

# A Drift-Free Decadal Climate Prediction System for the Community Earth System Model

YOSHIMITSU CHIKAMOTO

*Department of Plants, Soils and Climate, Utah State University, Logan, Utah*

AXEL TIMMERMANN

*Institute for Basic Science Center for Climate Physics, and Pusan National University, Busan, South Korea*

MATTHEW J. WIDLANSKY

*Joint Institute for Marine and Atmospheric Research, University of Hawai'i at Mānoa, Honolulu, Hawaii*

SHAOQING ZHANG

*Physical Oceanography Laboratory/Cooperation and Innovation Center for Marine Science and Technology, College of Atmosphere and Ocean, Ocean University of China, and Function Laboratory for Ocean Dynamics and Climate/Qingdao National Laboratory for Marine Science and Technology, Qingdao, China*

MAGDALENA A. BALMASEDA


*European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*

(Manuscript received 14 November 2018, in final form 4 April 2019)

## ABSTRACT

Performance of a newly developed decadal climate prediction system is examined using the low-resolution Community Earth System Model (CESM). To identify key sources of predictability and determine the role of upper and deeper ocean data assimilation, we first conduct a series of perfect model experiments. These experiments reveal the importance of upper ocean temperature and salinity assimilation in reducing sea surface temperature biases. However, to reduce biases in the sea surface height, data assimilation below 300 m in the ocean is necessary, in particular for high-latitude regions. The perfect model experiments clearly emphasize the key role of combined three-dimensional ocean temperature and salinity assimilation in reproducing mean state and model trajectories. Applying this knowledge to the realistic decadal climate prediction system, we conducted an ensemble of ocean assimilation simulations with the fully coupled CESM covering the period 1960–2014. In this system, we assimilate three-dimensional ocean temperature and salinity data into the ocean component of CESM. Instead of assimilating direct observations, we assimilate temperature and salinity anomalies obtained from the ECMWF Ocean Reanalysis version 4 (ORA-S4). Anomalies are calculated relative to the sum of the ORA-S4 climatology and an estimate of the externally forced signal. As a result of applying the balanced ocean conditions to the model, our hindcasts show only very little drift and initialization shocks. This new prediction system exhibits multiyear predictive skills for decadal climate variations of the Atlantic meridional overturning circulation (AMOC) and North Pacific decadal variability.

---

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

---

*Corresponding author:* Yoshimitsu Chikamoto, [yoshi.chikamoto@usu.edu](mailto:yoshi.chikamoto@usu.edu)

## 1. Introduction

Skillful forecasts on multiyear to decadal time scales using a dynamical climate prediction system have potential to provide useful information for sectors such as agriculture, renewable energy, water resources, coastal

DOI: 10.1175/JCLI-D-18-0788.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](https://www.ametsoc.org/PUBSReuseLicenses).

communities, forestry, and fisheries (Meehl et al. 2014; Meehl and Teng 2014; M. Chikamoto et al. 2015; Y. Chikamoto et al. 2015a). Potential sources of multiyear skill are often associated with slow ocean dynamical processes (Teng and Branstator 2011; Robson et al. 2012; Yeager et al. 2012; Chikamoto et al. 2013; Y. Chikamoto et al. 2015a), as documented in recent studies focusing on the Atlantic meridional overturning circulation (AMOC; Keenlyside et al. 2008; Teng et al. 2011; Matei et al. 2012; Doblas-Reyes et al. 2013; Ham et al. 2014; Karspeck et al. 2014) and the Pacific decadal oscillation (PDO; Mochizuki et al. 2010, 2012; Chikamoto et al. 2013; Kim et al. 2014).

Despite recent advances in initialization techniques, many dynamical climate prediction systems still suffer from model biases and initialization shocks, which can severely contaminate model forecasts and may substantially reduce predictive skills (Doblas-Reyes et al. 2011; Magnusson et al. 2013; Smith et al. 2013; Hazeleger et al. 2013; Hawkins et al. 2014). In particular, systems that rely on full data assimilation are likely to experience such shocks, because models adjust in two ways to the full initial conditions: 1) by responding to the difference in mean state between observation and model simulation (mean state bias) and 2) by adjusting to the unresolved physical processes due to model deficiencies (systematic error). Because the mean state biases present in all climate models, the prediction systems will show seasonally dependent drifts (Doblas-Reyes et al. 2011; Kim et al. 2012; Balmaseda 2017), which then needs to be removed by empirically developed methods. Generally, the predictive skill for the remaining anomalies is a function of the drift estimation method, ensemble size, and the number of initial conditions tested in addition to the climate model and assimilation performance (Kim et al. 2012; Kharin et al. 2012; Narapusetty et al. 2014; Kim et al. 2014; Fučkar et al. 2014; Choudhury et al. 2017). Such a posteriori correction can be particularly problematic for non-Gaussian distributed variables, such as rainfall, river runoff, and sea ice. Yet, full-field assimilation systems are widely employed and have provided initial conditions to skillful forecast systems, at least for seasonal outlooks (Saha et al. 2014; Vialard et al. 2005). More recent seasonal prediction studies use an online bias adjustment method that estimates and corrects the bias during the data assimilation process (Dee and Da Silva 1998; Bell et al. 2004; Balmaseda et al. 2007). However, such methods would be computationally too demanding to be applied for decadal climate prediction experiments.

Another approach to manage model drift is anomaly assimilation, in which the climate model is initialized with the observed anomaly, instead of the fully observed field. In this process, the model maintains its own

climatology during the assimilation process (Smith et al. 2007; Tabebe et al. 2012; Chikamoto et al. 2013). This approach requires knowledge of both observed and model simulated climatological conditions during the assimilation process. The goal is to minimize the drift that is otherwise induced by an initial condition mismatch between observed and simulated mean climatologies (Doblas-Reyes et al. 2011; Meehl et al. 2014). This approach shows a higher fidelity for non-Gaussian distributed climate variables (Y. Chikamoto et al. 2015b; Chikamoto et al. 2017) compared to the full data assimilation method with a posteriori drift removal. However, the anomaly assimilation approach also has disadvantages, associated with the fact that atmospheric sensitivities to ocean forcing are calculated relative to a biased model state, rather than the observed mean state. Furthermore, the anomaly assimilation approach can show localized model drifts during the predictions (Chikamoto et al. 2012; Balmaseda 2017).

Decadal climate prediction systems also differ from seasonal forecasting frameworks in the type of variables that are being assimilated. Given the extensive spatiotemporal coverage of high-quality sea surface temperature (SST) observations, it is appropriate to use these observations as a key component of a climate prediction system. In fact, seasonal predictions using SST nudging show a reasonable fidelity with skills that sometimes comparable with more complex climate prediction and assimilation systems (e.g., Alves et al. 2004; Luo et al. 2005). Whether SST nudging is sufficient to initialize decadal climate states and ocean circulation processes, such as the AMOC, is still a matter of debate (Keenlyside et al. 2008; Swingedouw et al. 2013; Dunstone and Smith 2010). The assumption is that SST contains enough information to initialize even subsurface ocean processes. If the atmospheric momentum, freshwater, and heat flux anomalies were only a function of the sea surface temperature anomaly, this approach would be appropriate. However, a large fraction of these fluxes, which also serve as drivers of the wind and thermohaline circulation, is internally generated in the atmosphere and is unrelated to the prevailing sea surface temperature anomaly. To capture this variability, either atmospheric data assimilation needs to be applied (Karspeck et al. 2014) or the three-dimensional ocean temperature and salinity fields that result from this forcing need to be assimilated (Zhang et al. 2010).

The goal of this paper is 1) to determine the effect of two- and three-dimensional temperature and salinity assimilation on coupled model biases, model drift, and predictability using a perfect modeling framework, and 2) to use this information to develop a new realistic decadal climate prediction system for the Community Earth System Model (CESM), which minimizes model

TABLE 1. Assimilation experiments in the perfect model framework.

Name	Depth	Variables	Ensemble	Period
SST	Surface	Temperature	1	100–150
SST–SSS	Surface	Temperature and salinity	1	100–200
TS upper	0–300 m	Temperature and salinity	1	100–200
TS deeper	300–3000 m	Temperature and salinity	1	100–200
T full	0–3000 m	Temperature	1	100–150
TS full	0–3000 m	Temperature and salinity	1	100–200

drift and initialization shocks by assimilating three-dimensional temperature and salinity anomaly fields from the ECMWF Ocean Reanalysis version 4 (ORA-S4) (Balmaseda et al. 2013). Section 2 of the paper describes the model setup and the key experiments. Section 3 assesses the effect of assimilating temperature and salinity at different depths on climate biases, drift, and predictability. In section 4, we introduce a new method to minimize the effect of model biases on multiyear predictions. Section 5 discusses key results from our perfect model and realistic hindcast experiments conducted with CESM and a simple ocean data assimilation system for the period from 1960–2014. We provide conclusions in section 6.

## 2. Model and experiment

### a. Climate model

In this study, we use a low-resolution version of the fully coupled climate model CESM 1.0.3 (Shields et al. 2012). The model physical component is largely based on the Community Climate System Model version 4 (Gent et al. 2011). The atmospheric and land components have T31 spectral resolution with hybrid sigma pressure coordinate consisting of 26 atmospheric levels from sea level to 1 hPa and 15 soil levels from the surface to a depth of 35 m. The ocean component consists of POP2 with 60 levels, which is coupled with a sea ice module and is used with horizontal resolution of approximately 3° latitude–longitude grid. A configuration with a displaced North Pole is chosen to avoid pole singularities and short time steps. Details of the basic model performance in this configuration can be found in results of CESM described in Shields et al. (2012) and Y. Chikamoto et al. (2015b).

### b. Perfect model experiment

After a 100-yr coupled model spinup, we perform a 900-yr-long preindustrial control simulation (CTL) using constant greenhouse gas conditions for year 1850 CE, which was introduced in Y. Chikamoto et al. (2015b). Following the *perfect* modeling framework idea, we interpret this CTL simulation as a “surrogate reality” or “surrogate observations.” For the assimilation experiment

(ASSM) in the perfect modeling framework, we conduct a climate model integration from an arbitrary initial condition (e.g., 1 January of model year 500). We then assimilate monthly ocean data from the CTL simulation (model years 100–200) into POP2. The polar sea ice regions are excluded from the assimilation.

These so-called monthly surrogate observations were then linearly interpolated to the daily fields. Our assimilation is based on the incremental analysis update scheme (Bloom et al. 1996; Huang et al. 2002). Analysis increments are estimated from a temporally, spatially, and vertically invariant model-to-observation ratio in analysis errors (1/2) and added as forcing into the model’s temperature and salinity tendency equations during an analysis interval of one day (Mochizuki et al. 2010; Tabebe et al. 2012). Using this simple approach, we evaluate the impacts of different ocean variable configurations on decadal climate initialization by partially assimilating the oceanic temperature and salinity at the surface, upper (>300 m) ocean, and deeper (<300 m) ocean. The model experiments conducted in this study are summarized in Table 1.

### c. Realistic decadal forecasting experiments

Based on the assimilation system developed and tested for the perfect model configuration, we conducted decadal climate predictions using estimates of global three-dimensional ocean temperature and salinity data. These data are obtained from the ECMWF ORA-S4 ocean reanalysis during 1958–2014 period. Overall our decadal prediction runs follow the experimental design of CMIP5 (Taylor et al. 2009; Murphy et al. 2010).

Our decadal climate predictions are based on three model experiments: the uninitialized, assimilation, and initialized runs. In the 10-member CESM ensemble simulations of the uninitialized run, we prescribe natural and anthropogenic radiative forcings for the period 1850–2030. Initial conditions for the uninitialized run are obtained from 10 random years of the preindustrial control simulation (i.e., the CTL run). The uninitialized but radiatively forced run is used to determine the model climatological biases relative to the observations and to identify the externally forced component  $\mathbf{X}_{\text{ext}}$  of the three-dimensional temperature and salinity fields.

The ensemble also allows us to obtain estimates of the model's internally generated climate variability  $\mathbf{X}_{\text{int}}$ . Estimates of the observed internal variability  $\mathbf{X}_{\text{int}}^o$ , added onto the model climatology  $\mathbf{X}_{\text{clm}}^m$  and an estimate of model externally forced signal  $\mathbf{X}_{\text{ext}}^m$  are then assimilated back into POP2 (i.e., the ocean component of CESM). This method is similar to the anomaly assimilation conducted by previous studies (Mochizuki et al. 2010; Chikamoto et al. 2013) but takes into account the fact that the model's externally forced component may differ from the real anthropogenic signal. Details for this estimation of internal variability and externally forced component are described in section 4.

The estimated three-dimensional monthly ocean temperature and salinity fields, obtained as the sum of the internal variability in the observations  $\mathbf{X}_{\text{int}}^o$ , the estimate of the externally forced signal  $\mathbf{X}_{\text{ext}}^m$ , and the model climatology  $\mathbf{X}_{\text{clm}}^m$ :  $\mathbf{X}^a = \mathbf{X}_{\text{int}}^o + \mathbf{X}_{\text{ext}}^m + \mathbf{X}_{\text{clm}}^m$  (outside sea ice regions) are referred to as the assimilated signal  $\mathbf{X}^a$ . Then  $\mathbf{X}^a$  is infused into POP2 using the incremental analysis update scheme (Bloom et al. 1996; Huang et al. 2002). We apply strong and weak model-to-observation ratios (1/2 and 1/3, respectively) in analysis errors with a temporally and spatially invariant value and add analysis increments as forcing terms into the model's temperature and salinity tendency equations during an analysis interval of one day (Mochizuki et al. 2010; Chikamoto et al. 2013). The assimilation runs include five ensemble members with strong and five with weak model-to-observation ratios and cover the period from 1958–2014. From the assimilation run, we first obtain a pair of atmospheric and oceanic initial conditions from the 10-member ensemble and then conduct a 10-yr-long ensemble hindcast experiment starting from 1 January every year from 1960 to 2014, which we call the initialized run. By assimilating anomalies relative to the model climatology, our initialized run has a negligible climate drift during the prediction period, as will be documented in subsequent sections. Therefore, no posteriori corrections of model drift need to be applied to the actual predictions.

### 3. Effect of different assimilation types in a perfect model framework

To identify the impacts of surface and subsurface temperature and salinity conditions on climatological biases and initial errors in SST, sea surface height (SSH), and the AMOC, we devised a set of perfect model experiments:

- assimilation of SST field,
- assimilation of SST and SSS fields,

- assimilation of upper (0–300 m) ocean temperature and salinity fields,
- assimilation of deeper (300–3000 m) temperature and salinity fields,
- assimilation of full-depth three-dimensional temperature field, and
- assimilation of full-depth three-dimensional temperature and salinity fields.

The experiments are summarized in Table 1. Here we demonstrate that the assimilation of insufficient information cannot only affect the predictability and error growth, but can also influence the model mean state through dynamical inconsistencies.

#### a. Climatological biases

Climatological biases during the assimilation are calculated here as the SST difference between the control simulation and the assimilation experiments averaged for 40 years after the 10-yr model spinup (Fig. 1). When we assimilate the 3D temperature and salinity from the surface to 3000 m (TS full run), the climatological SST bias is smaller than 0.1 K almost everywhere (Fig. 1f). A comparable result is also obtained by assimilating upper temperature and salinity from the surface to 300 m (TS upper run; Fig. 1c). In contrast, the climatological SST bias is larger when assimilating only deeper ocean data (TS deeper run; Fig. 1d), indicating a minor role of slow adjustment processes in the deep ocean for affecting surface conditions. Assimilating the SST field in the surface assimilation experiments (SST and SST–SSS runs; Figs. 1a,b), we still find considerable SST biases, which suggests that the thermodynamical structure of the mixed layer and thermocline play a crucial role in constraining the dynamical and thermodynamical state of the surface. The climatological SST biases increases further when we assimilate the 3D temperature data only (T full run; Fig. 1e). These results imply that assimilation of not only upper temperature but also salinity data is critical to accurately initializing the coupled model.

