

# Strongly Coupled Data Assimilation Using Leading Averaged Coupled Covariance (LACC). Part III: Assimilation of Real World Reanalysis

JINGZHE SUN

*College of Computer, National University of Defense Technology, Changsha, and Open Studio for Ocean-Climate-Isotope Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, and Atmospheric Science Program, Department of Geography, The Ohio State University, Columbus, Ohio, and Nelson Institute Center for Climatic Research, and Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, Wisconsin*

ZHENGYU LIU

*Atmospheric Science Program, Department of Geography, The Ohio State University, Columbus, Ohio*

FEIYU LU

*Atmospheric and Oceanic Sciences Program, Princeton University, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*

WEIMIN ZHANG

*College of Meteorology and Oceanography, National University of Defense Technology, and Laboratory of Software Engineering for Complex Systems, Changsha, China*

SHAOQING ZHANG

*Physical Oceanography Laboratory, College of Oceanic and Atmospheric Sciences, and Institute for Advanced Ocean Study, Ocean University of China, and Function Laboratory for Ocean Dynamics and Climate, Qingdao National Laboratory for Marine Science and Technology, and International Laboratory for High-Resolution Earth System Prediction, Qingdao, China*

(Manuscript received 12 September 2019, in final form 25 February 2020)

## ABSTRACT

Recent studies proposed leading averaged coupled covariance (LACC) as an effective strongly coupled data assimilation (SCDA) method to improve the coupled state estimation over weakly coupled data assimilation (WCDA) in a coupled general circulation model (CGCM). This SCDA method, however, has been previously evaluated only in the perfect model scenario. Here, as a further step toward evaluating LACC for real world data assimilation, LACC is evaluated for the assimilation of reanalysis data in a CGCM. Several criteria are used to evaluate LACC against the benchmark WCDA. It is shown that despite significant model bias, LACC can improve the coupled state estimation over WCDA. Compared to WCDA, LACC increases the globally averaged anomaly correlation coefficients (ACCs) of sea surface temperature (SST) by 0.036 and atmosphere temperature at the bottom level ( $T_s$ ) by 0.058. However, there also exist regions where WCDA outperforms LACC. Although the reduction in the anomaly root-mean-square error (RMSE) is not as consistently clear as the increase in ACC, LACC can largely correct the biased model climatology.

## 1. Introduction

Coupled data assimilation (CDA) is considered an effective initialization approach for coupled Earth system models (Zhang et al. 2007; Sugiura et al. 2008; Saha

et al. 2010; Dee et al. 2011). CDA assimilates observations into one or more model components and allows the exchange of information between different model components dynamically and statistically. Therefore, it is expected to produce more self-consistent state estimation for coupled models (Zhang et al. 2005, 2007; Sugiura et al. 2008). Generally, CDA can be classified

Corresponding author: Jingzhe Sun, sunjingzhe13@nudt.edu.cn

DOI: 10.1175/MWR-D-19-0304.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy ([www.ametsoc.org/PUBSReuseLicenses](http://www.ametsoc.org/PUBSReuseLicenses)).

into two categories (Han et al. 2013; Liu et al. 2013): weakly coupled DA (WCDA) and strongly coupled DA (SCDA). In WCDA, assimilation is applied independently to each model component in the coupled model, and the coupling between different components is accomplished only dynamically through cross-component fluxes during the forecast stage. In SCDA, assimilation is simultaneously applied to the fully coupled system, such that the coupled system is treated as a single integrated system. Thus, observations in one model component can be used to update other model components via the cross-component error covariance (hereafter coupled covariance). As a result, the interactions between different model components are accomplished not only dynamically via cross-component fluxes during the forecast stage but also statistically through the coupled covariance during the analysis phase.

SCDA is, in principle, a better CDA method than WCDA because the former makes full use of the observations and the coupled model to provide the more consistent and balanced initial conditions for coupled Earth system models. In practice, however, despite some progress in the development of SCDA in recent years (Liu et al. 2013; Lu et al. 2015b; Frolov et al. 2016; Sluka et al. 2016; Smith et al. 2017, 2018), many scientific and technical challenges remain to be solved before the SCDA system becomes operational (Penny et al. 2017; Penny and Hamill 2017). Sampling error is a serious problem for ensemble-based SCDA implementation in a complex model such as a coupled general circulation model (CGCM) because of the limited ensemble size (usually dozens at most) constrained by computation resources. Therefore, a simple application of the coupled covariance may introduce more noise than the signal, deteriorating SCDA relative to the corresponding WCDA (e.g., Han et al. 2013; Lu et al. 2015a,b). Recently, we proposed the SCDA scheme called leading averaged coupled covariance (LACC) in which the signal-to-noise ratio is increased by using the atmospheric observations of the lead times in time averaging (Lu et al. 2015a, hereafter Part I). It has been shown that LACC can lead to significant improvements in the analysis over WCDA in a CGCM using a small ensemble size (Lu et al. 2015b, hereafter Part II). However, LACC has been tested only in the perfect model scenario in a CGCM. In the real world scenario, model bias becomes one main issue and may provide a significant challenge for the successful implementation of the SCDA scheme. It therefore remains unclear whether LACC can improve the analysis in the presence of significant model bias. In a simple model, Part I showed that LACC can indeed improve the state estimation relative to WCDA in a biased model framework. This

result suggests that LACC can possibly improve the state estimation despite model bias. This implication is confirmed in this study.

As an extension of Part I and Part II, this study addresses the impact of model bias on LACC in a CGCM. As a further step toward the real world scenario, we evaluate LACC using reanalysis data as observations. To the best of our knowledge, this is the first application of a SCDA system assimilating real world reanalysis data in a CGCM. Our study shows that, relative to WCDA, LACC can indeed improve the anomaly correlation coefficient (ACC) of the coupled state in both the ocean and atmosphere. Although model bias complicates the evaluation in terms of the root-mean-square error (RMSE), LACC improves the anomaly RMSE over significant areas while correcting the biased model climatology. This paper is organized as follows. Section 2 describes the CGCM, our SCDA system, and the LACC method. The experiments and results are reported in section 3. More specifically, section 3a shows the benchmark WCDA experiment, sections 3b–3e show the results of four comparison schemes, section 3f shows the sensitivity experiments of the LACC method, and section 3g shows the correction effect of LACC to the biased model climatology. Section 4 discusses the results and summarizes the paper.

## 2. Model and methods

The Fast Ocean Atmosphere Model (FOAM, version 1.5) is a fully coupled general circulation model (Jacob 1997; Jacob et al. 2001). The atmosphere component is a spectral model with an R15 horizontal resolution and an 18-level hybrid vertical coordinate. The ocean component has a  $2.8^\circ \times 1.4^\circ$  horizontal resolution and a 24-level  $z$  coordinate. The land and sea ice components are based on those of Community Climate Model 2 and CSM Sea Ice Model 2, respectively. FOAM is able to capture most major features of the observed global climatology and climate variability as in some more advanced CGCMs (e.g., Liu et al. 2004, 2007a,b).

The SCDA system in FOAM is based on the ensemble adjustment Kalman filter (EAKF; Anderson 2001, 2003; Zhang et al. 2007), a variant of the ensemble Kalman filter (EnKF; Evensen 1994). It consists of the atmosphere DA (ADA), ocean DA (ODA), and a cross-domain update (hereafter cross update) from the atmosphere to the ocean. Given that the essence of LACC is to take advantage of the increased cross-domain correlations between the atmosphere of the lead times and the current state of the ocean, only the atmosphere-to-ocean cross update is considered in this study. Considering that the observations are more numerous in the atmosphere than in

the ocean, the atmosphere-to-ocean cross update is likely to play a more important role in practice than the ocean-to-atmosphere update. Both localization (Houtekamer and Mitchell 1998) and covariance inflation (Zhang et al. 2004) are included in the system. The flexibility of this DA system provides us with a SCDA system when all three DA components are activated and a WCDA system when the cross update is shut off.

