Development of Coupled Data Assimilation With the BCC Climate System Model: Highlighting the Role of Sea-Ice Assimilation for Global Analysis

X. Liu, J. Yao, T. Wu, S. Zhang, F. Xu, L. Zhang, W. Jie, W. Zhou, Q. Li, X. Liang, M. Chu, J. Yan, S. Nie, and Y. Cheng

1National Climate Center, China Meteorological Administration, Beijing, China, 2Key Laboratory of Physical Oceanography, Ministry of Education/Institute for Advanced Ocean Study/Center for Deep Ocean Multispheres and Earth System (DOMES), Ocean University of China, Qingdao, China, 3Pilot National Laboratory for Marine Science and Technology (QNLM), Qingdao, China, 4International Laboratory for High-Resolution Earth System Model and Prediction (iHESP), Qingdao, China, 5Ministry of Education Key Laboratory for Earth System Modeling, and Department of Earth System Science, Tsinghua University, Beijing, China, 6Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), China, 7School of Atmospheric Science, Nanjing University, Nanjing, China, 8South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China

Abstract The coupled data assimilation (CDA) system consisting of ocean, sea-ice, and atmosphere data assimilation components with the Beijing Climate Center (BCC) Climate System Model has been developed to provide reliable analyses of the atmosphere, ocean, and sea-ice states. It incorporates ocean temperature/salinity profiles, sea surface temperature, sea level height, and sea-ice concentration observations at a daily frequency, and atmosphere reanalysis at a 6-hourly frequency. Results show that the system is capable of realistically reproducing the climatology and variability of ocean, sea-ice, and atmosphere. The performances in analyzing ocean component are comparable to those of well-known ocean reanalyses that were once used to initialize the BCC model for climate predictions. A series of experiments with and without sea-ice observations in the CDA framework are designed to explore the role of sea-ice data assimilation (DA). The addition of sea-ice DA exerts very small influence to the analysis of upper ocean temperature over the Arctic area, but leads to an evident reduction of temperature error in the upper 1000 m of ocean south of 60°S. Particularly, only the inclusion of sea-ice DA can make the ocean temperature/salinity profiles reliable. On the other hand, adding sea-ice DA on the basis of ocean DA can improve the analysis of tropical troposphere variability in the tropical troposphere and mid- and high-latitude stratospheric atmosphere. These results address the importance of coordination of sea-ice observations and ocean observations in CDA.

Plain Language Summary Developing coupled data assimilation (CDA) technique has become an important task for many operational and research centers. Although CDA consisting of multiple assimilation components has made great progress in the past years, the impacts of each individual component, especially the sea-ice data assimilation (DA), are not fully addressed yet. In this study, we have developed a CDA system and implemented a coordinated assimilation scheme of ocean, sea-ice, and atmosphere data. This scheme shows reliable performances in analyzing the states of the ocean, sea-ice, and atmosphere. A series of experiments are conducted to examine the impacts of sea-ice DA on climate analysis. We stress that the importance of sea-ice DA can be highlighted only in a multicomponent coordinated assimilation framework. On one hand, on the basis of ocean/ocean-atmosphere DA, addition of sea-ice DA can decrease the ocean temperature error in the upper 1000 m of the high-latitude Southern Ocean. On the other hand, adding sea-ice DA on the basis of ocean DA can improve the analysis of atmospheric variability in the tropical troposphere and mid- and high-latitude stratosphere. These findings call for more attentions of operational model developers to comprehensively describe the role of sea-ice DA in climate analysis and forecast.
1. Introduction

Coupled climate models have been widely used in predicting climate variabilities from subseasonal to decadal scales. A main challenge for coupled model forecast is the coordinated initialization of multiple components in the coupled model. Originally, separated data assimilation (DA) schemes are applied on uncoupled simulation systems and their products are used to initialize the corresponding components in a coupled model. This strategy was extensively used by many operational climate forecast centers, such as the European Center of Medium-Range Weather Forecast (ECMWF) prediction system (Molteni et al., 2011), the National Center for Environmental Prediction (NCEP) climate forecast system (Saha et al., 2006), the Met Office global prediction system (MacLachlan et al., 2015), the Japan Meteorological Agency coupled prediction system (Takaya et al., 2018), and the Australia seasonal forecast system (Alves et al., 2003; Yin et al., 2011). In the absence of a consistent assimilation scheme, some models used uncoupled assimilation reanalysis products from external sources to initialize individual model components (e.g., Liu et al., 2015; Merryfield et al., 2013; Xin et al., 2018).

To reduce the possible inconsistency arising from uncoupled DA and initialize the multiple components in a more coordinated way, DA methods based on a coupled model, referred to as coupled DA (CDA), are proposed. They have shown advantages in improving the analysis of the assimilation component and its interaction with other components in coupled model (e.g., Chang et al., 2013; Fuji & Kamachi, 2003; Zhang et al., 2007). In extensive practices, developing CDA system based on a complex coupled model and applying it in numerical prediction have become an important task for many operational and research centers. Generally, assimilation is either applied to each component of a coupled model independently (i.e., weakly CDA), or applied to several components simultaneously and treats these components as one single integrated system (i.e., strongly CDA) (Penny et al., 2017; Zhang et al., 2020). The Japan Agency for Marine-Earth Science and Technology (JAMSTEC) developed an ocean-atmosphere coupled four-dimensional variational DA system and used it for seasonal and decadal predictions (Mochizuki et al., 2016; Sugiuura et al., 2008). The NCEP built a weakly CDA system in which ocean and atmosphere observations were independently assimilated using variational methods and sea-ice and land data were introduced using simple regridding strategy (Saha et al., 2010). This system was used to initialize the NCEP subseasonal and seasonal forecast operational model. Meanwhile, a strongly CDA framework was developed and tested in the NCEP (Sluka et al., 2016). The U.K. Met Office also made separate ocean-sea ice and atmosphere-land variational analyses in a coupled climate system model, and thus constructed a weakly CDA system for predicting multiscale climate (Lea et al., 2015). The ECMWF used increment variational algorithm to estimate ocean and atmospheric increments separately, and then applied additional iterations of the coupled model to increase the strength of the coupling in the analysis update, and thus built a weakly or quasi-strongly CDA system (Laloyaux et al., 2016, 2018). Using the fully coupled models at the National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory (NOAA/GFDL) and the Chinese Academy of Science/Institute of Atmospheric Physics (CAS/IAP), weakly CDA schemes based on different assimilation methods were also built for decadal climate predictions (e.g., He et al., 2017; Wu et al., 2018; Yang et al., 2013). In addition, weakly CDA schemes based on ocean-atmosphere coupled model were also developed to initialize the numerical weather forecast at the U.S. Naval Research Laboratory and ECMWF, respectively (e.g., Browne et al., 2019; Holt et al., 2011).

Although the CDA is developing toward an incorporation of various earth system components including atmosphere, ocean, sea-ice, and land, understanding the contributions and limitations of assimilation of each individual component to the coupled multicomponent analysis and forecast is not yet comprehensive. Especially, relative to much attentions to impacts of ocean and atmosphere DA (e.g., Balmaseda & Anderson, 2009; Chang et al., 2013; Fuji & Kamachi, 2003; Zhang et al., 2007), the role of sea-ice DA in fully coupled models is little explored, although sea-ice has been regarded as an important component for CDA systems (e.g., Lea et al., 2015; Saha et al., 2010). Using ocean-sea ice models, sea-ice data were often assimilated to improve the analysis and prediction of sea-ice state (e.g., Fritzner et al., 2019; Yang et al., 2015). Several studies also tried to assimilate sea-ice observations into fully coupled models to improve prediction of Arctic sea-ice (e.g., Chen et al., 2017; Yang et al., 2020). With weakly/strongly coupled ocean-sea ice DA scheme in a fully coupled earth system model, Kimmritz et al. (2018, 2019) revealed the obvious advantages of sea-ice DA for sea-ice analysis and prediction. Meanwhile, they found that sea-ice DA could clearly improve the
ocean temperature in the mixed layer in the idealized perfect-model experiments, but have little improvement of the upper ocean in the real-observation assimilation experiments. Till now, the impacts of sea-ice DA on the analysis and prediction of ocean and atmosphere components in coupled global climate models are still not fully understood.

