESS-S models, again regardless of the use of altimetry assimilation. We note that the OHC-300 verification data set is from the multi-model reanalysis (GREP) that includes ORAS5, which is also very similar to the SEAS5-Control at the lead-0 month. This choice of verification data may explain some of the ACC differences between the SEAS5 and ACCESS-S models. Thus the focus of assessing forecast skill differences should mostly be directed to considering the SEAS5 and ACCESS-S models separately.

We quantify the effect of altimetry assimilation on forecast skill, independent of any differences between SEAS5 and the ACCESS-S systems, using row-wise differences of the results in Figures 4 and 5. The assessment reveals

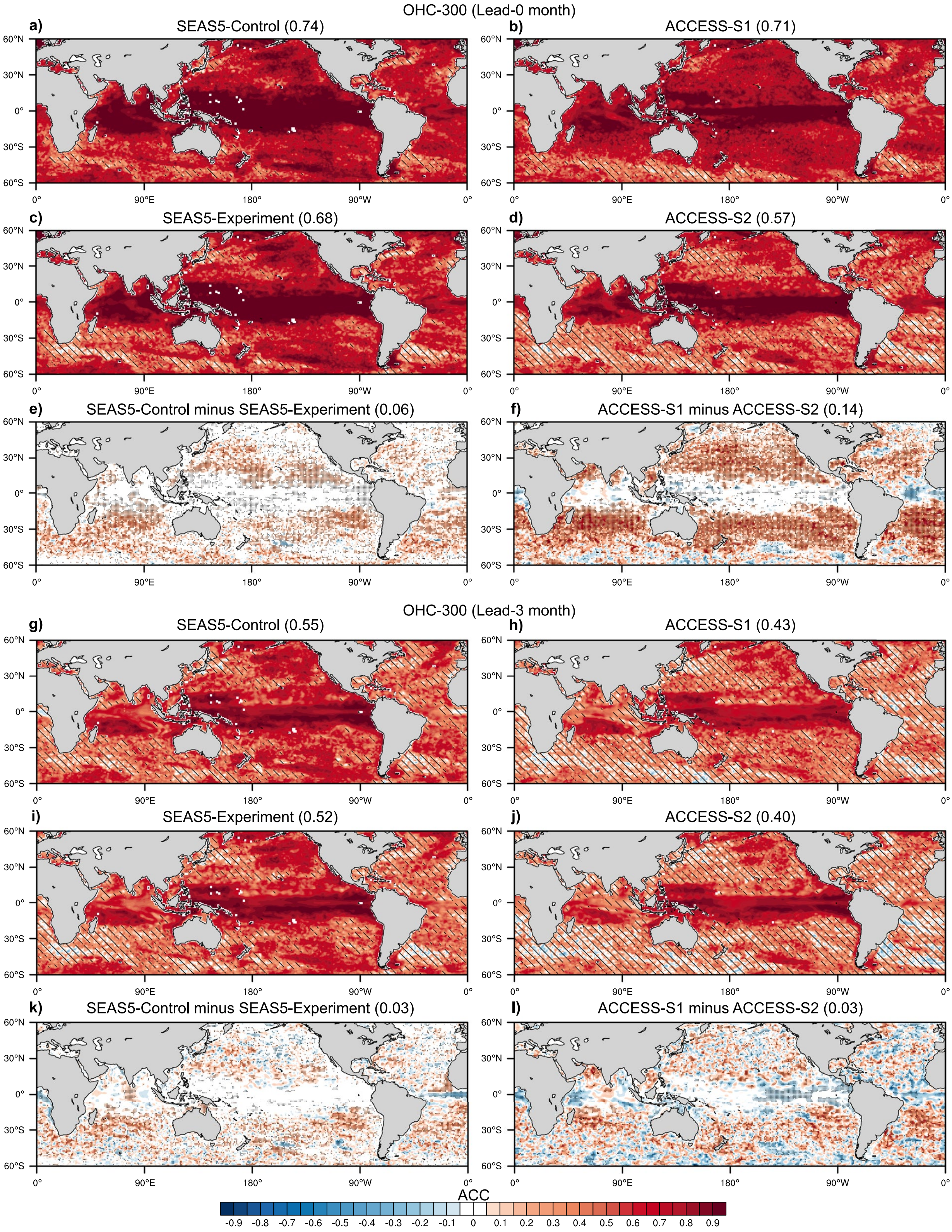

Figure 4. OHC-300 forecast skill measured by the ACC between GREP reanalysis OHC-300 (proxy for observations) and forecast OHC-300 at lead-0 and 3 months (a–d and g–j, respectively). Hatching indicates non-significant ACC values for each of SEAS5-Control, ACCESS-S1, SEAS5-Experiment, and ACCESS-S2. e–f and k–l) Differences between the forecasts with and without altimetry assimilation. Stippling indicates significant ACC differences.

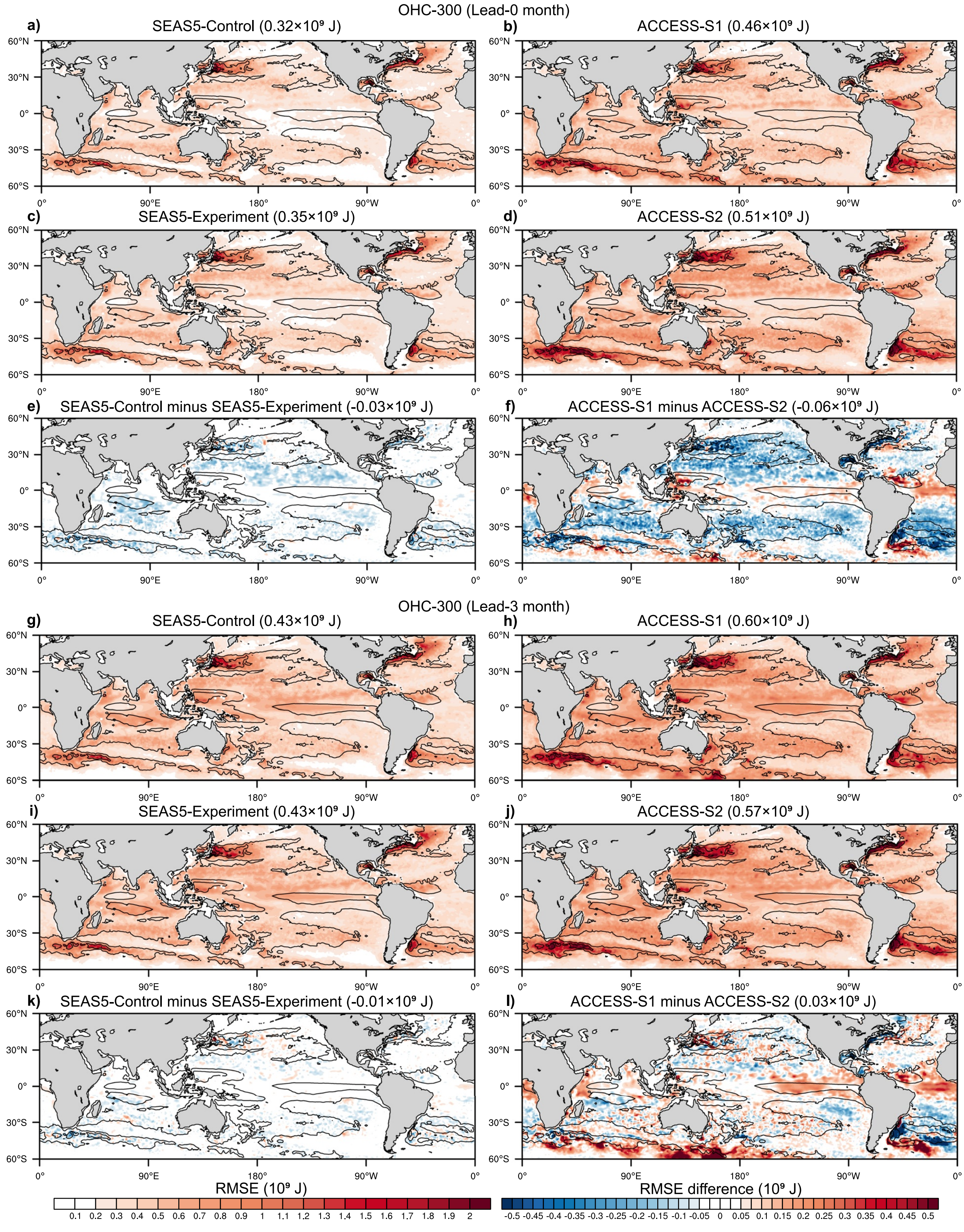

Figure 5. OHC-300 forecast skill measured by the RMSE between GREP reanalysis OHC-300 (proxy for observations) and forecast OHC-300 at lead-0 and 3 months (a–d and g–j, respectively). Black contours indicate the observed standard deviation of OHC-300 (0.5×10^9 J interval). (e–f and k–l) Differences between the forecasts with and without altimetry assimilation.

10.1029/2022JC019342

that the models including altimetry assimilation (i.e., SEAS5-Control and ACCESS-S1) have better forecast skill overall, compared to the corresponding model not using altimetry assimilation (i.e., SEAS5-Experiment and ACCESS-S2). At the lead-0 month, the patterns and amplitudes of the ACC differences for OHC-300 forecast skill (Figures 4e and 4f) closely resemble the previously noted ACC differences between observed SSH and the same OHC-300 model data (Figures 3g and 3h). In particular, the improved OHC-300 forecast using altimetry assimilation is most noticeable in the subtropics. At lead-0 month, there are few places where either model without altimetry assimilation has the higher forecast skill (i.e., blue shading is limited in Figures 4e and 4f). The difference between the two ACCESS-S models is noticeably larger than in the SEAS5 experiment, which suggests that altimetry assimilation may have a larger effect on the ocean in ACCESS-S1, at least relative to the influence from assimilation of in-situ observations (e.g., of subsurface temperature). However, there are some places where the ACCESS-S2 model has higher ACC values than ACCESS-S1, most noticeably in the equatorial regions. The RMSE results closely mirror the ACC assessment with few exceptions during the lead-0 month (Figures 5a–5f). Including altimetry assimilation does seem to improve the ACC and RMSE values overall for both SEAS5 and ACCESS-S1 at the lead-0 month, but it is clearly not a guarantee of improved forecasting ability everywhere, especially in the latter model.

