Evaluation on data assimilation of a global high resolution wave-tide-circulation coupled model using the tropical Pacific TAO buoy observations

SHI Junqiang[1, 2], YIN Xunqiang[2, 3, 4], SHU Qi[2, 3, 4], XIAO Bin[2, 3, 4], QIAO Fangli[2, 3, 4*]

[1] College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao 266100, China

[2] The First Institute of Oceanography, State Oceanic Administration, Qingdao 266061, China

[3] Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China

[4] Key Laboratory of State Oceanic Administration for Marine Sciences and Numerical Modeling, The First Institute of Oceanography, State Oceanic Administration, Qingdao 266061, China

Abstract

In order to evaluate the assimilation results from a global high resolution ocean model, the buoy observations from tropical atmosphere ocean (TAO) during August 2014 to July 2015 are employed. The horizontal resolution of wave-tide-circulation coupled ocean model developed by The First Institute of Oceanography (FIOCOM model) is 0.1°×0.1°, and ensemble adjustment Kalman filter is used to assimilate the sea surface temperature (SST), sea level anomaly (SLA) and Argo temperature/salinity profiles. The simulation results with and without data assimilation are examined. First, the overall statistic errors of model results are analyzed. The scatter diagrams of model simulations versus observations and corresponding error probability density distribution show that the errors of all the observed variables, including the temperature, isotherm depth of 20°C (D20), salinity and two horizontal component of velocity are reduced to some extent with a maximum improvement of 54% after assimilation. Second, time-averaged variables are used to investigate the horizontal and vertical structures of the model results. Owing to the data assimilation, the biases of the time-averaged distribution are reduced more than 70% for the temperature and D20 especially in the eastern Pacific. The obvious improvement of D20 which represents the upper mixed layer depth indicates that the structure of the temperature after the data assimilation becomes more close to the reality and the vertical structure of the upper ocean becomes more reasonable. At last, the physical processes of time series are compared with observations. The time evolution processes of all variables after the data assimilation are more consistent with the observations. The temperature bias and RMSE of D20 are reduced by 76% and 56% respectively with the data assimilation. More events during this period are also reproduced after the data assimilation. Under the condition of strong 2014/2016 El Niño, the Equatorial Undercurrent (EUC) from the TAO is gradually increased during August to November in 2014, and followed by a decreasing process. Since the improvement of the structure in the upper ocean, these events of the EUC can be clearly found in the assimilation results. In conclusion, the data assimilation in this global high resolution model has successfully reduced the model biases and improved the structures of the upper ocean, and the physical processes in reality can be well produced.

Key words: tropical Pacific, tropical atmosphere ocean, data assimilation, evaluation

---

1 Introduction

Since the distribution of the current observation network is quite sparse in the global ocean, using the observation data alone to study an ocean physical process is not sufficient in most cases. In recent decades, the data assimilation is popularly used to construct the data set which can include observation information to provide a better understanding of ocean phenomenon. For example, Moore (1991) applied an adjoint data assimilation scheme in a quasi-geostrophic (QG) ocean model of the Gulf Stream region, and the error of simulated speed and position of the Gulf Stream jet was largely corrected. Masuda et al. (2003) used a four-dimension variational data assimilation system to

reestablish the seasonal variation in the North Pacific. The results exhibited realistic features of the ocean circulation and a water mass movement process was clearly depicted. Shu et al. (2014) designed a series of data assimilation experiments based on an ensemble adjustment Kalman filter (EAKF) and The First Institute of Oceanography-earth system model (FIO-ESM) and the climatology and long-term trend of Arctic sea ice were well reproduced.