In this study, we aim to build a weakly CDA scheme with the Beijing Climate Center (BCC) Climate System Model (BCC-CSM), and further address the impact of DA of individual component, particularly the sea-ice, on the analysis of climate system. The out-of-domain influences of sea-ice DA in the framework with and without other assimilation components in coupled model are mainly explored. Also, the BCC-CSM used in climate prediction operation is still greatly limited by immature initialization scheme that adopts uncoupled DA or inconsistent-with-model reanalysis products. Thus another goal of this study is to build a CDA scheme to provide a reliable analysis that is comparable to other reanalysis products, and further replace the original initialization scheme.

The rest of this paper is organized as follows. Details of the model, assimilation scheme, experiment designs and validation data are provided in Section 2. In Sections 3, we show ocean-sea-ice analysis and the role of sea-ice DA; and in Section 4, we explore the role of sea-ice DA in atmosphere analysis. Conclusions are given in Section 5.

2. Model, Assimilation Scheme, and Experimental Design

2.1. Model

The model used in this study is the BCC-CSM version 2 with medium resolution. It is an atmosphere-land-ocean-sea ice coupled model. The atmosphere component is the BCC Atmospheric General Circulation Model version 3 with T106 triangular truncation (resolution of approximate 110 km) and 46 vertical hybrid sigma/pressure layers. The land component is the BCC Atmosphere and Vegetation Interaction Model version 2 with T106 horizontal resolution and 10 soil layers. The ocean component is the GFDL Modular Ocean Model version 4 with varying horizontal resolution of 1/3° (at the equator) to 1° (at the pole), and the sea-ice component is the GFDL Sea Ice Simulator with the same resolution as the ocean component. All components are coupled at a frequency of 30 min without any flux adjustment. Details of the coupled model and its general performance in the Coupled Model Intercomparison Project Phase 6 (CMIP6) have been documented in the study by Wu et al. (2019).

The BCC-CSM has been widely used in climate prediction operation. During the process when the model is participating in the subseasonal to seasonal (S2S) prediction project and in the meanwhile providing routine climate forecasts, the ocean component of the model is initialized by several sets of external-source reanalysis products, including the BCC uncoupled ocean assimilation reanalysis product for the S2S prediction (Liu et al., 2017, 2019), the NCEP Ocean Data Assimilation System (GODAS) product for the season-to-interannual prediction (Liu et al., 2015), and the Simple Ocean Data Assimilation (SODA) product for the decadal prediction (Xin et al., 2018). These practices may produce inconsistency in initial conditions among various components of the model and further give a negative impact on the performance of coupled forecast. Therefore, it is quite necessary to develop more consistent data assimilation scheme and produce more reliable analysis than the above reanalysis products, and thus improve the model initialization and prediction.

2.2. Data and Preprocesses

Multisource observations from different datasets are assimilated in this study. They include: (1) ocean temperature/salinity (T/S) profile observations from the Array for Real-Time Geostrophic Oceanography (ARGO) at the ARGO data assembly center in France; (2) ocean T/S profile observations from the Global Temperature and Salinity Profile Program (GTSSP; Sun et al., 2010); (3) sea surface temperature (SST) and sea-ice concentration (SIC) observations from the NOAA daily Optimum Interpolation Sea Surface Temperature dataset (OISST; Reynolds et al., 2007); and (4) sea level anomaly (SLA) observation from the AVISO (Archiving, Validation, and Interpretation of Satellite Oceanographic data). In addition, the atmospheric
temperature, zonal wind, meridional wind, and humidity fields from the ERA-Interim reanalysis dataset (Dee et al., 2011) are used.

To make an effective assimilation of the observations, a series of preprocesses including observation data quality control, data merging and conversion are first conducted. Some main steps are as follows:

1. Procedures of quality control of T/S profiles are conducted for the ARGO and GTSPP data respectively. They include the reliability judgment on observation date, depth, geographical location and value range, reasonableness check on the disparity of T/S observations relative to climatological data, and abnormality screening of vertical gradients of T/S observations in adjacent layers. Through these steps, all unreliable observation records are removed.

2. To get coordinated observational T/S fields and assimilate them into the model all at once, effective merging of the ARGO and GTSPP data is necessary. These two kinds of profiles, except for the ones with less-than-4 levels, are combined and interpolated onto the vertical layers of the ocean model. Then, at each level the observation records with extremely close locations to other observations, or with considerably large differences with the average of several surrounding observations, are filtered out. The assimilation scheme in our study allows assimilating the original profiles, but the vertical interpolation is done here. This preprocessing step can eliminate the conflict of different data on the same date and assure the coherence of observations in the horizontal space as much as possible, so that we can better assimilate all the observations simultaneously without worrying about the data sequence.

3. Given that the ocean and sea-ice model resolution over the extratropics is 1°, a thinning of observation for the gridded daily SST, SIC and SLA products at 1/4° resolution is done. Also, the SLA field is converted to absolute sea surface height (SSH) by adding the climatological mean of SSH from a multiyear free run of the model, in order to avoid the mismatch of reference height between model and observation. During this process, the high-resolution observations are unified using the model resolution for the following consideration: The assimilation analysis needs to transit the model background field from model space into observation space for obtaining the observation increment. The assimilation scheme in this study allows a bilinear interpolation or an inverse distance weighting interpolation, which may not be the best choice for interpolation of irregular grid, e.g., the tripolar geographical coordinate with a curvilinear grid beyond 60°N in the ocean model we used. Thus, the gridded observations are preprocessed onto the model resolution using the software of NCAR Command Language (NCL) version 6.5, so as to simplify the interpolation computation in the assimilation module and assure the precision of observation increment.

### 2.3. Assimilation Scheme

We designed different approaches for assimilating ocean, sea-ice, and atmosphere data, respectively. Each approach operates on one component model only, and exerts its influence on other component models through the coupling process of different components in the climate system model, thus constructing a weakly CDA system (Figure 1).
2.3.1. Ocean Assimilation

We design an ocean ensemble assimilation approach that consists of 100 static ensemble members. It is based on a combination of the Ensemble Optimal Interpolation (EnOI; Oke et al., 2005, 2008) and the Local Ensemble Transform Kalman Filter (LETKF; Hunt et al., 2007; Miyoshi et al., 2010); both methods are based on the Ensemble Kalman Filter (EnKF; Evensen, 2003). The EnOI is cost-effective in estimating background error covariance using static ensemble, and the LETKF provides efficiency in the amount of computation and memory usage of the analysis. Thus the ocean assimilation scheme in this study is efficient in computation and easy to use operationally.

The analysis equations are given below:

\[
x^a = x^b + K \left[ y^o - Hx^b \right]
\]

(1)

\[
K = P^b H^T \left[ H P^b H^T + R \right]^{-1}
\]

(2)

where \( x = \left[ T_{mk}, S_{mk}, u_{mk}, v_{mk}, \eta_{mk} \right]^T \) denotes the state vector of multiple variables, including temperature, salinity and currents at all grid points \((m)\) and vertical levels \((k)\), and SSH at horizontal grid points \((m)\). \( x^a \) is the estimate of analysis state, \( x^b \) is model background state, \( y^o \) is observation, \( H \) is an operator that interpolates model space into observation space, and \( K \) is a gain matrix that is determined by the background error covariance matrix \( P^b \) and the observation error covariance matrix \( R \). Here the observation errors are considered to be spatially uncorrelated, and thus \( R \) is a diagonal matrix. Its diagonal elements are given with fixed observation error variance based on instrumental errors.