The overall improved OHC-300 forecast skill in the models using altimetry assimilation mostly continues to the 3-month lead, although the improvement (i.e., positive ACC differences; negative RMSE differences) tends to become smaller at longer leads (Figures 4k and 4l, versus Figures 4e and 4f; Figures 5k and 5l, versus Figures 5e and 5f). As we expect that forecast skill usually decreases with increasing lead time, regardless of model or the initialization, it is not surprising that the ACC differences become smaller at the longer lead. The fact that there are any significant positive ACC differences at lead-3 month supports the hypothesis that altimetry assimilation improves the seasonal forecast skill related to OHC-300. The SEAS5 experiment difference is conclusive in support for this hypothesis, since altimetry assimilation was the only altered parameter in that model. Whereas the ACCESS-S1 minus ACCESS-S2 result can only be considered suggestive of the same, as there are other differences between these two models besides altimetry assimilation. Interestingly, the improved lead-3 month forecast skill of OHC-300 appears more uniform in the SEAS5 model using altimetry assimilation, compared to in ACCESS-S1 (e.g., see the slightly lower ACC in some equatorial regions and parts of the North Atlantic, as well as the overall more heterogeneous ACC differences; Figure 4l), which is unlike the lead-0 month differences that were less robust in the SEAS5 experiment. The only place where the SEAS5-Control and ACCESS-S1 both have a lower ACC during the lead-3 month (Figures 4k and 4l) is in a narrow region of the equatorial Atlantic that extends into the subtropics along the southwestern African Coast (see discussion in Section 4). The RMSE result at the 3-month lead (Figures 5g–5l) again mostly mirrors the ACC assessment, except that ACCESS-S1 appears to perform worse than ACCESS-S2 in each of the equatorial ocean regions, as well as the Southern Ocean, which contributes to the overall higher RMSE in the former model.

3.2.2. SST

Whether or not altimetry assimilation is used appears to have almost no effect on the SST forecast skill at the lead-0 month (Figures 6a–6f and 7a–7f). This is not a surprising result because the SST in the initial analysis is strongly nudged to independent SST observations. In the ACCESS-S1 and ACCESS-S2 comparison, there are some minor differences in the ACC values for SST (Figure 6f), although the differences are typically much smaller than the result for OHC-300 (Figure 4f). The RMSE global-average difference for ACCESS-S is also small at this lead (Figure 7f). We note again though that there are other differences between the ACCESS-S models besides whether or not they use altimetry assimilation, which manifests in the ACC and RMSE differences that are evident at particular locations such as in the Southern Ocean. Most importantly, the SEAS5-Experiment shows no sensitivity of the lead-0 month SST to inclusion of altimetry assimilation (Figures 6e and 7e).

The SST lead-3 month forecasts also do not show an overall improvement by using altimetry assimilation (Figures 6g–6l and 7g–7l). In the comparisons of both the SEAS5-Experiment and ACCESS-S models, the ACC differences for SST at the lead-3 month are as likely to be either positive or negative (i.e., the global averages are near zero), and there is limited regional coherence to the differences. The RMSE global-average difference is also close to zero for the SEAS5-Experiment (Figure 7k), although this error metric does suggest that the ACCESS-S1 forecast skill for SST is worse than ACCESS-S2 overall (Figure 7l). The positive RMSE difference in the ACCESS-S comparison consists of an error in the Southern Ocean persisting since the lead-0 month as well as an error appearing in the equatorial central Pacific by the lead-3 month. A limitation of our assessment is

10.1029/2022JC019342

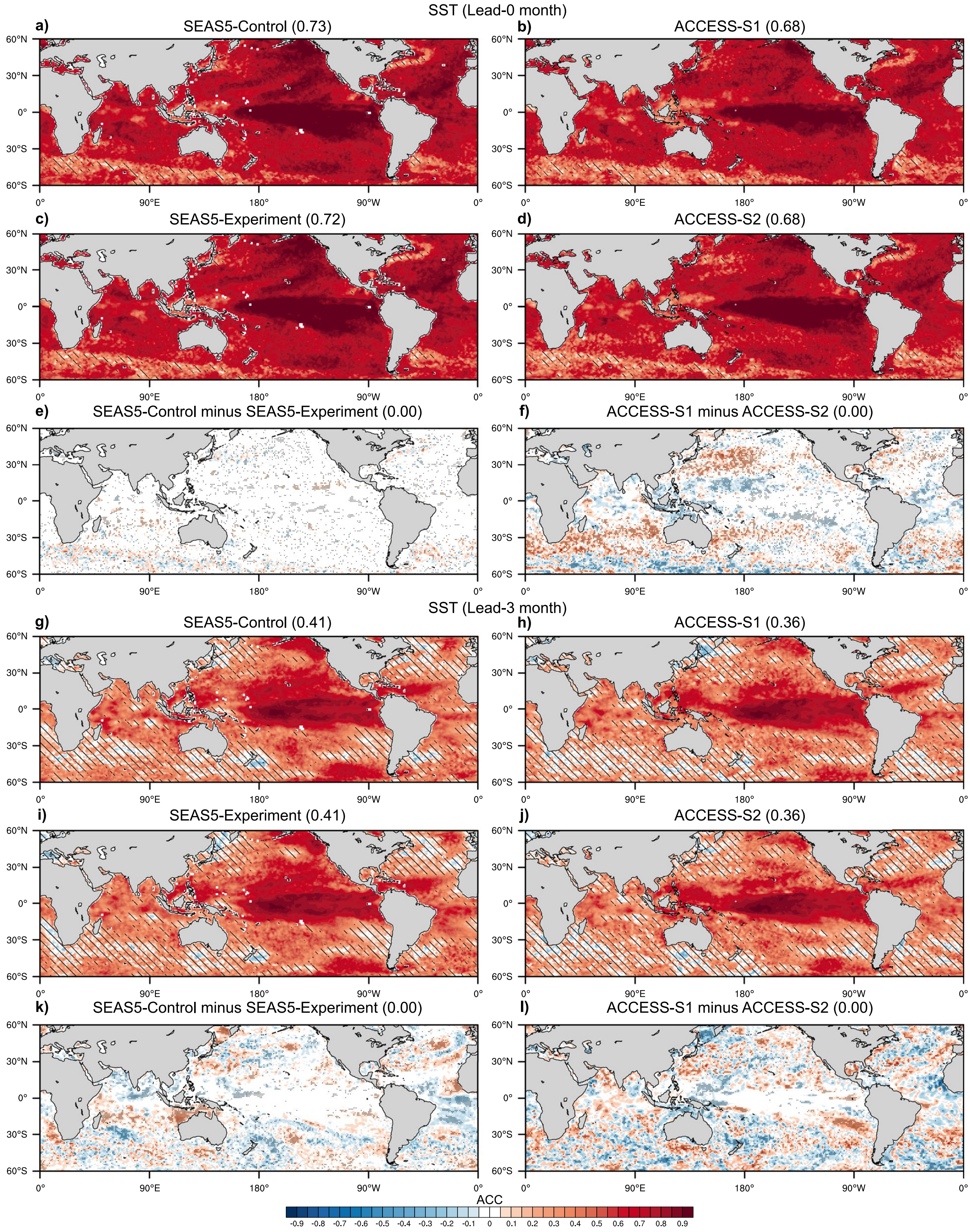

Figure 6. Similar to Figure 4 but for the SST forecast skill measured by the ACC between satellite-based observations of SST and the forecasts at lead-0 and 3 months (a–d and g–j, respectively), as well as their respective differences (e–f and k–l).

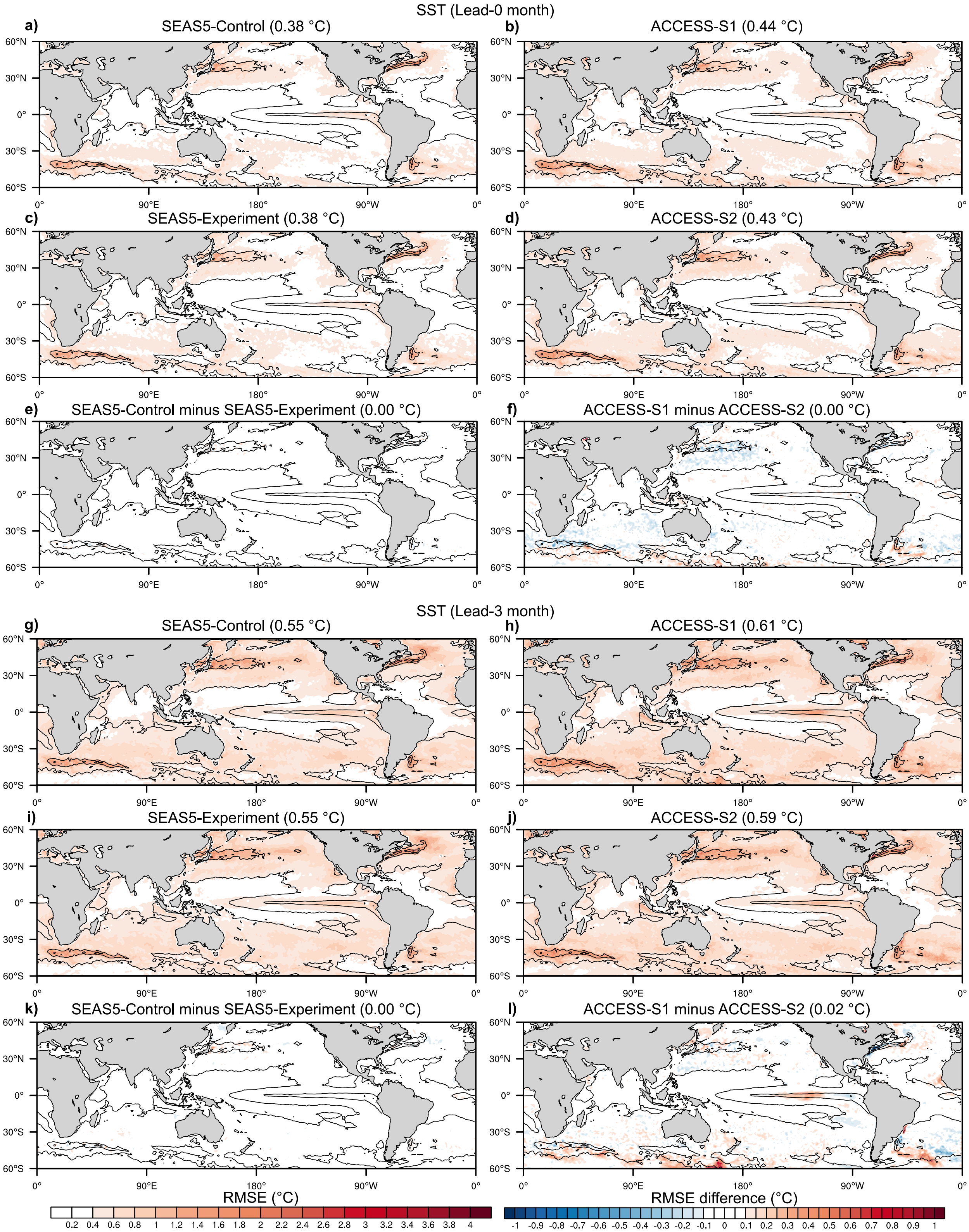

Figure 7. Similar to Figure 5 but for the SST forecast skill measured by the RMSE between satellite-based observations of SST and the forecasts at lead-0 and 3 months (a–d and g–j, respectively), as well as their respective differences (e–f and k–l). Black contours indicate the observed standard deviation of SST (0.5°C interval).

stopping at the lead-3 month, as this prevents us from assessing if the models with improved ocean physics early on (i.e., as inferred from the higher ACC and lower RMSE for OHC-300; Figures 4 and 5, respectively) have higher SST forecast skill at some much longer lead time (e.g., seasonal-to-decadal forecasts; see discussion in Section 4).