However, the credibility of the data set from the data assimilation is the most important bases for applying this data set to a series of scientific issues. Before this kind of data is used, the assessment for credibility should be done carefully. The performance evaluation to a global assimilation data set is usually carried out in some representative regions. For the global ocean, the tropical Pacific is one of the most important region needs to be concerned. In this region, physical processes are more complex than other waters (McPhaden, 2002; Wang et al., 2004; Henocq et al., 2010; Zuo et al., 2014; Anderson and Riser, 2014). Specifically, there are some unique but very significative phenomena existing in this area. In the case of El Niño and La Niña, the mechanism and pattern of manifestation of the periodically abnormal warming and abnormal cooling remains to be further researched (Gao and Zhang, 2017; Paek et al., 2017). Although these phenomena or events occur in the tropical Pacific, they have critical global influence on the climate change and ocean physical processes (Salau and Akinyemi, 2015; Wen et al., 2015; Bhowmick et al., 2016; Chowdary et al., 2017). Furthermore, as one of the most important datum in the global ocean, tropical atmosphere ocean (TAO) observations are widely distributed in this region and usually used for the global ocean and climate research. For example, TAO observations are commonly used to an ocean state estimation (Stammer et al., 2010) and provide the important information for a seasonal and climatological forecasting (Vidard et al., 2007; Balmaseda and Anderson, 2009). As for the assessment to assimilation data set, the TAO is also an important source of an independent observation. The performance of the assimilation results in the tropical Pacific could largely represent the overall performance of this global data set.

The error statistics is a primary step to understand the overall degree of improvement in the simulation results after assimilation. For example, Sun et al. (2007) summarized the overall performance of the Global Modeling and Assimilation Office (GMAO) and the Geophysical Fluid Dynamics Laboratory (GFDL) global ocean data assimilation systems in terms of a temperature error, a salinity error and a relaxation time scale. Oke et al. (2015) calculated the percentage improvement of the simulated temperature and salinity due to assimilating Argo profiles in the region of the eastern equatorial Pacific, the western equatorial Pacific, the subtropical North Pacific and the mid-latitude North Pacific through the methods of mathematical statistics.

Furthermore, the horizontal and vertical structures of the ocean in the research region also should be carefully examined to confirm if these spatial structures are more reasonable in the assimilation results. As in previous studies, the spatial distribution of an equatorial temperature was investigated by Keppenne and Rienecker (2003). The distribution characteristics of a temperature vertical structure that produced by a regional assimilation system was more reliable. Parent et al. (2003) applied a reduced-order Kalman filter method to an ocean general circulation model. An assimilated thermal and zonal current structure along the equator was well reconstructed compared with the independent data from the TAO. Xuan et al. (2012) established a four-dimensional data assimilation system. The structure distributions of the

temperature and front got from this system were more likely to exhibit the real condition. The improvement of these structures is a key factor for reasonability of the data assimilation method.

The isotherm depth of 20°C (D20) is a measurement of the depth of thermocline in general in the tropical Pacific region. D20 is a very important indicator because it represents the most active part in thermocline. Physical processes such as El Niño would be revealed by D20 analysis. The Equatorial Undercurrent (EUC) is another important phenomenon. The interannual variation of the EUC nearby D20 is closely connected with the ENSO. So it is necessary to apply accurate estimate of D20 and the EUC for revealing relevant processes in the tropical Pacific region.

Under the condition of well performance of the assimilation data set, it is the most important that the physical processes of time series and important historical ocean events should be well reproduced by means of the high resolution the assimilation data set. For example, the time evolution of the EUC under the background of El Niño events is an important indicator for checking the creditability of the assimilation data set (Kimoto et al., 1997; Chen et al., 2005; Wang et al., 2009). Since the strong 2014/2016 El Niño event happened in our study period, this paper has paid a special attention to the dynamic evolution process of the EUC in the assimilation data set which is an important aspect for evaluating the performance of this data set.

In current study, a comprehensive comparison with the independent TAO observations has been carried out to evaluate the performance of the assimilation data set based on a global high resolution ocean model. The overall error statistics, horizontal and vertical structures and time evolved physical process analysis is given in this paper. Follows the introduction in this section, Section 2 describes the data set from a global high resolution with and without the data assimilation, and the TAO observations are used as an independent data set for comparison. Section 3 presents the overall comparison, analysis of spatial and vertical structures, and dynamic analysis of physical processes along time. Finally, the summary and discussion are given in Section 4.