The cross update in FOAM is based on the LACC approach, which takes advantage of the enhanced coupled covariance between the leading averaged atmosphere and the current state of the ocean. In the EnKF-based cross update of LACC, the ensemble ocean analysis at time  $t$ ,  $x_{\text{ocn}}^a(t)$ , can be represented as follows:

$$x_{\text{ocn}}^a(t) = x_{\text{ocn}}^f(t) + \bar{K} \times [\overline{x_{\text{atm}}^o(\tau_2, \tau_1)} - \overline{x_{\text{atm}}^f(\tau_2, \tau_1)}]$$

$$\overline{x_{\text{atm}}^o(\tau_2, \tau_1)} = \overline{x_{\text{atm}}^{\text{obs}}(\tau_2, \tau_1)} + N(0, \sigma_{\text{atm}}/\sqrt{\tau})$$

$$\bar{K} = \frac{\text{cov}\langle x_{\text{ocn}}^f(t), \overline{x_{\text{atm}}^f(\tau_2, \tau_1)} \rangle}{\text{var}\langle \overline{x_{\text{atm}}^f(\tau_2, \tau_1)} \rangle + \sigma_{\text{atm}}^2/\tau},$$

where  $x_{\text{ocn}}^f(t)$  is the ensemble ocean forecast at time  $t$ ,  $\overline{x_{\text{atm}}^{\text{obs}}(\tau_2, \tau_1)}$  is the leading averaged atmosphere observation between  $\tau_2$  and  $\tau_1$  ( $\tau_2 \leq \tau_1 \leq t$ ) with error  $\sigma_{\text{atm}}$ ,  $\overline{x_{\text{atm}}^o(\tau_2, \tau_1)}$  is the leading averaged ensemble atmosphere observation,  $\overline{x_{\text{atm}}^f(\tau_2, \tau_1)}$  is the corresponding ensemble atmosphere forecast, and  $\bar{K}$  is the Kalman gain. Here we note that the error of the time-averaged observation  $\overline{x_{\text{atm}}^{\text{obs}}(\tau_2, \tau_1)}$  is reduced to  $\sigma_{\text{atm}}/\sqrt{\tau}$ , where  $\tau$  is the number of observations between  $\tau_2$  and  $\tau_1$ . More details of the LACC method can be found in Part I. More specifically in this study, observations of atmosphere temperature in the bottom four levels are used to directly adjust the sea surface temperature (SST) during the analysis phase. Here, the use of the observations in the bottom four levels is consistent with the choice in Part II. This usage indicates the trade-off between the fact that the cross correlations of the atmospheric column with SST decrease rapidly with altitude and the desire to use as many observations as possible in the cross update. Compared with the straightforward simultaneous coupled covariance (SimCC) method (Han et al. 2013; Liu et al. 2013), LACC can significantly increase the cross correlation between the atmosphere and ocean and therefore enhance the signal-to-noise ratio during assimilation.

For simplicity here, reanalysis data instead of raw observational data are used as observations for assimilation. The reanalysis data are first interpolated onto the FOAM grids before assimilation. Furthermore, our

LACC is mainly compared with the corresponding WCDA. Since the reanalysis data are used in both cases, it provides a fair comparison. The assimilated reanalysis datasets include the daily mean air temperature and wind components from the NCEP–NCAR Reanalysis I (Kalnay et al. 1996) and the monthly mean SST from the NOAA Extended Reconstructed Sea Surface Temperature (ERSST) version 5 (Huang et al. 2017). The observational uncertainties used in this study are 1 K for monthly mean SST and daily mean atmosphere temperature and  $1 \text{ m s}^{-1}$  for daily mean wind components, similar to previous studies (Zhang et al. 2007; Liu et al. 2014; Lu et al. 2017a,b). Experiments with different levels of uncertainties show that our conclusions are robust. For the verification of the atmosphere analysis, the NCEP–DOE Reanalysis II (Kanamitsu et al. 2002) is used. Reanalysis II is recognized as an improved version of Reanalysis I and thus can be used as a better approximation to the real world than the previous version. To provide independent reference data for the ocean, the EN4.2.1 quality-controlled ocean temperature and salinity analysis from the Met Office Hadley Centre (Good et al. 2013) as well as the SODA3.4.2 reanalysis from the University of Maryland (Carton et al. 2018) are used for validation. EN4.2.1 is a statistical objective analysis without model influence, while SODA3.4.2 is one of the latest reanalyses available. Together these two data sources can serve as reasonable references. All data are interpolated onto the FOAM grids before comparison.

Both ACC and RMSE are used for assessing the analysis. ACC and RMSE, with different emphases, are two of the most widely used statistical measures in the verification and evaluation of experimental results. In climate system analyses, more attention is usually paid to the anomaly than to the climatology. ACC is the correlation between anomalies of the modeled data and the verifying data. A higher ACC implies a higher similarity of evolving variability between two time series of anomalies. RMSE is the standard deviation of the prediction errors, which measures how much the simulated data differ from the reference data. A CGCM, as in our case, inevitably has bias. Given the model bias from the real world, the full-value RMSE is further divided into the anomaly RMSE and the climatology RMSE, which are the RMSEs in temporal anomalies and seasonal climatology, respectively.

A series of experiments are carried out in FOAM to systematically evaluate our SCDA method LACC against the corresponding WCDA. All experiments use 16 ensemble members, as in previous studies (Part II; Lu et al. 2017a,b). The initial ensemble starts from the initial conditions of the atmosphere for 16 consecutive years.

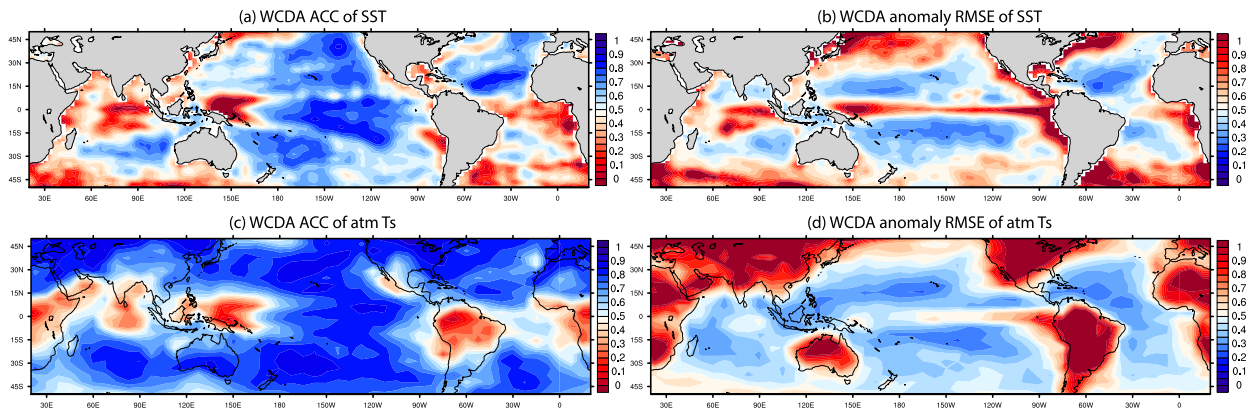


FIG. 1. ACC and anomaly RMSE of (a),(b) SST and (c),(d)  $T_s$  from the WCDA experiment.

The experiments are conducted for 28 years from January 1990 to December 2017, and the most recent 25 years are used for the final comparison and evaluation. All experiments are evaluated by the ACC and RMSE of the monthly ensemble mean. The monthly output averages the model states at all time steps. They provide a good analysis product because the incremental analysis update (IAU) scheme has been used to allow small increments to be added to the model states at every time step. Similar to Part I and Part II, this study uses the monthly values to conduct a fair comparison between WCDA, SimCC, and the LACC method of various averaging intervals. The frequencies for ADA and ODA are every day and every month, respectively. The frequency for the cross update of LACC is every 7 days (hereafter Ave7) unless otherwise specified. Here “Ave7” indicates that the cross correlation is calculated between the current state of SST and the average of 7 daily mean atmosphere temperature in the bottom four levels from 6 days ago to the current day; thus, a shorter averaging time for LACC requires more frequent cross updates than a longer averaging time. The cross update is applied between 50°S and 50°N at all ocean grid points, and the model output between 50°S and 50°N is used in the evaluation.

Localization is applied for all three assimilation components with the widely used scheme of Gaspari and Cohn (1999). A horizontal influence radius of 1000 km is used in both ADA and ODA. Vertically, the atmosphere observations can affect three levels both above and below the observed level, and the SST observations are able to adjust the ocean temperature and salinity down to a depth of 300 m. Considering the relatively low cross correlation between the atmosphere and the ocean, a smaller horizontal influence radius of 500 km is used for the cross update, and the atmosphere observations are only allowed to directly adjust the SST. Covariance inflation is not applied in ADA and ODA. However, given the

larger noise from sampling the coupled covariance, a relaxation factor of 0.5 is set in the relax-to-prior scheme (Zhang et al. 2004) for the cross update.

### 3. Experiments and results

Before beginning the evaluation, we need to first focus on the specific but important issue with nonidealized assimilation: there is no truth. In the perfect model framework, the “truth” can be easily obtained from the model control run as the precise reference for assessing the DA performance. In the real world, observations are the closest samples of the truth that we can obtain. For simplicity here, gridded reanalysis or analysis data instead of raw observations are used as references in the evaluation. These widely used data are statistically and/or dynamically processed products based on raw observations; thus, they can also serve as reliable references to the real world. To provide a reliable performance evaluation of the SCDA system for the assimilation of reanalysis data, we evaluate LACC using four comparison schemes: (i) using the assimilated variables, (ii) using the unasimilated variables, (iii) ensemble forecast verification, and (iv) innovation analysis. As WCDA is used as the benchmark, the performance of the WCDA system is first examined in section 3a, and the LACC performance is analyzed thereafter.