Assimilating deeper ocean conditions also contributes toward reducing SSH climatological biases (Fig. 2). In the TS full run, the climatological SSH bias is mostly smaller than 1 cm, except for the Kuroshio–Oyashio Extension region and high latitudes of the Southern Ocean (Fig. 2f). Sea surface height biases in the TS deeper run are comparable to the TS full run (Fig. 2d), whereas the TS upper run shows a larger SSH bias in the Atlantic Ocean (Fig. 2c). The SST–SSS run (Fig. 2b) also shows a similar SSH bias pattern compared to the TS upper run, indicating that upper ocean data are not sufficient to reduce the SSH biases. Consistent with

## SST climatological bias (110-150)

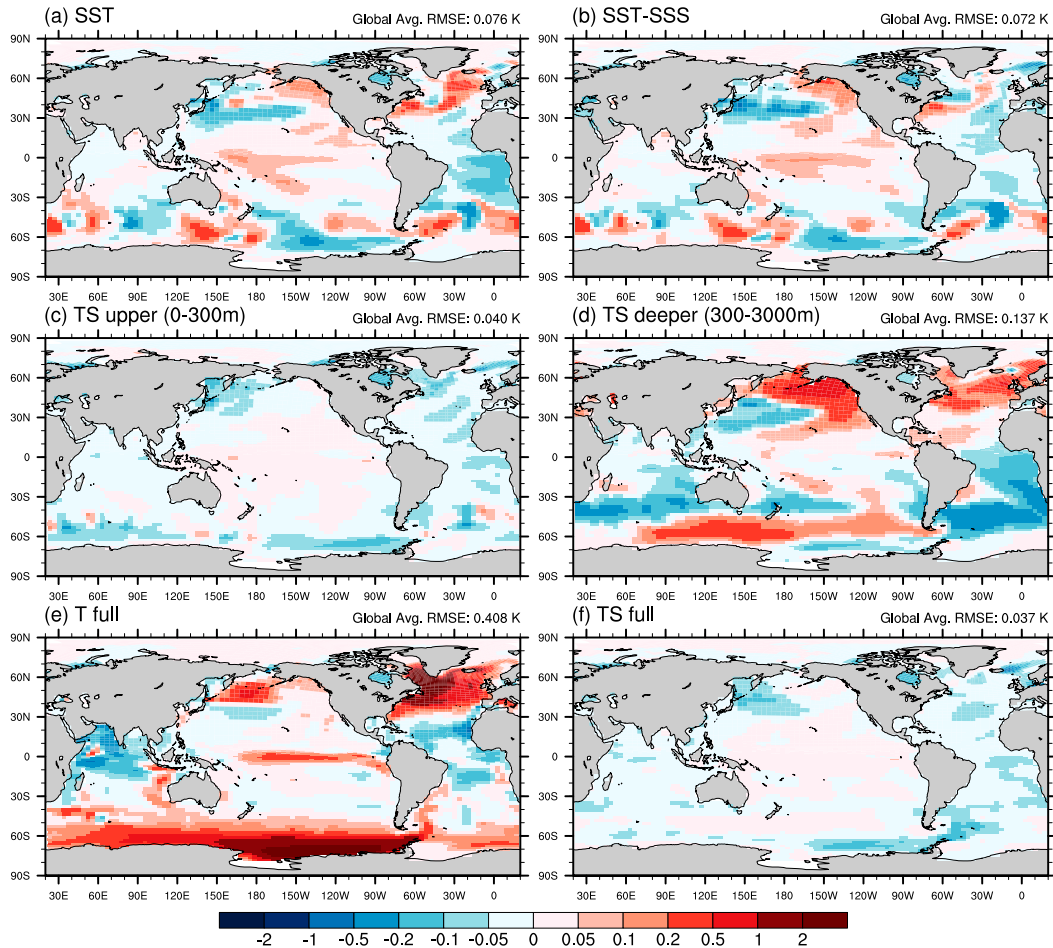


FIG. 1. Climatological biases of annual mean SST in partial assimilation experiments of (a) SST, (b) SST-SSS, (c) TS upper, (d) TS deeper, (e) T full, and (f) TS full runs. Climatology is defined as the average from model year 110 to 150. Globally averaged RMSEs of climatological bias are denoted in upper-right corners.

these global maps, globally averaged root-mean-square errors (RMSEs) in TS full and TS deeper runs relative to the perfect model control run are less than half of those in the TS upper and the SST-SSS runs (0.673, 0.944, 1.908, and 1.973 cm, respectively). The RMSE values for SSH in the TS full and TS deeper runs are on average smaller compared to those of the TS upper and SST-SSS runs, implying that the density fields in the deeper ocean are essential to capture the large-scale barotropic flows. In addition to the deeper ocean assimilation run, we find larger SSH biases in the SST and the T full runs (Figs. 2a,e) compared to the other runs. This result also highlights the importance of salinity initialization for resolving large-scale ocean circulations as well as decadal climate predictions. Specifically, the SSH bias becomes worse when we assimilate temperature without salinity in the deeper ocean (T full run), compared to assimilations with only the surface fields (SST and

SST-SSS runs). This deterioration indicates that accurate 3D salinity observations are required to realistically initialize the seasonal to multiyear predictions.

### b. Initial error

To evaluate how well the assimilated runs can reproduce the perfect model trajectory of the control run, we estimate the RMSE skill score between the control and assimilated experiments as follows:

$$\text{RMSE skill score (\%)} = \left[ 1 - \frac{\sqrt{\sum_{i=1}^N (x'_i - y'_i)^2}}{\sqrt{\sum_{i=1}^N y_i'^2}} \right] \times 100, \quad (1)$$

where  $x'$  and  $y'$  are anomalies in the assimilation and control runs with a time  $i = 1, \dots, N$ . Anomalies are

## SSH climatological bias (110-150)

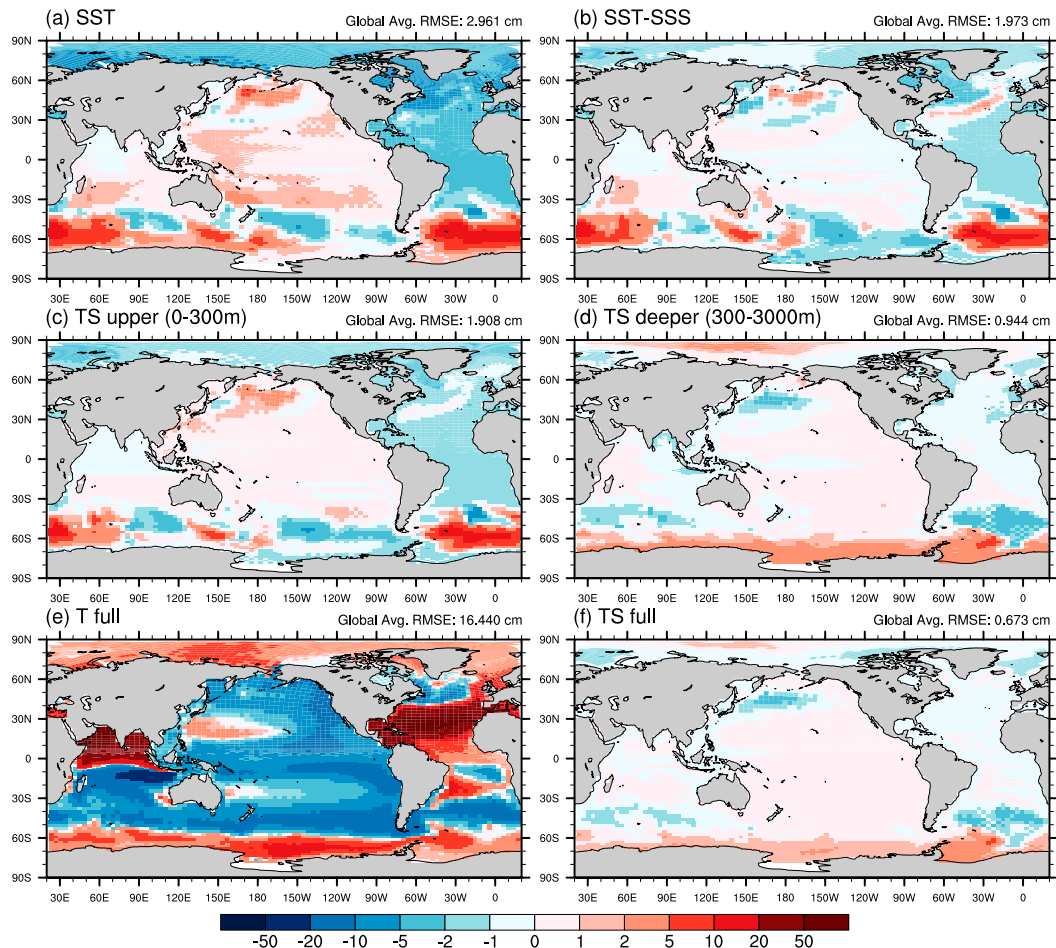


FIG. 2. Climatological biases of annual mean SSH in partial assimilation experiments of (a) SST, (b) SST-SSS, (c) TS upper, (d) TS deeper, (e) T full, and (f) TS full runs. Climatology is defined as the average from model year 110 to 150. Globally averaged RMSEs of climatological bias are denoted in upper-right corners.

defined as deviations from the climatological mean for 40 years after the 10-yr spinup in each experiment. To focus on the error of temporal variability, rather than on the mean bias, we neglect the climatological biases introduced by the assimilation and discussed in the previous section. Higher (lower) skill RMSE score values indicate high (low) levels fidelity in reproducing the trajectory of control run. These values are an indication about how well certain physical variables constrain the assimilation.

Figure 3 shows RMSE skill score of SST anomalies in each assimilation experiment. Similar to the result for the climatological SST biases, TS full and TS upper (0–300 m) runs show the highest RMSE skill scores of SST anomalies in our experiments (Figs. 3c,f), which suggest that upper ocean temperature variability contributes to initializing SST anomalies through thermodynamical

processes, although SST anomalies are often disturbed by atmospheric high-frequency noise. This upper thermodynamical contribution is confirmed by the T full run (i.e., the 3D temperature without salinity assimilation): high RMSE skill scores of SST anomalies, similar to those found in the TS full run, except for the Oyashio region and over the Labrador Sea (Fig. 3e). The surface ocean assimilation also contributes to initializing SST anomalies by buffering against atmospheric noise, particularly over the tropical Pacific and Atlantic (Figs. 3a,b), but in general the constraints are relatively weak. Understandably, the assimilation of deep ocean TS does not provide any constraints on the evolution of SST anomalies (Fig. 3d).

Comparing the SSH skill score with the SST skill, we find a stronger contribution from deeper ocean variability, as well as from the upper ocean (Fig. 4). Similar

## RMSE skill score of SST (12-mon, 110-150)

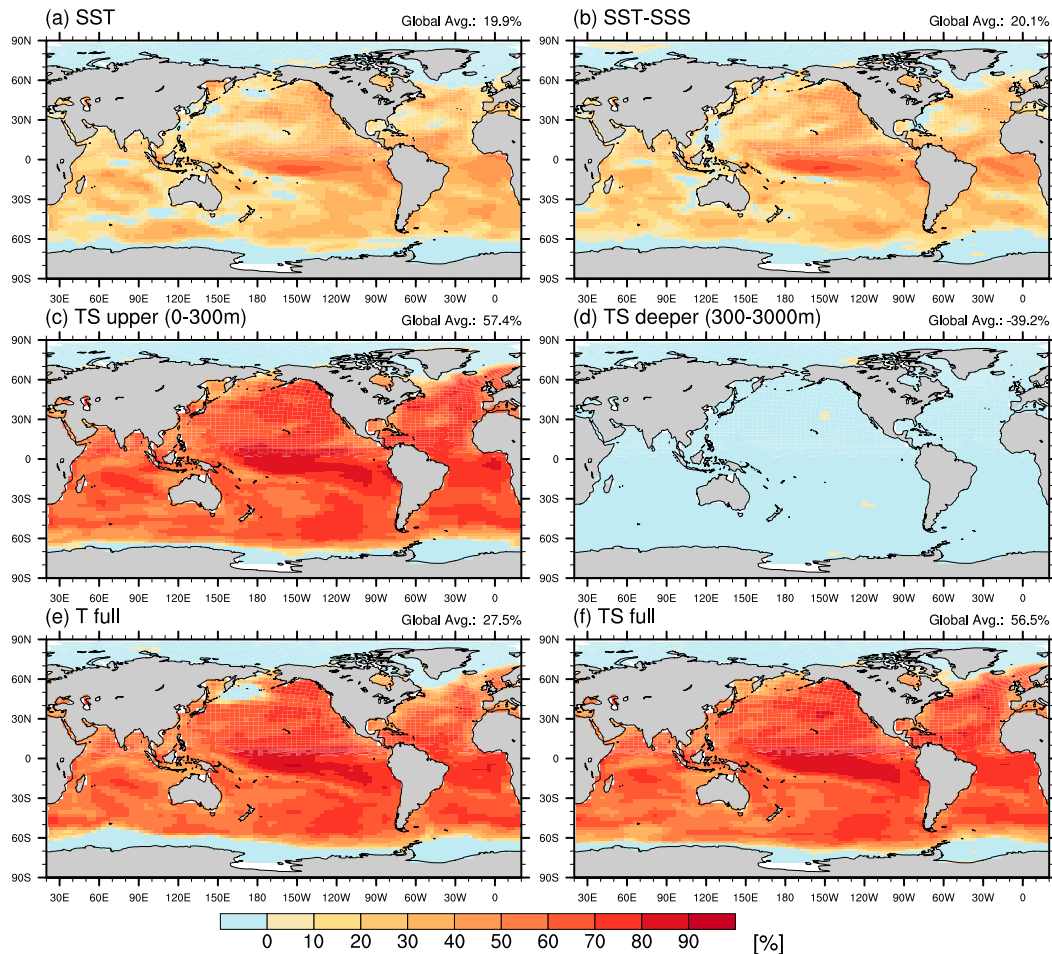


FIG. 3. RMSE skill score of SST anomalies in assimilation experiments of (a) SST, (b) SST-SSS, (c) TS upper, (d) TS deeper, (e) T full and (f) TS full runs. A 12-month running mean is applied to anomalies. Anomalies are calculated by deviations from the climatological mean for the model years 110–150. Globally averaged RMSE skill scores are denoted in upper-right corners.

to our previous results, the TS full run shows the highest skill of SSH anomalies globally (Fig. 4f). The TS upper run also has the higher skill score, particularly in the tropical oceans, whereas it is lower in the Southern Ocean and the northwestern Pacific (Fig. 4c). Similar features are found in the SST and SST-SSS runs: skill scores are higher in the tropics and lower in the extratropics (Figs. 4a,b). In contrast, the TS deeper run shows higher skill scores outside the tropics (Fig. 4d). This difference suggests that the deeper ocean initialization contributes to improving decadal climate predictability in the extratropical regions similarly to where deep ocean assimilation is most beneficial. We note that the SSH skill score deteriorates in the tropical eastern Pacific when assimilating the 3D temperature without salinity (T full run; Fig. 4e) whereas the surface assimilation runs perform

better (Figs. 4a,b). Clearly, we should assimilate the 3D temperature and salinity fields to most accurately initialize ocean density structures, which are critical for successful decadal climate prediction. Impacts of temperature and salinity assimilation on AMOC and Pacific decadal variability are examined in the following two subsections.

### c. Atlantic meridional overturning circulation

Previous studies have demonstrated decadal predictive skill for the AMOC (Griffies and Bryan 1997; Grötzner et al. 1999; Branstator and Teng 2010; Keenlyside et al. 2008; Matei et al. 2012; Chikamoto et al. 2013). The predictability originates from the long-term adjustment processes of the thermohaline circulation, which range from baroclinic wave to diffusive time

## RMSE skill score of SSH (12-mon, 110-150)

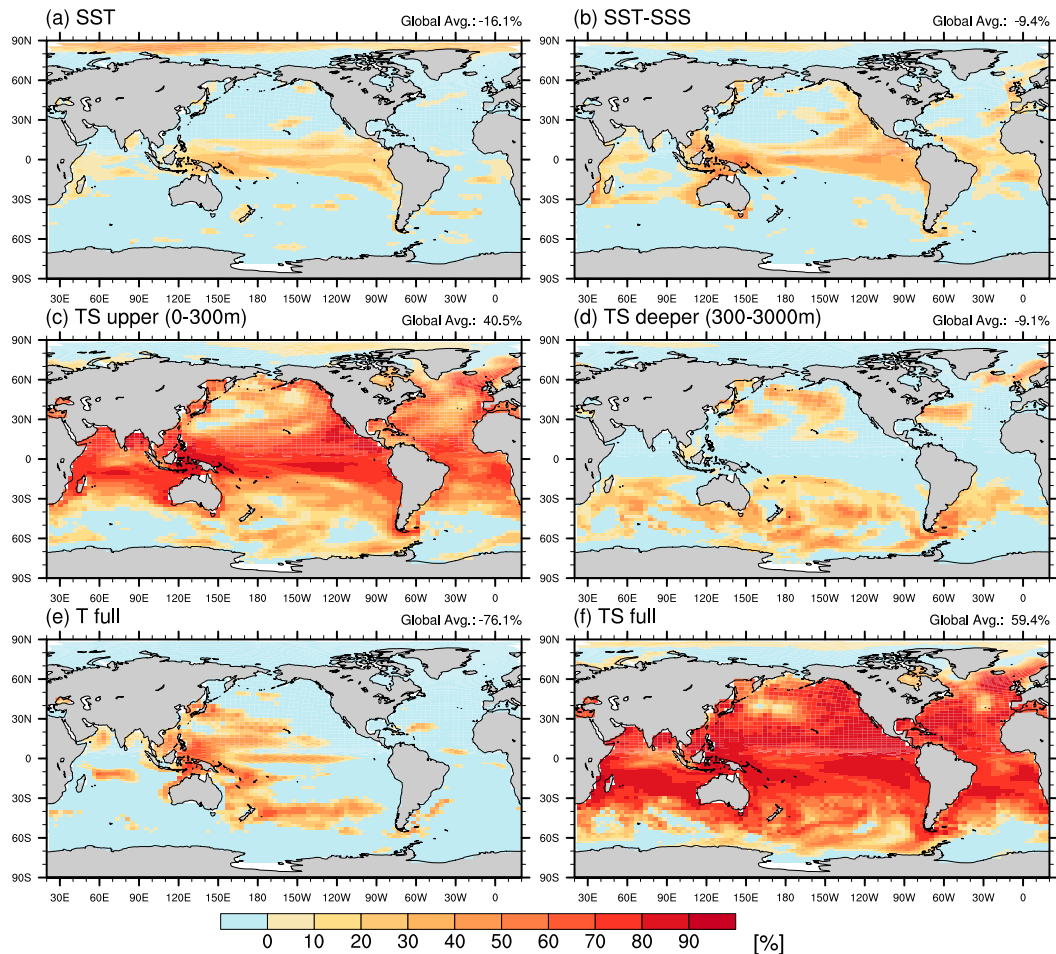


FIG. 4. RMSE skill score of SSH anomalies in assimilation experiments of the (a) SST, (b) SST-SSS, (c) TS upper, (d) TS deeper, (e) T full, and (f) TS full runs. A 12-month running mean is applied to anomalies. Anomalies are calculated by deviations from the climatological mean for the model years 110–150. Globally averaged RMSE skill scores are denoted in the upper-right corners.

scales. Still there is no consensus on how to optimally initialize the AMOC state. Several mechanisms have been proposed as key controlling factors for AMOC variability: atmospheric wind forcing (Danabasoglu et al. 2016), atmosphere–ocean interactions attributed to SST variability (Keenlyside et al. 2008), 3D ocean states (Dunstone and Smith 2010), and external forcing such as aerosol concentration and volcanic eruption (Booth et al. 2012; Swingedouw et al. 2013). In addition to the issue of initializing the AMOC, there remains a large model spread in simulating the AMOC and its variability (Gregory et al. 2005; Cheng et al. 2013). This situation is further exacerbated by incomplete observational constraints on the AMOC state (Keller et al. 2007). Using the partial assimilation approach and following the perfect model framework, we ascertain here

the effect of deeper ocean temperature and salinity data in initializing AMOC variability.

Figure 5 shows the climatology and RMSE of annual mean AMOC variations in each assimilation experiment relative to the perfect model control simulation.<sup>1</sup> The RMSE is obtained from difference of the annual mean streamfunction between control and assimilation runs, which include climatological biases. When we assimilate 3D temperature and salinity data from the control simulation from the surface to 3000 m, the annual mean AMOC variability is simulated well (Fig. 5f). A similar

<sup>1</sup> The RMSE here is different from RMSE skill score in Figs. 4 and 5. High values of the RMSE indicate low reproducibility of the control trajectory through assimilation.