2 Model and data

2.1 *FIOCOM model*

A numerical model used in this paper is wave-tide-circulation coupled ocean model developed by The First Institute of Oceanography, China (FIOCOM). The FIOCOM is a state-of-the-art primitive equation general model based on ocean circulation model of MOM5 (modular ocean model, Version 5) and surface wave model of MASNUM (Yuan et al., 1991). In the FIOCOM, the non-breaking-surface-wave-induced mixing is incorporated into the K-profile parameterization (KPP) vertical mixing scheme. Qiao et al. (2004, 2008, 2010, 2016) proposed a parameterization of Bv, which can effectively improve the simulation ability of the upper ocean (Shu et al., 2011; Wu et al., 2015) and even processes in lake (Torma and Krámer, 2017). In order to simulate the mixing effects of tides, the parameterized tide-induced mixing (Qiao et al., 2006) is also introduced into the FIOCOM. Bv and tide-induced mixing make the FIOCOM a wave-tide-circulation fully coupled ocean model. The FIOCOM has a global tripolar grid as MOM5 and with a horizontal resolution of 0.1°×0.1°. There are 54 vertical layers with separations varying smoothly from about 2 m on the surface to 366 m on the bottom. The initial field was derived from the integral results of MOM5 forcing by Version 2 forcing for coordinated ocean-ice reference experiments (CORE2) data. The sea surface forcing field was derived from a global forecasting system (GFS). The forcing variables include the

air temperature, atmospheric pressure on the sea surface, ocean surface vector winds at 10 m height, the precipitation, specific humidity and heat flux with horizontal resolution of 0.5° and temporal resolution of 3 h. Furthermore, the EAKF data assimilation scheme (Yin et al., 2010, 2011, 2012) is employed to assimilate an available sea surface temperature (SST), sea level anomaly (SLA) and Argo observations into the model.

2.2 *Global high resolution FIOCOM simulation and assimilation data set*

No_assim data set was produced by a global high resolution FIOCOM with horizontal resolution of 0.1°. We used the EAKF data assimilation method to produce the Assim data set. Observation data used for assimilation in this paper include global SST, SLA, and Argo *T*/*S* profiles. The global SST that merged satellite microwave radiometers and infrared radiometer data was provided by the remote sensing systems (RSS) and released nearly real time every day with high spatial resolution of 9 km. The SLA data are a global, nearly real time and multiple platform fusion products provided by the archiving, validation and interpretation of satellite oceanographic data (AVISO) with spatial resolution of 25 km. The Argo temperature and salinity profiles were provided by Coriolis Global Data Center. Variables to be analyzed include temperature (*T*), salinity (*S*), current and the

depth of 20°C isotherm (D20) that was obtained by the interpolation of temperature. More information are shown in Table 1.

2.3 *TAO observation array*

In this study, the independent tropical Pacific TAO buoy array data, the global high resolution FIOCOM simulation data (No_assim) and the corresponding assimilation data (Assim) were used to assess the effectiveness of assimilation system. The specified simulated variables before and after assimilation in the tropical Pacific were compared with the corresponding TAO observation variables to analyze the credibility of the Assim data set. The time scope was from August 1, 2014 to July 31, 2015.

As one of the most representative data in the global ocean, the TAO has played the vital role in ocean and climate research (Hayes et al., 1991; Bennett et al., 1998; Fu et al., 2009; Xue et al., 2017). The horizontal location and observation number distribution of the observed temperature, D20, salinity, current are shown in Fig. 1. It is clear that the location distribution of the current is much sparse with respect to other variables. Table 2 demonstrated the temporal distribution of the observation number of specified variables in research area. The observation number of different variables was quite different along time. In particular, the temperature observation number is larger than the salinity and D20. The current observation number is least.

Table 1. List of two data sets

Data set	Resolution			Time period	Assimilation method	Data for assimilation	Variables to be analyze
	Time	Horizontal	Vertical/m				
No_assim	daily	0.1°×0.1°	2.1–365.9	2014/08/01–2015/07/31	–	–	*T*/D20/*S*/*u*/*v*
Assim	daily	0.1°×0.1°	2.1–365.9	2014/08/01–2015/07/31	EAKF	SST/SLA/Argo profiles	*T*/D20/*S*/*u*/*v*

temperature

	150°	E180°W		150°		120°		
10°	364	361	229 289	361	308	327	361	310
5°	75	154 361	333 94	361	128	326	360	360
N 0°	201 365	365 361	362 278	361	311	325	361	302
S 5°	365	365 272	337 234	252	314	323	358	164
		365 332	361 284	230	315	322	361	261
		146 360	361 282	361	317	321	361	359
10°		365	359 282	255		321	92	361