\( P^b \) is estimated by multiple ensemble members and is defined as follows,

\[
P^b = \frac{1}{N-1} A^T A
\]

(3)

where \( A^b = A^b - \overline{A}^b \) is a matrix of ensemble perturbation relative to the ensemble average, \( A^b = \left[ x^b_1, x^b_2, \ldots, x^b_N \right] \) is the ensemble of model background state, \( x^b_i \) is \( i \)th member, and \( N \) is the number of ensemble members.

Combining Equations 1–3, the gain matrix \( K \) can be expressed by

\[
K = A^b A^T H^T \left( H A^b A^T H^T + (N-1)R \right)^{-1}
\]

(4)

The EnOI method uses a static ensemble of the members, instead of the dynamic ensemble of the members used in the EnKF method. In this study, 100 stationary members from a long-term free run of the model are adopted. To achieve this, using an initial condition at the beginning of this century output by the CMIP6 historical simulation experiment, the coupled model is integrated from 2000 to 2015, with prescribed external forcing used in the CMIP6 historical simulation before 2014 and SSP2-4.5 projection beyond 2014. We then select the continuous outputs of ocean temperature, salinity, currents, and SSH on the first day of each month during 2006–2015, and produce the anomaly fields after removing the monthly climatological mean of each variable over the 10 years. With a monthly interval, 100 static members are extracted from the anomaly fields in time sequence. These members cover all the seasons and do not vary with time during the assimilation integration.

As in Equations 1 and 2, we construct a whole background covariance matrix of multiple variables. Given that the background covariance is estimated from the model historical simulation and may comprise false multivariate relationship, the correlations between different ocean variables are checked. We find that the relationship between SSH and subsurface T/S is seriously overestimated, thus leading to a strong adjustment of T/S analysis field due to assimilation of SSH observations. Given that the accuracy of T/S analysis
is mainly contributed by the assimilation of T/S profile observations and the reference height of SSH observations does not necessarily agree with the current reality, we choose not to update the T/S analysis by SSH observation assimilation in the ocean DA scheme. For T/S and currents, the correlations between them are weak in most oceans except for some limited areas, e.g., the central equatorial Pacific, where an apparent positive correlation between surface temperature and zonal current reasonably exists because of the joint occurrence of eastward/westward current and warming/cooling sea water during the El Niño/La Niña phase. Finally, we adopt a background covariance matrix of ocean temperature, salinity, and currents to jointly update these variables, but update the SSH independently. On this basis, we still adopt a localization scheme to limit the impacts of possible false covariance and meanwhile reduce the cost of assimilation computation. Here we use a similar strategy for localization of observation data and solution of gain matrix as the LETKF scheme in Miyoshi et al. (2010), Miyazawa et al. (2012) and Xu et al. (2013). The localization is done on the horizontal and vertical directions of model coordinates. The horizontal localization scale \( L_h \) is measured by the multiples of horizontal grid spacing, and the vertical localization scale \( L_v \) is measured by the actual distance in the vertical direction. Then, the two localization terms can be defined as follows:

\[
\text{Dist}_h = L_h \ast \sqrt{10 / 3} \ast 2, \text{Dist}_v = L_v \ast \sqrt{10 / 3} \ast 2
\]

For the analysis done on each model grid point, the observations far from the target grid with horizontal distance \( d_h \) larger than \( \text{Dist}_h \) or vertical distance \( d_v \) larger than \( \text{Dist}_v \) are not used. Moreover, the observation errors of data far from the target grid are multiplied by a weighting function \( C \), which is defined as follows,

\[
C = \exp \left[ 0.5 \ast \left( \frac{d_h^2}{L_h^2} + \frac{d_v^2}{L_v^2} \right) \right]
\]

By testing the localization parameters for fixed \( L_h \) and increasing \( L_v \) from 0.5 to 2.5 at an interval of 0.5 or fixed \( L_v \) and adjusting \( L_h \) from 50 to 250 m at an interval of 50 m, and further examining the temporal and spatial root mean square error (RMSE) of temperature and salinity at various depths in 2-years-integration assimilation experiments with the parameter values used, we choose multiple localization scales for various observation data with different characteristics. For the gridded SST observation data that has been interpolated onto the model grid, we set \( L_h = 0.5 \) and \( L_v = 50 \) m, which allows the SST observation points far from the target model grid with a distance of approximate 1.8 times grid spacing (about 180 km) in the horizontal and 180-m depth in the vertical to affect the analysis at the target grid. This localization limits the relatively dense SST observations to exhibit distant (nonlocal) influences on the model analysis. For the ARGO and GTSSP T/S profiles, \( L_h \) and \( L_v \) are both set to three times of those for SST observation data. For the AVISO SLA data, the same \( L_h \) as that for SST observation data is used.

The assimilation module is coupled with the ocean component of BCC-CSM and is able to assimilate T/S profiles, SST and SLA observations at a daily frequency. During the assimilation analysis process, a direct replacement strategy is used to add the analysis increment to the background state at once when one assimilation cycle is finished. For efficient computation, we adopt an ocean point-parallel strategy, in which all the ocean points are divided into approximately equal parts and each processor solves one part for enhancing the load balance of all processors as much as possible.

### 2.3.2. Sea-Ice Assimilation

To assimilate the gridded SIC data, we use an Optimal Interpolation (OI) algorithm in the sea-ice component of the coupled model. The analysis equations are the same as Equations 1 and 2, but the background error covariance matrix \( P_b \) is simply estimated by a parameterization, rather than by an ensemble method. Given that SIC is only a single-level variable, the elements of \( P_b \) that represents the covariance between two points of \((x_1, y_1)\) and \((x_2, y_2)\) in the two-dimensional space can be expressed as follows,

\[
P_b(x_1, y_1, x_2, y_2) = \sigma^2 \exp \left( -\frac{(x_1 - x_2)^2}{L_x^2} - \frac{(y_1 - y_2)^2}{L_y^2} \right)
\]
where \( \sigma^2 \) is error variance, which is set to 2.5e−3; \( L_x \) and \( L_y \) are the correlation lengths in zonal and meridional directions, respectively, and they are both set to five times the grid spacing (about 500 km) in the module. The OI scheme adopts a 1-day window for assimilating the daily SIC data and uses a localization strategy, in which the distance \( d \) between the model point \((x_m, y_m)\) and observation point \((x_o, y_o)\) is computed by the formula

\[
\frac{(x_o - x_m)^2}{L_x^2} + \frac{(y_o - y_m)^2}{L_y^2},
\]

and the 30 nearest observations surrounded the given model grid point are then chosen to affect the analysis field.

### 2.3.3. Atmosphere Nudging

In the atmosphere component, to input the multilevel air temperature, zonal wind, meridional wind, and humidity fields from the 6-hourly ERA-Interim reanalysis, we adopt a simple nudging method. It is expressed as follows,

\[
V = V_m + \left( \frac{V_{\text{obs}} - V_m}{\tau} \right) \times \Delta t
\]

where \( V_m \) is model background field, and \( V_{\text{obs}} \) is observation. \( \Delta t \) is integration step of the model, which is set to 300 s. The relaxation time scale \( \tau \) is assigned to 30 min. To operate a grid nudging in the model spectral space, the multiple variables in reanalysis products are interpolated onto the horizontal and vertical resolution of the model. The nudging method may exert excessive restraint on the model state, but it is simple to apply and efficient in computation, especially for the use of gridded products with relatively dense records in a three-dimensional space. Liu et al. (2015, 2017) adopted a similar nudging scheme to effectively initialize the BCC coupled models using other different reanalysis products and achieved reasonable subseasonal and seasonal forecast skills.