## AMOC climatology (contour) &amp; RMSE (shaded)

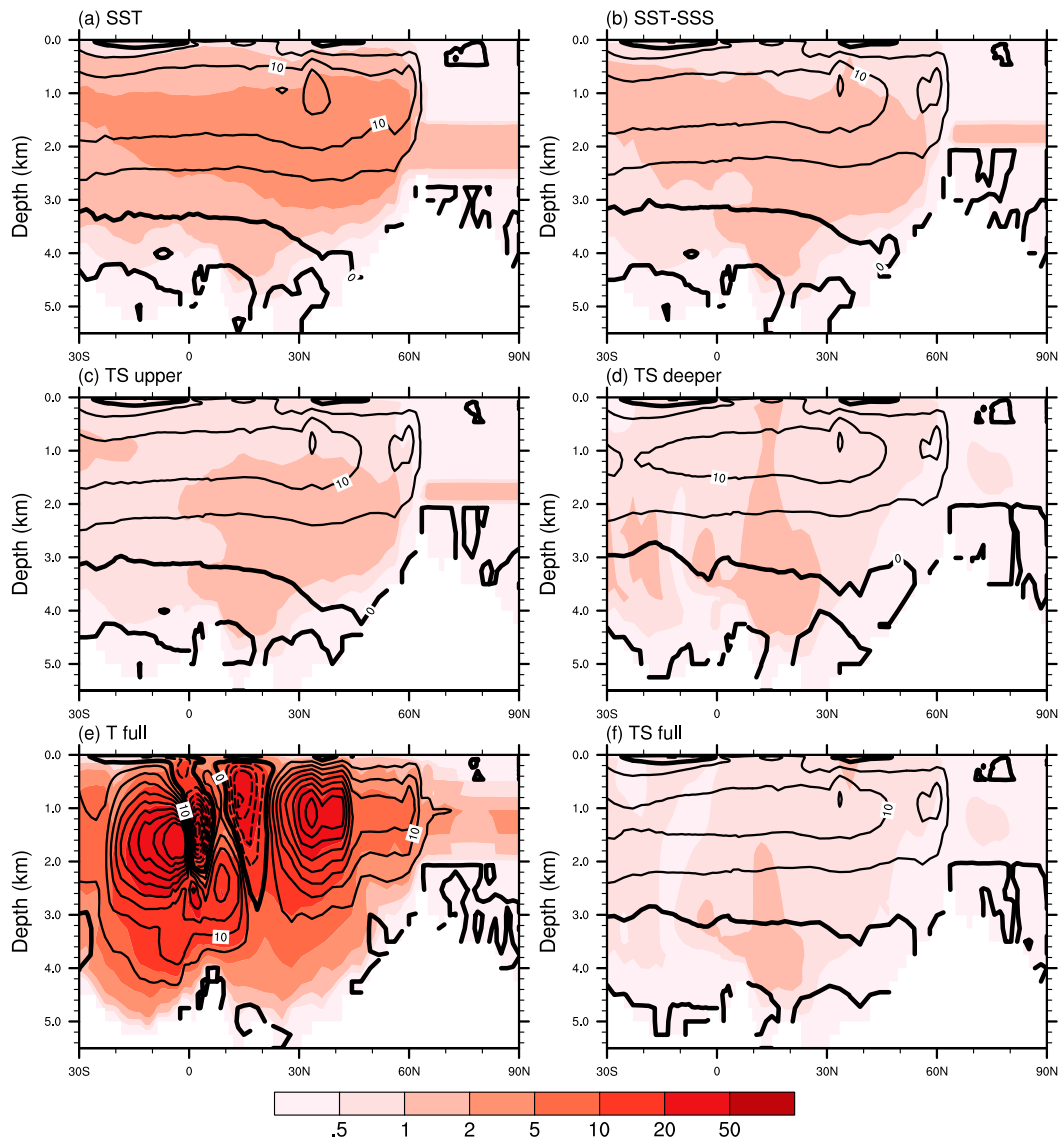


FIG. 5. Annual mean climatology (contours) and RMSE (shaded) of AMOC in (a) SST, (b) SST-SSS, (c) TS upper, (d) TS deeper, (e) T full, and (f) TS full assimilation experiments. AMOC is derived from the zonal mean streamfunction ( $S_v$ ) in the Atlantic. Line contours are at 5- $S_v$  intervals, and zero contours are thickened.

result (i.e., relatively small RMSE) is found in the TS deeper run (Fig. 5d). When we assimilate the 3D temperature without salinity, by contrast, the AMOC structure is completely destroyed by unrealistic deep convection, especially in low-latitude regions (Fig. 5e), even though much of the decadal variability of the AMOC is related to temperature variations. In particular by ignoring salinity anomalies, the density structure of the ocean is severely altered and isopycnals will shift dramatically, which will also lead to unrealistic adjustment processes. Comparison of surface

assimilation with (Fig. 5b; lower RMSE) or without salinity (Fig. 5a; higher RMSE) confirms the importance of salinity and freshwater fluxes in properly initializing the AMOC.

To further evaluate the role of surface versus subsurface temperature and salinity assimilation in reproducing the trajectory of the AMOC in the control simulation, we calculate the empirical orthogonal function (EOF) of the annual mean streamfunction anomalies in the control run. The first leading EOF mode represents changes in deep convection strength around

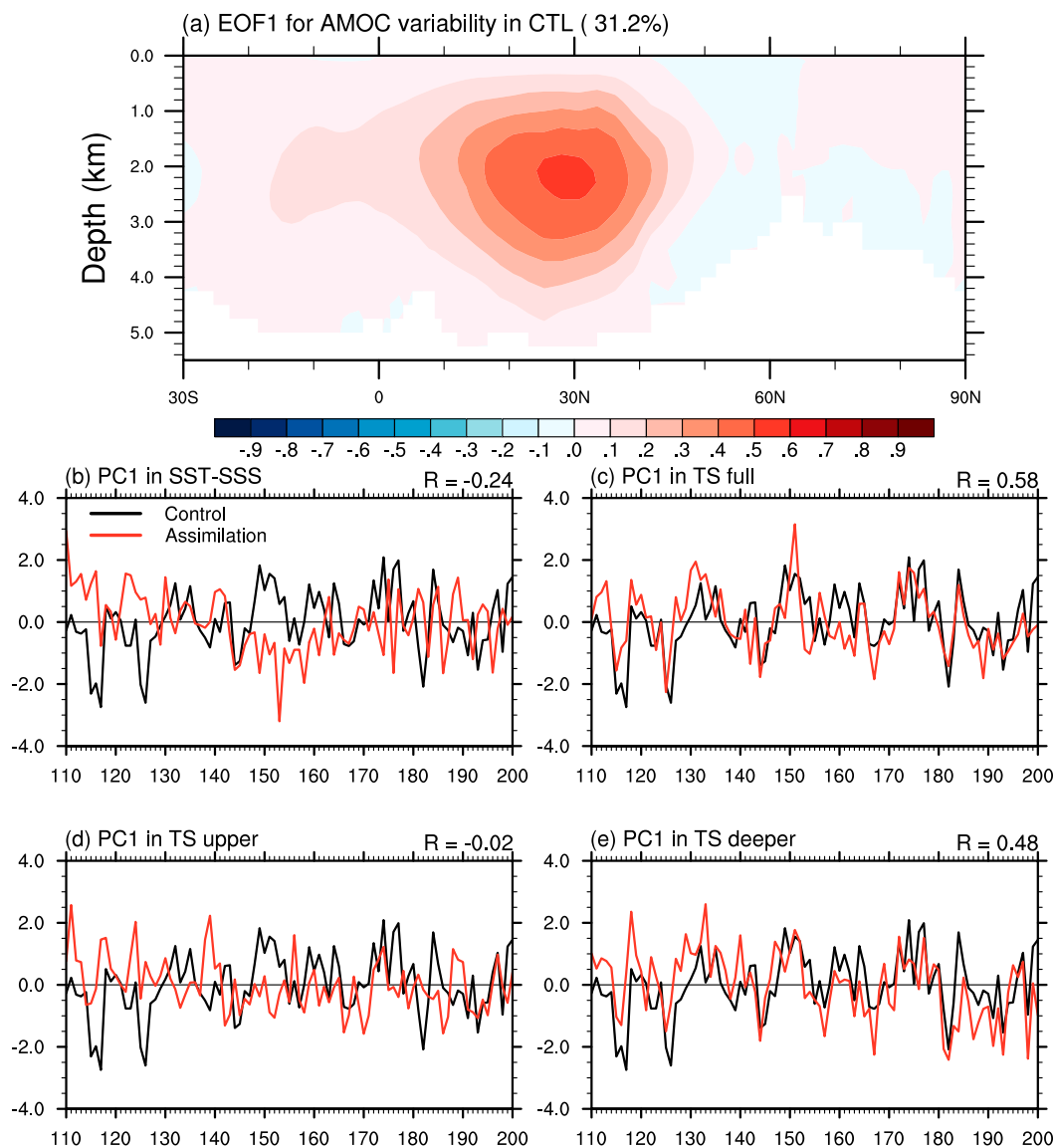


FIG. 6. (a) Regression map associated with the first leading EOF mode of annual mean AMOC variability in the control simulation. Time series projected onto this eigenvector of first EOF mode in the control (black lines) and assimilation experiments (red lines) for the (b) SST-SSS, (c) TS full, (d) TS upper, and (e) TS deeper runs. Correlation coefficients of time series in the assimilation experiments with the control are denoted in the upper-right corner of (b)–(e).

45°N and explains 31.2% of the streamfunction variance (Fig. 6a). Principal components projected onto this pattern show that the decadal AMOC variability is simulated well in the TS full and TS deeper runs (correlation coefficients  $R$  are 0.58 and 0.48, respectively; Figs. 6c,e). In contrast, the principal component in the control run is poorly correlated with those in the SST-SSS or TS upper runs ( $R = -0.24$  and  $-0.02$ , respectively; Figs. 6b,d). These partial assimilation experiments in the perfect model framework clearly show

that deeper ocean density assimilation constrains decadal AMOC variability.

#### d. Pacific decadal variability

The North Pacific Ocean is home to considerable amount of decadal climate variability, which can be partly attributed to the slow Rossby wave adjustment processes of the oceanic gyre circulation (Trenberth and Hurrell 1994; Mantua et al. 1997; Di Lorenzo et al. 2008; Meehl et al. 2009). By applying EOF analysis to the

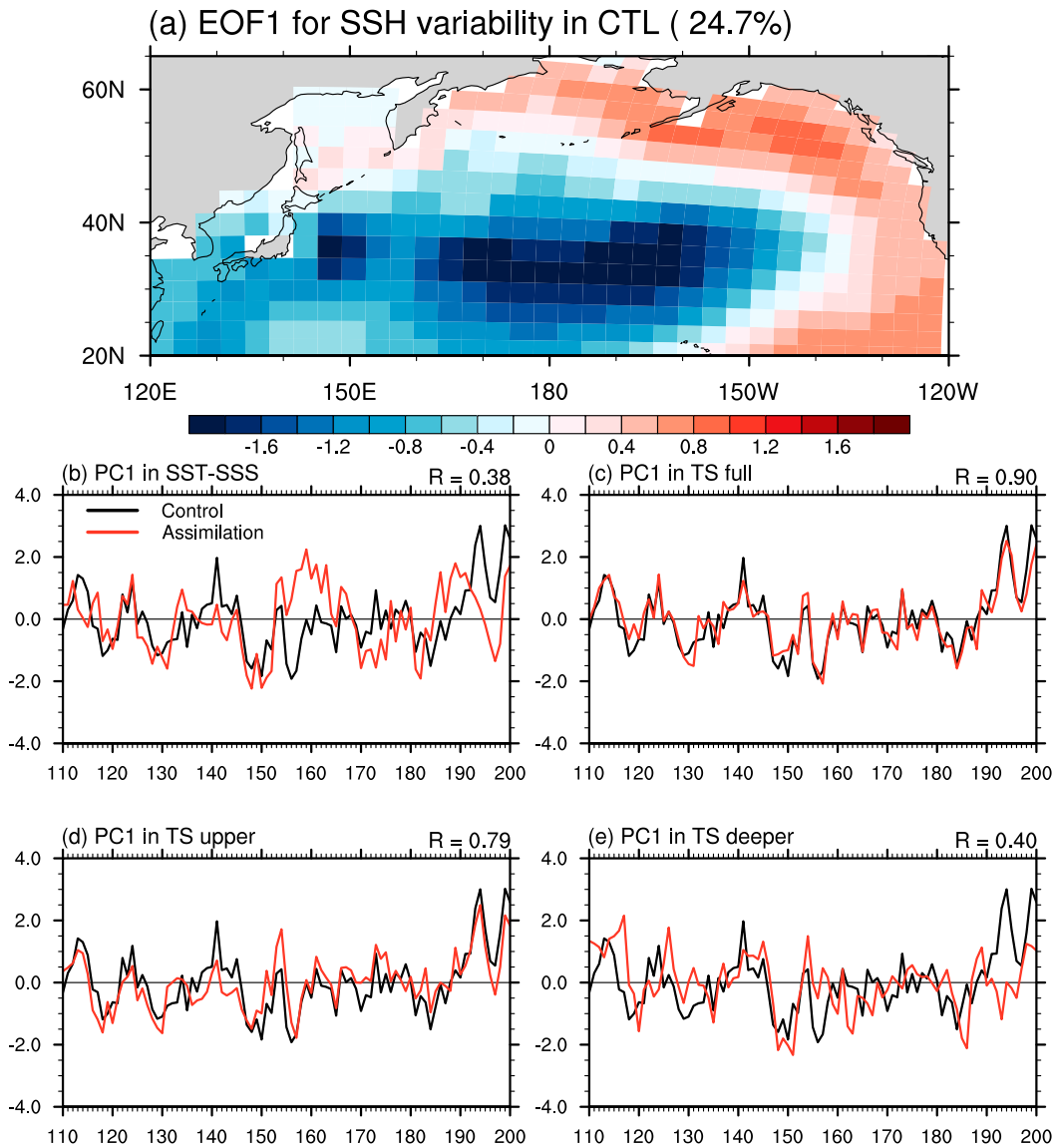


FIG. 7. (a) Regression map associated with the first leading EOF mode of North Pacific SSH variability in the control simulation. Time series projected onto this eigenvector of first EOF mode in the control (black lines) and assimilation experiments (red lines) for the (b) SST-SSS, (c) TS full, (d) TS upper, and (e) TS deeper runs. Correlation coefficients of time series in the assimilation experiments with the control are denoted in the upper-right corner of (b)–(e).

North Pacific SSH variability in the control run (Figs. 7 and 8), we identify two leading modes, which have patterns similar to the observed Pacific decadal oscillation (PDO; Mantua et al. 1997) and the North Pacific Gyre Oscillation (NPGO; Di Lorenzo et al. 2008). These two modes explain together 43.2% of the total annual mean SSH variance. To evaluate the impact of ocean assimilation on the simulation of Pacific decadal variability, we projected the SSH anomalies in the TS full run onto these EOF patterns. Consistent with the RMSE skill score of SSH anomalies in Fig. 4, the TS full

run simulates temporal variations of these two leading modes (Figs. 7c and 8c) well. Correlation coefficients of the first and second EOF modes between CTL and TS full runs attain values of  $R = 0.90$  and  $0.96$ , respectively.

To further evaluate impacts of surface, upper, and deeper ocean assimilation of temperature and salinity on the generation of Pacific decadal variability, we compared partial assimilation experiments in SST-SSS, TS upper, and TS deeper runs. Although the SST-SSS run captures some decadal variations in these two leading modes, correlations of the first and second

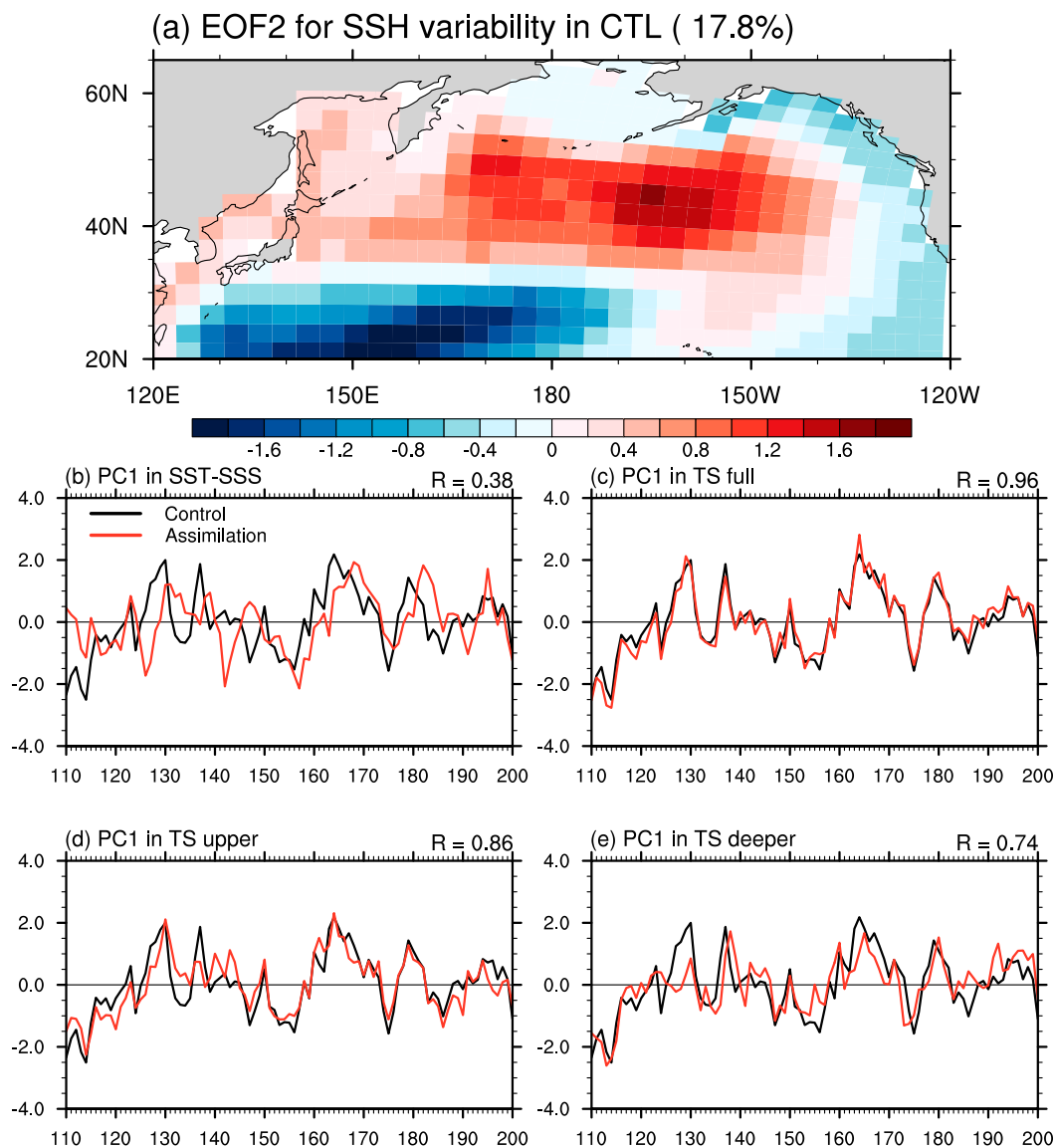


FIG. 8. (a) Regression map associated with the second leading EOF mode of North Pacific SSH variability in the control simulation. Time series projected onto this eigenvector of first EOF mode in the control (black lines) and assimilation experiments (red lines) for the (b) SST-SSS, (c) TS full, (d) TS upper, and (e) TS deeper runs. Correlation coefficients of time series in the assimilation experiments with the control are denoted in the upper-right corner of (b)–(e).