#### a. Benchmark WCDA experiment

The performance of the WCDA system is first examined to provide a reliable benchmark for comparison with LACC. Both SST and the atmosphere temperature at the bottom level ( $T_s$ ) are well constrained by the WCDA system across the globe. Figure 1 shows the ACC and anomaly RMSE of SST and  $T_s$  from the WCDA experiment, using EN4.2.1 analysis and Reanalysis II as references. The RMSE patterns of SST and  $T_s$  are quite



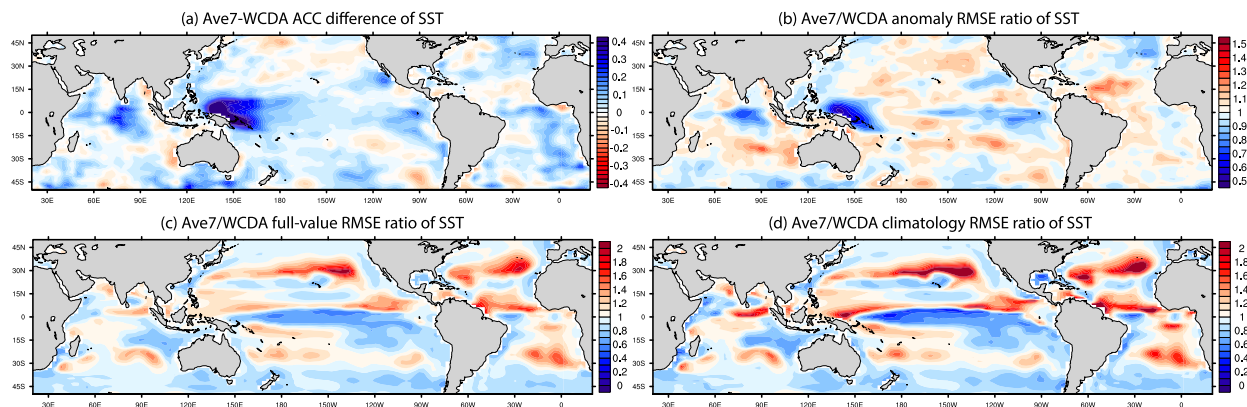


FIG. 2. (a) Ave7-WCDA ACC difference, and Ave7/WCDA (b) anomaly RMSE ratio, (c) full-value RMSE ratio, and (d) climatology RMSE ratio of SST. “Ave7” indicates that the cross correlation is calculated between SST and the average of 7 daily mean atmosphere temperature values in the bottom four levels from 6 days ago to the current day in LACC.

similar to those in the perfect model scenario (Fig. 1 in Part II) but have larger values, especially in the equatorial Pacific for both SST and  $T_s$  and in the eastern equatorial Indian Ocean, the northwestern Pacific and the Southern Ocean to the southwest of Australia for SST. Lower SST ACCs are found in a few equatorial and midlatitude regions with high natural variability, such as the equatorial Indian Ocean, the equatorial Pacific Ocean near the western boundary, the Southern Ocean to the southwest of Australia, the western North Pacific at approximately 50°N, and most areas of the South Atlantic Ocean. The anomaly RMSE of SST shows a pattern similar to that of ACC, except for a larger RMSE belt across almost the entire equatorial Pacific. Over the ocean, ACC and the anomaly RMSE for  $T_s$  are comparable to those of the underlying SST but with higher ACC and smaller RMSE than for SST. The comparison between SST and  $T_s$  indicates that the atmosphere is better constrained by the observations than the ocean, likely because of the higher frequency of DA in the atmosphere (1 day) than in the ocean (1 month). There is no data assimilation for any variables in the land model, so  $T_s$  over the land is affected by poor boundary conditions from the land and the anomaly RMSE is relatively large. Meanwhile, small ACCs of  $T_s$  are confined mainly to the low-latitude regions. In comparison with an ensemble of control simulations without data assimilation, the ACCs of both SST and  $T_s$  from the WCDA experiment are greatly improved with the RMSEs reduced significantly. Between WCDA and control, the differences of the globally averaged ACC are 0.49 for SST and 0.70 for  $T_s$ , and the improvement percentages of the full-value RMSE are 28% for SST and 21% for  $T_s$ .

The ensemble calibration ratio (Bonavita et al. 2012), which is defined as the ratio between the ensemble mean analysis error and the ensemble spread, is also examined

for the WCDA experiment. The globally averaged ratios of SST and  $T_s$  are both significantly larger than 1, which is not a desired characteristic for a well-behaved operational DA system. This feature may be caused by the existence of significant model bias and the insufficient covariance inflation used in the experiment. However, this is unlikely to affect our comparison of the relative performance of LACC against WCDA, because the settings of the covariance inflation in both ADA and ODA remain the same among all experiments.

### b. Comparison of assimilated variables

A comparison between the LACC method and WCDA is first conducted for the SST and atmosphere temperature, which are assimilated variables in the CDA system. We first evaluate the analysis of the SST field. With the LACC of 7-day averaging (i.e., Ave7), the improvement in the ACC of SST is clear, while the improvements in the RMSEs are less clear. Figure 2 shows the ACC difference of Ave7-minus-WCDA (hereafter Ave7-WCDA) and the RMSE ratios of Ave7-to-WCDA (hereafter Ave7/WCDA) for the anomaly, full-value, and climatology of SST, using the EN4.2.1 analysis as the reference. The globally averaged ACC of SST is improved by 0.036, especially in the regions of low ACC in the WCDA experiment (Fig. 2a versus Fig. 1a), such as the tropical Indian Ocean, the western tropical Pacific, to the south of Australia, and most areas of the South Atlantic. The RMSE ratio for the anomaly shows an opposite pattern to that of ACC (Fig. 2a versus Fig. 2b, with a pattern correlation of  $-0.74$ ), with the greatest reductions also in the regions of high RMSE for WCDA, such as the tropical Indian Ocean and the western tropical Pacific. However, the overall improvement is not as clear as in the ACC, with areas of decreased RMSE (blue regions) almost comparable

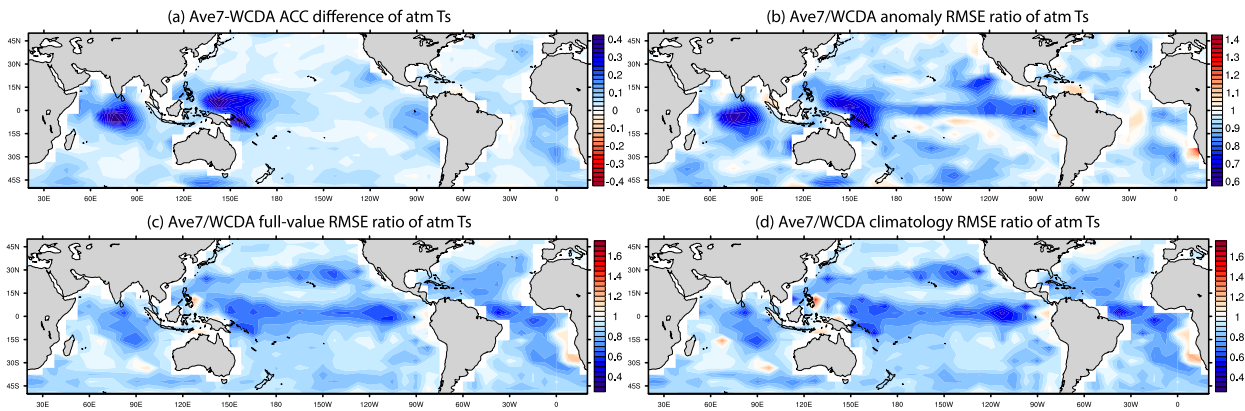


FIG. 3. As in Fig. 2, but for  $T_s$  over the ocean.

with areas of increased RMSE (red regions), leading to a globally averaged RMSE ratio of 1.009. In the meantime, the pattern of the full-value RMSE is quite similar to that of the climatology RMSE (Fig. 2c versus Fig. 2d, with a pattern correlation of 0.95), implying that the impact of cross update on RMSE is dominated by its influence on the mean error. Here the climatology indicates the seasonal climatology. This large correction to climatology is partly because the model usually has a large bias relative to the real world. This climatology-dominated impact of DA may be specific to the climate models with large model biases.