### 2.4. Experimental Design

Using the CDA scheme, several sets of experiments (Figure 1) are conducted to reveal the reliability and impact of assimilation: (1) control experiment (EXP-CTL): model simulation without assimilation of any observational data; (2) ocean DA experiment (EXP-OCN): the ocean observations including ARGO and GTSPP T/S, OISST SST, and AVISO SLA data are assimilated at a daily frequency; (3) sea-ice DA experiment (EXP-ICE): the OISST SIC observations are assimilated at a daily frequency; (4) ocean-sea ice DA experiment (EXP-OCNICE): the ocean and sea-ice observations are simultaneously assimilated at a daily frequency; (5) atmosphere-ocean DA experiment (EXP-ATMOCN): on the basis of EXP-OCN, the atmosphere component is nudged toward the 6-hourly ERA-Interim reanalysis; and (6) atmosphere-ocean-sea ice DA experiment (EXP-ATMOCNICE): on the basis of EXP-OCNICE, the atmosphere component is nudged toward the 6-hourly ERA-Interim reanalysis. The above experiments are all integrated from January 2000 to December 2015, and the results during the last 15 years are verified in this study. During the integration, the model adopts time-varying external forcing including volcanoes, solar variability, greenhouse gases, aerosol, ozone, and land-use, which are the same as those used in the CMIP6 historical simulation before 2014 and in the CMIP6 SSP2-4.5 projection beyond 2014. The initial condition on 1 January 2000 is from a long-term simulation of the model, which had a spin-up of thousands of years, was then run for 500 years (preindustrial control experiment of the CMIP6), and finally integrated for another 150 years (historical experiment of the CMIP6).

### 2.5. Validation Data

To evaluate the assimilation and prediction, we use SST and SIC fields from the OISST dataset and atmospheric variable fields from the ERA-Interim reanalysis for a dependent verification. Also, the following observations are adopted for an independent evaluation: (1) monthly ocean temperature and salinity objective analyses from the Hadley Center EN4.2.1 dataset (Good et al., 2013); (2) monthly climatologies of ocean temperature and salinity from the World Ocean Atlas 2013 (WOA13; Locarnini et al., 2013); and (3)
monthly SIC observations from the Hadley Center Sea Ice and Sea Surface Temperature dataset (HadISST; Rayner et al., 2003).

In addition, the ERA5 atmosphere reanalysis (Hersbach et al., 2020) and two ocean reanalysis products including the SODA version 3 (Carton et al., 2018) and NCEP GODAS (Behringer & Xue, 2004) are also compared with the CDA results in section 3.

3. Analyses of Ocean and Sea-Ice, and the Role of Sea-Ice DA

In this section, the performances of CDA scheme in ocean and sea-ice analyses and impacts of sea-ice DA are examined using the experiments listed in section 2.4.

Figure 2 shows the depth-latitude cross-section of climatological biases of zonally averaged ocean temperature/salinity with respect to WOA13 data. Compared to EXP-CTL, EXP-OCN improves the temperature climatology from surface to 1000-m depth in most areas, with decrease of RMSE from 0.47 to 0.27 and mean absolute bias from 0.29 to 0.18. In contrast, EXP-CTL gives a relatively reasonable simulation of salinity climatology, which is not further improved in EXP-OCN. Although with regionally different biases, the various experiments assimilating different datasets show overall similar magnitudes of global mean bias and RMSE of temperature/salinity climatology. The mean bias in EXP-ATMOCNICE is comparable to that in the SODA reanalysis. Note that the water in the upper 400 m is apparently colder and fresher than observations near 70°N in all the experiments, as well as in the SODA reanalysis, because of the commonly existing upper-ocean simulation biases in the North Atlantic in most ocean models (Danabasoglu et al., 2014). Also, in EXP-OCN and EXP-ATMOCN, cold biases are clear below 150 m of the high-latitude Southern Ocean, along with high salinity biases above that depth. This may be because the high salinity bias in the mixed layer can increase the water density, strengthen the mixing, and further contribute to formation of the above cold bias (Figure 2a2, 2a7, 2b7). In addition, overestimated near-surface temperatures over the tropics are found in all the assimilation experiments and SODA reanalysis. This systematically warm bias may be due to that the WOA13 climatology data used in this study is based on sufficient ocean observations in multiple decades during 1955–2012, whereas the assimilation experiments only use the observations in the recent two decades.

To examine the monthly-to-interannual variation of ocean temperature in 60°S–60°N and the polar areas in assimilation experiments and reanalysis products, we show the profiles of temporal RMSE with respect to the EN4 dataset in Figure 3. Among the various experiments, EXP-ATMOCNICE gives the lowest RMSE, which is even smaller than those of SODA and GODAS products once used to initialize the BCC model. In 60°S–60°N (Figure 3a), similar profiles of RMSE are found for EXP-OCN and EXP-OCNICE, as well as for EXP-ATMOCN and EXP-ATMOCNICE, denoting little influence of sea-ice DA on the analysis of tropical and midlatitude ocean temperature. This may be due to the fact that the tropical to midlatitude oceans are strongly constrained by plenty of daily observations when ocean DA is operated and that the model integrations are too short to fully show the impacts of sea-ice DA. In the polar areas (Figure 3b), sea-ice DA leads to a reduction of temperature RMSE from surface to depth of 1000 m. Especially, below the mixed layer, the RMSE declines by about 0.1 °C on the basis of ocean DA, and by about 0.2 °C on the basis of ocean-atmosphere DA. This is different from the finding of Kimmritz et al. (2018), in which the impact of sea-ice DA on ocean temperature analysis is predominantly shown in the ocean mixed layer. We find that the above result is mainly contributed by the DA's effects in the high-latitude Southern Ocean. It corresponds to the feature shown in Figure 2 that, in the high-latitude Southern Ocean, the high salinity bias in the mixed layer and the cold bias below are more remarkable in the experiments without sea-ice DA than in those with sea-ice DA. It implies that the inclusion of sea-ice DA is absolutely crucial for the skillful analysis of Antarctic oceans.

The causes for such notable impacts of sea-ice DA in our experiments deserve consideration. The high-latitude Southern Ocean is weakly stratified with small salinity difference across the pycnocline (Figure 2b1). Many studies indicated that insufficient vertical mixing, sea-ice or freshwater forcing over that area can cause a high salinity bias in the mixed layer which enhances mixing, erodes the stratification, and induces excessive deep convection in ocean models (e.g., Heuzé et al., 2015; Kjellsson et al., 2015; Timmermann & Beckmann, 2004). In this study, high salinity bias in the mixed layer exists in EXP-CTL and further
Figure 2. Depth-latitude cross-section of zonally averaged ocean (a) temperature (units: °C) and (b) salinity (units: psu) climatology for (a1, b1) WOA13, (a2, b2) EXP-ATMOCN, and (a3, b3) EXP-ATMOCNICE. Also shown are the biases of climatological mean (experiment minus WOA13) during 2001–2015 for (a4–a8, b4–b8) various experiments and (a9, b9) SODA reanalysis. The decimals shown at the top right of each panel are successively the RMSE and mean absolute bias of cross-section temperature/salinity between the experiment and WOA13.
strengthens in EXP-OCN and EXP-ATMOCN (Figure 2b4, 2b5, 2b7), associated with a severe underestimation of sea-ice extent over the high-latitude Southern Ocean (figure not shown). This may lead to excessive deep convection allowing more mixing of cold surface water and warm circumpolar deep water, thus produce lower-than-observed ocean temperature in deep ocean (Figure 2a2, 2a7). In contrast, EXP-OCNICE and EXP-ATMOCNICE skillfully reproduce the sea-ice distribution and do not show high salinity bias in the mixed layer, thus give a more reasonable description of deep-ocean temperature. In this context, the sea-ice DA may play an important role in providing a reliable analysis of the deep-ocean structure in the high-latitude Southern Ocean through improving the upper ocean mixing and the deep convection.