principal components with the control run are markedly diminished ( $R = 0.38$  and  $0.38$ , respectively; Figs. 7b and 8b) compared to the TS full run. Better performance is found when assimilating upper temperature and salinity in addition to the surface observations ( $R = 0.79$  and  $0.86$ , respectively; Figs. 7d and 8d), supporting previous findings that upper ocean thermodynamics strongly influence the Pacific decadal variability (Schneider et al. 1999; Alexander et al. 1999; Xie et al. 2000; Schneider and Cornuelle 2005). In contrast, there are notable differences between the two leading EOF modes for the deeper ocean

assimilation. The first principal component in the TS deeper run is weakly correlated with that in the control run ( $R = 0.40$  in Fig. 7e), which corresponds to that in the SST-SST run. On the other hand, the second principal component exhibits a high correlation between the control and TS deeper runs ( $R = 0.74$  in Fig. 8e), which is almost comparable to the TS upper run ( $R = 0.86$  in Fig. 8d). Results of the partial assimilation experiment indicate that accurate initialization for Pacific decadal variability requires deeper ocean dynamics as well as upper ocean thermodynamical processes. This finding is

also consistent with previous observational and modeling studies that suggested a physical interpretation of the connection between North Pacific SSH variability and equivalent barotropic Rossby waves (Qiu 2003; Qiu and Chen 2005; Taguchi and Schneider 2014).

#### e. Summary of the perfect model experiments

To identify the impacts of surface and subsurface temperature and salinity conditions on climatological biases and initial errors in SST, SSH, and AMOC, we examined six partial assimilation experiments under the perfect model framework (Table 1). By assimilating the 3D temperature and salinity from surface to 3000 m into the model (i.e., TS full run), we minimized climatological biases and initial errors in SST, SSH, and AMOC. Whereas the upper ocean data assimilation (0–300 m) is crucial for initializing SST and SSH conditions in most of the global ocean, the deeper ocean data assimilation (300–3000 m) also plays a role for initializing SSH variability, especially in higher-latitude regions and related to AMOC variability. From the perspective of which ocean observations are available, SST is the most prevalent variable in terms of spatial coverage, number of samplings, and data period. However, our perfect model experiments suggest that the SST-only assimilation is not good enough to initialize ocean conditions for skillful decadal climate predictions. Furthermore, salinity assimilation is also required to stabilize the model simulation. When we assimilated 3D temperature without salinity (i.e., the T full run), the simulated SSH and AMOC conditions were completely destroyed. Therefore, our results in perfect model experiment indicate that the 3D ocean temperature and salinity observations are required to initialize the climate model for skillful decadal climate predictions.

## 4. Realistic decadal climate prediction experiments using CESM

For the perfect modeling framework, we demonstrated that combined 3D temperature and salinity assimilation provides more realistic initial conditions for ensemble predictions. In reality, we must often deal with considerable biases between the observed climatology and the model simulated climatology (see section 1). Assimilating full observed fields into a global climate model and then launching predictions leads to the unfortunate situation that the model has to adjust to both a new climatological state and the actual initial errors. Attempts to overcome this issue using observed anomaly assimilation (relative to the simulated model climatology) have proven promising (Doblas-Reyes et al. 2011; Chikamoto et al. 2012; Magnusson et al. 2013).

However, anomaly assimilation is not sufficient to reduce artificial climate drift during the actual decadal prediction (Chikamoto et al. 2012; Smith et al. 2013), because additional uncertainties exist with respect to the model sensitivity to natural and anthropogenic radiative forcings (Andrews et al. 2012). In this section, we describe a new initialization approach for decadal climate prediction, which tries to address this issue by assimilating climate anomalies in the context of a gradually changing and externally forced climate background field.

#### a. The externally forced component and model drift

Here we estimate the externally forced component in the following way: We first decompose the observed 3D temperature and salinity fields  $\mathbf{X}^o$  into a mean climatology  $\mathbf{X}_{\text{clm}}^o$  (reference period: 1971–2000), internally generated variability  $\mathbf{X}_{\text{int}}^o$ , and an externally forced component  $\mathbf{X}_{\text{ext}}^o$ . Thus, we have  $\mathbf{X}^o = \mathbf{X}_{\text{clm}}^o + \mathbf{X}_{\text{int}}^o + \mathbf{X}_{\text{ext}}^o$ . Instead of using raw observational data, we have decided to use an existing ocean data assimilation product as an estimate of the gridded observational state. We have chosen the ORA-S4 reanalysis conducted by the ECMWF (Balmaseda et al. 2013). To estimate  $\mathbf{X}_{\text{ext}}^o$ , we perform a singular value decomposition (SVD) of the ORA-S4 monthly mean temperature and salinity anomalies (i.e.,  $\mathbf{X}_{\text{ann}}^o = \mathbf{X}^o - \mathbf{X}_{\text{clm}}^o = \mathbf{X}_{\text{int}}^o + \mathbf{X}_{\text{ext}}^o$ ) and the ensemble mean anomalies from the uninitialized run (i.e.,  $\mathbf{X}_{\text{ann}}^m = \mathbf{X}^m - \mathbf{X}_{\text{clm}}^m = \mathbf{X}_{\text{int}}^m + \mathbf{X}_{\text{ext}}^m$ ) at each depth for the 1960–2005 period. The two fields of the leading SVD mode are then used as an estimate for  $\mathbf{X}_{\text{ext}}^o$  and  $\mathbf{X}_{\text{ext}}^m$  because the global warming trend associated with the external forcings is the dominant variation in the ensemble mean of the uninitialized run (Mochizuki et al. 2012; Chikamoto et al. 2013). Our estimation assumes that the internally generated variability is negligible for the multimember ensemble mean of the uninitialized run (i.e.,  $\mathbf{X}_{\text{int}}^m \approx 0$ ) and that covariance between internally generated variability and the externally forced component is much smaller than the covariance of the externally forced components between observation and simulation of the uninitialized run (i.e.,  $[\mathbf{X}_{\text{int}}^o, \mathbf{X}_{\text{ext}}^m] \ll [\mathbf{X}_{\text{ext}}^o, \mathbf{X}_{\text{ext}}^m]$  where  $[a, b]$  denotes symbolically the matrix norm of the covariance tensor between  $a$  and  $b$ ).

Figure 9 shows the first SVD modes of monthly SST and SSS anomalies between the reanalysis (ORA-S4) and the 10-member ensemble mean of the uninitialized run, which explains 92.2% and 56.5% of total covariances, respectively. We observe a dominant warming trend almost globally in both reanalysis and the model, although SST cooling events briefly occur after major volcanic eruptions (particularly the 1963/64 Agung, 1982 El Chichon, and 1991 Mount Pinatubo events) (Figs. 9e,f). Temporal evolutions of the first

## First SVD mode

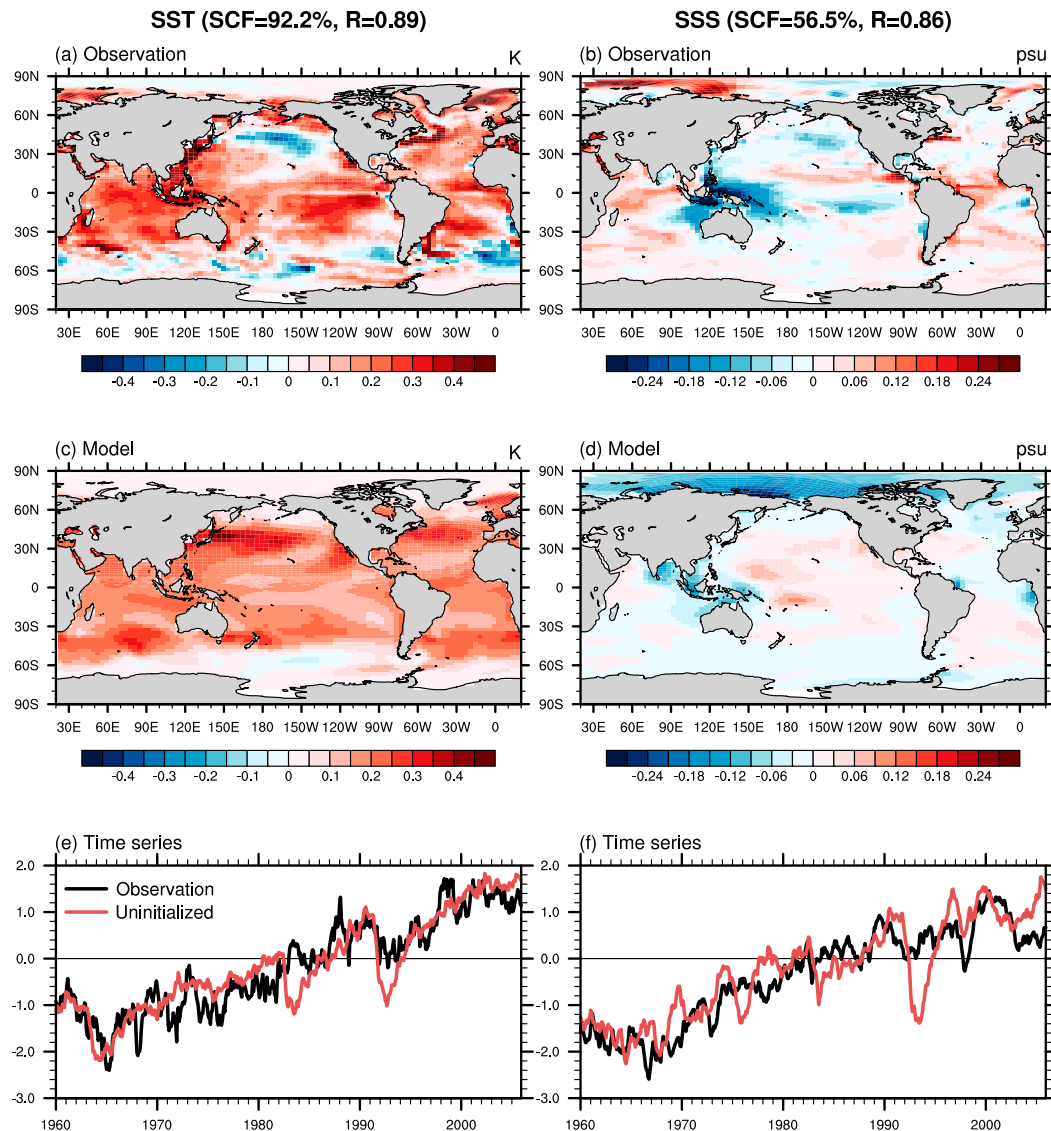


FIG. 9. First leading SVD modes of (left) SST and (right) SSS anomalies between (a),(b) observations and (c),(d) the ensemble mean of the uninitialized run. The squared covariance fractions (SCF) explained by the first leading SVD modes of SST and SSS and their temporal correlation coefficients  $R$  between the expansion coefficients for the observation and the model are indicated above (a) and (b). Anomalous patterns are estimated from regression maps with the expansion coefficients. Color units are denoted in the upper-right corner. (bottom) Time coefficients for the first SVD modes of (e) SST and (f) SSS in the observation (black) and the ensemble mean of uninitialized runs (red).

leading SVD modes for SSS also show positive trends in both reanalysis and the model. This first SSS SVD mode also captures a freshening in the tropical western Pacific as reported by observational analysis of a 50-yr record (Durack and Wijffels 2010; Pierce et al. 2012). These results suggest that our approach is reasonable to estimate the observed and simulated externally forced components of climate variability. Results for other

depth levels are not shown here but the first SVD modes of temperature and salinity at each depth between the reanalysis and the ensemble mean of the historical run explain more than 80% of total covariances in deeper ocean ( $<1$  km).

In spite of overall good correspondence in estimated trend amplitudes between the uninitialized ensemble mean model simulation and observations represented

## Magnitude of drift (drift/rmse)

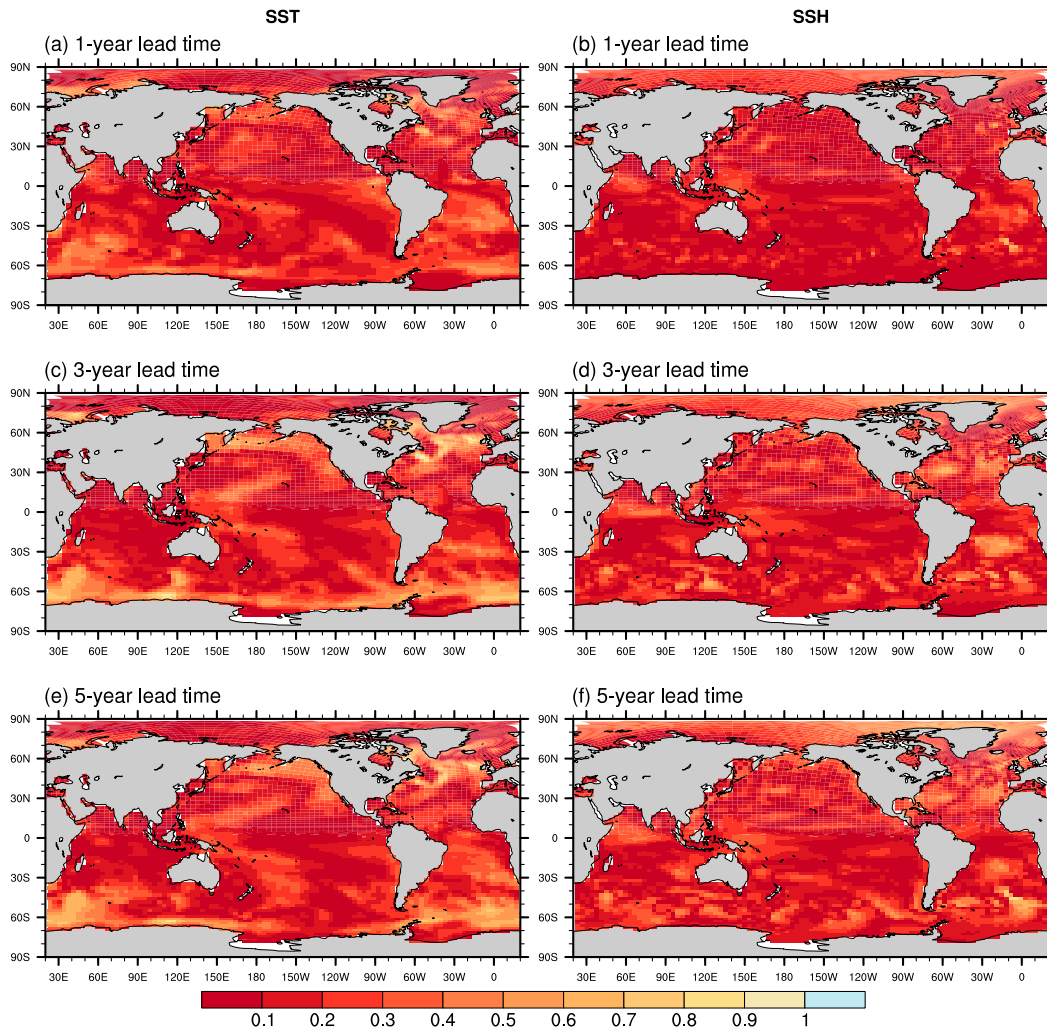


FIG. 10. Ratios between the magnitude of model climate drift during prediction and the RMSE of the uninitialized run relative to the ASSM run in (left) SST and (right) SSH for (a),(b) 1-, (c),(d) 3-, and (e),(f) 5-yr lead time.

here by the ORA-S4 state estimate, we find also important discrepancies. In the North Pacific, particularly, the first SVD mode of SST anomalies shows cooling in the reanalysis but large warming in the model (Figs. 9a,c). In addition to this SST discrepancy, SSS anomalies in the North Pacific and large areas of the Arctic show opposite signs between the reanalysis and the model (Figs. 9b,d). Such externally forced differences due to  $\mathbf{X}_{\text{ext}}^o - \mathbf{X}_{\text{ext}}^m$  may still represent a potential source of prediction errors. To mitigate this effect, our approach removes these biases estimated from the first SVD modes of anomalous temperature and salinity fields at each depth between the reanalysis and the ensemble mean of the uninitialized runs. From the difference between  $\mathbf{X}_{\text{ann}}^o$  and our SVD-based analysis of  $\mathbf{X}_{\text{ext}}^o$ , we

then calculate  $\mathbf{X}_{\text{int}}^o$ . The final data assimilated into the CESM ocean model using the incremental analysis update technique described above are calculated from  $\mathbf{X}_{\text{int}}^o + \mathbf{X}_{\text{ext}}^m + \mathbf{X}_{\text{clm}}^m$ .