We now examine the analysis for the atmosphere using Reanalysis II as the reference. Given the absence of land data assimilation in our DA system, to focus on the addition of cross update in LACC, we consider only the atmosphere above the ocean. Overall, LACC improves on WCDA for  $T_s$  more than for SST. The ACC of  $T_s$  is improved across the globe (Fig. 3a), with a globally averaged ACC difference of 0.058 and maximum improvements in the regions of low ACC in WCDA (Fig. 1c). In comparison, the ACC of SST is still lower in LACC than in WCDA in some regions, such as the midlatitudes of the North Pacific, to the west of Australia, and to the southwest of South America (Fig. 2a). Similarly, the anomaly RMSE is improved across most of the globe, with a globally averaged ratio value of 0.928 and maximum improvements in the regions of high RMSE in WCDA (Fig. 1d). Different from SST which shows mixed regions of reduced and increased RMSEs, all three RMSEs are reduced for  $T_s$  over almost the entire globe. The full-value RMSE ratio is also dominated by that in the climatology, implying the tendency to correct climatology with the model bias. The relatively large reductions in full-value and climatology RMSEs in the tropics of the three oceans, the northern subtropical Atlantic, and the narrow belt in the North

Pacific centered at approximately 25°N indicate that the strong ocean–atmosphere coupling in these regions may further facilitate the advantages of cross update. Furthermore, the improvements in the ACC and anomaly RMSE of air temperature are not limited to the surface field; rather, they extend vertically into the upper atmosphere, with the maximum improvements located in the low-latitude regions (with a globally averaged ACC difference of 0.038 and a RMSE ratio of 0.951, Fig. 4). In addition, comparison between LACC and SimCC shows that LACC can improve the ACC and anomaly RMSE over SimCC in most regions for both SST and  $T_s$  (with a globally averaged ACC difference of 0.018 and a RMSE ratio of 0.977 for SST and an ACC difference of 0.011 and a RMSE ratio of 0.982 for  $T_s$ ), especially in the western and eastern tropical Pacific, the Southern Ocean to the southwest of Australia, and the midlatitude North Atlantic (Fig. 5).

### c. Comparison of unassimilated variables

In this section, subsurface ocean temperature and salinity from EN4.2.1 and SODA3.4.2 are used as independent verifying data to evaluate LACC. Only the results for EN4.2.1 are shown because the results with SODA3.4.2 have the same conclusion. Here the unassimilated variables are the variables whose observations are not assimilated in the experiments. In the ocean component of our SCDA system, only SST is directly cross updated by atmospheric observations. Therefore, subsurface ocean temperature and salinity are influenced indirectly through the corrected SST during the forecast phase of a DA cycle. Thus, these variables can provide a fair comparison between WCDA and SCDA.

LACC improves the subsurface ocean temperature and salinity in terms of the globally averaged ACC but not clearly in terms of the anomaly RMSE. Figure 6 shows the Ave7-WCDA ACC difference and the Ave7/WCDA

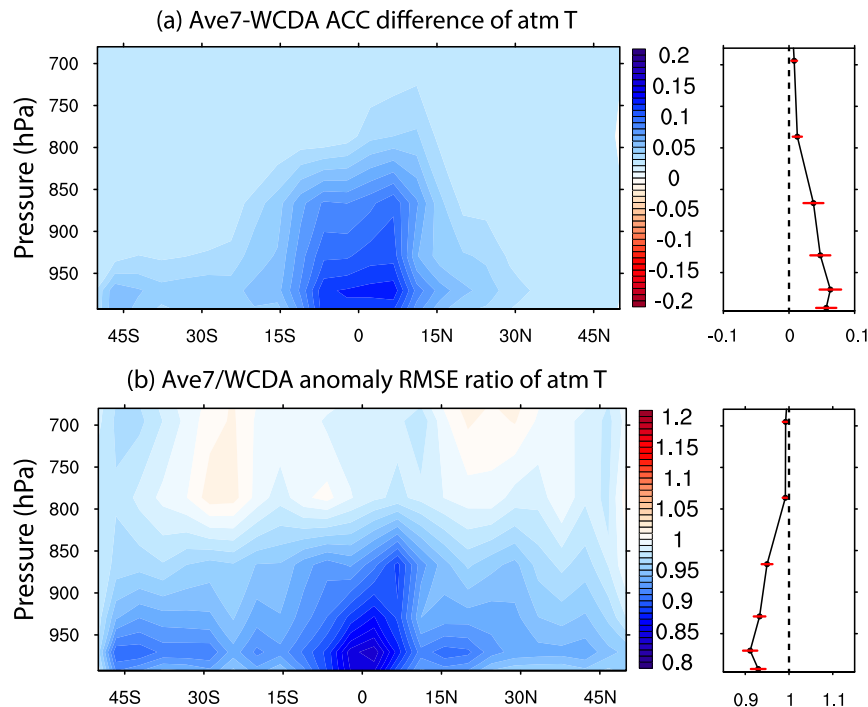


FIG. 4. Zonally averaged (a) Ave7-WCDA ACC difference and (b) Ave7/WCDA anomaly RMSE ratio of atmosphere temperature. The right profiles correspond to the global mean values at different levels, and the short red lines indicate 0.5 standard deviation in the meridional direction.

anomaly RMSE ratio of the 50 m ocean temperature and sea surface salinity (SSS). LACC increases the ACCs of 50 m temperature and SSS over most regions, especially in the Pacific Ocean and the Southern Ocean around Australia for both fields and in the North Atlantic for SSS. The improvement in the anomaly RMSE is not as clear as that in the ACC, with a globally averaged RMSE ratio of 1.0 for both variables. Because of the existence of model bias, the cross update may tend to correct the biased model climatology most. For a biased model, the error in climatology is usually larger than that in anomaly. We note that

there exist regions where WCDA outperforms LACC, particularly in some sensitive regions such as the California Current, the Kuroshio, and the coastal Pacific upwelling and stratus-covered region off South America. The improvements in the ACCs of subsurface temperature and salinity are not limited to the near-surface levels; instead, they extend to the deeper ocean (Fig. 7). For example, the maximum improvements in the salinity ACC are restricted mostly to the upper ocean, but they penetrate from the surface downward at approximately 45°S and 40°N (Fig. 7b). In addition, there are also regions where the

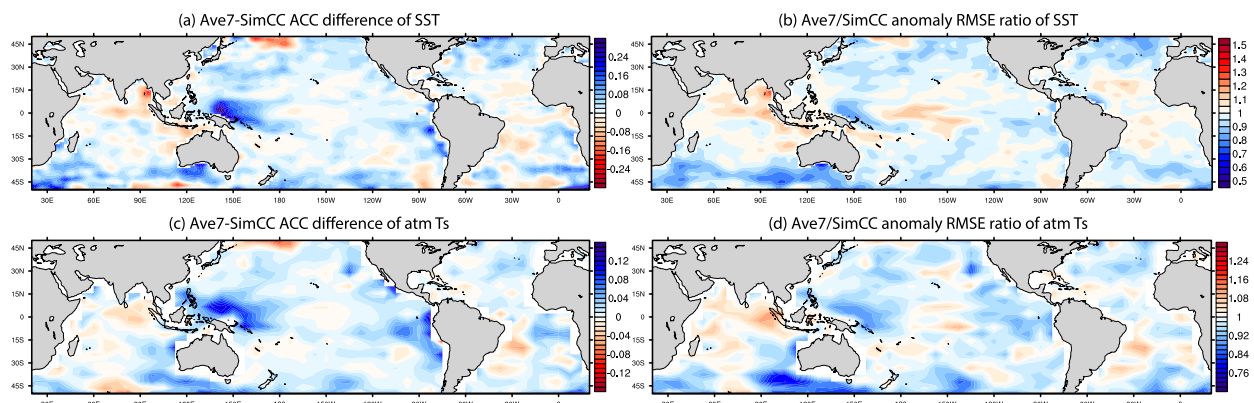


FIG. 5. Ave7-SimCC ACC difference and Ave7/SimCC anomaly RMSE ratio of (a), (b) SST and (c), (d)  $T_s$ .

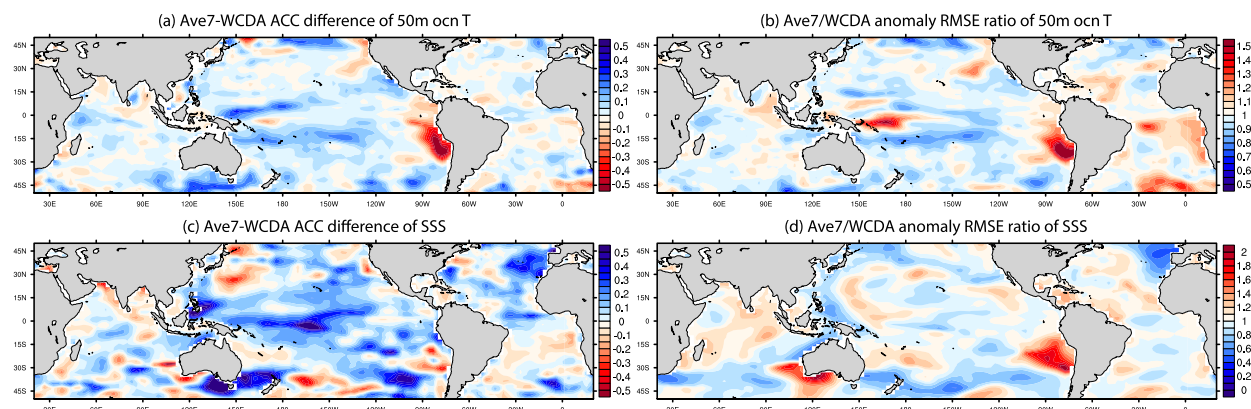


FIG. 6. Ave7-WCDA ACC difference and Ave7/WCDA anomaly RMSE ratio of (a),(b) 50-m ocean temperature and (c),(d) SSS.