Figure 4 further shows the depth-time variations of spatial RMSE of ocean temperature in 60°S–60°N, 60°S–90°S, and 60°N–90°N. In 60°S–60°N, the ocean DA shows relatively evident errors in the thermocline most of the time, especially the years before 2007. These errors are mostly decreased by inclusion of atmosphere nudging, but unchanged by addition of sea-ice DA. In the Arctic area (60°N–90°N), the temperature errors are large in the upper 150 m and show apparent seasonal disparity. The differences between EXP-OCN and EXP-OCNICE and between EXP-ATMOCN and EXP-ATMOCNICE are overall small in subsurface ocean, with maximum error declines of about 0.1 °C near the depth of 100 m in several years due to addition of sea-ice DA. In the Antarctic area (60°S–90°S), because of the inclusion of sea-ice DA, the temperature in upper 1000 m are entirely improved from EXP-OCN/EXP-ATMOCN to EXP-OCNICE/EXP-ATMOCNICE. Especially, the temperature errors below the depth of 150 m are evidently decreased in many years, with maximum declines of about 0.4 °C and 0.8 °C when adding sea-ice DA to the ocean DA and ocean-atmosphere DA, respectively. Moreover, the comparisons among various experiments indicate that in the multicomponent DA framework with the model, only the addition of sea-ice DA can result in a reliable analysis of ocean temperature in the high-latitude Southern Ocean. As aforementioned, such impacts of sea-ice DA may be due to the sea-ice’s effect on deep ocean through changing the upper ocean mixing and the resulting deep convection (e.g., Heuzé et al., 2015; Kjellsson et al., 2015). This effect is more sensitive in the Antarctic oceans than in the Arctic oceans because of the clearly weaker stratification with smaller salinity difference across the pycnocline in the former than in the latter (Figure 2b1–b3). Moreover, the model deficiency in SIC simulation and the improvement of SIC due to sea-ice DA are more apparent in the Antarctic oceans than in the Arctic oceans. These may jointly account for the more evident impacts of
sea-ice DA in the high-latitude Southern Ocean than in the Arctic area. With incorporation of multicomponent data, EXP-ATMOCNICE shows the most skillful analysis of ocean temperature variability, with RMSE magnitudes smaller than those of SODA reanalysis in all the above three areas.

For the analysis of sea-ice, climatological mean and bias of Arctic sea-ice concentration in March and September in various experiments are shown in Figure 5. EXP-CTL is featured by apparent biases over the Okhotsk Sea, Greenland Sea, Barents Sea, and Labrador Sea in March when the Arctic sea-ice area
Figure 5. Climatological mean of Arctic sea-ice concentration in (a) March and (b) September during 2001–2015 for (a1, b1) OISST and (a2, b2) EXP-ATMOCNICE. Also shown are (a3–a8, b3–b8) the differences between various experiments and OISST, (a9, b9) between EXP-ICE and HadISST, and (a10, b10) between EXP-ATMOCNICE and HadISST.
(SIA) reaches its maximum, and by biases over the Arctic Ocean, Siberian Sector and Beaufort Sea in September when the Arctic SIA reaches its minimum. Although with different bias features, both EXP-OCN and EXP-ATMOCN show no appreciable improvement in simulating Arctic SIC. With assimilation of SIC observations, EXP-ICE reduces the biases and is thus very close to the observed feature in most Arctic area. On this basis, EXP-OCNICE and EXP-ATMOCNICE further decrease the bias over the Barents Sea in March and the bias over the Arctic Ocean in September. Compared with different observational datasets, the Arctic SIC biases in sea-ice DA experiments differ little in March, but are different over the Siberian Sector and Beaufort Sea in September. Similar results are found for simulating the Antarctic SIC (figure not shown), namely, extensive and severe underestimation in both March and September in EXP-CTL, EXP-OCN and EXP-ATMOCN are reduced in most high-latitude Southern Ocean in EXP-ICE and further improved locally in EXP-OCNICE and EXP-ATMOCNICE. This indicates that, for a reliable analysis of SIC, the sea-ice DA is decisive and the ocean or ocean-atmosphere DA also plays a positive role.

Focusing on the Arctic sea-ice in September and Antarctic sea-ice in March, Figure 6 illustrates the interannual evolutions of SIA in observation and various experiments. The OISST and HadISST observation datasets show similar interannual variations of SIA, although the magnitudes are slightly different. Compared to the observation, EXP-CTL often overestimates the Arctic SIA but severely underestimates the Antarctic SIA. Most of time both EXP-OCN and EXP-ATMOCN underestimate the magnitude of SIA in polar areas when giving an improved depiction on the interannual variation of SIA. With assimilation of SIC observations in EXP-OCNICE and EXP-ATMOCNICE, the temporal RMSEs of SIA decrease about 4 times in the Arctic area and above 10 times in the Antarctic area, and the temporal correlations increase to above 0.9, confirming the sea-ice DA’s decisive role in realistically reproducing the SIA change and especial importance in diminishing the analysis error in the high-latitude Southern Ocean. Moreover, the evolution of SIA is skillfully reproduced by EXP-ICE and further improved slightly by EXP-OCNICE and EXP-ATMOCNICE, suggesting the importance of CDA of multiple components. Particularly, for the notable events of extreme changes of Arctic SIA around 2007 and 2012 and Antarctic SIA around 2002, 2006, and 2011, the multicomponent DA can well capture the observed features.

Further, focusing on the period during 2011–2012, daily variations of Arctic/Antarctic SIC and SST RMSE between observation and various experiments are given in Figure 7, to explore the effects of coordinated assimilation of these two variables and the impacts of sea-ice DA. For analysis of SST in the Arctic area (Figure 7a), the notable RMSE in EXP-CTL can be decreased only when the ocean DA and atmosphere nudging are successively included. Note that the addition of sea-ice DA slightly decreases the averaged RMSE of SST on the basis of EXP-CTL or EXP-OCN, but increases the SST RMSE in July–October on the basis of EXP-ATMOCN, suggesting a possible disagreement among observations of multiple components in that season. In contrast, in the high-latitude Southern Ocean (Figure 7b), DA of only sea-ice can obtain a nearly same magnitude of SST RMSE as those in EXP-OCN and EXP-ATMOCN, and the addition of sea-ice DA to the ocean or ocean-atmosphere DA can further reduce SST RMSE by half, indicating an important role of sea-ice DA for the Antarctic SST analysis. For analysis of SIC (Figures 7c and 7d), compared to EXP-CTL, mostly EXP-OCN and EXP-ATMOCN show similar SIC RMSEs in the Arctic area but reduced RMSEs in the high-latitude Southern Ocean. It means that without assimilation of sea-ice observations, the roles of incorporating atmosphere and ocean data are slightly more important for the Antarctic SIC analysis than for the Arctic SIC analysis. When assimilating sea-ice observations, the SIC analysis is clearly improved in EXP-ICE, with an especially more remarkable decline of SIC RMSE in the Southern Ocean than in the Arctic area. This is probably due to the large errors of the model itself in simulating the Antarctic SIC.
in coordination with ocean DA or ocean-atmosphere DA, the SIC RMSE is somewhat reduced during September–December in the Arctic area and during March–June in the high-latitude Southern Ocean. This indicates that for improving the SIC analysis, multicomponent DA is necessary in the ice-freezing season while a separate sea-ice DA may be sufficient in other seasons.

4. Role of Sea-Ice DA in Improving Atmosphere Analysis

Figure 8 gives the temporal correlation coefficients (TCCs) between experiments and reanalysis datasets for monthly evolution of 2-m air temperature (T2M) anomaly during 2001–2015. Without inclusion of atmosphere nudging, the coupled ocean DA and sea-ice DA exert apparent impacts on the T2M analysis. Specifically, EXP-ICE shows significant skills over the Arctic area from the Barents Sea to the Beaufort Sea and over the high-latitude Southern Ocean (Figure 8b), and EXP-OCN is highly skillful in most oceans between 60°S–60°N (Figure 8c). These features are jointly reproduced in most areas of EXP-OCNICE (Figure 8d). Particularly, the ocean DA leads to especially high TCCs over the eastern tropical and subtropical Pacific but relatively low TCCs over the western tropical Pacific, indicating an obvious regional difference of its impacts. When incorporating only the multilevel atmosphere data, the model can skillfully capture the variability of T2M over most ocean and land area in the globe, implying that the atmosphere nudging is generally sufficient for a reliable T2M analysis (figure not shown). Nevertheless, comparisons with both ERA-Interim and ERA5 reanalyses show that TCCs over the high-latitude Southern Ocean are relatively low in EXP-ATMOCN but enhanced in EXP-ATMOCNICE (Figures 8e–8h). This indicates that sea-ice DA is indispensable to the accurate analysis of near-surface atmosphere over the polar areas even the ocean and atmosphere data have been incorporated in the CDA system.