Our decadal forecasts are then initialized once a year on 1 January, from the bias-adjusted trajectory for the period 1960–2014 (see section 2). Figure 10 shows SST and SSH ratios of the model drifts in the initialized run to the RMSE between the assimilation and the uninitialized runs for various lead times. Small values (indicated by darker red values) represent a small forecast drift. Because of our bias adjusted assimilation method, the RMSE should be comparable in magnitude to the standard deviation of observed internal variability. The model drift is estimated from the initialized run on

## ACC skill: SAT

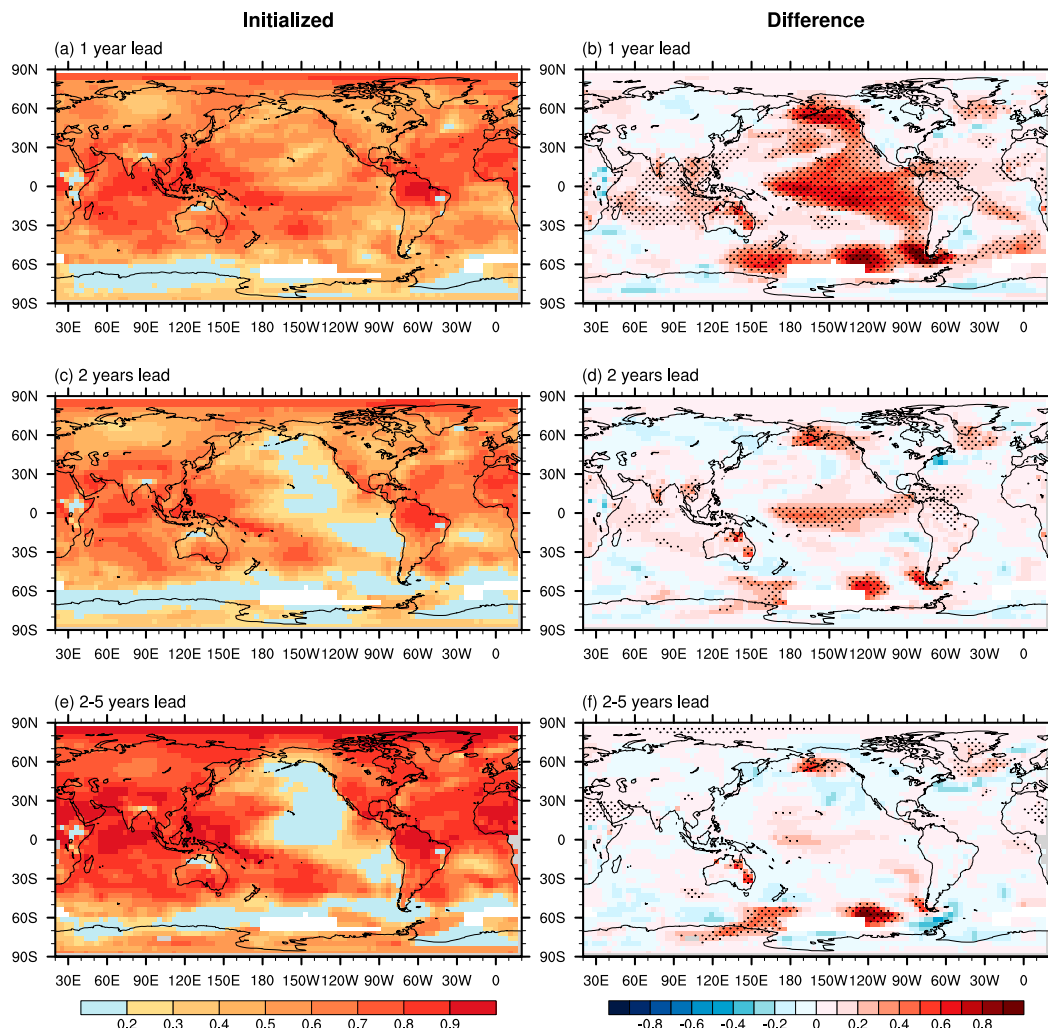


FIG. 11. (left) Predictive skills of surface air temperature anomalies and (right) the impact of initialization for (a),(b) 1-, (c),(d) 2-, and (e),(f) 2–5-yr lead time. Predictive skills are measured by anomaly correlation coefficients of the initialized and the uninitialized runs compared to the observation (NASA GISTEMP v3; Hansen et al. 2010). Correlation coefficients of 0.18 correspond to the statistical significance at 90% levels with 50 degrees of freedom on the basis of a one-sided Student's  $t$  test. The impact of initialization is estimated by the difference in predictive skills between the initialized and the uninitialized runs. Dotted areas indicate the statistically significant ( $\geq 90\%$ ) difference in the anomaly correlation skill between the initialized and the uninitialized runs, as measured by Fisher's  $z$  score.

the basis of Chikamoto et al. (2012), that is, by calculating the sum of anomaly differences between the assimilation and initialized runs (ensemble mean) for all hindcast cases. The ratio of SST and SSH drifts in Fig. 10 is less than one everywhere at all lead times, indicating that drift in our system for various lead times is less than the magnitude of observed internal variability.

#### b. Predictive skills in the drift-free prediction system

To assess the predictive skills in our newly developed drift-free prediction system, we evaluate the anomaly correlation coefficient (ACC) of surface air temperature

(SAT), ocean heat content from the surface to 300 m (OHC300), and SSH between the observations and the ensemble mean of initialized/uninitialized runs (Figs. 11–13). Observation-based datasets for SAT, OHC300, and SSH are obtained from NASA GISTEMP version 3 (Hansen et al. 2010), an ocean objective analysis named ProjD version 7.2 (Ishii et al. 2017), and ORA-S4 (Balmaseda et al. 2013), respectively. Anomalies are defined as deviation from the climatological mean for 1960–2014 in both observations and the uninitialized run, whereas for the initialized run, the climatological mean is calculated based on a function of lead time. The impact of



## ACC skill: OHC300

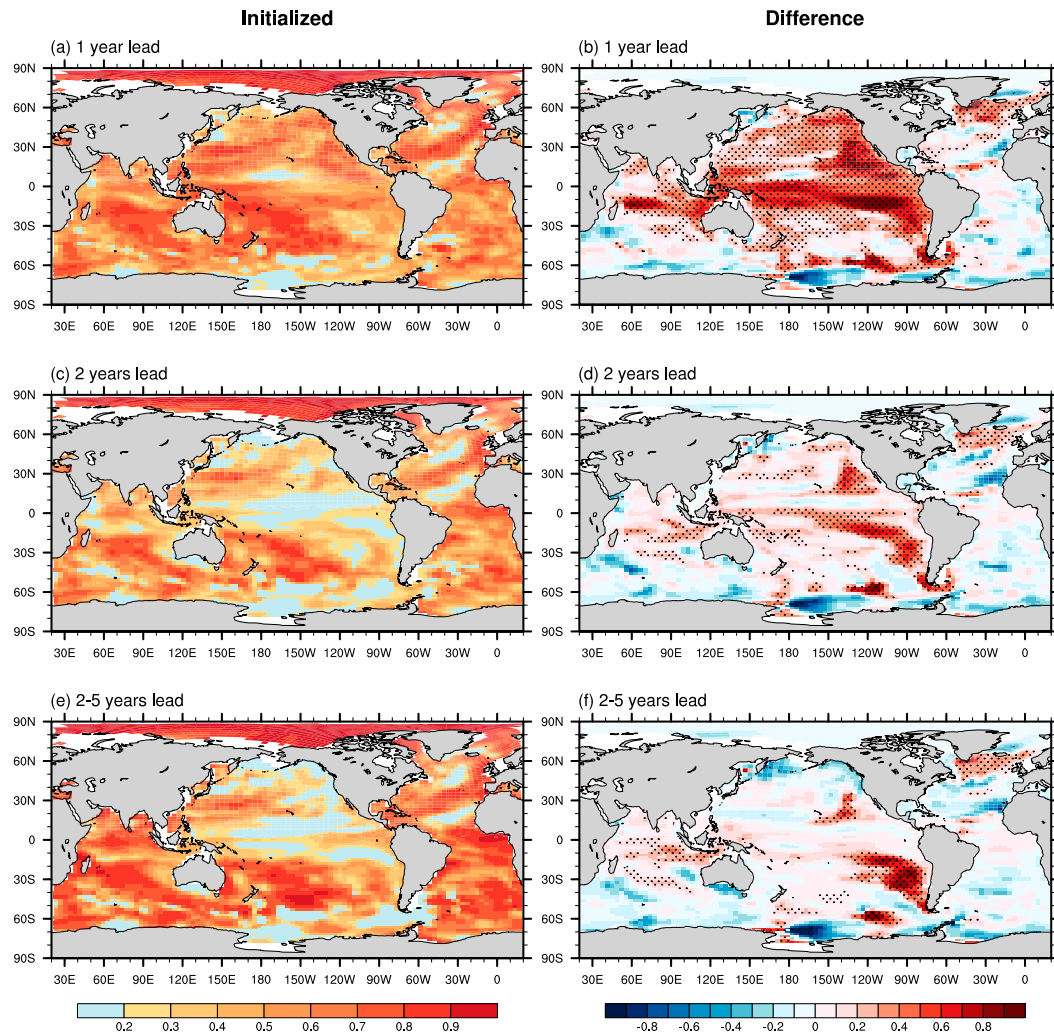


FIG. 12. (left) Predictive skills of ocean heat content averaged from the surface to 300 m and (right) the impact of initialization for (a),(b) 1-, (c),(d) 2-, and (e),(f) 2–5-yr lead times. Predictive skills are measured by anomaly correlation coefficients of the initialized and the uninitialized runs compared to the observation (ProjDv7.2; Ishii et al. 2017). Correlation coefficients of 0.18 correspond to the statistical significance at 90% levels with 50 degrees of freedom on the basis of a one-sided Student's  $t$  test. The impact of initialization is estimated by the difference in predictive skills between the initialized and the uninitialized runs. Dotted areas indicate the statistically significant ( $\geq 90\%$ ) difference in the anomaly correlation skill between the initialized and the uninitialized runs, as measured by Fisher's  $z$  score.

initialization is estimated by assessing the difference of ACC between initialized and uninitialized runs.

Our initialized run shows significant predictive skills of SAT anomalies across the entire globe at the first year lead time, except for the Antarctic region (Fig. 11a), and becomes gradually lower at the second year lead time (Fig. 11c). Whereas this significant predictive skill of SAT is mostly contributed from the long-term warming trend associated with the externally forced component, we still find the impact of ocean initialization for skill enhancement particularly over the ocean, such as in the

tropical central and eastern Pacific, the higher-latitude North and South Pacific, the North Atlantic around the Labrador Sea, and the Indian Ocean (Figs. 11b,d). Because of the large contribution from the long-term warming trend, the ACC predictive skills in the initialized run become much higher at 2–5-yr lead time (Figs. 11e,f) compared to the annual mean predictions. These characteristics of SAT predictive skills in our initialized run are consistent with results from decadal prediction experiments from phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Boer et al. 2016).

## ACC skill: SSH (detrended)

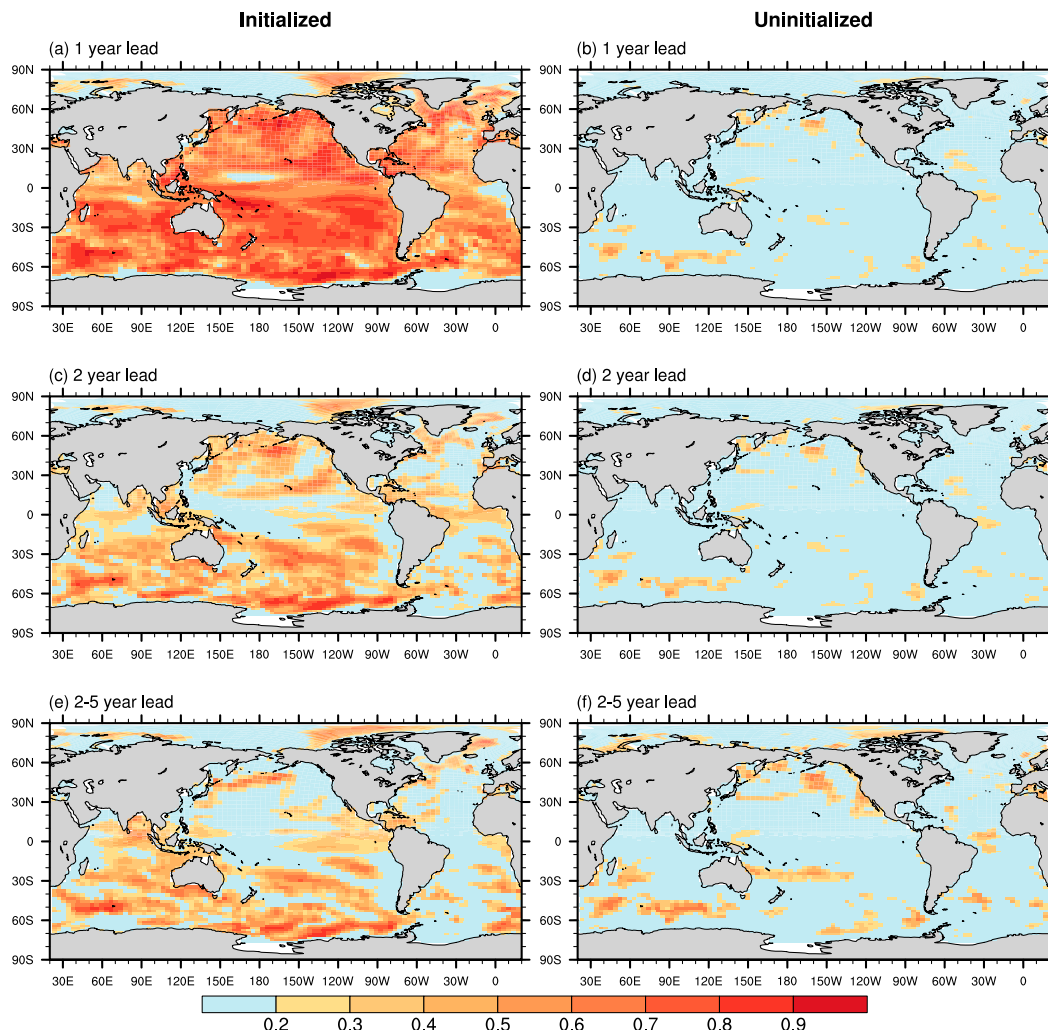


FIG. 13. Predictive skills of SSH anomalies in the (left) initialized and (right) uninitialized runs for (a),(b) 1-, (c),(d) 2-, and (e),(f) 2–5-yr lead times. Anomalies are defined from the deviation of the climatological mean for 1960–2015, and the estimated trends by the least squares quadratic are removed in each grid point. Predictive skills are measured by anomaly correlation coefficients of the initialized and the uninitialized runs compared to the reanalysis (ORA-S4; Balmaseda et al. 2013). Correlation coefficients of 0.18 correspond to the statistical significance at 90% levels with 50 degrees of freedom on the basis of a one-sided Student's  $t$  test.

Our ocean assimilation approach also exhibits significant predictive skills in the OHC300 (Fig. 12). Consistent with the SAT predictive skills, the initialized run shows high ACC predictive skills of OHC300 in the global ocean at the first year lead time (Fig. 12a). For these significant predictive skills, initialization enhances the skill particularly in the Pacific Ocean, the North Atlantic Subpolar Gyre region, and the Indian Ocean (Fig. 12b). The ACC predictive skills and the initialization impacts at 2- and 2–5-yr lead times also show similar patterns to those at 1-yr lead time, although the correlations have smaller amplitude (right panels in Fig. 12).

Whereas predictive skill of the OHC300 in the equatorial Pacific diminishes greatly beyond 1-yr lead time, primarily due to the predictable limit of El Niño Southern Oscillation (Timmermann et al. 2018), the slow ocean dynamics may be prolonging predictive skills in the mid- and high-latitude regions. In the North Atlantic, especially, the high predictive skills of OHC300 around the Labrador Sea may be related to predictability of AMOC as described in the next subsection.

Figure 13 shows the ACC predictive skills of SSH in the initialized and the uninitialized runs compared to the detrended SSH anomalies in the ORA-S4. Because of

## EOF1 for annual mean AMOC

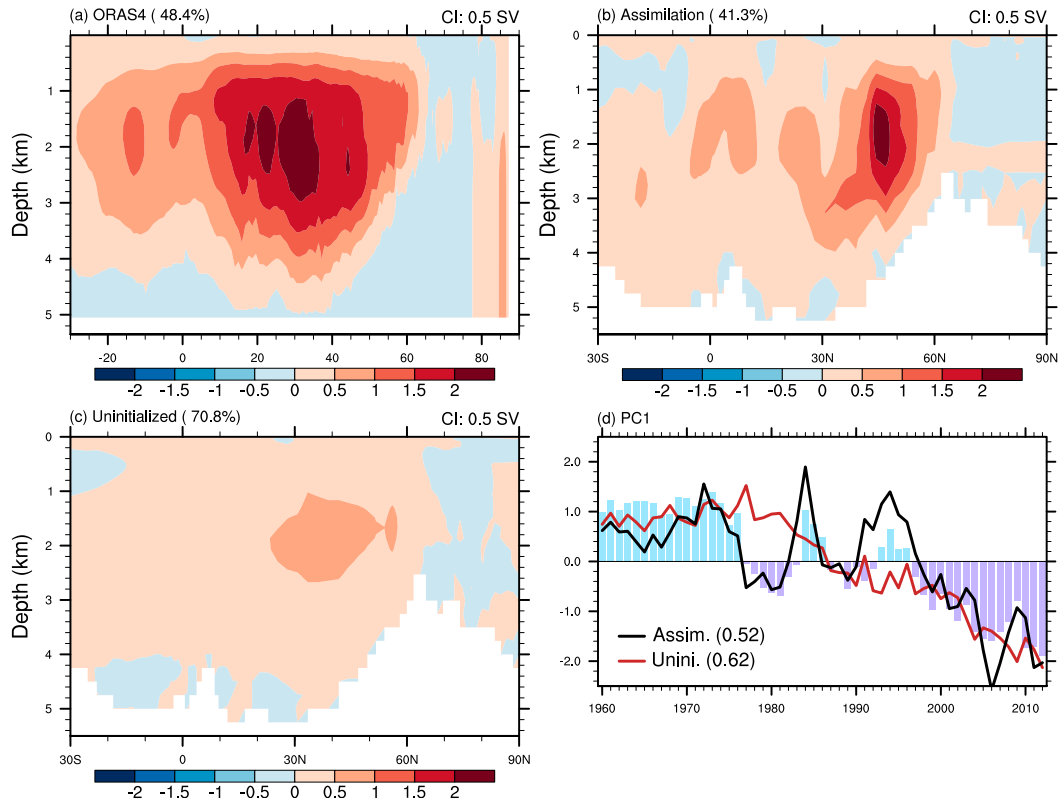


FIG. 14. Regression maps associated with the first leading EOF mode of annual mean streamfunction anomalies in the Atlantic for (a) reanalysis ORA-S4, (b) assimilation, and (c) uninitialized run. (d) Principal components in observation (color bars), assimilation (black line), and uninitialized run (red line). Parentheses in (d) denote RMSE of time series compared to the observation.

the Boussinesq approximation in the POP2, our model cannot account for the observed long-term SSH trend (Jordà and Gomis 2013). Therefore, the SSH predictive skills estimated here relate to the internally generated climate predictability rather than the combination of the internally generated variability and the externally forced component. As a result, SSH predictive skills in the uninitialized run are almost negligible (right panels in Fig. 13), whereas the initialized run exhibits significant predictive skills globally at 1-yr lead time (Fig. 13a). Skills are persistent at 2- and 2–5-yr lead times in the North Pacific, the Southern Ocean, and parts of the Atlantic Ocean (Figs. 13c,e). The longer predictive skills in the North Atlantic and North Pacific imply that the AMOC and the North Pacific decadal variability, respectively, are predictable in our system.