3.2.3. SSH

Results of the SSH forecast skill assessment (Figures 8 and 9) are mostly similar to what we showed for OHC-300, which is not surprising considering the strong physical relationship between the two variables (Figure 1a). Our hypothesis is that using altimetry assimilation will improve the SSH forecast if, and only if, the ocean density structure is improved; since we expect SSH monthly anomalies to respond to buoyancy variability via primarily a thermosteric sea level response (e.g., Widlansky et al., 2020). We already showed that the OHC-300 forecast skill is generally improved (i.e., higher ACC and lower RMSE) in the models including such initial information about the subsurface ocean from the assimilation of altimetry (Figures 4 and 5), and so the expectation is that these same models will better predict SSH. Despite this seemingly obvious result, an advantage of specifically assessing the SSH forecast skill is that the observations for SSH verification are completely independent of the models (unlike our choice of OHC-300 data from the multi-model reanalysis). Independent SSH observations for forecast verification allows fairer cross-model comparisons to be made (e.g., considering differences between the SEAS5-Control and ACCESS-S1), which we will also assess here in addition to the same-model differences.

For SSH at lead-0 month, ACCESS-S1 has a much higher global-average ACC value (Figure 8b; 0.66) compared to the SEAS5-Control (Figure 8a; 0.56), despite both models assimilating altimetry. A similar comparison between these two models was previously noted by Long et al. (2021). Interestingly, the RMSE for SSH of SEAS5 is somewhat lower than ACCESS-S1 (Figures 9a and 9b), with much of the difference being in the Southern Ocean. The SEAS5 and ACCESS-S1 forecast systems share many similarities in addition to both assimilating altimetry (e.g., both ocean model components have similar 0.25° nominal resolutions), although the way of assimilating altimetry is different between them: in SEAS5, the altimeter observations are not assimilated poleward of 50°N/S and are given lesser weight overall (Zuo et al., 2019), compared to ACCESS-S1. This difference in usage of altimetry data during the assimilation process may explain the difference in ACC at the lead-0 month (see discussion in Section 4). Interestingly, in the comparison of the two models without altimetry assimilation, the SEAS5-Experiment has slightly higher ACC overall compared to ACCESS-S2 (Figures 8c and 8d; 0.51 vs. 0.45, respectively) and the global-average RMSE is lower in the former model (Figures 9c and 9d; 5.12 vs. 6.11 cm). Although altimetry assimilation has a much smaller effect in SEAS5 than in the ACCESS-S systems (Figures 8e and 8f and 9e,f), in both cases the forecast skill at lead-0 month of SSH in the subtropics is clearly higher if altimetry assimilation is used in the initialization. We continue to note that ACCESS-S2 employs a different data assimilation system than what was used in ACCESS-S1; hence, this comparison on its own is inconclusive as far as hypothesis testing.

At the lead-3 month forecast for SSH (Figures 8g–8l), the models including altimetry assimilation continue to have somewhat higher ACC values overall (i.e., global-average differences of 0.03 and 0.04 for SEAS5 and ACCESS-S, respectively). These ACC differences (Figures 8k and 8l) are smaller compared to at the lead-0 month. Also, there is less of a difference between SEAS5 and ACCESS-S forecast skill at the lead-3 month, regardless of whether or not altimetry assimilation was used. At this lead, we see that ACCESS-S1 shows no advantage for higher forecast skill compared to the SEAS5-Control (Figures 8g and 8h), which was noted by Long et al. (2021). Here, we also compare the SEAS5-Experiment with ACCESS-S2 and see that the former has a minor forecast skill advantage (Figures 8i and 8j; ACC global averages are 0.37 vs. 0.31, respectively). Regardless of these cross-model differences, we see that the models including altimetry assimilation tend to have the highest ACC values in most places, which is especially the case outside of the tropical regions where the forecast skill in ACCESS-S1 appears to be less than in ACCESS-S2 (Figures 8h, 8j, and 8l). The RMSE analysis of the lead-3 month forecast for SSH (Figures 9g–9l) provides a similar result, although skill in the Southern Ocean again appears to be substantially degraded in ACCESS-S1.

3.3. Global and Regional Summary of Forecast Skill Changes

We now consider the ACC values for each ocean grid point within the near-global domain (i.e., everywhere within 60°N/S). Figure 10 presents a sample of these values as a function of whether or not altimetry assimilation is included in the models, and allows for direct comparison of the OHC-300, SST, and SSH results for the two

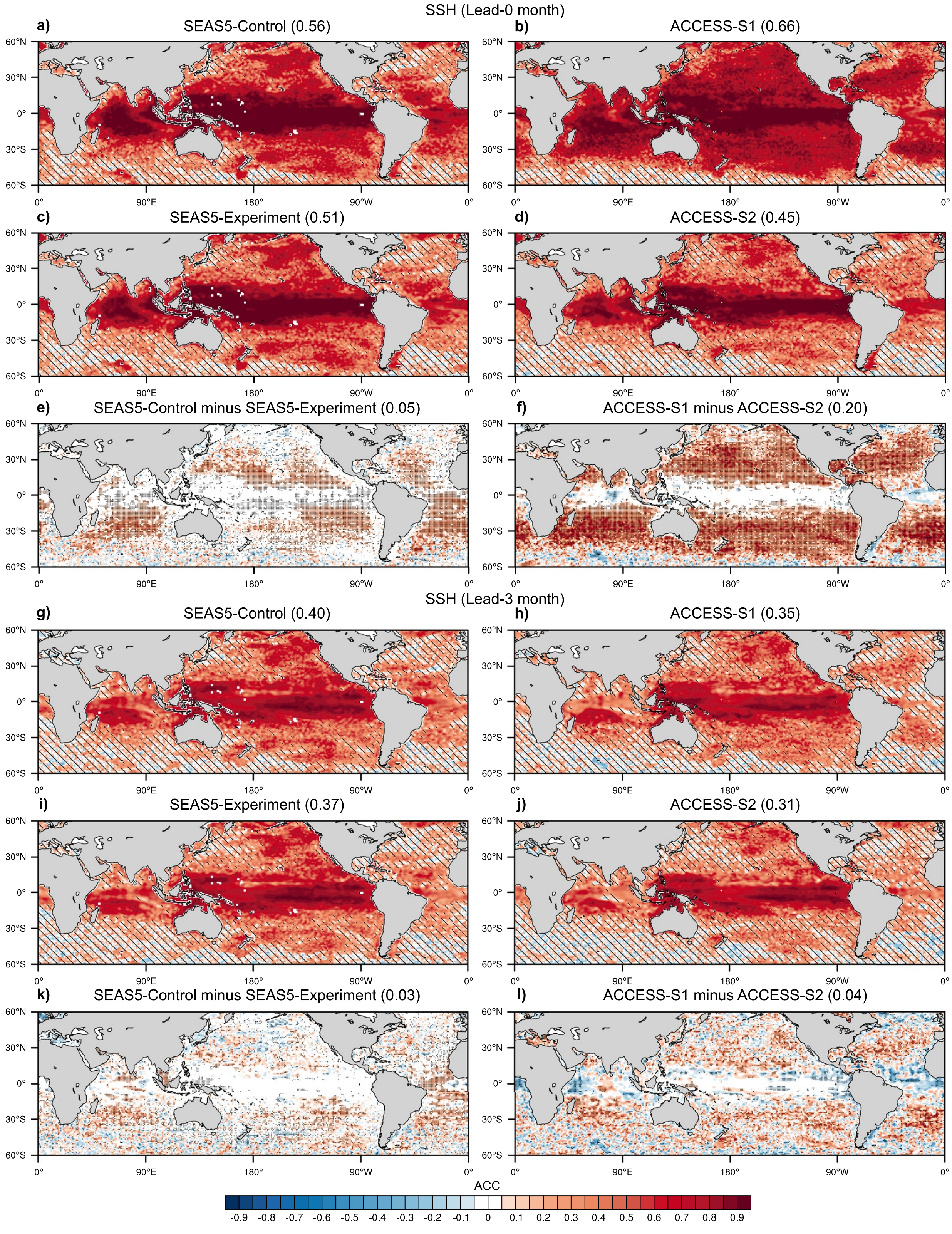

Figure 8. Similar to Figures 4 and 6 but for the SSH forecast skill measured by the ACC between altimetry observations of SSH with the forecasts at lead-0 and 3 months (a–d and g–j, respectively), as well as their respective differences (e–f and k–l).