D20

	150°	E180°W		150°		120°		
10°	364	278	225 162	57	288	327	360	175
5°	73	154 359	331 93	348	127	312	166	360
N 0°	198 365	365 360	361 281	336	240	281	361	213
S 5°	365	364 271	335 165	234	294	276	272	164
		365 332	356 283	220	263	234	361	261
		146 271	361 281	351	267	294	361	358
10°		365	359 156	240		319	92	288

salinity

	150°	E180°W		150°		120°		
10°	364	361	9 37	361	308	158	361	310
5°	75	154 314	15 94	129	8	326	360	360
N 0°	201 365	365 72	362 287	361	311	325	354	43
S 5°	365	365 272 S2	337 218	252	186 SC1	323	268 SC2 164	
		365 325	126 284	230	296	267	247	135
		146 274 S1	361 271	361	317	321	268	328
10°		142	360 282	255		321	92	361

current

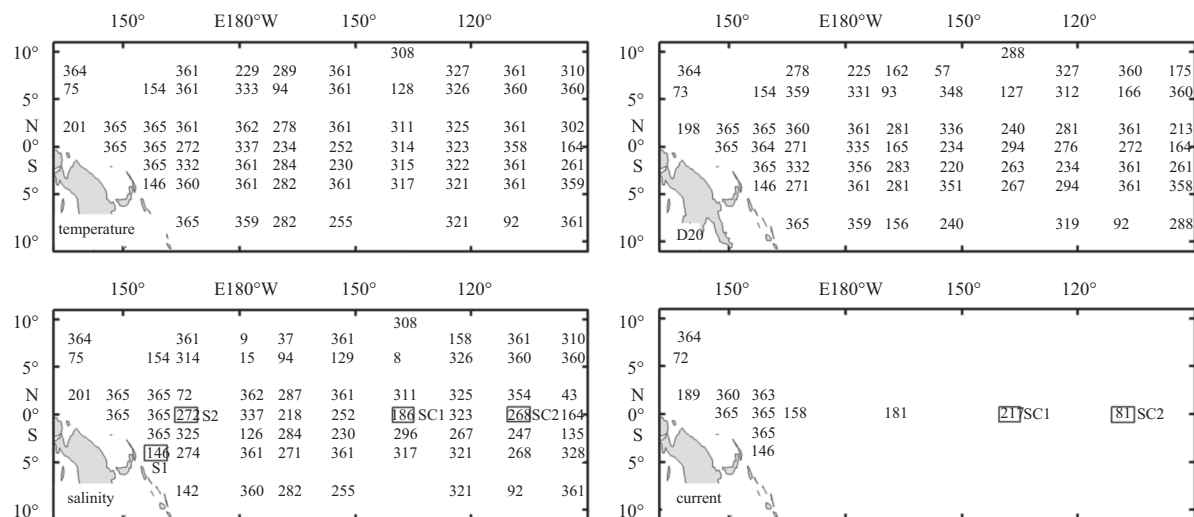|	150°	E180°W		150°		120°		
10°	364							
5°	72							
N 0°	189 360	363 158		181		217 SC1		81 SC2
S 5°	365	365						
		146						

Fig. 1. The distributions of available data quantity for the temperature, D20, salinity and current. Numbers in the figure represent the total available observation number for each variable during the selected time scope (August 1, 2014 to July 31, 2015). In the salinity and current panels, the locations with the black box represent the profile sites of S1, S2, SC1 and SC2, respectively.

Table 2. Time series of observation number

Variable	Number of observation											
	Aug. 2014	Sep.	Oct.	Nov.	Dec.	Jan. 2015	Feb.	Mar.	Apr.	May	Jun.	Jul.
Temperature	1 178	1 370	1 667	1 782	1 854	1 819	1 652	1 823	1 716	1 818	1 664	1 438
D20	1 110	1 300	1 584	1 690	1 770	1 695	1 496	1 581	1 349	1 538	1 466	1 284
Salinity	1 016	1 175	1 459	1 630	1 637	1 532	1 373	1 539	1 417	1 563	1 365	1 129
Current	248	228	258	292	290	281	280	317	269	288	243	232

All salinity sites except S1, S2, SC1 and SC2 have records at 1 m underwater. The number of the current data was much smaller than other variables. Especially, all current sites have records at 10 m underwater except SC1 and SC2. Relatively intact vertical current observations were recorded by only two sites SC1 and SC2.