TCCs of sea level pressure (SLP) and 500-hPa geopotential height (GH500) anomalies between assimilation experiments and observations and the TCC differences between EXP-OCN and EXP-OCNICE are given in Figure 9. The significance of the TCC difference in Figures 9g and 9h is verified using the Steiger’s Z test (Raghunathan et al., 1996). For analysis of both SLP and GH500, EXP-ICE is basically unskillful over most areas, but EXP-OCN shows relatively significant skills over the tropics (Figures 9a–9d). Further, EXP-OCNICE shows increased skills over most of the tropics, especially in the Pacific, compared to EXP-OCN,
although the differences between them do not pass the 95% confidence level of the Z test in many regions (Figures 9e–9h). These results may imply a positive impact of sea-ice DA on tropical atmosphere analysis that can only be reproduced by the ocean and sea ice coordinated DA, but not by the separate sea-ice DA. When further incorporating atmosphere data, both EXP-ATMOCN and EXP-ATMOCNICE give highly skillful analyses of SLP and GH500, with TCCs of above 0.9 over most areas except the tropics (figure not shown). This makes the impacts of sea-ice DA indistinguishable.

To further explore the seasonal difference and magnitude of impact of sea-ice DA, we examine the skills of SLP and GH500 anomalies and the skill differences between EXP-OCN and EXP-OCNICE in various

Figure 8. Temporal correlation coefficients (a–f) between various experiments and ERA-Interim reanalysis and (g, h) between the experiments and ERA5 reanalysis for monthly anomalies of 2-m air temperature during January 2001–December 2015. Stippling indicates the 95% confidence level.
It is found that the differences of TCC over the tropics are evident in the boreal spring and small in the other seasons. Figure 10 gives the TCCs of SLP and GH500 anomalies between analyses and observations during February to June of 2001–2015 in EXP-OCN and EXP-OCNICE, and the TCC differences between these two experiments. For the SLP field, the high TCC areas are mainly located in the western tropical Pacific, eastern tropical Pacific, and the tropical Atlantic (Figures 10a and 10c). The addition of sea-ice DA in EXP-OCNICE increases the SLP TCC over the above three tropical areas. Especially, over the Maritime Continent and western tropical Pacific, the increase of TCC passes the 95% confidence level of the Z test, showing an obvious impact of sea-ice DA on analysis of SLP over these areas (Figure 10e). There are also increases of TCC over most Antarctic areas and the Arctic area from the Barents Sea to the Beaufort Sea, but they are not significant. Note that the SLP TCCs over several land areas and sea-land boundary areas are somewhat reduced by adding sea-ice DA, but the TCCs themselves are mostly insignificant over these areas in EXP-OCN and EXP-OCNICE. For the GH500 field, the skills of analyses are often significant.
in the whole tropics, with especially higher TCCs in EXP-OCNICE than in EXP-OCN (Figures 10b and 10d). The TCC differences between these two experiments are significant over the tropical areas from the western Indian Ocean to the western Pacific and from the eastern Pacific to the Atlantic, suggesting an evident influence of sea-ice DA on analysis of tropical GH500 (Figure 10f). Moreover, from EXP-OCN to EXP-OCNICE, the TCCs over a large proportion of the polar areas and some small midlatitude regions are also enhanced, although their differences may not always be significant (Figure 10f). Nevertheless, both the TCCs in EXP-OCNICE and their differences from those in EXP-OCN are significant over most Arctic areas during the seasons of March–May and April–June, and over most Antarctic areas during February–April (figures not shown). This indicates the impacts of sea-ice DA on tropical atmosphere analysis and on polar atmosphere analysis are simultaneously visible and possibly closely linked. This link may not be through the internal process in the ocean, because the two experiments show no clear differences of surface and subsurface ocean states in tropics and subtropics (figure not shown), partially due to the strong constraint by ocean DA at a daily frequency.

Focusing on the key regions where analysis skills and their differences between EXP-OCN and EXP-OCNICE are found to be significant, Figure 11 further shows the height-longitude cross-section of TCC for the meridionally averaged geopotential height (GH) anomalies in 10°S–10°N and height-latitude cross-section of TCC for the zonally averaged GH anomalies in 90°–150°E. In 10°S–10°N, the analysis of GH shows skillful results in most troposphere and midlower stratosphere, indicating a strong impact of ocean DA on tropical atmosphere analysis (Figures 11a and 11c). From EXP-OCN to EXP-OCNICE, the TCCs of GH from surface to 30 hPa are mostly enhanced, showing especially significant improvement in the midupper troposphere.
particularly, the range of TCC difference passing the 95% confidence level of the Z test can extend to the lower stratosphere near the western Indian Ocean (Figures 11e). For the section of 90°–150°E, compared to EXP-OCN, EXP-OCNICE not only shows more significant skills in the tropical troposphere, but also shows higher skills in the subtropical and midlatitude stratosphere (Figures 11b and 11d). As a result, several centers with significant increase of TCC are found near 30°S/N and 60°S/N in the stratosphere (in both hemispheres). Correspondingly, the TCCs of GH in the troposphere near these locations are also enhanced (Figures 11f). Local change of sea-ice may be partially responsible for the variations of TCC centers. Particularly, in the boreal spring the Arctic sea-ice extent can extend to several regions near 60°N, including the Bering Sea, Beaufort Sea, and Labrador Sea. These areas are the major places where the Arctic sea-ice concentration is apparently improved from EXP-OCN to EXP-OCNICE (Figure 5), and may thus be connected with the improvement of midlatitude atmosphere analysis. The above results indicate that the better analysis of tropical tropospheric atmosphere is accompanied by an improved description of mid- and high-latitude stratospheric atmosphere. A number of studies documented the impacts of sea-ice change on stratospheric atmosphere through upward propagating planetary waves (e.g., Kim et al., 2014; Zhang et al., 2016) and the feedback of polar stratospheric anomaly to midlatitude tropospheric atmosphere (e.g., Cohen et al., 2014; Francis et al., 2017). Thus, we speculate in this study that the positive role of sea-ice DA on tropical atmosphere analysis may be linked by

Figure 11. (a–d) Temporal correlation coefficients between EXP-OCN/EXP-OCNICE and ERA-Interim reanalysis and (e, f) correlation differences between those two experiments for monthly anomalies of geopotential height averaged over 10°S–10°N (left panel) and 90°–150°E (right panel) during February to June of 2001–2015. Stippling in (a–d) denotes where correlations are significant at 95% confidence level of Student's t test and that in (e, f) denotes where correlation differences are significant at 95% confidence level of Steiger's Z test.
a complex succession of the planetary wave’s propagation into the stratosphere due to sea-ice change and the stratospheric anomaly signal’s backward propagation and influence on the three-ring circulations in the troposphere. The comprehensive physical mechanism needs an in-depth study.