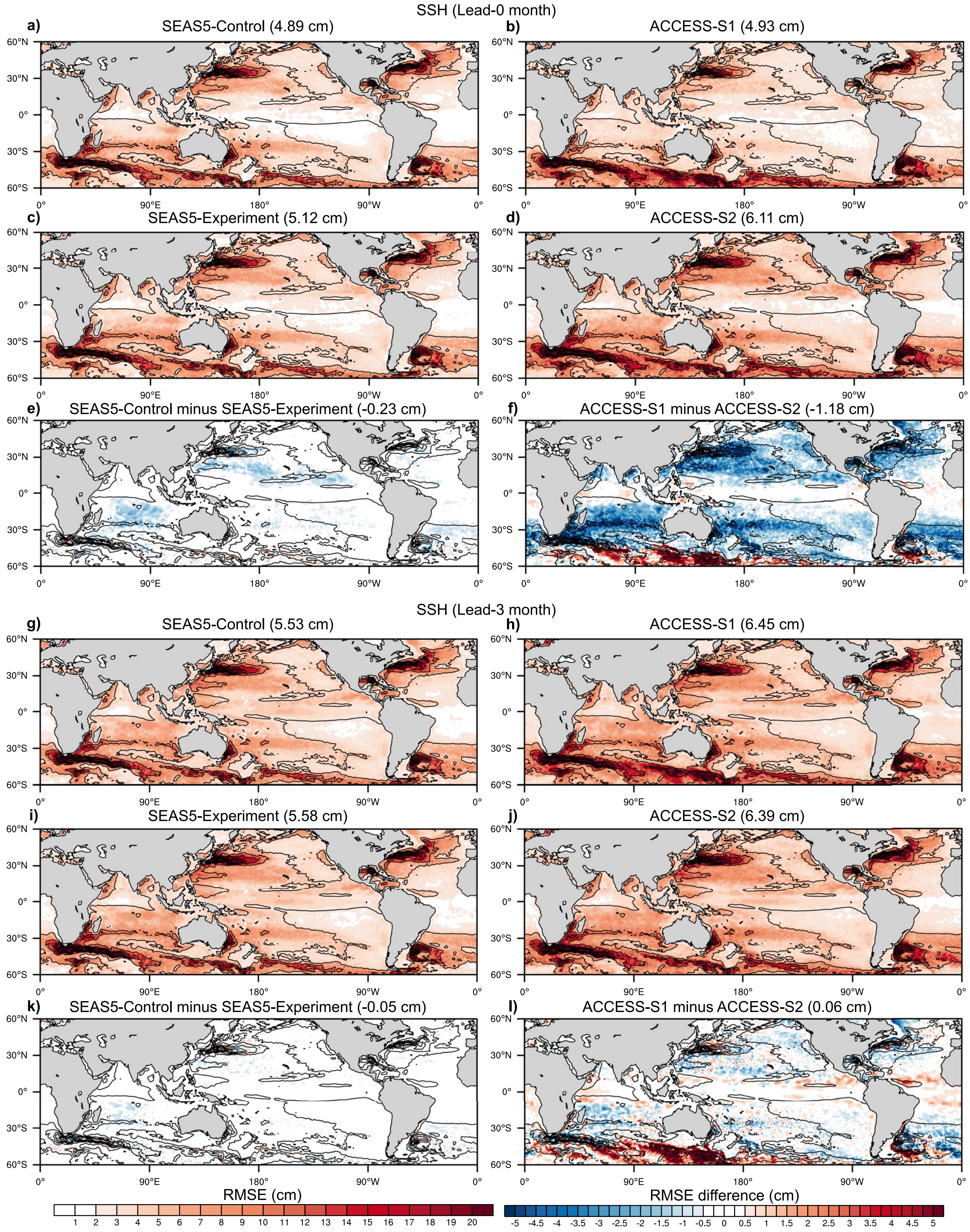

Figure 9. Similar to Figures 5 and 7 but for the SSH forecast skill measured by the RMSE between altimetry observations of SSH and the forecasts at lead-0 and 3 months (a–d and g–j, respectively), as well as their respective differences (e–f and k–l). Black contours indicate the observed standard deviation of SSH (5 cm interval).

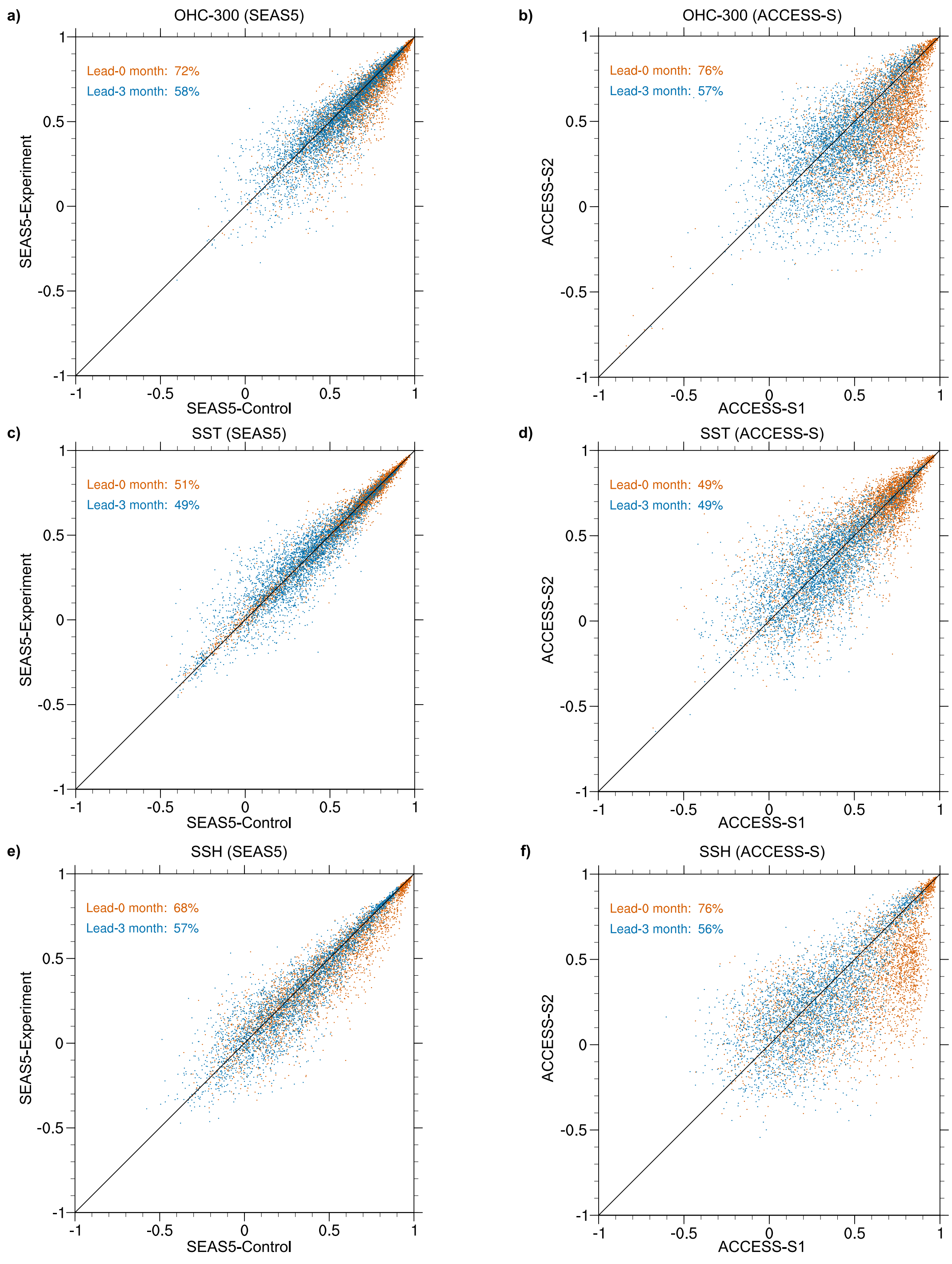

Figure 10. Global interpretation of forecast skill. ACC values calculated at a random 10% of grid cells (60°S–60°N) are shown for OHC-300 (a and b), SST (c and d), and SSH (e and f). Percentages of total grid points ($n = 43{,}200$) are shown for forecasts at lead-0 (orange) and 3 (blue) months where the ACC for the SEAS5-Control or ACCESS-S1 (altimetry assimilated, x-axes) are, respectively, higher than the SEAS5-Experiment or ACCESS-S2 (no altimetry assimilated, y-axes).

lead times that we assessed. By presenting the ACC values for the experiment models with no altimetry assimilation (y-axes in Figure 10) as a function of the control models (i.e., with altimetry assimilation; x-axes), we see a clear tendency for the OHC-300 and SSH forecasts initialized using altimetry assimilation to have higher skill since there are more points below each of the diagonal lines (see percentage values in Figures 10a, 10b, 10e, and 10f legends). In contrast, the SST forecasts show no tendency for differing skill depending on inclusion of altimetry assimilation in the models (Figures 10c and 10d). Quantifying the effect of altimetry assimilation in this manner shows how strikingly similar the results are for OHC-300 and SSH. In fact, the only situation where the percentages are different between the OHC-300 and SSH variables by more than 1% is for the SEAS5 lead-0 month (Figures 10a and 10e).

Figure 10 also shows that there is consistency between the SEAS5 and ACCESS-S models as far as the number of grid cells where altimetry assimilation is associated with higher ACC values for the OHC-300 and SSH forecasts. The cross-model consistency for the OHC-300 and SSH variables is greater for the lead-3 month forecast (percentages are within 1%; Figures 10a, 10b, 10e, and 10f) compared to at the lead-0 month (percentages differ between the SEAS5 and ACCESS-S models by 4% and 8% for OHC-300 and SSH, respectively). At least initially in the models, including altimetry assimilation seems to have a larger effect on ACCESS-S1 compared to SEAS5, which is also evident in the spatial ACC maps (Figures 4 and 8). Whereas, by lead-3 month the results are much more consistent across models. At the longer lead time, approximately 6%–8% of the globe has higher forecasting skill for OHC-300 and SSH in the two models that use altimetry assimilation. For the SST forecasts, the SEAS5 and ACCESS-S models are likewise consistent, although the result for this variable is that there is minimal sensitivity to altimetry assimilation (Figures 10c and 10d).

Figure 11 shows the ACC differences (i.e., SEAS5-Control minus SEAS5-Experiment and ACCESS-S1 minus ACCESS-S2) for the lead-0 and 3 month forecasts (*x*- and *y*-axes, respectively). The results are composited globally and also shown for five regions (i.e., as specified by the latitudinal bands in the legend). Compared to some of the regional composites, the global density distribution of ACC differences is centered fairly close to zero (see the shaded contours in Figure 11). However, the tails of the density distributions clearly extend toward positive ACC differences, at least for OHC-300 and SSH. Also, the lead-0 month differences are usually larger than for the lead-3 month (i.e., the maximum densities of the ACC differences are below the diagonal lines in Figures 11a, 11b, 11e, and 11f). Both of these characteristics about the OHC-300 and SSH density distributions are to be expected given the spatial maps in Figures 4 and 8. For the SST forecasts, neither the ACC maps (Figure 6) nor the presentation in Figures 11c and 11d suggest that altimetry assimilation has much of an effect on skill, except that the ACCESS-S1 model at lead-0 month does have somewhat higher skill in the subtropics compared to ACCESS-S2. Lastly, and likewise to be expected based on the other results, the global density distributions of the differences between ACCESS-S1 and ACCESS-S2 are shifted further toward positive values compared to those in the SEAS5 experiment.

As we saw in the ACC maps for OHC-300 and SSH (Figures 4 and 8), nearly all of the improved forecast skill associated with altimetry assimilation occurs in the subtropics. The regionally composited results represented by colored dots in Figures 11a, 11b, 11e, and 11f are similar (i.e., the largest changes are in the subtropics). It is in the subtropics that we previously noticed the models using altimetry assimilation usually had the greater ACC between SSH and OHC-300 at the lead-0 month (Figure 3) as well as the greater lead-3 month forecasting skill for both variables (Figures 4 and 8). Figure 11 also shows that the mid-latitude average ACC difference at lead-0 month is clearly positive for ACCESS-S1 minus ACCESS-S2. For the lead-3 month comparison of ACCESS-S, and the SEAS5-Experiment at both leads, the mid-latitude differences are typically smaller. In the tropics, the ACC differences are starkly different from the other regions in that the values are either near zero (SEAS5 experiment for both leads) or negative (ACCESS-S1 minus ACCESS-S2 for the lead-3 month forecast). To summarize briefly, whereas globally the overwhelming majority of ACC differences for OHC-300 and SSH are positive for the models that include altimetry assimilation, and the differences are clearly largest in the subtropical regions, these same models show no improvement of ocean seasonal forecast skill in the tropics.