ACC is reduced by LACC, such as between 50 and 200 m at approximately 30°N and below 50 m at approximately 40°N for temperature, and below 50 m between 15° and 32°N and below 200 m poleward of 44°N for salinity. When LACC extracts additional information from the lower levels of atmosphere temperature observations to correct SST, the corrected SST further affects subsurface temperature and salinity. This downward and cross-variable impact may be accomplished via oceanic dynamics during the forecast phase, which transfers the corrected surface

signal downward. LACC not only improves the ocean state over WCDA but also can lead to higher ACCs for ocean temperature and salinity than SimCC over significant areas (Fig. 8). Maximum improvements are located mostly in the midlatitudes of both hemispheres and are consistent with the regions where the maximum ocean-atmosphere correlations occur when the atmosphere leads the ocean. The higher global ACC averages of Ave7 than SimCC fail to hold for salinity below 200 m, which is caused mainly by the poorer performance of LACC at approximately 40°S,

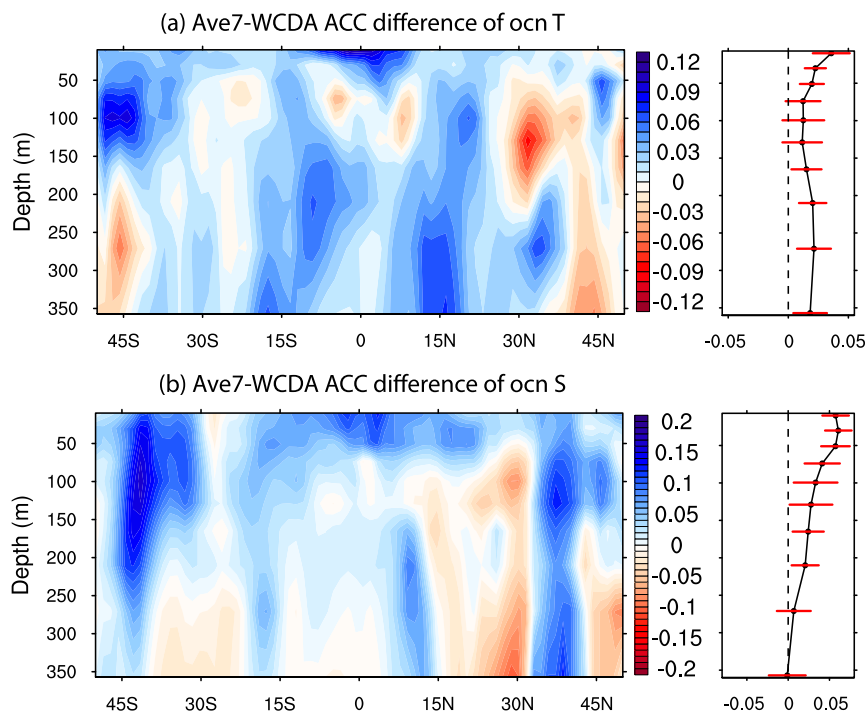


FIG. 7. Zonally averaged Ave7-WCDA ACC difference of ocean (a) temperature and (b) salinity. The right profiles correspond to the global mean values at different levels, and the short red lines indicate 0.5 standard deviation in the meridional direction.



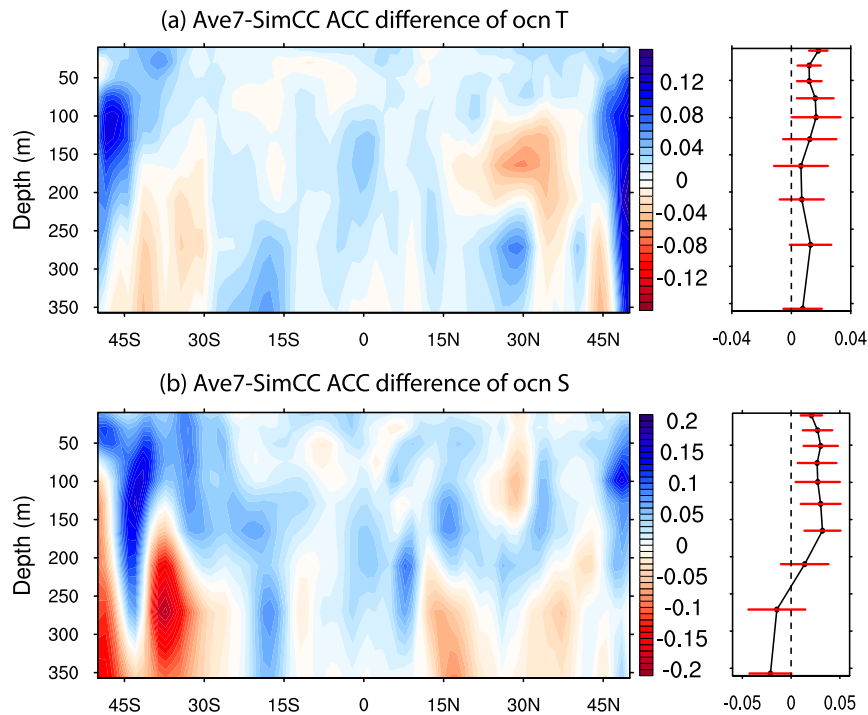


FIG. 8. As in Fig. 7, but for Ave7-SimCC.

details of which may need to be further explored in the future.

#### d. Ensemble forecast verification

Forecast is another effective approach for verifying a DA system without the truth (Laloyaux et al. 2016). Indeed, one ultimate purpose of CDA is to improve the model forecast by providing better initial conditions. We perform 6-month ensemble forecasts initialized from the ensemble analyses of WCDA and LACC every month within 10 consecutive years. The forecast statistics are then evaluated against the unassimilated (future) reanalysis data (i.e., Reanalysis II for the atmosphere and EN4.2.1 for the ocean).

Our experiments show that LACC provides better initial conditions than WCDA and subsequently improves the forecasts. Figure 9 shows the globally averaged ACC and anomaly RMSE for the ensemble forecasts of SST and  $T_s$  over the ocean up to 6 months of forecast lead time. First, the ensemble mean forecasts start from better initial conditions in LACC than in WCDA, with a higher ACC and lower RMSE at the forecast horizon. Second, the improved initial conditions lead to improved forecasts of both SST and  $T_s$  for the first several months. After 3 (4) months, however, the improvements of LACC over WCDA diminish for SST ( $T_s$ ) in terms of RMSE. In addition, improvements in the ACC of SST are almost imperceptible after 3 months. Although the confidence intervals show that the differences between Ave7 and

WCDA are not statistically significant at the 95% level, the ensemble means exhibit consistent improvements for the initialization and short-term forecasts. In addition, the ACCs of WCDA are already high, and our model has a poor forecast ability compared with more advanced operational models, which may make the improvements in LACC even harder to achieve. In all forecasts, the ensemble mean ACCs and anomaly RMSEs are better than those of the means of all 16 members, confirming the merit of averaging various forecasts as is possible with ensemble forecasting. The improvements in the  $T_s$  forecasts indicate the impacts of corrected SST on the atmosphere through dynamic coupling. The improvements are also made in LACC over WCDA for the subsurface ocean temperature and salinity, which can last with longer lead times for the ensemble mean forecasts than the surface. Therefore, an improved analysis of SST by cross update can also benefit the subsurface ocean states through oceanic dynamics. Zonally averaged ACCs and the anomaly RMSEs for both SST and  $T_s$  show that the prominent improvements in ACC are located mostly in the tropics, while improvements are more spatially distributed for RMSE with maximum reductions located mainly in the deep tropics and the midlatitudes of the Southern Hemisphere.