### c. The Atlantic meridional overturning circulation

To quantify the forced and internally generated variability of the AMOC, we calculated the leading EOF of the annual mean Atlantic meridional streamfunction

(Fig. 14) using the reanalysis (ORA-S4) and the assimilated and uninitialized CESM simulations. The first mode for the reanalysis and the corresponding principal component (Fig. 14a, and colored bars in Fig. 14d) document a slowing down of the large-scale AMOC structure over the observational period. The assimilation run captures these temporal variations well. However, the structure of the corresponding overturning circulation changes is more confined to higher latitudes of the North Atlantic. Yet, we see that the 3D temperature and salinity assimilation is successful in reproducing key feature of the AMOC in the ORA-S4. Interestingly, we find also a weakening trend of the AMOC in the uninitialized run (Figs. 14c,d), thus suggesting a potential, albeit weak, contribution from external forcings to the dynamics of the AMOC during the past 60 years. External forcing of the AMOC has also been found in previous studies (Cheng et al. 2013).

To quantify the potential multiyear predictability of AMOC variations, we project the predicted streamfunction anomalies simulated by the initialized and

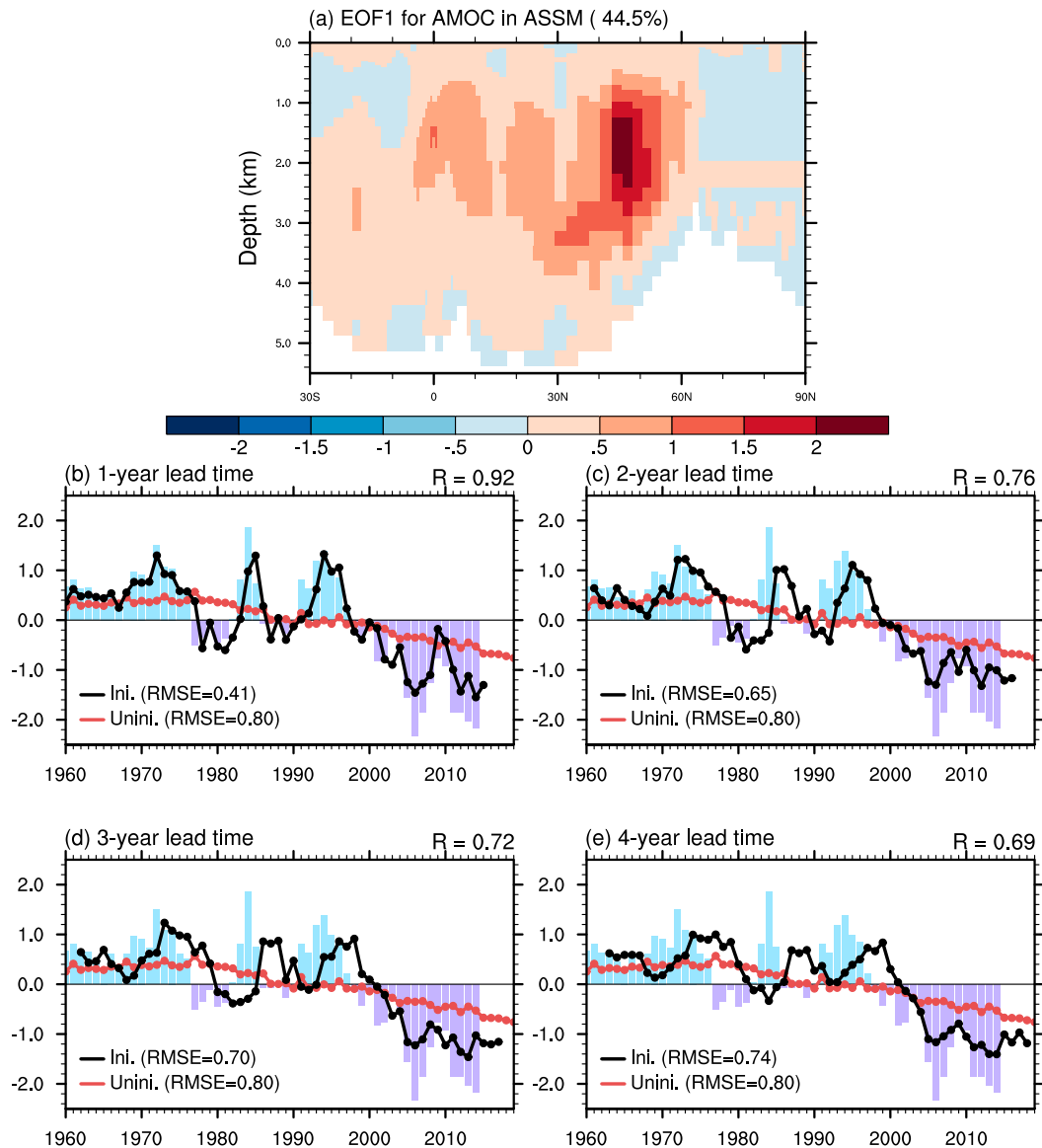


FIG. 15. (a) Regression map associated with the first EOF mode for AMOC variability in the assimilation run and time coefficients projected onto that pattern in the assimilation (color), uninitialized (red), and initialized runs (black) for (b) 1-, (c) 2-, (d) 3-, (e) 4-yr lead times. Correlation coefficient between the assimilation and initialized runs is denoted at the upper-right corner of (b)–(e). Parentheses in legends are RMSE of time coefficients with the assimilation run (see Table 2).

uninitialized runs onto the EOF1 pattern in the assimilation run (Fig. 15). Even though the uninitialized run shows the weakening AMOC trend (red lines in Figs. 15b–e), its contribution to the AMOC predictability is quite limited, yielding only 20% improvement (RMSE = 0.80). The initialized run outperforms the predictive skills in the uninitialized run considerably during the initial 4-yr lead time (black lines in Figs. 15b–e), and we see a high degree of predictability and persistence in the AMOC index used here. In particular, the

initialized run captures the pronounced AMOC weakening since 2005, even at a 4-yr lead time, which is consistent with Keenlyside et al. (2008). Our multiyear AMOC predictability originates from the slow baroclinic adjustment process of the AMOC in response to wind and density forcings. Some previous studies have suggested that external forcings related to aerosol emissions and volcanic eruptions can further extend the long-term predictability of the AMOC (Booth et al. 2012; Swingedouw et al. 2013). Regardless of the underlying

TABLE 2. Potential predictive skills of the first EOF mode in the AMOC, the first EOF mode in the North Pacific SSH (PDO), and the second EOF mode in the North Pacific SSH (NPGO). The predictive skills are measured by the ACC and RMSE between the assimilation and the uninitialized/initialized runs.

		Uninitialized run	Initialized run			
			1 yr	2 yr	3 yr	4 yr
AMOC	ACC	0.71	0.92	0.76	0.72	0.69
	RMSE	0.80	0.41	0.65	0.70	0.74
PDO	ACC	0.56	0.93	0.81	0.62	0.56
	RMSE	0.85	0.37	0.64	0.81	0.85
NPGO	ACC	0.31	0.95	0.79	0.59	0.35
	RMSE	0.95	0.31	0.61	0.77	0.91

mechanisms, our CESM decadal prediction runs indicate that the 3D ocean temperature and salinity initialization is important to skillfully predict AMOC variability.

Consistent with the potential AMOC predictability, our initialized run also shows comparable predictive skills compared to the observational estimates for AMOC index and OHC300 variability in the North Atlantic Subpolar Gyre (SPG) (Tables 2 and 3). Here, the AMOC index and SPG OHC300 are defined as the maximum streamfunction at 45°N ( $Sv$ ;  $1 Sv \equiv 10^6 m^3 s^{-1}$ ) and the area average of OHC300 (45°–65°N 60°–10°W), respectively. Note that the strengthened (weakened) AMOC index accompanies the upper ocean temperature cooling (warming) in the North Atlantic, which is simulated well in the assimilation run (Figs. 16a,b). For these observation-based AMOC and SPG OHC300 indices, our initialized run outperforms the ACC predictive skills in the uninitialized run for 3-yr lead time (Figs. 16c–f and Table 3). These results also verify the importance of 3D ocean temperature and salinity initialization for AMOC predictability.

#### d. North Pacific decadal variability

To determine decadal predictability in the North Pacific Ocean, we first calculate the EOFs of annual mean SSH anomalies in the assimilation runs (Figs. 17 and 18). Consistent with the CTL run, the first and second leading EOF modes show PDO-like and NPGO-like SSH patterns, respectively (Figs. 17a and 18a). These two modes together explain 40.5% of total variance. In the first EOF mode, the uninitialized run captures a large phase change from 1991 to 1992 associated with the Mt. Pinatubo eruption (red lines in Figs. 17b–e), indicating the contribution of externally forced radiative changes to PDO-like SSH variability. However, the low predictive skills in the uninitialized run (15% and 5% skill improvements in the first and second EOF modes compared to the standard deviation, respectively) suggest that the long-term trend associated with the

TABLE 3. Predictive skills of OHC300 anomalies in the North Atlantic SPG region and AMOC index measured by the anomaly correlation coefficient between the observation-based estimates and the uninitialized/initialized run.

Variable	Uninitialized run	Initialized run			
		1 yr	2 yr	3 yr	4 yr
SPG OHC300	0.36	0.62	0.48	0.37	0.40
AMOC index	0.61	0.91	0.83	0.72	0.58

externally forced component plays only a minor role in the generation of Pacific decadal variability (red lines in Figs. 17 and 18).

In contrast to the uninitialized run, the initialized run shows higher predictive skill for the two leading EOF modes, even at several years lead time. We also identify a typical 2–4-yr damping time scale of the predicted anomalies in the initialized run toward the uninitialized prediction (black lines in Figs. 17b–e and 18b–e). This damping toward the uninitialized simulation is much larger for the first EOF mode compared to the second mode. In summary, the initialized run, compared to the uninitialized run, exhibits higher predictive skills for the first EOF mode for lead time up to 3 years (black lines in Figs. 17b–e) and for the second mode up to 4 years (black lines in Figs. 18b–e).

## 5. Discussion

We have demonstrated that our decadal climate prediction system exhibits multiyear predictive skills for variations of the AMOC and for North Pacific decadal SSH variability. Because our skill assessment is based on the comparison between initialized prediction runs and the assimilation run, we expect the true skill (comparison against observations) to be less. In fact, the ACC predictive skills of the AMOC index in the initialized run compared to the reanalysis are slightly less than the EOF-based predictive skills, except for 2-yr lead time (Tables 2 and 3). The reduction of skill when evaluating against observations would originate from additional discrepancies between observations and the initialized runs, such as differences in the externally forced component, resolution, spatiotemporal sampling, data coverage, period, and other inconsistencies. Therefore, our skill estimates indicate the current upper limit of predictive skills in our system, although the current predictive skills could be further enhanced by improving our prediction system (e.g., increasing the ensemble size and applying a more sophisticated assimilation scheme). Nevertheless, our results in the perfect model and the real hindcast experiments provide an encouraging perspective on the feasibility of decadal climate predictions

## Correlation map: OHC300 vs AMOC index

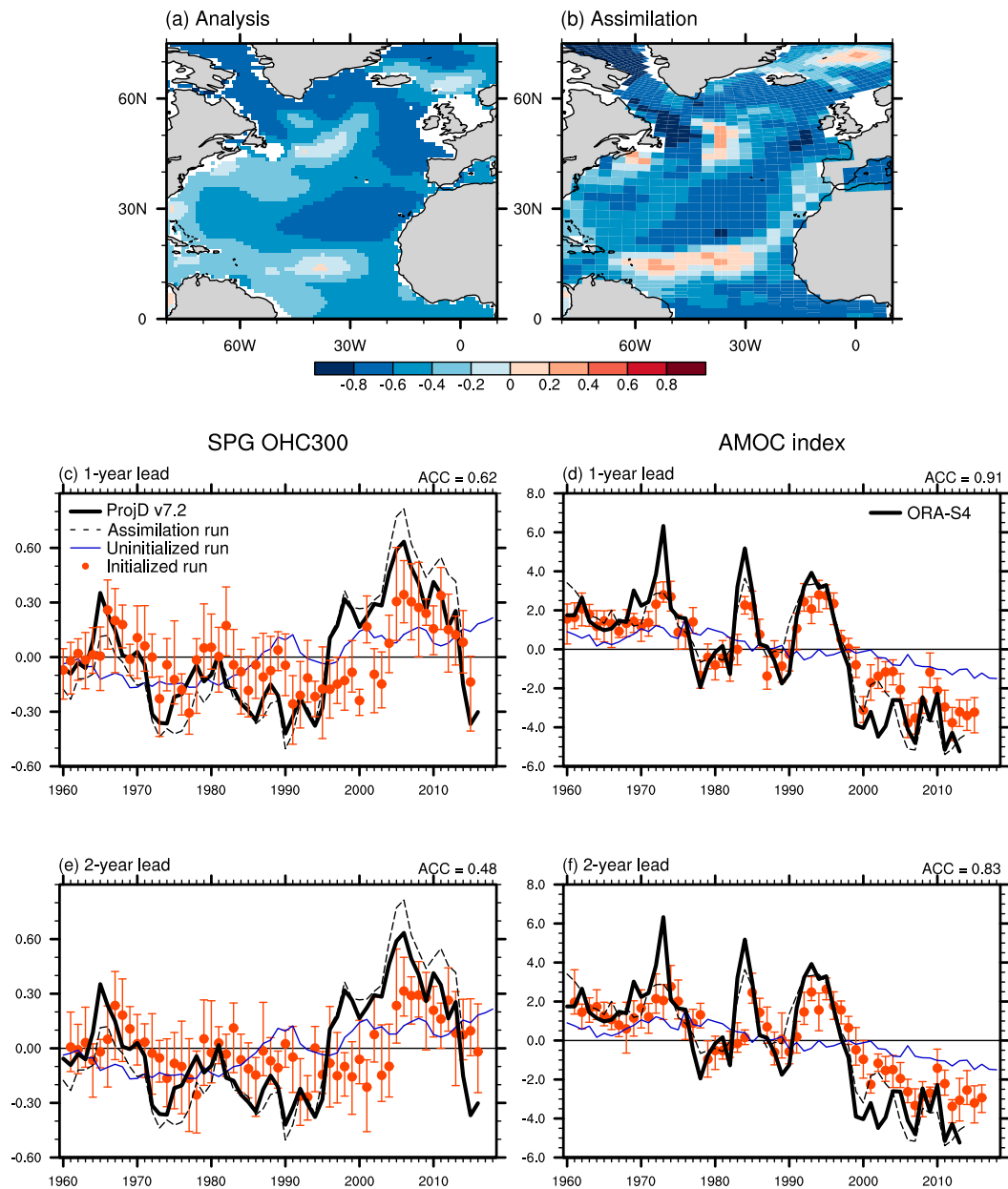


FIG. 16. Correlation maps of OHC300 with the AMOC index (defined as the maximum of the streamfunction at  $45^{\circ}\text{N}$ ) in (a) the observation-based estimate and (b) the assimilation run. The observation-based estimates of OHC300 and AMOC index are obtained from ProjDv7.2 and ORA-S4. Also shown are temporal evolutions and predictions for (c),(e) the SPG OHC300 ( $45^{\circ}$ – $65^{\circ}\text{N}$ ,  $60^{\circ}$ – $10^{\circ}\text{W}$ ) and (d),(f) the AMOC index for 1- and 2-yr lead times. Black solid and dashed lines denote the observation-based estimates and the assimilation runs, respectively. The blue line and red marks correspond to the uninitialized and initialized runs, respectively. Red error bars represent the minimum to maximum range of ensemble members in the initialized run. The ACC predictive skills are denoted at the upper-right corner in each panel (see also Table 3).

that are achievable even with lower-resolution climate models, such as the CESM configuration used here.

Recent studies confirm that AMOC variations are predictable on interannual to decadal time scales (Griffies

and Bryan 1997; Grötzner et al. 1999; Keenlyside et al. 2008; Pohlmann et al. 2009; Teng and Branstator 2011; Branstator and Teng 2010; van Oldenborgh et al. 2012; Teng et al. 2011; Robson et al. 2012; Chikamoto et al.