4. Summary and Discussion

This study quantified the effect on ocean prediction skill of assimilating satellite-altimetry measurements of SSH in two operational climate forecasting systems (SEAS5 and ACCESS-S). We assessed a pair of retrospective forecasts from both SEAS5 and ACCESS-S, which either used altimetry assimilation or did not. Combined, the different assimilation treatments in these four models provide a set of forecasts that are well targeted for determining the importance of including altimetry assimilation in climate forecasting systems.

The forecast skill assessment focused on the variability of upper-ocean physical properties, which we described using monthly anomalies of OHC-300, SST, and SSH. Of these variables, OHC-300 is the most uncertain because

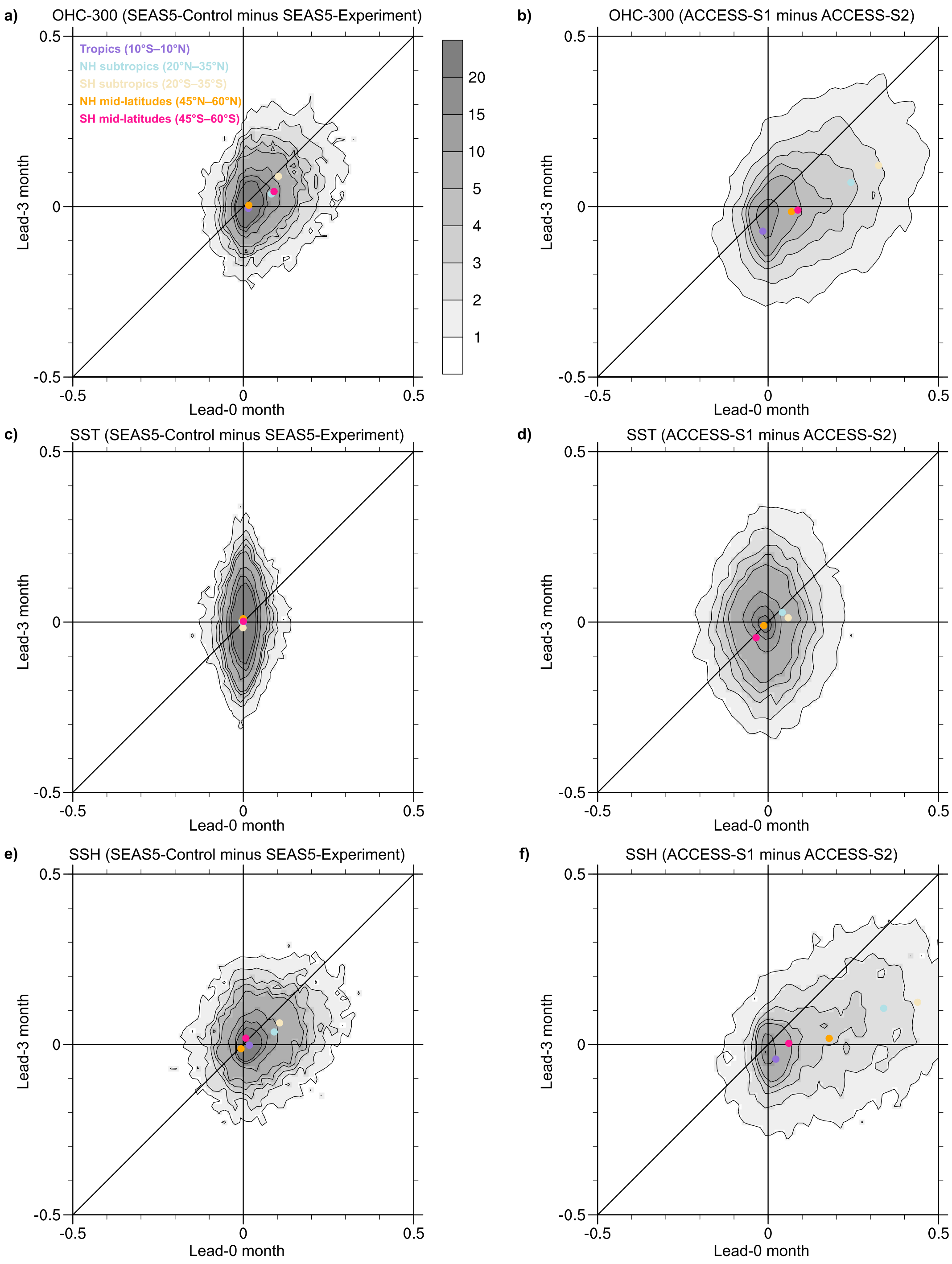

Figure 11. Regional interpretation of forecast skill. ACC differences for SEAS5-Control minus SEAS5-Experiment and ACCESS-S1 minus ACCESS-S2 are shown for OHC-300 (a and b), SST (c and d), and SSH (e and f) at lead-0 and 3 months (x and y axes, respectively). Regional ACC averages (circles; see latitude domains in legend of panel (a) are compared to the global ACC differences using kernel density estimates of the results shown in Figures 4, 6 and 8 (non-uniform gray scale; see color bar in panel (a).

direct observations of the subsurface temperature do not exist uniformly over the near-global domain. Such observational gaps are typically filled using either objective interpolation of in-situ measurements, or dynamical analyses that are initialized after assimilation of the best available data to describe the ocean state. Overall, we found that the latter method produces OHC-300 monthly anomalies that are more consistent with the expected physics (i.e., having the highest correlation with the closely-related and satellite-observed SSH variable; Figure 1). The strong correlation between OHC-300 and SSH is especially true for the multi-model reanalysis (GREP), which we chose for verifying the OHC-300 forecasts. We showed that OHC-300 and SST are also positively correlated in most places (Figure 2), which supports the expectation that all three of these variables are physically related to some extent.

The main result of this study is that including altimetry assimilation in the climate forecasting systems is associated with an improvement of the subsurface ocean initial conditions and also higher prediction skill at the lead-3 month. We found this result to be generally the case for OHC-300 and SSH, but not SST. Our hypothesis that including altimetry assimilation is conducive to having a more realistic ocean initial state is supported by Figure 3, which shows that during the lead-0 month the ACC between observed SSH and the simulated OHC-300 is higher in the models with altimetry assimilation (i.e., closer to what is expected to be the case based on the NEMO model realization shown in Figure 1a). The improvement (i.e., higher ACC values associated with altimetry assimilation) is clear in the subtropics and some mid-latitude locations, whereas there appears to be little or no effect near the equator. For the equatorial Pacific, in particular, Zuo et al. (2017) showed that SSH and OHC-300 variability are well simulated by the NEMO model because of its initialization with data from a robust network of in-situ observations along with forcing by the ERA-Interim atmospheric fluxes.

Likewise, our quantification of the effect of altimetry assimilation on forecasting skill (Figures 4–11) showed that the largest improvements (i.e., higher ACC and lower RMSE values for the verifications of OHC-300 and SSH with the reanalysis or satellite observations, respectively) are generally in these same regions (i.e., larger changes in the subtropics and mid-latitudes compared to the tropics). These results mostly support the second part of our hypothesis that improvements to the ocean initial conditions are associated with better forecast skill (i.e., higher ACC and lower RMSE values). In the tropics, we found either no difference in retrospective forecast skill between assimilation systems, or perhaps a slight worsening of the predictions when altimetry assimilation was included (e.g., ACCESS-S1 in the equatorial regions for all three variables at all leads as well as SEAS5-Control but only in the equatorial Atlantic for OHC-300 and SST at the lead-3 month). Reasons for apparent worsening forecast skill in the equatorial Atlantic are unknown, however a possibility is that the local balance between salinity and temperature nudging during the altimetry assimilation may need to be adjusted. In contrast to the forecasting improvements noted for OHC-300 and SSH outside of the tropics, assessment of the SST forecasts showed no widespread sensitivity to whether or not altimetry assimilation was used. The limited effect of altimetry assimilation on the SST, especially for the lead-0 month of SEAS5 (Figures 6e and 7e), perhaps somewhat contradicts the previous finding of M. A. Balmaseda and Anderson (2009), but our result is not unexpected considering how well SST is observed and assimilated into current-generation climate forecasting models (e.g., McAdam et al., 2022). Collectively, the results of this study support our expectation that altimetry assimilation mostly improves the temperature and density structure of the subsurface ocean, which would more strongly affect OHC-300 and SSH compared to SST.

We address a specific limitation about our study, which is the focus on only the lead-0 and 3 months, prior to suggesting opportunities for further research. The focus on these early-month lead times is primarily because of the limited lead time for the retrospective forecasts from some of the models (i.e., out to lead-5 months for ACCESS-S and the SEAS5-Experiment). We assessed the ACC differences at every available lead time and generally found that the effect of including altimetry assimilation decreased at the longer leads compared to at the lead-0 month (Figure 12; RMSE differences decrease similarly to the ACC result and are not shown), which was expected because seasonal forecast skill typically decreases toward zero with increasing lead time (e.g., Jacox et al., 2022). For where there are the largest responses to including altimetry assimilation (i.e., OHC-300 and SSH in the subtropics), we note that most of the decay with increasing lead time occurs between 0 and 3 months. However, it is perceivable that longer retrospective forecasts (i.e., beyond lead-6 months) reveal a stronger effect of including altimetry assimilation on the SST forecast skill in particular, since subsurface anomalies sometimes take many months before influencing the surface (e.g., Deser et al., 2003). Hence, it would be interesting to assess the effect of altimetry assimilation on climate forecasts at much longer timescales (e.g., interannual to multi-decadal), but this would require substantially longer retrospective forecasts using each of the models. Considering any seasonal dependency of the response to altimetry assimilation would also be interesting, however this too would require more retrospective forecasts to have a sufficient degree of freedom necessary to minimize random noise (i.e., a longer forecasting epoch would be needed, which is feasible now that about 30 years of altimetry observations exist).