#### e. Innovation analysis

The advantage of LACC is also supported by the improvement in atmospheric prior innovation. Prior

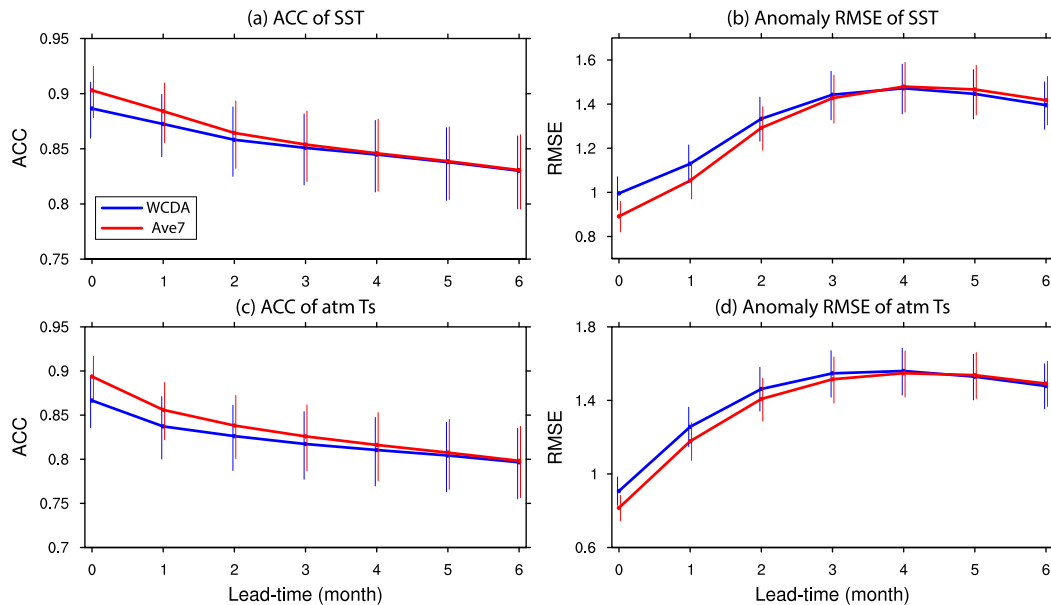


FIG. 9. Globally averaged ACC and anomaly RMSE of (a),(b) SST and (c),(d)  $T_s$  for the ensemble mean forecast (bold lines) and 95% confidence intervals obtained from bootstrap estimates (vertical lines). Blue lines indicate the ensemble mean of WCDA, and red lines indicate Ave7.

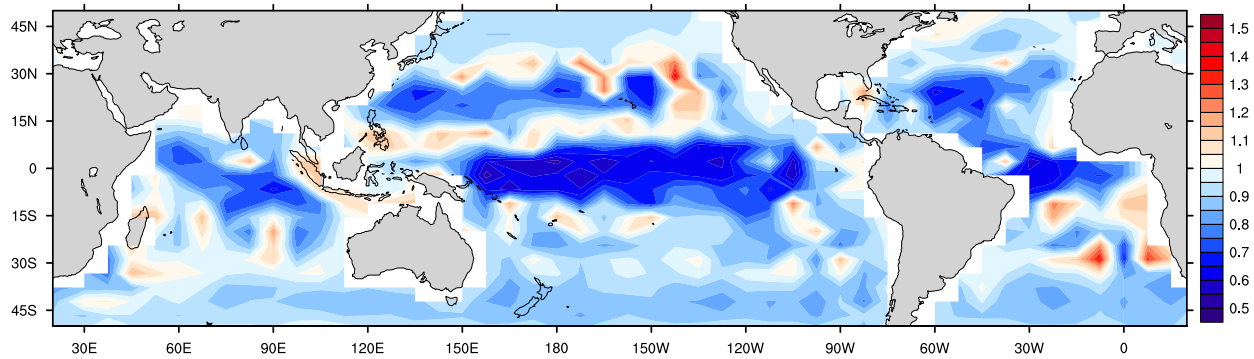
innovation is typically defined as observation minus prior background forecast at the analysis step. Given the atmosphere–ocean cross update in LACC, the innovation for cross update is in the atmosphere. Additionally, atmospheric innovation can reflect the subsequent impact of the cross update on the overlying atmosphere. The RMS ratio of  $T_s$  innovation shows an improvement pattern for LACC over WCDA similar to that in Fig. 3c, indicating that a smaller difference between the observation and the model background can be obtained by LACC. The LACC method reduces the globally averaged RMS of  $T_s$  innovation by 9.7% compared to WCDA, especially in the equatorial regions of the three oceans, in the northern subtropical Atlantic, and in the Southern Ocean (Fig. 10). There also exist regions where the RMS of the atmospheric innovation is increased by LACC, such as the narrow belts of the North Pacific at approximately  $10^\circ$  and  $30^\circ\text{N}$  and the southern subtropical Atlantic. We note that the innovation calculated here is not exactly the same as that defined above. Our model outputs the analyses, instead of the forecasts, at the analysis steps. Therefore, the atmospheric innovation used here is calculated from the time-averaged observations and daily mean forecasts between every two cross update steps. This result should be quantitatively close to the theoretical innovation because only the days on which the cross update is conducted are excluded from the calculation. In addition, the RMS ratio of the SST innovation, which has a globally averaged value of 0.984, is also calculated (not shown). It shows a

quite similar pattern to the full-value RMSE ratio of SST in Fig. 2c, which agrees with the climatology-dominated feature mentioned above.

#### f. Sensitivity experiments

The averaging length is a centrally important parameter to the performance of the LACC method. To provide a thorough evaluation of our SCDA method LACC, sensitivity experiments for different averaging lengths are also performed. Figure 11 shows the zonally averaged SCDA–WCDA ACC differences in SST,  $T_s$ , vertically averaged ocean temperature and salinity for different averaging lengths in LACC. “AveX” means that the ocean–atmosphere cross correlations are calculated between SST and the average of  $X$  daily mean atmosphere temperature from  $X - 1$  days ago to the current day. For the Ave1 and SimCC experiments in this study, the same atmosphere temperature observations of the current day are used. The prior ensemble of the current time step is used in SimCC, while that of the current daily mean is used in Ave1.

LACC can improve the ACC of SST over the benchmark WCDA and even SimCC across almost the entire domain when an appropriate averaging length (e.g., larger than 3) is selected (Fig. 11a). The positive impact of LACC shows a consistent meridional distribution as in the perfect model framework, where maximum improvements in LACC are located in the low-latitude regions (Fig. 11a versus Fig. 5 in Part II). The SimCC method performs better than WCDA mostly in the

FIG. 10. Ave7/WCDA RMS ratio of  $T_s$  innovation.

low-latitude regions, where the simultaneous cross correlations are large, but poorly in the midlatitude regions, where the maximum cross correlations occur when the atmosphere leads the ocean. LACC benefits from this asymmetric feature of cross correlation and provides better performance than SimCC in the midlatitudes. When a small averaging length is chosen (e.g., 1), LACC still performs poorly in the midlatitudes, and its meridional distribution is quite similar to that of SimCC. As the averaging length increases, LACC starts to exhibit a consistent improvement over other DA configurations. The improvements in  $T_s$  with LACC versus WCDA are significant and robust across all latitudes for all the  $X$ s adopted here (Fig. 11b). Compared with  $T_s$ , SST is more sensitive to the choice of  $X$ , implying that the cross

update of SST is more sensitively affected by the lead time of the atmosphere.

Subsurface ocean temperature and salinity are also compared with EN4.2.1 for different averaging lengths. LACC shows consistent improvements for the globally averaged ACC over WCDA and SimCC at the subsurface but has a more complicated meridional distribution (Figs. 11c,d). The complex patterns of the improvements in subsurface ocean temperature and salinity may be caused by the indirect influence of the adjusted SST and the complexity of oceanic dynamics. Here, we note that the ACCs of the vertically averaged ocean temperature and salinity are relatively small. This may be caused by the relatively poor simulation performance of FOAM in comparison with more advanced CGCMs. In addition,

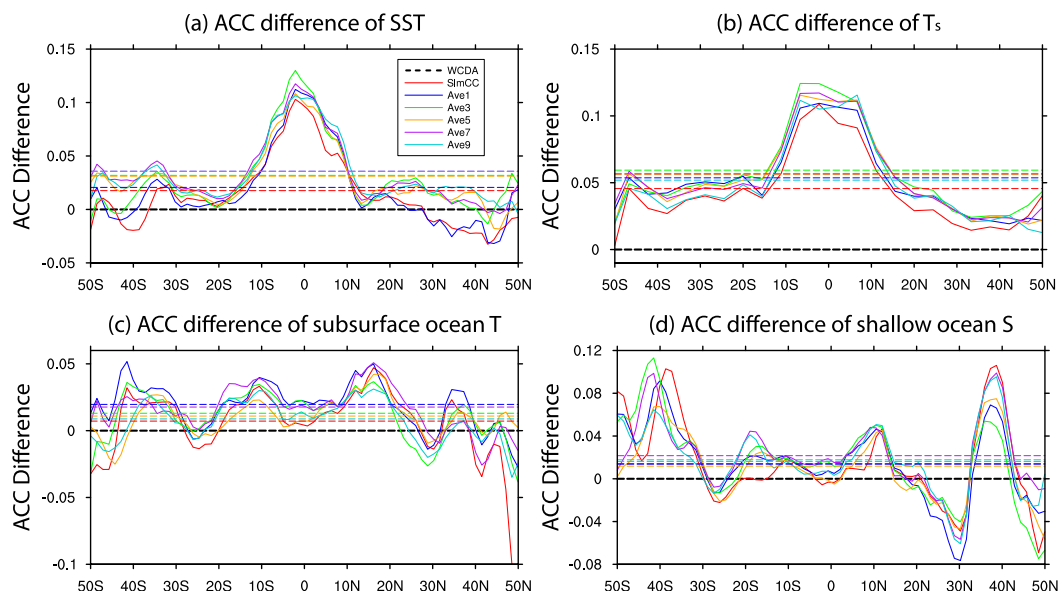


FIG. 11. Zonally averaged SCDA-WCDA ACC difference of (a) SST, (b)  $T_s$ , (c) vertically averaged subsurface ocean temperature from the second model level to 350 m, and (d) vertically averaged ocean salinity from the surface to 350 m. “Ave $X$ ” indicates that the cross correlation is calculated between SST and the average of  $X$  daily mean atmosphere temperature values in the bottom four levels from  $X - 1$  days ago to the current day in LACC. The corresponding global mean values are illustrated as horizontal dashed lines with the same colors.