Note that the above-mentioned impacts of sea-ice DA to tropical atmosphere analysis are only shown in an ocean and sea-ice coordinated DA framework without considering the influence of atmosphere nudging. When the atmosphere nudging is included, the atmospheric state will obey a strong constraint by the atmosphere reanalysis data and its changes due to the impacts of ocean and sea-ice DA are not easily distinguishable as shown in Figure 9. Nevertheless, if we use a better method to assimilate sparse atmosphere observations at a fixed frequency rather than use the strong nudging of gridded reanalysis data at each step, the results may be different. In the short-term climate prediction, the atmospheric initial condition often shows a gradually weakening impact with increasing forecast lead time, but is still very influential at subseasonal to seasonal scales and its impacts on the forecast results are quite uncertain under the complex interaction between initial error and model error. In this context, it may be difficult to distinguish the impacts of sea-ice DA on tropical atmosphere forecast. Here, we conduct two sets of reforecast experiments using initial conditions from different CDA experiments (i.e., EXP-ATMOCNICE and EXP-ATMOCN) to simply explore the possible impacts of sea-ice DA on subseasonal prediction. The reforecasts are started on the first day of March, April, and May during 2001–2015. The correlation coefficients in EXP-ATMOCN and EXP-ATMOCNICE not shown here are mostly above 0.8 in the mid-upper troposphere and the stratosphere over the tropics.

Figure 12. The differences of forecast skill between the reforecast initialized by EXP-ATMOCNICE and that initialized by EXP-ATMOCN at the third (left panel) and fourth (right panel) weeks of forecast. The forecast skills are measured by temporal correlation coefficients between ensemble mean forecasts and ERA-Interim reanalysis for weekly anomalies of geopotential height averaged over (a), (b) 10°S–10°N, (c) 120°E–180°, and (d) 120°–60°W. Stippling indicates the 95% confidence level of Steiger’s Z test. The reforecasts consisting of 4 ensemble members are started on the first day of March, April and May during 2001–2015. The correlation coefficients in EXP-ATMOCN and EXP-ATMOCNICE not shown here are mostly above 0.8 in the mid-upper troposphere and the stratosphere over the tropics.
inclusion of sea-ice DA in initialization, the week-3 and week-4 forecast skills of GH are improved over most of the tropics. Especially, relatively larger improvements appear over the western tropical Pacific in week 3 and the eastern tropical Pacific in week 4. These two centers of maximum difference, although not presented at the same forecast time, are somewhat similar to the features shown in Figure 11. In correspondence, for the zonally averaged GH over 120°E–180° and 120°–60°W, TCC skills are enhanced in most parts of troposphere and stratosphere over both the tropics and high latitudes of the Northern Hemisphere, with especially significant increases in week 4 in the midtroposphere over the tropics and in the stratosphere near 30°N and 60°N. These features also bear some similarities to those shown in Figure 11. They suggest that the sea-ice DA may play a positive role in both analysis and forecast of tropical atmospheric variability during spring. Nevertheless, the uncertainties of the above results should be pointed out. On one hand, the number of forecast cases and ensemble members may be inadequate for providing a highly definite conclusion. On the other hand, the ensemble forecast members show very small spread over the tropics but considerable spread over the mid and high latitudes (figure not shown), which makes the extratropical atmosphere forecast more uncertain. These uncertainties may be partially responsible for the reduced GH skills over the midlatitudes of the Southern Hemisphere when adding sea-ice DA in initialization (Figure 12). Thus, more forecast experiments and verifications are needed in future for fully revealing the role of sea-ice DA in atmosphere forecast.

5. Summary

We have designed a weakly CDA scheme with the BCC-CSM and implemented a coordinated assimilation of multisource data of ocean, sea-ice, and atmosphere. A series of assimilation experiments are conducted to explore the impacts of sea-ice DA on the analyses of ocean, sea-ice, and atmosphere in the coupled model.

We develop different assimilation approaches for ocean, sea-ice, and atmosphere components of the coupled model. Based on the combination of EnOI and LETKF algorithms, an ocean ensemble assimilation approach with low computational cost is built to assimilate ocean T/S profiles, SST, and SLA observations at a daily frequency. Also, OI-based sea-ice assimilation and atmosphere nudging are implemented to incorporate daily SIC observation and 6-hourly atmosphere multivariable reanalysis data. Thus, a weakly CDA scheme consisting of the ocean, sea-ice, and atmosphere components is established. The multicomponent CDA system can reasonably estimate the observed climatology and variability of ocean and sea-ice. Particularly, it gives a reliable analysis of the monthly-to-interannual variation of ocean temperature in the upper 1000 m, which is better than the GODAS and SODA reanalyses that were used to initialize the BCC-CSM for climate predictions before.

The impacts of sea-ice DA on the analyses of ocean, sea-ice, and atmosphere are explored in detail by several sets of assimilation experiments. On the basis of ocean or ocean-atmosphere DA, inclusion of sea-ice DA exerts very small influence on analysis of upper ocean over the Arctic area, but leads to a clear reduction of temperature RMSE in the upper 1000 m of ocean over the high latitudes of Southern Hemisphere. For the latter, the temperature RMSEs below the mixed layer show maximum declines of about 0.4 °C and 0.8 °C when adding sea-ice DA onto the ocean DA and ocean-atmosphere DA, respectively. Particularly, in the multicomponent DA framework, only the addition of sea-ice DA can make the ocean or ocean-atmosphere DA effective in providing a reliable analysis of ocean temperature in the high-latitude Southern Ocean. Such notable impacts of sea-ice DA may be attributed to the sea-ice’s effects on deep ocean through changing the upper ocean mixing and the deep convection in the Antarctic oceans. For analysis of SIC, the sea-ice DA plays a decisive role, but its coordination with ocean DA and atmosphere nudging can further decrease the SIC RMSE. To obtain low SIC RMSE during the subseasonal and seasonal evolution of sea-ice, the sea-ice DA’s cooperation with ocean or ocean-atmosphere DA is necessary in the ice-freezing season while a separate sea-ice DA may be sufficient in the other seasons.

We give a particular focus on the impacts of sea-ice DA on atmosphere analysis. In the CDA system, the assimilation of sea-ice observations is helpful for an accurate analysis of atmospheric T2M over the polar regions. The sea-ice DA often shows no clear influence on analysis of mid- and low-latitude atmosphere by itself, but can improve the analysis of tropical atmosphere on the basis of ocean DA. Especially during February–June, from ocean DA to ocean and sea-ice DA, the analysis skills of GH500 over the tropical
areas from the western Indian Ocean to the western Pacific and from the eastern Pacific to the Atlantic are significantly improved. In correspondence, the skills over several polar areas and some sparse midlatitude regions are also enhanced. We find that the significant improvement of atmospheric GH mainly appears in midupper troposphere over the tropics, along with a clear increase of skill in midupper stratosphere over the mid and high latitudes. This suggests that the sea-ice DA's impacts on tropical atmosphere analysis may be through the interactions among sea-ice, stratosphere, and troposphere, given that the tropical ocean states are similar in CDAs with and without sea-ice DA. Using a limited number of forecast experiments, we also briefly discuss the similarity of impact between forecast and assimilation, and suggest the positive role of sea-ice DA in both analysis and subseasonal forecast of tropical atmosphere.

To provide more comprehensive analysis of the state of climate system, the CDA scheme presented in this study needs to add assimilation of some important observation data, such as sea-ice thickness, land surface temperature, and soil moisture, etc. In addition, a number of studies have indicated the superiority of strongly CDA to weakly CDA using idealized experiments of ocean-atmosphere coupled models (e.g., Frolov et al., 2016; Lu et al., 2013; Sluka et al., 2016); thus the strongly CDA scheme at the air-sea interface is also a useful research topic. These will be gradually addressed in our future studies.

Data Availability Statement

This work used the ARGO (available at ftp://ftp.ifremer.fr/ifremer/argo/geo), GTSSPP (available at https://www.nodc.noaa.gov/GETSSPP/access_data/index.html), OISST (available at https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.highres.html), AVISO (available at https://www.aviso.altimetry.fr/en/data/data-access.html), ERA-Interim (available at https://apps.ecmwf.int/datasets/data/interim-full-daily/lev-type=ml/), ERA5 (available at https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels?tab=overview), EN4.2.1 (available at https://www.metoffice.gov.uk/hadobs/en4/download.html), WOA13 (available at https://www.nodc.noaa.gov/OC5/woa13/woa13data.html), HadISST (available at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html), SODA version 3 (available at http://dsrs.atmos.umd.edu/DATA/soda3.7.2/REGRIDED/ocean/), and GODAS (available at https://cfs.ncep.noaa.gov/cfs/godas/monthly/) data sets. The authors thank their providers. The coupled assimilation experiment results used in the study can be downloaded from the website (ftp://esgf.bcccsn.ncc-cma.net/0622/bcc-exp-data/bcc-cda/).