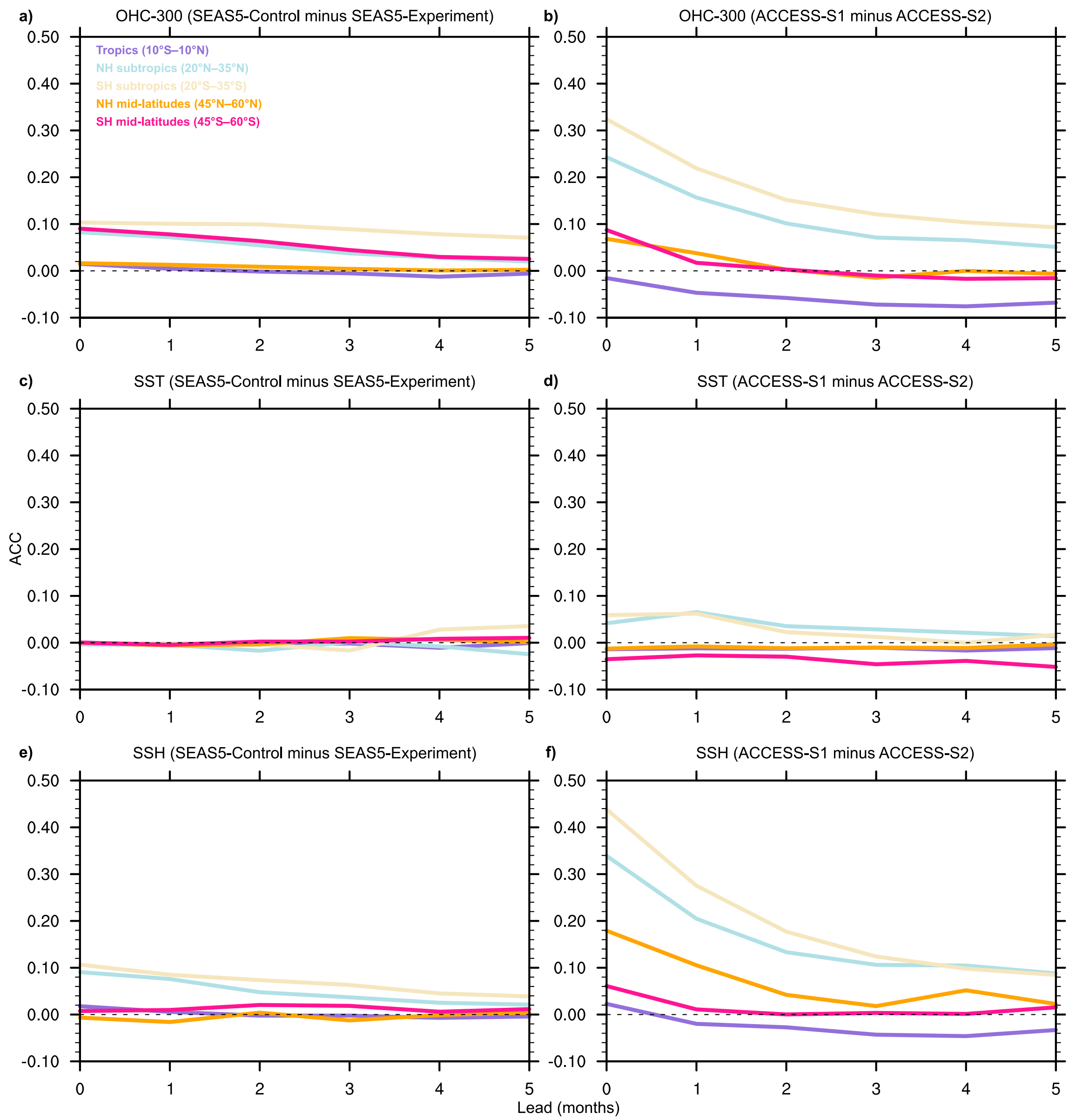

Figure 12. Similar to Figure 11 but for the regional ACC differences (*y* axes) as a function of lead month (*x* axes). Results for SEAS5-Control minus SEAS5-Experiment and ACCESS-S1 minus ACCESS-S2 are shown for OHC-300 (a and b, SST (c and d), and SSH (e and f).

In the remainder of this section, we discuss three specific research questions motivated by assessment of the effect of altimetry assimilation on initializing and forecasting the upper-ocean monthly variability in climate models. The first question concerns why the climate models assimilating altimetry seem to produce a more realistic analysis of OHC-300 in most but not all places, compared to the objective analyses. The second question specifically relates to the forecast skill differences in the tropics between ACCESS-S1 and ACCESS-S2 (i.e., are there reasons besides not including altimetry assimilation in the latter model that explains the improvement in some equatorial regions?). The third question, which is motivated by all of these regional as well as inter-model differences in the results, asks how could the altimetry assimilation procedure be improved. Collectively, we think that answers to these questions may lead toward improved ocean seasonal forecasting capabilities and associated applications. We will conclude the discussion by mentioning immediate implications of this study.

There are presumably two reasons why the climate model analyses (GREP and ORAS5) have a more realistic OHC-300 compared to the objective analyses (IAP and EN4), which we saw is the case outside of the tropics at least for the assumption of the NEMO model having the most realistic association between SSH and

OHC-300 (Figure 1). First, GREP and ORAS5 both assimilate altimetry, which is an additional source of observations not used by IAP or EN4, and Figure 3 shows that the forecast models without altimetry assimilation do have a less realistic OHC-300 as represented by the reduction of ACC of that variable with the observed SSH. Second, the climate models, by design, dynamically merge all available observations (i.e., of temperature, circulation, and salinity) into a physically-consistent simulation of the ocean, which is an advantage that the observation-only analyses of OHC-300 lack. Regardless of including altimetry assimilation, the climate models are likely to produce improvements in the OHC-300 patterns associated with variations in the large-scale ocean circulation (e.g., associated with ENSO or the Atlantic Meridional Overturning Circulation, AMOC). However, climate models can also introduce errors, as we saw in parts of the tropics in ACCESS-S1 (Figures 4–8) as well as ORAS5 in the North Atlantic Subpolar Gyre (Figure 1c). In SEAS5, the relatively weak correlation in the North Atlantic between simulated OHC-300 and observed SSH is evident whether or not altimetry assimilation is used (Figures 3a and 3c). Interestingly, OHC-300 from the GREP multi-model reanalysis does not appear to have such a diminished ACC with observed SSH in the North Atlantic, or anywhere else equatorward of 60°N/S (Figure 1b). In fact, the degradation in ORAS5 (and SEAS5) in the North Atlantic has been reported (Tietsche et al., 2020), and this issue has been corrected in the new ECMWF ocean reanalysis system (ORAP6; see Zuo et al., 2021).

In spite of not including altimetry assimilation, ACCESS-S2 has somewhat higher forecast skill in the tropics compared to ACCESS-S1, whereas we showed that the opposite is true almost everywhere else. Recently, Wedd et al. (2022) assessed in detail the data assimilation procedures of the two ACCESS-S forecasting systems, and a brief review of that study is helpful to answer why we see these regional differences. They similarly found that ACCESS-S2 outperforms ACCESS-S1 near the equator, whereas ACCESS-S1 does better in the subtropics and mid-latitudes. Wedd et al. (2022) identified spurious upward velocities in the deep ocean of the ACCESS-S1 reanalysis, which they suggest results from a dynamical imbalance during the data assimilation. It remains unknown how directly the altimetry assimilation is involved with this error. ACCESS-S2 seems to perform better in the equatorial oceans due to more dynamically consistent increments of the data assimilation procedure. Wedd et al. (2022) also found that a better ENSO prediction skill was achieved with ACCESS-S2 compared to ACCESS-S1. Although we did not find significant differences in the SST forecast skill for the equatorial Pacific according to the ACC metric (Figure 6), the RMSE for ACCESS-S2 is lower there at the lead-3 month (Figure 7l). Also, the OHC-300 and SSH results for the lead-3 month (Figures 4 and 8, respectively) do show ACCESS-S2 having slightly higher ACC values in the off-equatorial eastern Pacific, which is consistent with that model perhaps having improved ENSO prediction capability.

It has been determined that the ACCESS-S1 data assimilation system does in fact have a deficiency, mainly due to spurious vertical velocity fields near the equator caused by dynamical imbalance after data assimilation. According to Wedd et al. (2022), such imbalance can degrade the performance of the ocean model and lead to unrealistic subsurface temperature and current fields, as seen in previous studies by Gasparin et al. (2021), Park et al. (2018), and Waters et al. (2016). The problem of dynamical imbalance is one of the unsolved challenges in ocean reanalyses, as noted by Storto, Alvera-Azcárate, et al. (2019). However, the more dynamically consistent initial conditions of ACCESS-S2 have been found to perform better than ACCESS-S1 in the equatorial oceans, as reported by Wedd et al. (2022).

Based on the regional and inter-model differences that we identified in how altimetry assimilation seems to affect the subsurface ocean, we lastly ask if the altimetry assimilation procedure can be improved. At the lead-0 month, there is clearly an overall higher ACC of the SSH from ACCESS-S1 with observations compared to SEAS5 (Figures 8a and 8b) despite both models assimilating altimetry. Likewise, the difference at this lead between ACCESS-S1 and ACCESS-S2 is much larger than for the SEAS5-Experiment, both for SSH (Figures 8e and 8f) as well as OHC-300 (Figures 4e and 4f). These inter-model differences are largest at lead-0 month in the mid-latitudes and especially near the continental coasts (e.g., around North America) where altimetry observations are not weighted as heavily into the SEAS5 assimilation (Zuo et al., 2019). By the lead-3 month, we see no widespread differences between SEAS5 and ACCESS-S as far as how much altimetry assimilation is affecting the forecast skill, at least according to the ACC differences (Figures 4k and 4l and 8k and 8l). Interestingly, the RMSE assessment (Figures 5 and 9) mostly mirrors the ACC result, except that both ACCESS-S models have overall larger error amplitudes at all leads compared to SEAS5.

It remains unknown what is the ideal weighting to give altimetry observations in initializing climate forecasts, especially since suboptimal results are perceivable if the data assimilation is poorly performed. Besides potential errors near the equator in the ACCESS-S1 data assimilation noted above, there are many other challenging regions in all current-generation climate models (e.g., around the Gulf Stream where ocean-atmosphere interactions and horizontal mixing are important; respectively, Frankignoul et al., 2001; Wenegrat et al., 2020). In such places, weighting altimetry observations too heavily could perhaps worsen errors in the ocean or atmosphere, and actually degrade the seasonal forecasting skill. In fact, Long et al. (2021) found that the seasonal forecasting skill for SSH of ACCESS-S1 was no higher than SEAS5, despite the former model more closely matching the phase of observations initially (see also Figure 8). Targeted testing of different weighting procedures for altimetry assimilation are needed as a way forward to developing optimal forecasting systems.