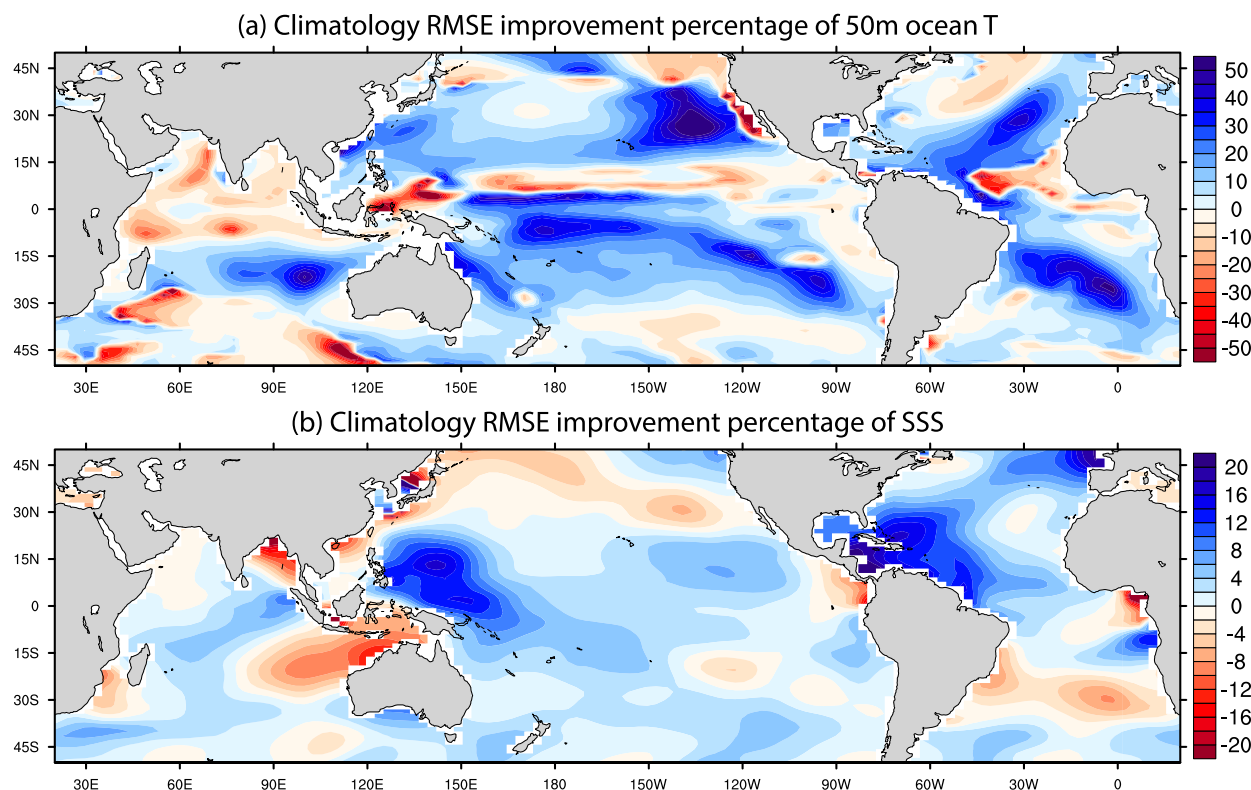


FIG. 12. Improvement percentage in Ave7 climatology RMSE of (a) 50-m ocean temperature and (b) SSS. The improvement is calculated with respect to WCDA.

no subsurface observation is assimilated in the experiments, which also weakens the correction to the subsurface. Overall, the improvements of LACC compared with WCDA and SimCC are still robust in terms of the globally averaged ACC (the horizontal dashed lines). Experiments using different values of atmospheric and oceanic observation errors are also conducted, with consistent results here, ensuring that our conclusions are not sensitive to the choice of observation errors (not shown).

#### g. Correction of climatology

The SCDA system using LACC can correct the biased model climatology as well. In the real world scenario, model climatology often drifts away from the observed climatology to converge toward its own climatology. Figure 12 shows the improvement percentage in the Ave7 climatology RMSE of 50m ocean temperature and SSS with respect to the WCDA experiment. LACC can correct the climatology RMSE of WCDA over most regions, especially in the Pacific and Atlantic Oceans. Compared with WCDA, the globally averaged climatology RMSEs of the 50m temperature and SSS are reduced by 7.4% and 1.8%, respectively. Analyses of ocean temperature, salinity, and air temperature at

different model levels also confirm the robustness of the correction effect to the biased model climatology. As illustrations of the correction to the biased climatology by LACC, Fig. 13 shows the seasonal cycle climatology of 50m ocean temperature at four model grid points. The first two points show clearly positive corrections to the biased climatology, while the third and fourth points show neutrally positive and negative corrections, respectively. Overall, LACC can reduce the error of WCDA climatology in most regions, leading to a reduced globally averaged climatology RMSE.

## 4. Summary and discussion

Our study confirms the effectiveness of the SCDA scheme LACC in the assimilation of real world reanalysis data. To the best of our knowledge, this is the first successful application of a SCDA system with a CGCM for the assimilation of reanalysis data. Both ACC and RMSE are used to evaluate the DA results. The LACC scheme is evaluated in four comparison schemes: using the assimilated variables, using the unassimilated variables, ensemble forecast verification, and prior innovation analysis. LACC is shown to significantly improve the ACCs of model variables in both



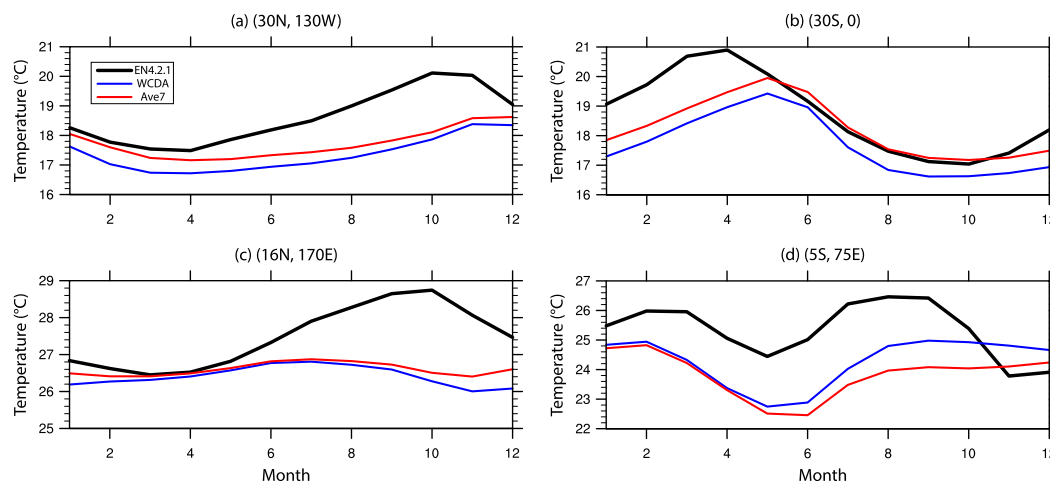


FIG. 13. Seasonal climatology of 50-m ocean temperature at (a) (30°N, 130°W), (b) (30°S, 0°), (c) (16°N, 170°E), and (d) (5°S, 75°E). Black lines indicate the climatology of EN4.2.1 analysis data; blue lines, WCDA; and red lines, Ave7.

the ocean and atmosphere in comparison with WCDA as well as SimCC. The biased model climatology is also improved over WCDA. In the meantime, there are also regions where the analysis is deteriorated by the cross update of LACC, such as in the western and eastern equatorial Pacific, the subtropical Pacific centered at approximately 30°N, and the tropical to subtropical Atlantic of the SST climatology, and in broader regions with a rather uniform distribution of the SST anomaly. These negative impacts may be caused by an inappropriate averaging length of LACC for such regions. A better analysis may be expected if an adaptive averaging length can be adopted for different regions. Overall, our evaluation in this study demonstrates the potential effectiveness of our SCDA system using the LACC approach in the presence of significant model bias.