Acknowledgments

The authors thank the editors and four anonymous reviewers for their fruitful comments that were helpful in improving the overall quality of the manuscript. This study was jointly supported by the National Key R&D Program of China (Grant 2016YFA0602102) and the National Natural Science Foundation of China (Grant Nos. 42075161, 41675090, 41775100, 41830964).

References

Alves, O., Wang, G., Zhong, A., Smith, N., Tsetikitin, F., Warren, G., et al. (2003). POAMA: Bureau of Meteorology operational coupled model seasonal forecast system. In Science for Drought: Proc. National Drought Forum (pp. 49–56). Brisbane, Australia: Department of Primary Industries.

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

Behringer, D., & Xie, Y. (2004). Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. In Eighth Symposium on integrated observing and assimilation systems for atmosphere, oceans, and land surface, AMS 84th Annual Meeting (pp. 11–15). Seattle, Washington: Washington State Convention and Trade Center.

Browne, P., de Rosnay, P., Zuo, H., Bennett, A., & Dawson, A. (2019). Weakly Coupled Ocean-Atmosphere Data Assimilation in the ECMWF NWP System. Remote Sensing, 11(3), 234. https://doi.org/10.3390/rs11030234

Carton, J. A., Chepurin, G. A., & Chen, L. (2018). SODA3: A new ocean climate reanalysis. Journal of Climate, 31(17), 6967–6983. https://doi.org/10.1175/jcli-d-18-0149.1

Chang, T.-S., Zhang, S., Rosati, A., Delworth, T. L., & Stern, W. F. (2013). An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation. Climate Dynamics. 40(3–4), 775–803. https://doi.org/10.1007/s00382-012-1412-2

Chen, Z., Liu, J., Song, M., Yang, Q., & Xu, S. (2017). Impacts of assimilating satellite sea ice concentration and thickness on Arctic sea ice prediction in the NCEP Climate Forecast System. Journal of Climate, 30(21), 8429–8446. https://doi.org/10.1175/jcli-d-17-0093.1

Cohen, J., Screen, J. A., Furtado, J. C., Barlow, M., Whittleston, D., Coumou, D., et al. (2014). Recent Arctic amplification and extreme mid-latitude weather. Nature Geoscience, 7(9), 627–637.

Danabasoglu, G., Yeager, S. G., Bailey, D., Behrens, E., Bentsen, M., et al. (2014). North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part I: Mean states. Ocean Modelling, 73, 76–107.

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

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dynamics, 53(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Francis, J. A., Avravus, S. J., & Cohen, J. (2017). Amplified Arctic warming and mid-latitude weather: New perspectives on emerging connections. Wiley Interdisciplinary Reviews: Climate Change, 8(5). https://doi.org/10.1002/wcc.474
Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, 20(22), 5473–5496. https://doi.org/10.1175/2007jcli1824.1

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8), 1015–1058. https://doi.org/10.1175/2010bams3001.1

Saha, S., Nadiga, S., Thiau, C., Wang, J., Wang, W., Zhang, Q., et al. (2006). The NCEP climate forecast system. *Journal of Climate*, 19(15), 3483–3517. https://doi.org/10.1175/jcli3812.1

Sliuka, T. C., Penny, S. G., Kalnay, E., & Miyoshi, T. (2016). Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation. *Geophysical Research Letters*, 43, 752–759. https://doi.org/10.1002/2015gl067238

Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., et al. (2008). Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *Journal of Geophysical Research: Oceans*, 113, C10017. https://doi.org/10.1029/2008jc004741

Sun, L. C., Thresher, A., Keeley, R., Hall, N., Hamilton, M., Chinn, P., et al. (2010). The data management system for the Global Temperature and Salinity Profile Program (GTSPP). *Paper presented at the Proceedings of the "OceanObs'09: Sustained ocean observations and Information for Society" Conference* (Vol. 2, ESA Publ. WPP-306), Venice, Italy, September 21–25, 2009.

Takaya, Y., Hirahara, S., Yasuda, T., Matsueda, S., Toyoda, T., Fujiy, Y., et al. (2018). Japan Meteorological Agency/Meteorological Research Institute-Coupled Prediction System version 2 (JMA/MRI-CPS2): Atmosphere–land–ocean–sea ice coupled prediction system for operational seasonal forecasting. *Climate Dynamics*, 50(3–4), 751–765. https://doi.org/10.1007/s00382-017-3638-5

Timmermann, R., & Beckmann, A. (2004). Parameterization of vertical mixing in the Weddell Sea. *Ocean Modelling*, 6(1), 83–100. https://doi.org/10.1016/s1463-5003(02)00061-6

Wu, B., Zhou, T., & Zheng, F. (2018). EnOIAU Initialization Scheme Designed for Decadal Climate Prediction System IAP-DecPreS. *Journal of Advances in Modeling Earth Systems*, 10(2), 342–356. https://doi.org/10.1029/2017ms001132

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., et al. (2019). The Beijing Climate Center Climate System Model (BCC-CSM): The progress from CMIP5 to CMIP6. *Geoscientific Model Development*, 12, 1573–1600. https://doi.org/10.5194/gmd-12-1573-2019

Xin, X., Gao, F., Wei, M., Wu, T., Fang, Y., & Zhang, J. (2018). Decadal prediction skill of BCC-CSM1.1 climate model in East Asia. *International Journal of Climatology*, 38(2), 584–592. https://doi.org/10.1002/joc.5195

Xu, F.-H., Oey, L.-Y., Miyazawa, Y., & Hamilton, P. (2013). Hindcasts and forecasts of Loop Current and eddies in the Gulf of Mexico using local ensemble transform Kalman filter and optimum-interpolation assimilation schemes. *Ocean Modelling*, 69, 22–38. https://doi.org/10.1016/j.ocemod.2013.05.002

Yang, C. Y., Liu, J., & Xu, S. (2020). Seasonal Arctic sea ice prediction using a newly developed fully coupled regional model with the assimilation of satellite sea ice observations. *Journal of Advances in Modeling Earth Systems*, e2019MS001938.

Yang, Q., Losa, S. N., Losch, M., Liu, J., Zhang, H., & Yang, H. (2015). Assimilating summer sea-ice concentration into a coupled ice-ocean model using a LSEIK filter. *Annals of Glaciology*, 56(69), 38–44. https://doi.org/10.3189/2015aog69a740

Yang, X., Rosati, A., Zhang, S., Delworth, T. L., Gudgel, R. G., Zhang, R., et al. (2013). A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *Journal of Climate*, 26(2), 650–661. https://doi.org/10.1175/jcli-d-12-00231.1

Yin, Y., Alves, O., & Oke, P. R. (2011). An ensemble ocean data assimilation system for seasonal prediction. *Monthly Weather Review*, 139(3), 786–808. https://doi.org/10.1175/2010mwr3419.1

Zhang, J., Tian, W., Chipperfield, M. P., Xin, F., & Huang, J. (2016). Persistent shift of the Arctic polar vortex toward the Eurasian continent in recent decades. *Nature Climate Change*, 6(12), 1094–1099. https://doi.org/10.1038/nclimate3136

Zhang, S., Harrison, M. J., Rosati, A., & Wittenberg, A. (2007). System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Monthly Weather Review*, 135(10), 3541–3564. https://doi.org/10.1175/mwr3466.1

Zhang, S., Liu, Z., Zhang, X., Wu, X., Han, G., Zhao, Y., et al. (2020). Coupled data assimilation and parameter estimation in coupled ocean–atmosphere models: a review. *Climate Dynamics*, 54(11–12), 5127–5144. https://doi.org/10.1007/s00382-020-05275-6