To conclude, we showed that including altimetry assimilation in operational climate forecasting systems does appear to improve the overall ocean prediction skill at least for OHC-300 and SSH, and especially in the subtropics as well as parts of the mid-latitudes. In the tropics, for the models without altimetry assimilation, we did not find diminished forecast skill of any variable. This result may be important to consider in the context of choosing which models are used for ENSO forecasting and making other predictions about the tropical ocean, such as producing sea level outlooks for the Pacific Islands (Widlansky et al., 2017) or early warnings about marine heatwaves (Jacox et al., 2022; C. M. Spillman & Smith, 2021; C.M. Spillman et al., 2021), in which it may be advantageous to include models regardless of whether altimetry assimilation is included. For the regions where we did find evidence of higher seasonal forecasting skill in the models that included altimetry assimilation, it remains to be determined whether such improvements are meaningful for seasonal forecasting applications. Such information is critical to support the development of next-generation forecasting systems (Becker et al., 2022) such as the forthcoming ACCESS-S3 from the Australian Bureau of Meteorology or the Unified Forecast System from NOAA. Sensitivity experiments using these new systems regarding altimetry assimilation are warranted.

Data Availability Statement

The ECMWF modeling data used in this study are available from https://uhslc.soest.hawaii.edu/opendap/AltimetryAssimilationExperiment/ (M. A. Balmaseda & Zuo, 2023). Retrospective forecasts from ACCESS-S are available from the Australian National Computing Infrastructure (NCI) by following the instructions for acquiring research data at http://poama.bom.gov.au/general/hindcast_data.html (C. M. Spillman et al., 2023). The analysis and observation data used in this study are available from the following sources: GREP, https://resources.marine.copernicus.eu/product-detail/GLOBAL_REANALYSIS_PHY_001_031 (Global Ocean Ensemble Physics Reanalysis, 2023); ORAS5, https://www.cen.uni-hamburg.de/icdc/data/ocean/easy-init-ocean/ecmwf-oras5.html (Zuo & Balmaseda, 2023); EN4, https://www.metoffice.gov.uk/hadobs/en4/ (EN.4.2.2, 2023); IAP, http://www.ocean.iap.ac.cn/ftp/cheng/IAP_Ocean_heat_content_0_2000m/ (IAP Ocean Heat Content, 2023); SST, https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html (NOAA Optimum Interpolation (OI) SST V2, 2023); and SSH, https://resources.marine.copernicus.eu/product-detail/SEALEVEL_GLO_PHY_L4_MY_008_047/INFORMATION (Global Ocean Gridded L 4 Sea Surface Heights And Derived Variables Reprocessed, 1993 Ongoing, 2023).

Acknowledgments
This study was primarily supported by the NOAA Climate Program Office's Modeling, Analysis, Predictions, and Projections (MAPP) program through Grant NA17OAR4310110. The study was partially funded by the European Union's Horizon 2020 research and innovation programme under grant agreement 862626, as part of the EuroSea project. The authors thank the ECMWF and Australian Bureau of Meteorology for their production and sharing of the model data assessed here. The authors also acknowledge the insightful comments from four anonymous reviewers, which helped to improve this study.

References

Alves, J. O. S., Haines, K., & Anderson, D. L. T. (2001). Sea level assimilation experiments in the tropical Pacific. *Journal of Physical Oceanography*, *31*(2), 305–323. https://doi.org/10.1175/1520-0485(2001)031<0305:SLAEIT>2.0.CO;2

Anderson, D. (2012). The development of seasonal forecasting. *Paper presented at the ECMWF Seminar on Seasonal Prediction*. Retrieved from https://www.ecmwf.int/sites/default/files/elibrary/2013/7718-development-seasonal-prediction.pdf

Balmaseda, M. A. (2017). Data assimilation for initialization of seasonal forecasts. *Journal of Marine Research*, *75*(3), 331–359. https://doi.org/10.1357/002224017821836806

Balmaseda, M. A., & Anderson, D. (2009). Impact of initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters*, *36*(1), L01701. https://doi.org/10.1029/2008GL035561

Balmaseda, M. A., Mogensen, K., & Weaver, A. T. (2013). Evaluation of the ECMWF Ocean Reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1132–1161. https://doi.org/10.1002/qj.2063

Balmaseda, M. A., Vidard, A., & Anderson, D. L. T. (2008). The ECMWF ocean analysis system: ORA-S3. *Monthly Weather Review*, *136*(8), 3018–3034. https://doi.org/10.1175/2008MWR2433.1

Balmaseda, M. A., & Zuo, H. (2023). The ECMWF modeling data used in this study. [Dataset]. Retrieved from https://uhslc.soest.hawaii.edu/opendap/AltimetryAssimilationExperiment/

Becker, E. J., Kirtman, B. P., L'Heureux, M., Muñoz, Á. G., & Pegion, K. (2022). A decade of the North American Multimodel Ensemble (NMME): Research, application, and future directions. *Bulletin America Meteorology Social*, *103*(3), E973–E995. https://doi.org/10.1175/BAMS-D-20-0327.1

Behrens, E., Fernandez, D., & Sutton, P. (2019). Meridional oceanic heat transport influences marine heatwaves in the Tasman Sea on interannual to decadal timescales. *Frontiers in Marine Science*, *6*(28). https://doi.org/10.3389/fmars.2019.00228

Bulgin, C. E., Merchant, C. J., & Ferreira, D. (2020). Tendencies, variability and persistence of sea surface temperature anomalies. *Scientific Reports*, *10*(1), 7986. https://doi.org/10.1038/s41598-020-64785-9

Chelton, D. B., Ries, J. C., Haines, B. J., Fu, L. L., & Callahan, P. S. (2001). Satellite altimetry. In L. L. Fu & A. Cazenave (Eds.), *Satellite altimetry and Earth sciences: A handbook of techniques and applications* (Vol. 69, pp. 1–131). International Geophysics. https://doi.org/10.1016/S0074-6142(01)80146-7

Chelton, D. B., & Schlax, M. G. (1996). Global observations of oceanic Rossby waves. *Science*, *272*(5259), 234–238. https://doi.org/10.1126/science.272.5259.234

Cheng, L., Trenberth, K. E., Fasullo, J., Boyer, T., Abraham, J., & Zhu, J. (2017). Improved estimates of ocean heat content from 1960 to 2015. *Science Advances*, *3*(3), e1601545. https://doi.org/10.1126/sciadv.1601545

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

Deser, C., Alexander, M. A., & Timlin, M. S. (2003). Understanding the persistence of sea surface temperature anomalies in midlatitudes. *Journal of Climate*, *16*(1), 57–72. https://doi.org/10.1175/1520-0442(2003)016<0057:UTPOSS>2.0.CO;2

Dusek, G., Sweet, W. V., Widlansky, M. J., Thompson, P. R., & Marra, J. J. (2022). A novel statistical approach to predict seasonal high tide flooding. *Frontiers in Marine Science*, *9*, 1073792. https://doi.org/10.3389/fmars.2022.1073792

EN.4.2.2 (2023). [Dataset]. Retrieved from https://www.metoffice.gov.uk/hadobs/en4/

Feder, T. (2000). Argo begins systematic global probing of the upper oceans. *Physics Today*, *53*(7), 50–51. https://doi.org/10.1063/1.1292477

Feng, W., & Zhong, M. (2015). Global sea level variations from altimetry, GRACE and Argo data over 2005–2014. *Geodesy and Geodynamics*, *6*(4), 274–279. https://doi.org/10.1016/j.geog.2015.07.001

Frankignoul, C., de Coëtlogon, G., Joyce, T. M., & Dong, S. (2001). Gulf Stream variability and ocean–atmosphere interactions. *Journal of Physical Oceanography*, *31*(12), 3516–3529. https://doi.org/10.1175/1520-0485(2002)031<3516:GSVAOA>2.0.CO;2

Fu, L., & Smith, R. D. (1996). Global ocean circulation from satellite altimetry and high-resolution computer simulation. *Bulletin America Meteorology Social*, *77*(11), 2625–2636. https://doi.org/10.1175/1520-0477(1996)077<2625:GOCFSA>2.0.CO;2

Fu, L. L., & Davidson, R. A. (1995). A note on the barotropic response of sea level to time-dependent wind forcing. *Journal of Geophysical Research*, *100*(C12), 24955–24963. https://doi.org/10.1029/95JC02259

Fukumori, I., Raghunath, R., & Fu, L. (1998). Nature of global large-scale sea level variability in relation to atmospheric forcing: A modeling study. *Journal of Geophysical Research*, *103*(C3), 5493–5512. https://doi.org/10.1029/97JC02907

García, D., Ramillien, G., Lombard, A., & Cazenave, A. (2007). Steric sea-level variations inferred from combined Topex/Poseidon altimetry and GRACE gravimetry. *Pure and Applied Geophysics*, *164*(4), 721–731. https://doi.org/10.1007/s00024-007-0182-y

Gasparin, F., Cravatte, S., Greiner, E., Perruche, C., Hamon, M., Gennip, S. V., & Lellouche, J.-M. (2021). Excessive productivity and heat content in tropical Pacific analyses: Disentangling the effects of in situ and altimetry assimilation. *Ocean Modelling*, *160*, 101768. https://doi.org/10.1016/j.ocemod.2021.101768

Global Ocean Ensemble Physics Reanalysis. (2023). [Dataset]. Retrieved from https://data.marine.copernicus.eu/product/GLOBAL_REANALYSIS_PHY_001_031/description

Global Ocean Gridded L 4 Sea Surface Heights And Derived Variables Reprocessed 1993 Ongoing. (2023). [Dataset]. Retrieved from https://data.marine.copernicus.eu/product/SEALEVEL_GLO_PHY_L4_MY_008_047/description

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, *118*(12), 6704–6716. https://doi.org/10.1002/2013JC009067

Gurvan, M., Bourdallé-Badie, R., Chanut, J., Clementi, E., Coward, A., Ethé, C., et al. (2022). *NEMO ocean engine (Version v4.2)*. Zenodo. https://doi.org/10.5281/zenodo.6334656

Hayes, S. P., Mangum, L. J., Picaut, J., Sumi, A., & Takeuchi, K. (1991). Toga TAO: A moored array for real-time measurements in the tropical Pacific ocean. *Bulletin America Meteorology Social*, *72*(3), 339–347. https://doi.org/10.1175/1520-0477(1991)072<0339:TTAMAF>2.0.CO;2

Hudson, D., Alves, O., Hendon, H. H., Lim, E.-P., Liu, G., Luo, J.-J., et al. (2017). ACCESS-S1: The new Bureau of Meteorology multi-week to seasonal prediction system. *J. South. Hemisph. Earth Syst. Sci.*, *67*(3), 132–159. https://doi.org/10.22499/3.6703.001

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

IAP Ocean Heat Content. (2023). [Dataset]. Retrieved from http://www.ocean.iap.ac.cn/ftp/cheng/IAP_Ocean_heat_content_0_2000m/