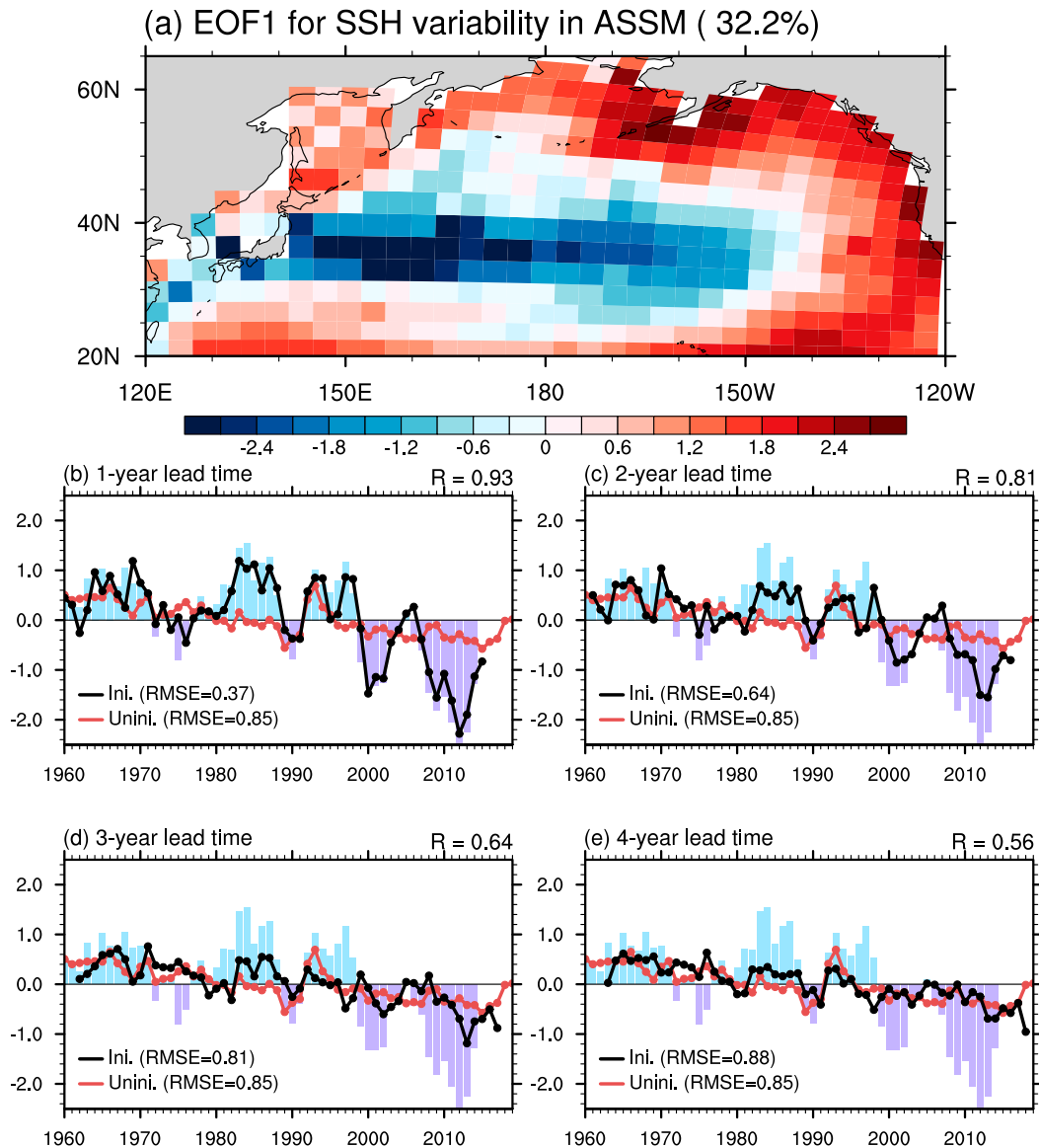


FIG. 17. (a) Regression map associated with the first EOF mode for North Pacific SSH variability in the assimilation run and time coefficients projected onto that pattern in the assimilation (color), uninitialized (red), and initialized runs (black) for (b) 1-, (c) 2-, (d) 3-, (e) 4-yr lead times. Correlation coefficient between the assimilation and initialized runs is denoted at the upper-right corner of (b)–(e). Parentheses in legends are RMSE of time coefficients with the assimilation run (see Table 2).

2013; Matei et al. 2012; Yeager et al. 2012; Doblas-Reyes et al. 2013; Swingedouw et al. 2013; Ham et al. 2014; Karspeck et al. 2014). However, it still remains unclear how to best initialize AMOC variations in global climate models (Dunstone and Smith 2010; Meehl et al. 2014). This debate arises from insufficient observations, lack of knowledge of the actual drivers of AMOC variability, large model diversity in representing the AMOC, and short data coverage to evaluate the AMOC decadal variability. According to our 90-yr-long perfect model

experiments, the SST assimilation is insufficient to initialize AMOC states appropriately. In fact, our SST assimilation run in the perfect model framework shows large climatological biases in both SST and SSH fields in the North Atlantic (Figs. 1a and 2a). Furthermore, our surface and upper ocean data assimilation runs (above 300 m) fail to capture the AMOC variability (Figs. 6b,d). In contrast to these upper ocean assimilation runs, the complete 3D ocean temperature and salinity assimilation run simulates well the targeted and observed AMOC

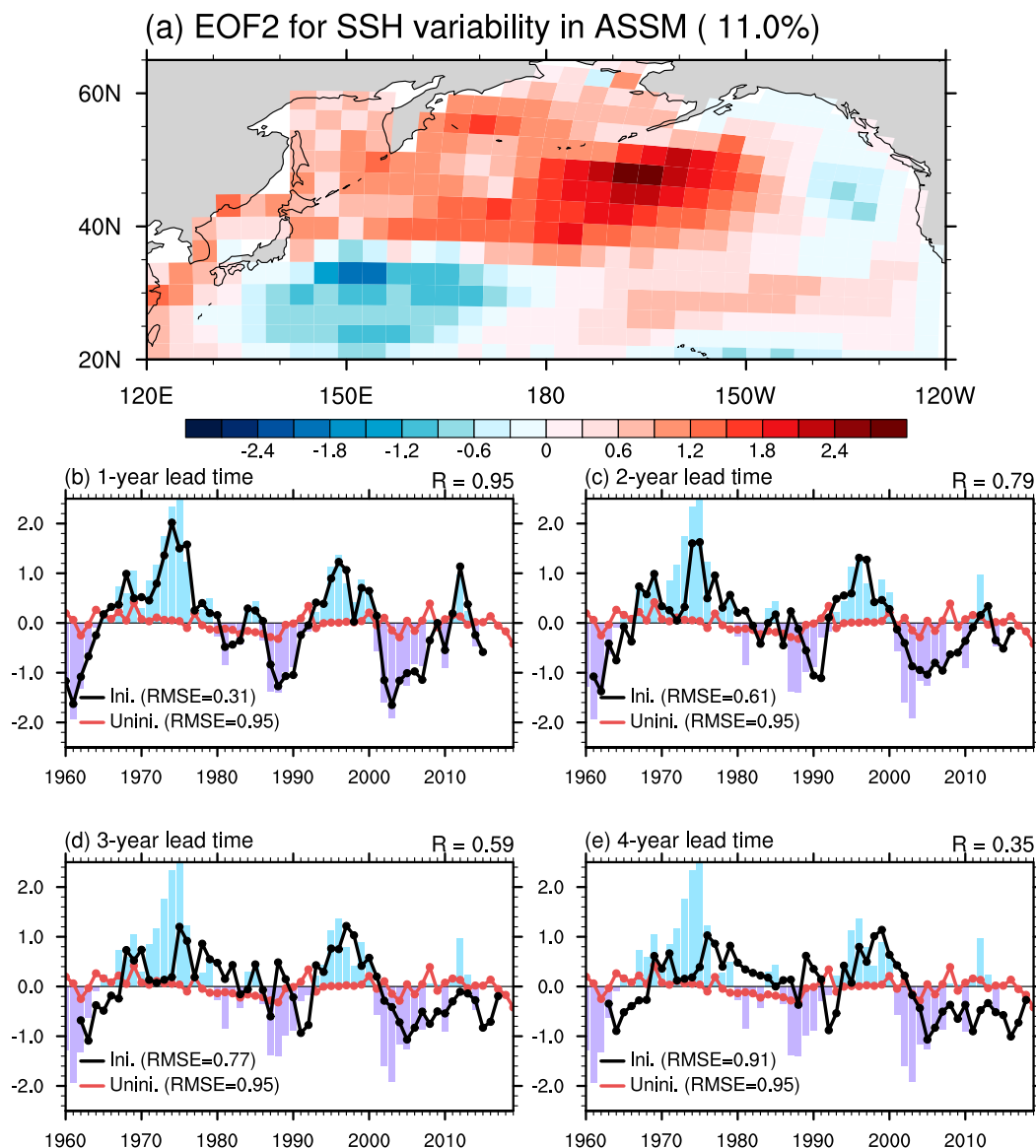


FIG. 18. (a) Regression map associated with the second EOF mode for North Pacific SSH variability in the assimilation run and time coefficients projected onto that pattern in the assimilation (color), uninitialized (red), and initialized runs (black) for (b) 1-, (c) 2-, (d) 3-, (e) 4-yr lead times. Correlation coefficient between the assimilation and initialized runs is denoted at the upper-right corner of (b)–(e). Parentheses in legends are RMSE of time coefficients with the assimilation run (see Table 2).

variability in the perfect model and the real world, as represented here by the ORA-S4 ocean reanalysis estimate (Figs. 6c and 15). These results support the hypothesis by Dunstone and Smith (2010), in which 3D ocean temperature and salinity assimilation is crucial for AMOC initialization. Using the SST nudging approach, Swingedouw et al. (2013) argued that a volcanic eruption contributes to the AMOC decadal predictability. However, our results show only a minor impact of volcanic eruptions on the AMOC predictive skill in the

uninitialized run (Figs. 14 and 15). Because of the prevailing model uncertainty in how climate models react to external radiative forcings, further studies need to be conducted to clarify the AMOC response to external forcing factors.

Our perfect model experiments demonstrate that assimilating salinity in addition to temperature is important for skillful initialization of the model, especially to resolve well the AMOC variability. When we assimilate the 3D ocean temperature without salinity, the model



simulated AMOC is severely degraded (Fig. 5e). Temperature constraint on the AMOC, in the absence of salinity data assimilation, is also reported by pacemaker experiments in the CMIP6 Decadal Climate Prediction Project-Component C (Boer et al. 2016). Model drift in the AMOC is sensitive to the strength of SST restoring time scales and it appears that the additional SSS restoring contributes to reducing the unrealistic model drift. Temperature and salinity assimilation helps to maintain the appropriate density condition.

The importance of temperature and salinity contributions to density implies a disadvantage of anomaly assimilation. Since ocean density is a nonlinear function of temperature and salinity, model mean-state biases of temperature and salinity may cause unrealistic simulation of the ocean density structure. As a result, the anomaly assimilation approach may have poorer predictive skills of density-driven ocean circulations compared to the prediction system that utilizes full-field assimilation. An example of the disadvantage of our anomaly assimilation approach is during the mid-1990s when the initialized run failed to predict the rapid warming of OHC300 in the SPG region (Figs. 16c,e) even though some prediction systems using full-field ocean assimilation successfully predicted this warming several years in advance (Robson et al. 2012; Yeager et al. 2012; Msadek et al. 2014).

Concerning North Pacific SSH variability, we identified two leading decadal modes that resemble the observed PDO and NPGO modes (Figs. 7, 8, 17, and 18). Our perfect model experiments indicate that the 3D ocean temperature and salinity assimilation reproduces the temporal variations of PDO, which is defined as the first EOF mode of SSH variability (Fig. 7c). This PDO-like SSH variability is mainly attributed to upper ocean thermodynamics and dynamics (Fig. 7d), whereas surface or deeper ocean data assimilation play only secondary roles in constraining the PDO-like SSH variability (Figs. 7b,e).

Because the full depth ocean assimilation includes upper ocean data, our real hindcasts showed multiyear predictive skills for PDO-like SSH variability (Fig. 17). Kim et al. (2014) assessed the PDO predictability using six prediction systems in the CMIP5 decadal hindcast experiments and found that half of those systems show high predictive skills over the North Pacific Ocean. Consistent with our results showing the importance of assimilating the full ocean depth, those three well-performing prediction systems incorporate not only SST, but also 3D ocean temperature and salinity observations, in their initialization approaches (Meehl et al. 2014). Moreover, Kim et al. (2014) demonstrated that PDO predictions further improve by correcting

model systematic errors. Our SVD analysis between observations and the uninitialized run also highlights large model biases of the externally forced component over the North Pacific Ocean (Fig. 9). Therefore, higher predictive skills of the PDO may be achieved by initializing 3D ocean data, as well as taking into account the model biases associated with the climate response to radiative forcings.

Our decadal prediction experiments also document multiyear predictive skill for NPGO variability. Compared to the PDO, NPGO SSH predictability is enhanced more by accounting for deeper ocean assimilation (Fig. 8), but we find a smaller role of the externally forced component (Fig. 18). Whereas PDO predictive skill is contaminated by many high-frequency atmosphere–ocean phenomena (Schneider and Cornuelle 2005; Mochizuki et al. 2014), our model experiments suggest that the long-term memory of NPGO is mainly attributed to low-frequency ocean dynamics in the North Pacific. Because of the strong connection between the NPGO and marine nutrient variations (Di Lorenzo et al. 2008), our successful NPGO prediction could be beneficial for marine ecosystem predictions (M. Chikamoto et al. 2015). However, our findings are inconsistent with Lienert and Doblus-Reyes (2013), who suggested that the long-term SST trend is the main contributor to the NPGO predictive skills, whereas the benefits of initialization are limited to two years lead. Further intermodel comparison of decadal climate prediction may yield better understanding of such discrepancies.

## 6. Conclusions

We developed a new decadal climate prediction system for CESM, which employs a new bias adjustment approach to reduce unrealistic climate drift and initialization shocks. Decadal climate predictions are performed in two steps: 1) we adjust the observations by removing model biases in the climatology and externally forced component, and 2) then initialize the climate model based on those adjusted observations. This paper focuses on the second step to explore how to initialize the climate model for decadal climate predictions using both a perfect model framework and a real world setting. By estimating model biases in climatological states and the externally forced component, our initialization approach minimizes climate drift during the prediction. This approach is readily adaptable to evaluating predictability of non-Gaussian distribution variables such as precipitation and soil water content (Chikamoto et al. 2017). Our technique to estimate the externally forced component is also applicable to evaluate uncertainties of the climate response

to radiative forcings in ocean reanalyses as well as the forthcoming CMIP6 climate models (Boer et al. 2016).

For the perfect model framework, our assimilation experiments demonstrate that initialization with the full depth ocean temperature and salinity is crucial for proper initialization of the model, particularly in the extratropics. Although some previous studies demonstrated considerable skill in regions such as the North Atlantic by nudging model data to the observed SST (Luo et al. 2005; Keenlyside et al. 2008; Swingedouw et al. 2013), our partial assimilation experiments indicate that 3D temperature and salinity initialization is superior to surface initialization methods, as it significantly reduces climatological and anomalous errors for a variety of ocean dynamical states. Specifically, deeper ocean temperature and salinity fields are required to initialize the low-frequency AMOC variations. We also find contributions of deeper ocean states to the dominant decadal SSH variability in the midlatitude North Pacific. Therefore, essential activities to improve decadal climate predictability include development, maintenance, and operational monitoring for the temperature and salinity observations from the surface to deeper ocean, such as the Moored Buoy Array and Argo floats. This result may also provide an implication for the design of ocean reanalyses: how to impose multivariate and spatial correlations under the limited observations as well as estimate the uncertainty associated with the lack of observations in order to avoid spurious circulations.

To conduct realistic decadal climate predictions of the observational period, we initialized the CESM climate model using 3D temperature and salinity fields estimated from ocean reanalysis data (i.e., ORA-S4). Even though there is a model discrepancy between the ocean reanalysis products and our assimilation system, our initialized experiments show multiyear predictive skills of ocean dynamical variations in the North Atlantic and North Pacific. We expect that decadal predictions could become beneficial in future for fisheries management and perhaps additional marine planning activities (Mantua et al. 1997; Di Lorenzo et al. 2008; M. Chikamoto et al. 2015).

A limitation of our decadal climate prediction approach is that we rely on an existing 3D ocean reanalysis dataset. To become operational, the data assimilation system will need to handle ocean profiles with both spatially and temporally inhomogeneous observational network, especially for the multidecadal time scales. It remains challenging to apply such an ocean assimilation system to fully coupled climate models such as CESM because of the large climatological biases and systematic errors in the models. By utilizing an anomaly assimilation approach, we provide an interim option to skillfully

predict decadal climate variability, which may be especially useful until climate model biases are improved.

*Acknowledgments.* This study benefited from the constructive comments of three anonymous reviewers. The CESM experiments were conducted with resources from the University of Southern California Center for High-Performance Computing and Communications (<http://hpcc.usc.edu>) and from supercomputer Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR's Computational and Information Systems Laboratory, which is sponsored by the National Science Foundation. This work was supported through NSF Award 1049219, Investigation of Decadal Climate Predictability and Hydroclimate impacts (IDCPI) on the Western U.S. YC is supported by the Utah Agricultural Experiment Station (project number UTA01373), Utah State University, and approved as journal paper number 9191. MJW is partially supported through the NOAA Award NA17OAR4310110, Multi-model seasonal sea level forecasts for the U.S. Coast. AT is supported by the Institute for Basic Science (IBS), South Korea, under IBS-R028-D1. SZ is supported by CNFS (41775100, 41830964), Key R and D program of Ministry of Science and Technology of China (2017YFC1404100, 2017YFC1404102).

#### REFERENCES

- Alexander, M. A., C. Deser, and M. S. Timlin, 1999: The reemergence of SST anomalies in the North Pacific Ocean. *J. Climate*, **12**, 2419–2433, [https://doi.org/10.1175/1520-0442\(1999\)012<2419:TROSAT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2419:TROSAT>2.0.CO;2).
- Alves, O., M. A. Balmaseda, D. Anderson, and T. Stockdale, 2004: Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Quart. J. Roy. Meteor. Soc.*, **130**, 647–667, <https://doi.org/10.1256/qj.03.25>.
- Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor, 2012: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere–ocean climate models. *Geophys. Res. Lett.*, **39**, L09712, <https://doi.org/10.1029/2012GL051607>.
- Balmaseda, M. A., 2017: Data assimilation for initialization of seasonal forecasts. *J. Mar. Res.*, **75**, 331–359, <https://doi.org/10.1357/002224017821836806>.
- , D. Dee, A. Vidard, and D. Anderson, 2007: A multivariate treatment of bias for sequential data assimilation: Application to the tropical oceans. *Quart. J. Roy. Meteor. Soc.*, **133**, 167–179, <https://doi.org/10.1002/qj.12>.
- , K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF Ocean Reanalysis System ORAS4. *Quart. J. Roy. Meteor. Soc.*, **139**, 1132–1161, <https://doi.org/10.1002/qj.2063>.
- Bell, M. J., M. Martin, and N. Nichols, 2004: Assimilation of data into an ocean model with systematic errors near the equator. *Quart. J. Roy. Meteor. Soc.*, **130**, 873–893, <https://doi.org/10.1256/qj.02.109>.
- Bloom, S. C., L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271, [https://doi.org/10.1175/1520-0493\(1996\)124<1256:DAUIAU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2).