Many questions remain to be further studied. For convenience, this study uses gridded reanalysis data as the assimilated observations. Liu et al. (2016) pointed out that the substitution of reanalysis data for raw observations could affect the analysis quality. Further studies are needed to test LACC using raw observations. In addition, our SCDA system using LACC allows only one-way cross update from the atmosphere to the ocean. It would be useful to further explore SCDA schemes in which ocean observations are used to update the atmosphere. Finally, our model involves substantial model bias and has a relatively low resolution. It would be highly desirable to repeat this study in more advanced coupled climate models.

**Acknowledgments.** We thank three anonymous reviewers for their insightful suggestions and comments.

We thank Drs. Chenyu Zhu, Chengfei He, Yuchu Zhao, and Yishuai Jin for their helpful discussions. We declare no conflicts of interest. NCEP–NCAR Reanalysis and NOAA\_ERSST\_V5 data are provided by the NOAA/OAR/ESRL PSD from their website at <https://www.esrl.noaa.gov/psd/>. We acknowledge high-performance computing support from Cheyenne (doi:10.5065/D6RX99HX) provided by NCAR's CISL. This research is sponsored by the National Key R&D Programs of China (2018YFC1406202 and 2018YFC1506704), Chinese MOST 2017YFA0603801, NSFC 41830964, NSFC 41605070, and U.S. NSF AGS-1656907.

## REFERENCES

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2).
- , 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, [https://doi.org/10.1175/1520-0493\(2003\)131<0634:ALLSFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2).
- Bonavita, M., L. Isaksen, and E. Hölm, 2012: On the use of EDA background error variances in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **138**, 1540–1559, <https://doi.org/10.1002/qj.1899>.
- Carton, J. A., G. A. Chepurin, and L. Chen, 2018: SODA3: A new ocean climate reanalysis. *J. Climate*, **31**, 6967–6983, <https://doi.org/10.1175/JCLI-D-18-0149.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, <https://doi.org/10.1029/94JC00572>.
- Frolov, S., C. H. Bishop, T. Holt, J. Cummings, and D. Kuhl, 2016: Facilitating strongly coupled ocean–atmosphere data

- assimilation with an interface solver. *Mon. Wea. Rev.*, **144**, 3–20, <https://doi.org/10.1175/MWR-D-15-0041.1>.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, <https://doi.org/10.1002/qj.49712555417>.
- Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Oceans*, **118**, 6704–6716, <https://doi.org/10.1002/2013JC009067>.
- Han, G., X. Wu, S. Zhang, and E. Al, 2013: Error covariance estimation for coupled data assimilation using a Lorenz atmosphere and a simple pycnocline ocean model. *J. Climate*, **26**, 10218–10231, <https://doi.org/10.1175/JCLI-D-13-00236.1>.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, [https://doi.org/10.1175/1520-0493\(1998\)126<0796:DAUAEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2).
- Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, <https://doi.org/10.1175/JCLI-D-16-0836.1>.
- Jacob, R., 1997: Low frequency variability in a simulated atmosphere ocean system. Ph.D. thesis, University of Wisconsin–Madison, 159 pp.
- , C. Schafer, I. Foster, M. Tobis, and J. Anderson, 2001: Computational design and performance of the Fast Ocean Atmosphere Model, version one. *Computational Science—ICCS 2001*, V. N. Alexandrov et al., Eds., Lecture Notes in Computer Science, Vol. 2073, Springer, 175–184, [https://doi.org/10.1007/3-540-45545-0\\_26](https://doi.org/10.1007/3-540-45545-0_26).
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, [https://doi.org/10.1175/1520-x0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-x0477(1996)077<0437:TNYRP>2.0.CO;2).
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1644, <https://doi.org/10.1175/BAMS-83-11-1631>.
- Laloyaux, P., M. Balmaseda, D. Dee, K. Mogensen, and P. Janssen, 2016: A coupled data assimilation system for climate reanalysis. *Quart. J. Roy. Meteor. Soc.*, **142**, 65–78, <https://doi.org/10.1002/qj.2629>.
- Liu, H., F. Lu, Z. Liu, Y. Liu, and S. Zhang, 2016: Assimilating atmosphere reanalysis in coupled data assimilation. *J. Meteor. Res.*, **30**, 572–583, <https://doi.org/10.1007/s13351-016-6014-1>.
- Liu, Y., Z. Liu, S. Zhang, R. Jacob, F. Lu, X. Rong, and S. Wu, 2014: Ensemble-based parameter estimation in a coupled general circulation model. *J. Climate*, **27**, 7151–7162, <https://doi.org/10.1175/JCLI-D-13-00406.1>.
- Liu, Z., Q. Zhang, and L. Wu, 2004: Remote impact on tropical Atlantic climate variability: Statistical assessment and dynamic assessment. *J. Climate*, **17**, 1529–1549, [https://doi.org/10.1175/1520-0442\(2004\)017<1529:RIOTAC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<1529:RIOTAC>2.0.CO;2).
- , and Coauthors, 2007a: Simulating the transient evolution and abrupt change of Northern Africa atmosphere–ocean–terrestrial ecosystem in the Holocene. *Quat. Sci. Rev.*, **26**, 1818–1837, <https://doi.org/10.1016/j.quascirev.2007.03.002>.
- , Y. Liu, L. Wu, and R. Jacob, 2007b: Seasonal and long-term atmospheric responses to reemerging North Pacific Ocean variability: A combined dynamical and statistical assessment. *J. Climate*, **20**, 955–980, <https://doi.org/10.1175/JCLI4041.1>.
- , S. Wu, S. Zhang, Y. Liu, and X. Rong, 2013: Ensemble data assimilation in a simple coupled climate model: The role of ocean–atmosphere interaction. *Adv. Atmos. Sci.*, **30**, 1235–1248, <https://doi.org/10.1007/s00376-013-2268-z>.
- Lu, F., Z. Liu, S. Zhang, and Y. Liu, 2015a: Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part I: Simple model study. *Mon. Wea. Rev.*, **143**, 3823–3837, <https://doi.org/10.1175/MWR-D-14-00322.1>.
- , —, —, —, and R. Jacob, 2015b: Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part II: CGCM experiments. *Mon. Wea. Rev.*, **143**, 4645–4659, <https://doi.org/10.1175/MWR-D-15-0088.1>.
- , —, Y. Liu, S. Zhang, and R. Jacob, 2017a: Understanding the control of extratropical atmospheric variability on ENSO using a coupled data assimilation approach. *Climate Dyn.*, **48**, 3139–3160, <https://doi.org/10.1007/s00382-016-3256-7>.
- , —, S. Zhang, and R. Jacob, 2017b: Assessing extratropical impact on the tropical bias in coupled climate model with regional coupled data assimilation. *Geophys. Res. Lett.*, **44**, 3384–3392, <https://doi.org/10.1002/2017GL072890>.
- Penny, S. G., and T. M. Hamill, 2017: Coupled data assimilation for integrated earth system analysis and prediction. *Bull. Amer. Meteor. Soc.*, **98**, ES169–ES172, <https://doi.org/10.1175/BAMS-D-17-0036.1>.
- , and Coauthors, 2017: Coupled data assimilation for integrated Earth system analysis and prediction: Goals, challenges and recommendations. WMO Tech. Rep. WWRP 2017-3, 59 pp., [https://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final\\_WWRP\\_2017\\_3\\_27\\_July.pdf](https://www.wmo.int/pages/prog/arep/wwrp/new/documents/Final_WWRP_2017_3_27_July.pdf).
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1058, <https://doi.org/10.1175/2010BAMS3001.1>.
- Sluka, T. C., S. G. Penny, E. Kalnay, and T. Miyoshi, 2016: Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophys. Res. Lett.*, **43**, 752–759, <https://doi.org/10.1002/2015GL067238>.
- Smith, P. J., A. S. Lawless, and N. K. Nichols, 2017: Estimating forecast error covariances for strongly coupled atmosphere–ocean 4D-Var data assimilation. *Mon. Wea. Rev.*, **145**, 4011–4035, <https://doi.org/10.1175/MWR-D-16-0284.1>.
- , —, and —, 2018: Treating sample covariances for use in strongly coupled atmosphere–ocean data assimilation. *Geophys. Res. Lett.*, **45**, 445–454, <https://doi.org/10.1002/2017GL075534>.
- Sugiura, N., T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa, 2008: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res.*, **113**, C10017, <https://doi.org/10.1029/2008JC004741>.
- Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, [https://doi.org/10.1175/1520-0493\(2004\)132<1238:IOIEAO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2).
- Zhang, S., M. J. Harrison, A. T. Wittenberg, A. Rosati, J. L. Anderson, and V. Balaji, 2005: Initialization of an ENSO forecast system using a parallelized ensemble filter. *Mon. Wea. Rev.*, **133**, 3176–3201, <https://doi.org/10.1175/MWR3024.1>.
- , —, A. Rosati, and A. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, **135**, 3541–3564, <https://doi.org/10.1175/MWR3466.1>.