Jacox, M. G., Alexander, M. A., Amaya, D., Becker, E., Bograd, S. J., Brodie, S., et al. (2022). Global seasonal forecasts of marine heatwaves. *Nature*, *604*(7906), 486–490. https://doi.org/10.1038/s41586-022-04573-9

Johnson, G. C., Lyman, J. M., Boyer, T., Cheng, L., Gilson, J., Ishii, M., et al. (2021). Ocean heat content [in "State of the Climate in 2020," Chapter 3]. *Bulletin America Meteorology Social*, *102*(8), S14–S17. https://doi.org/10.1175/BAMS-D-21-0083.1

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., III., Paolino, D. A., Zhang, Q., et al. (2014). The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bulletin America Meteorology Social*, *95*(4), 586–601. https://doi.org/10.1175/bams-d-12-00050.1

L'Heureux, M. L., Tippett, M. K., & Wang, W. (2022). Prediction challenges from errors in tropical Pacific Sea surface temperature trends. *Frontiers in Climate*, *4*. https://doi.org/10.3389/fclim.2022.837483

Long, X., Widlansky, M. J., Schloesser, F., Thompson, P. R., Annamalai, H., Merrifield, M. A., & Yoon, H. (2020). Higher sea levels at Hawaii caused by strong El Niño and weak trade winds. *Journal of Climate*, *33*(8), 3037–3059. https://doi.org/10.1175/JCLI-D-19-0221.1

Long, X., Widlansky, M. J., Spillman, C., Kumar, A., Balmaseda, M. A., Thompson, P., et al. (2021). Seasonal forecasting skill of sea level anomalies in a multi-model prediction framework. *J. Geophys. Res. Ocn.*, *126*(6), e2020JC017060. https://doi.org/10.1029/2020JC017060

McAdam, R., Masina, S., Balmaseda, M., Gualdi, S., Senan, R., & Mayer, M. (2022). Seasonal forecast skill of upper-ocean heat content in coupled high-resolution systems. *Climate Dynamics*, *58*(11–12), 3335–3350. https://doi.org/10.1007/s00382-021-06101-3

Meng, X.-l., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, *111*(1), 172–175. https://doi.org/10.1037/0033-2909.111.1.172

Nerem, R. S., Beckley, B. D., Fasullo, J. T., Hamlington, B. D., Masters, D., & Mitchum, G. T. (2018). Climate-change–driven accelerated sea-level rise detected in the altimeter era. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(9), 2022–2025. https://doi.org/10.1073/pnas.171731211

NOAA Optimum Interpolation (OI) SST V2. (2023). [Dataset]. Retrieved from: https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html

Palmer, T. N., & Anderson, D. L. T. (1994). The prospects for seasonal forecasting—A review paper. *Quarterly Journal of the Royal Meteorological Society*, *120*(518), 755–793. https://doi.org/10.1002/qj.49712051802

Park, J.-Y., Stock, C. A., Yang, X., Dunne, J. P., Rosati, A., John, J., & Zhang, S. (2018). Modeling global ocean biogeochemistry with physical data assimilation: A pragmatic solution to the equatorial instability. *Journal of Advances in Modeling Earth Systems*, *10*(3), 891–906. https://doi.org/10.1002/2017MS001223

Piecuch, C. G., Dangendorf, S., Ponte, R. M., & Marcos, M. (2016). Annual sea level changes on the North American Northeast Coast: Influence of local winds and barotropic motions. *Journal of Climate*, *29*(13), 4801–4816. https://doi.org/10.1175/jcli-d-16-0048.1

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite SST analysis for climate. *Journal of Climate*, *15*(13), 1609–1625. https://doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2

Rohith, B., Paul, A., Durand, F., Testut, L., Prerna, S., Afroosa, M., et al. (2019). Basin-wide sea level coherency in the tropical Indian Ocean driven by Madden–Julian Oscillation. *Nature Communications*, *10*(1), 1257. https://doi.org/10.1038/s41467-019-09243-5

Segschneider, J., Anderson, D. L. T., & Stockdale, T. N. (2000). Toward the use of altimetry for operational seasonal forecasting. *Journal of Climate*, *13*(17), 3115–3138. https://doi.org/10.1175/1520-0442(2000)013<3115:TTUOAF>2.0.CO;2

Shi, H., Jin, F.-F., Wills, R. C. J., Jacox, M. G., Amaya, D. J., Black, B. A., et al. (2022). Global decline in ocean memory over the 21st century. *Science Advances*, *8*(18). Retrieved from https://www.science.org/doi/abs/10.1126/sciadv.abm3468

Spillman, C. M., Smith, G., Yin, Y., & Alves, O. (2023). Retrospective forecasts from ACCESS-S. [Dataset]. Retrieved from http://poama.bom.gov.au/general/hindcast_data.html

Spillman, C. M., & Smith, G. A. (2021). A new operational seasonal thermal stress prediction tool for coral reefs around Australia. *Frontiers in Marine Science*, *8*(687833). https://doi.org/10.3389/fmars.2021.687833

Spillman, C. M., Smith, G. A., Hobday, A. J., & Hartog, J. R. (2021). Onset and decline rates of marine heatwaves: Global trends, seasonal forecasts and marine management. *Frontiers in Climate*, *3*. https://doi.org/10.3389/fclim.2021.801217

Stephens, S. A., Bell, R. G., Ramsay, D., & Goodhue, N. (2014). High-water alerts from coinciding high astronomical tide and high mean sea level anomaly in the Pacific Islands region. *Journal of Atmospheric and Oceanic Technology*, *31*(12), 2829–2843. https://doi.org/10.1175/JTECH-D-14-00027.1

Stockdale, T. N. (1997). Coupled ocean–atmosphere forecasts in the presence of climate drift. *Monthly Weather Review*, *125*(5), 809–818. https://doi.org/10.1175/1520-0493(1997)125<0809:COAFIT>2.0.CO;2

Stockdale, T. N., Anderson, D. L. T., Alves, J., & Balmaseda, M. (1998). Global seasonal forecasts using a coupled atmosphere-ocean model. *Nature*, *392*(6674), 370–373. https://doi.org/10.1038/32861

Storto, A., Alvera-Azcárate, A., Balmaseda, M. A., Barth, A., Chevallier, M., Counillon, F., et al. (2019). Ocean reanalyses: Recent advances and unsolved challenges. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00418

Storto, A., Masina, S., Simoncelli, S., Iovino, D., Cipollone, A., Drevillon, M., et al. (2019). The added value of the multi-system spread information for ocean heat content and steric sea level investigations in the CMEMS GREP ensemble reanalysis product. *Climate Dynamics*, *53*(1–2), 287–312. https://doi.org/10.1007/s00382-018-4585-5

Tietsche, S., Balmaseda, M., Zuo, H., Roberts, C., Mayer, M., & Ferranti, L. (2020). The importance of North Atlantic Ocean transports for seasonal forecasts. *Climate Dynamics*, *55*(7–8), 1995–2011. https://doi.org/10.1007/s00382-020-05364-6

Timmermann, A., McGregor, S., & Jin, F. F. (2010). Wind effects on past and future regional sea level trends in the southern Indo-Pacific. *Journal of Climate*, *23*(16), 4429–4437. https://doi.org/10.1175/2010JCLI3519.1

Vidard, A., Anderson, D. L. T., & Balmaseda, M. (2007). Impact of ocean observation systems on ocean analysis and seasonal forecasts. *Monthly Weather Review*, *135*(2), 409–429. https://doi.org/10.1175/MWR3310.1

Vidard, A., Balmaseda, M., & Anderson, D. (2009). Assimilation of altimeter data in the ECMWF ocean analysis system 3. *Monthly Weather Review*, *137*(4), 1393–1408. https://doi.org/10.1175/2008MWR2668.1

Waters, J., Bell, M. J., Martin, M. J., & Lea, D. J. (2016). Reducing ocean model imbalances in the equatorial region caused by data assimilation. *Quart. J. Roy. Meteor. Soc.*, *143*(702), 195–208. https://doi.org/10.1002/qj.2912

Webb, D. J., & Moore, A. (1986). Assimilation of altimeter data into ocean models. *Journal of Physical Oceanography*, *16*(11), 1901–1913. https://doi.org/10.1175/1520-0485(1986)016<1901:AOADIO>2.0.CO;2

Wedd, R., Alves, O., de Burgh-Day, C., Down, C., Griffiths, M., Hendon, H. H., et al. (2022). ACCESS-S2: The upgraded Bureau of Meteorology multiweek to seasonal prediction system. *J. South. Hemisph. Earth Syst. Sci.*, *72*(3), 218–242. https://doi.org/10.1071/ES22026

Wenegrat, J. O., Thomas, L. N., Sundermeyer, M. A., Taylor, J. R., D'Asaro, E. A., Klymak, J. M., et al. (2020). Enhanced mixing across the gyre boundary at the Gulf Stream front. *Proceedings of the National Academy of Sciences*, *117*(30), 17607–17614. https://doi.org/10.1073/pnas.2005558117

Widlansky, M. J., Long, X., & Schloesser, F. (2020). Increase in sea level variability with ocean warming associated with the nonlinear thermal expansion of seawater. *Communications Earth & Environment*, *1*, 1–12. https://doi.org/10.1038/s43247-020-0008-8

Widlansky, M. J., Marra, J. J., Chowdhury, M. R., Stephens, S. A., Miles, E. R., Fauchereau, N., et al. (2017). Multi-model ensemble sea level forecasts for tropical Pacific Islands. *Journal of Applied Meteorology*, *56*(4), 849–862. https://doi.org/10.1175/jamc-d-16-0284.1

Wilks, D. S. (Ed.) (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic Press.

Wunsch, C. (1991). Large-scale response of the ocean to atmospheric forcing at low frequencies. *Journal of Geophysical Research*, *96*(15), 15083–15092. https://doi.org/10.1029/91JC01457

Zuo, H., & Balmaseda, M. A. (2023). ECMWF ORAS5. [Dataset]. Retrieved from https://www.cen.uni-hamburg.de/icdc/data/ocean/easy-init-ocean/ecmwf-oras5.html

Zuo, H., Balmaseda, M. A., de Boisseson, E., Tietsche, S., Mayer, M., & de Rosnay, P. (2021). The ORAP6 ocean and sea-ice reanalysis: Description and evaluation. *Paper presented at the EGU General Assembly*.

Zuo, H., Balmaseda, M. A., & Mogensen, K. (2017). The new eddy-permitting ORAP5 ocean reanalysis: Description, evaluation and uncertainties in climate signals. *Climate Dynamics*, *49*(3), 791–811. https://doi.org/10.1007/s00382-015-2675-1

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., & Mayer, M. (2019). The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: A description of the system and assessment. *Ocean Science*, *15*(3), 779–808. https://doi.org/10.5194/os-15-779-2019