- Boer, G. J., and Coauthors, 2016: The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geosci. Model Dev.*, **9**, 3751–3777, <https://doi.org/10.5194/GMD-9-3751-2016>.
- Booth, B. B., N. J. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232, <https://doi.org/10.1038/nature10946>.
- Branstator, G., and H. Teng, 2010: Two limits of initial-value decadal predictability in a CGCM. *J. Climate*, **23**, 6292–6311, <https://doi.org/10.1175/2010JCLI3678.1>.
- Cheng, W., J. C. Chiang, and D. Zhang, 2013: Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. *J. Climate*, **26**, 7187–7197, <https://doi.org/10.1175/JCLI-D-12-00496.1>.
- Chikamoto, M. O., A. Timmermann, Y. Chikamoto, H. Tokinaga, and N. Harada, 2015: Mechanisms and predictability of multiyear ecosystem variability in the North Pacific. *Global Biogeochem. Cycles*, **29**, 2001–2019, <https://doi.org/10.1002/2015GB005096>.
- Chikamoto, Y., and Coauthors, 2012: Predictability of a stepwise shift in Pacific climate during the late 1990s in hindcast experiments using MIROC. *J. Meteor. Soc. Japan*, **90A**, 1–21, <https://doi.org/10.2151/jmsj.2012-A01>.
- , and Coauthors, 2013: An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC. *Climate Dyn.*, **40**, 1201–1222, <https://doi.org/10.1007/s00382-012-1351-y>.
- , and Coauthors, 2015a: Skillful multi-year predictions of tropical trans-basin climate variability. *Nat. Commun.*, **6**, 6869, <https://doi.org/10.1038/ncomms7869>.
- , A. Timmermann, S. Stevenson, P. DiNezio, and S. Langford, 2015b: Decadal predictability of soil water, vegetation, and wildfire frequency over North America. *Climate Dyn.*, **45**, 2213–2235, <https://doi.org/10.1007/s00382-015-2469-5>.
- , —, M. J. Widlansky, M. A. Balmaseda, and L. Stott, 2017: Multi-year predictability of climate, drought, and wildfire in southwestern North America. *Sci. Rep.*, **7**, 6568, <https://doi.org/10.1038/s41598-017-06869-7>.
- Choudhury, D., A. Sen Gupta, A. Sharma, R. Mehrotra, and B. Sivakumar, 2017: An assessment of drift correction alternatives for CMIP5 decadal predictions. *J. Geophys. Res. Atmos.*, **122**, 10 282–10 296, <https://doi.org/10.1002/2017JD026900>.
- Danabasoglu, G., and Coauthors, 2016: North Atlantic simulations in Coordinated Ocean-Ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability. *Ocean Modell.*, **97**, 65–90, <https://doi.org/10.1016/j.ocemod.2015.11.007>.
- Dee, D. P., and A. M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295, <https://doi.org/10.1002/qj.49712454512>.
- Di Lorenzo, E., and Coauthors, 2008: North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.*, **35**, L08607, <https://doi.org/10.1029/2007GL032838>.
- Doblas-Reyes, F., M. Balmaseda, A. Weisheimer, and T. Palmer, 2011: Decadal climate prediction with the European Centre for Medium-Range Weather Forecasts Coupled Forecast System: Impact of ocean observations. *J. Geophys. Res.*, **116**, D19111, <https://doi.org/10.1029/2010JD015394>.
- , and Coauthors, 2013: Initialized near-term regional climate change prediction. *Nat. Commun.*, **4**, 1715, <https://doi.org/10.1038/ncomms2704>.
- Dunstone, N. J., and D. M. Smith, 2010: Impact of atmosphere and sub-surface ocean data on decadal climate prediction. *Geophys. Res. Lett.*, **37**, L02709, <https://doi.org/10.1029/2009GL041609>.
- Durack, P. J., and S. E. Wijffels, 2010: Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. *J. Climate*, **23**, 4342–4362, <https://doi.org/10.1175/2010JCLI3377.1>.
- Fučkar, N. S., D. Volpi, V. Guemas, and F. J. Doblas-Reyes, 2014: A posteriori adjustment of near-term climate predictions: Accounting for the drift dependence on the initial conditions. *Geophys. Res. Lett.*, **41**, 5200–5207, <https://doi.org/10.1002/2014GL060815>.
- Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, <https://doi.org/10.1175/2011JCLI4083.1>.
- Gregory, J., and Coauthors, 2005: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO<sub>2</sub> concentration. *Geophys. Res. Lett.*, **32**, L12703, <https://doi.org/10.1029/2005GL023209>.
- Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate variability. *Science*, **275**, 181–184, <https://doi.org/10.1126/science.275.5297.181>.
- Grötzner, A., M. Latif, A. Timmermann, and R. Voss, 1999: Interannual to decadal predictability in a coupled ocean–atmosphere general circulation model. *J. Climate*, **12**, 2607–2624, [https://doi.org/10.1175/1520-0442\(1999\)012<2607:ITDPIA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2607:ITDPIA>2.0.CO;2).
- Ham, Y.-G., M. M. Rienecker, M. J. Suarez, Y. Vikhliayev, B. Zhao, J. Marshak, G. Vernieres, and S. D. Schubert, 2014: Decadal prediction skill in the geos-5 forecast system. *Climate Dyn.*, **42**, 1–20, <https://doi.org/10.1007/s00382-013-1858-x>.
- Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Rev. Geophys.*, **48**, RG4004, <https://doi.org/10.1029/2010RG000345>.
- Hawkins, E., B. Dong, J. Robson, R. Sutton, and D. Smith, 2014: The interpretation and use of biases in decadal climate predictions. *J. Climate*, **27**, 2931–2947, <https://doi.org/10.1175/JCLI-D-13-00473.1>.
- Hazeleger, W., V. Guemas, B. Wouters, S. Corti, I. Andreu-Burillo, F. Doblas-Reyes, K. Wyser, and M. Caian, 2013: Multiyear climate predictions using two initialization strategies. *Geophys. Res. Lett.*, **40**, 1794–1798, <https://doi.org/10.1002/grl.50355>.
- Huang, B., J. Kinter, and P. Schopf, 2002: Ocean data assimilation using intermittent analyses and continuous model error correction. *Adv. Atmos. Sci.*, **19**, 965–992, <https://doi.org/10.1007/s00376-002-0059-z>.
- Ishii, M., Y. Fukuda, S. Hirahara, S. Yasui, T. Suzuki, and K. Sato, 2017: Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, **13**, 163–167, <https://doi.org/10.2151/SOLA.2017-030>.
- Jordà, G., and D. Gomis, 2013: On the interpretation of the steric and mass components of sea level variability: The case of the Mediterranean basin. *J. Geophys. Res. Oceans*, **118**, 953–963, <https://doi.org/10.1002/jgrc.20060>.
- Karspeck, A., S. Yeager, G. Danabasoglu, and H. Teng, 2014: An evaluation of experimental decadal predictions using CCSM4. *Climate Dyn.*, **44**, 907–923, <https://doi.org/10.1007/s00382-014-2212-7>.
- Keenlyside, N. S., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, **453**, 84–88, <https://doi.org/10.1038/nature06921>.

- Keller, K., C. Deutsch, M. G. Hall, and D. F. Bradford, 2007: Early detection of changes in the North Atlantic meridional overturning circulation: Implications for the design of ocean observation systems. *J. Climate*, **20**, 145–157, <https://doi.org/10.1175/JCLI3993.1>.
- Kharin, V., G. Boer, W. Merryfield, J. Scinocca, and W.-S. Lee, 2012: Statistical adjustment of decadal predictions in a changing climate. *Geophys. Res. Lett.*, **39**, L19705, <https://doi.org/10.1029/2012GL052647>.
- Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys. Res. Lett.*, **39**, L10701, <https://doi.org/10.1029/2012GL051644>.
- , Y.-G. Ham, and A. A. Scaife, 2014: Improvement of initialized decadal predictions over the North Pacific Ocean by systematic anomaly pattern correction. *J. Climate*, **27**, 5148–5162, <https://doi.org/10.1175/JCLI-D-13-00519.1>.
- Lienert, F., and F. J. Doblas-Reyes, 2013: Decadal prediction of interannual tropical and North Pacific sea surface temperature. *J. Geophys. Res. Atmos.*, **118**, 5913–5922, <https://doi.org/10.1002/JGRD.50469>.
- Luo, J., S. Masson, S. Behera, S. Shingu, and T. Yamagata, 2005: Seasonal climate predictability in a coupled OAGCM using a different approach for ensemble forecasts. *J. Climate*, **18**, 4474–4497, <https://doi.org/10.1175/JCLI3526.1>.
- Magnusson, L., M. Alonso-Balmaseda, S. Corti, F. Molteni, and T. Stockdale, 2013: Evaluation of forecast strategies for seasonal and decadal forecasts in the presence of systematic model errors. *Climate Dyn.*, **41**, 2393–2409, <https://doi.org/10.1007/s00382-012-1599-2>.
- Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069–1079, [https://doi.org/10.1175/1520-0477\(1997\)078<1069:APICOW>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<1069:APICOW>2.0.CO;2).
- Matei, D., J. Baehr, J. H. Jungclaus, H. Haak, W. A. Müller, and J. Marotzke, 2012: Multiyear prediction of monthly mean Atlantic meridional overturning circulation at 26.5°N. *Science*, **335**, 76–79, <https://doi.org/10.1126/science.1210299>.
- Meehl, G. A., and H. Teng, 2014: Regional precipitation simulations for the mid-1970s shift and early-2000s hiatus. *Geophys. Res. Lett.*, **41**, 7658–7665, <https://doi.org/10.1002/2014GL061778>.
- , and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc.*, **90**, 1467–1485, <https://doi.org/10.1175/2009BAMS2778.1>.
- , and Coauthors, 2014: Decadal climate prediction: An update from the trenches. *Bull. Amer. Meteor. Soc.*, **95**, 243–267, <https://doi.org/10.1175/BAMS-D-12-00241.1>.
- Mochizuki, T., and Coauthors, 2010: Pacific decadal oscillation hindcasts relevant to near-term climate prediction. *Proc. Natl. Acad. Sci. USA*, **107**, 1833–1837, <https://doi.org/10.1073/pnas.0906531107>.
- , and Coauthors, 2012: Decadal prediction using a recent series of MIROC global climate models. *J. Meteor. Soc. Japan*, **90A**, 373–383, <https://doi.org/10.2151/jmsj.2012-A22>.
- , M. Kimoto, Y. Chikamoto, M. Mori, M. Watanabe, and M. Ishii, 2014: Error sensitivity to initial climate states in Pacific decadal hindcasts. *SOLA*, **10**, 39–44, <https://doi.org/10.2151/SOLA.2014-009>.
- Msadek, R., and Coauthors, 2014: Predicting a decadal shift in North Atlantic climate variability using the GFDL Forecast system. *J. Climate*, **27**, 6472–6496, <https://doi.org/10.1175/JCLI-D-13-00476.1>.
- Murphy, J., and Coauthors, 2010: Towards prediction of decadal climate variability and change. *Procedia Environ. Sci.*, **1**, 287–304, <https://doi.org/10.1016/j.proenv.2010.09.018>.
- Narapusetty, B., C. Stan, and A. Kumar, 2014: Bias correction methods for decadal sea-surface temperature forecasts. *Tellus*, **66A**, 23681, <https://doi.org/10.3402/tellusa.v66.23681>.
- Pierce, D. W., P. J. Gleckler, T. P. Barnett, B. D. Santer, and P. J. Durack, 2012: The fingerprint of human-induced changes in the ocean's salinity and temperature fields. *Geophys. Res. Lett.*, **39**, L21704, <https://doi.org/10.1029/2012GL053389>.
- Pohlmann, H., J. H. Jungclaus, A. Köhl, D. Stammer, and J. Marotzke, 2009: Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. *J. Climate*, **22**, 3926–3938, <https://doi.org/10.1175/2009JCLI2535.1>.
- Qiu, B., 2003: Kuroshio Extension variability and forcing of the Pacific decadal oscillations: Responses and potential feedback. *J. Phys. Oceanogr.*, **33**, 2465–2482, <https://doi.org/10.1175/2459.1>.
- , and S. Chen, 2005: Variability of the Kuroshio Extension jet, recirculation gyre, and mesoscale eddies on decadal time scales. *J. Phys. Oceanogr.*, **35**, 2090–2103, <https://doi.org/10.1175/JPO2807.1>.
- Robson, J., R. Sutton, K. Lohmann, D. Smith, and M. D. Palmer, 2012: Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s. *J. Climate*, **25**, 4116–4134, <https://doi.org/10.1175/JCLI-D-11-00443.1>.
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, <https://doi.org/10.1175/JCLI-D-12-00823.1>.
- Schneider, N., and B. D. Cornuelle, 2005: The forcing of the Pacific decadal oscillation. *J. Climate*, **18**, 4355–4373, <https://doi.org/10.1175/JCLI3527.1>.
- , A. J. Miller, M. A. Alexander, and C. Deser, 1999: Subduction of decadal North Pacific temperature anomalies: Observations and dynamics. *J. Phys. Oceanogr.*, **29**, 1056–1070, [https://doi.org/10.1175/1520-0485\(1999\)029<1056:SODNPT>2.0.CO;2](https://doi.org/10.1175/1520-0485(1999)029<1056:SODNPT>2.0.CO;2).
- Shields, C. A., D. A. Bailey, G. Danabasoglu, M. Jochum, J. T. Kiehl, S. Levis, and S. Park, 2012: The low-resolution CCSM4. *J. Climate*, **25**, 3993–4014, <https://doi.org/10.1175/JCLI-D-11-00260.1>.
- Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317**, 796–799, <https://doi.org/10.1126/science.1139540>.
- , R. Eade, and H. Pohlmann, 2013: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. *Climate Dyn.*, **41**, 3325–3338, <https://doi.org/10.1007/s00382-013-1683-2>.
- Swingedouw, D., J. Mignot, S. Labetoulle, E. Guilyardi, and G. Madec, 2013: Initialization and predictability of the AMOC over the last 50 years in a climate model. *Climate Dyn.*, **40**, 2381–2399, <https://doi.org/10.1007/s00382-012-1516-8>.
- Taguchi, B., and N. Schneider, 2014: Origin of decadal-scale, eastward-propagating heat content anomalies in the North Pacific. *J. Climate*, **27**, 7568–7586, <https://doi.org/10.1175/JCLI-D-13-00102.1>.
- Tatebe, H., and Coauthors, 2012: The initialization of the MIROC climate models with hydrographic data assimilation for decadal prediction. *J. Meteor. Soc. Japan*, **90A**, 275–294, <https://doi.org/10.2151/jmsj.2012-A14>.

- Taylor, K., R. Stouffer, and G. Meehl, 2009: A summary of the CMIP5 experiment design. PCDMI Rep., 33 pp., accessed 26 July 2010, [https://pcmdi.llnl.gov/mips/cmip5/docs/Taylor\\_CMIP5\\_22Jan11\\_marked.pdf?id=4](https://pcmdi.llnl.gov/mips/cmip5/docs/Taylor_CMIP5_22Jan11_marked.pdf?id=4).
- Teng, H., and G. Branstator, 2011: Initial value predictability in prominent modes of North Pacific subsurface temperature in a coupled GCM. *Climate Dyn.*, **36**, 1813–1834, <https://doi.org/10.1007/s00382-010-0749-7>.
- , —, and G. A. Meehl, 2011: Predictability of the Atlantic overturning circulation and associated surface patterns in two CCSM3 climate change ensemble experiments. *J. Climate*, **24**, 6054–6076, <https://doi.org/10.1175/2011JCLI4207.1>.
- Timmermann, A., and Coauthors, 2018: El Niño–Southern Oscillation complexity. *Nature*, **559**, 535–545, <https://doi.org/10.1038/s41586-018-0252-6>.
- Trenberth, K. E., and J. W. Hurrell, 1994: Decadal atmosphere–ocean variations in the Pacific. *Climate Dyn.*, **9**, 303–319, <https://doi.org/10.1007/BF00204745>.
- van Oldenborgh, G. J., F. J. Doblas-Reyes, B. Wouters, and W. Hazeleger, 2012: Decadal prediction skill in a multi-model ensemble. *Climate Dyn.*, **38**, 1263–1280, <https://doi.org/10.1007/s00382-012-1313-4>.
- Vialard, J., F. Vitart, M. Balmaseda, T. Stockdale, and D. Anderson, 2005: An ensemble generation method for seasonal forecasting with an ocean–atmosphere coupled model. *Mon. Wea. Rev.*, **133**, 441–453, <https://doi.org/10.1175/MWR-2863.1>.
- Xie, S.-P., T. Kunitani, A. Kubokawa, M. Nonaka, and S. Hosoda, 2000: Interdecadal thermocline variability in the North Pacific for 1958–97: A GCM simulation. *J. Phys. Oceanogr.*, **30**, 2798–2813, [https://doi.org/10.1175/1520-0485\(2000\)030<2798:ITVITN>2.0.CO;2](https://doi.org/10.1175/1520-0485(2000)030<2798:ITVITN>2.0.CO;2).
- Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng, 2012: A decadal prediction case study: Late twentieth-century North Atlantic ocean heat content. *J. Climate*, **25**, 5173–5189, <https://doi.org/10.1175/JCLI-D-11-00595.1>.
- Zhang, S., A. Rosati, and T. Delworth, 2010: The adequacy of observing systems in monitoring the Atlantic meridional overturning circulation and North Atlantic climate. *J. Climate*, **23**, 5311–5324, <https://doi.org/10.1175/2010JCLI3677.1>.