3 Results analysis

In this part, we focus on the region of the tropical Pacific. A comprehensive contrast and analysis using the above three data sets will be carried out in three aspects including the overall error statistics, time-averaged physics structure analysis and reconstruction of physical process analysis respectively.

3.1 *Overall error statistics*

The overall statistical characteristic of error of the temperature, D20, the salinity, the two horizontal components *u* and *v* of velocity was analyzed by the scatter diagrams of model simulations versus observations and corresponding probability density distribution (Fig. 2). Obviously, all variables were more agreement with the observations after assimilation. The scatter distributions of all variables were more centralized nearby the diagonal lines. According to the corresponding error probability density distribution, the mean error of all variables was more close to 0 and the standard deviation all became smaller after assimilation.

It was also obvious that different variables had different assimilation effects. In particular, there was the best improvement effect on the temperature, D20 and *u* component of velocity. The averaged absolute errors of the temperature, D20 and *u* component of velocity after assimilation have been reduced by 42.9%, 54% and 11.8%, respectively. Specifically, from the scatter diagrams of the temperature, we can clearly find that the majority of the temperature scatters before assimilation are located below the diagonal line which indicates that most of the simulated temperature before assimilation is larger than observation. Obviously, the distribution of the temperature scatters is more symmetric and centralized nearby the diagonal line after assimilation. Similarly, D20 is generally deeper before assimilation, but the error is great corrected after assimilation. As for the zonal current, the value is somewhat smaller in general before assimilation than observation. After assimilation, the zonal current is more close to observation especially in the interval with large value. On the contrary, for the salinity and the *v* component of velocity, the difference of scatter diagrams before and after assimilation was quite small though the improvement indeed happened according to the corresponding probability density distribution. It is noteworthy that the error of current especially zonal component in the Assim data set had reduced to some extent though this variable was not assimilated.

3.2 *Time averaged physical structure analysis*

In this part, time-averaged physical structure of the specified variables was analyzed. Figure 3 (left column) demonstrates the meridional and time-averaged vertical distribution of the temperature. Observation indicated that the position of 20°C isotherm first became gradually deeper and then became shallow from west to east. Although the simulated 20°C isotherm before assimilation could depict this phenomenon, its position was significantly deeper than the observation case. The location near the 20°C isotherm carried a significant positive deviation which could be up to 4°C before assimilation. In other words, the simulated temperature near the thermocline is much higher than the observation case especially on the eastern and western side of the

region. In addition, there is an area in the tropical eastern Pacific Ocean where the simulated temperature is somewhat lower than the observation below 300 m. In sharp contrast, the position of the 20°C isotherm after assimilation agrees well with the observation and the temperature deviation which varies between 0 and 0.5°C in most regions was also greatly reduced after assimilation.

Similarly, zonal- and time-averaged 20°C isotherm (Fig. 3, right column) before assimilation could also depict the observed phenomenon that the position of 20°C isotherm is gradually shallow from south to north but the simulated position is much deeper than the observation as well. The temperature deviation that is mainly concentrated near the thermocline could be up to 4°C at the north part of the region. After assimilation, the 20°C isotherm is more close to the observation, simultaneously the temperature mean absolute deviation is also largely reduced by 72.2%.

The improvement of the vertical structures shows that the assimilation effect is very significant, especially in the near 20°C isotherm which suggests that the ability of simulating upper mixed layer is relatively weak without assimilation. This may be a common problem in an ocean numerical simulation that the model without assimilation cannot accurately describe the complex structure of the temperature near the thermocline. It is no doubt that the assimilation largely alleviates this problem because the simulated structure of the upper mixed layer has been significantly improved after assimilation.

Furthermore, the horizontal distribution of the temperature deviation from the observation at three different layers (Fig. 4) was inspected. For convenience, we defined three different levels in vertical. The first 50 m level was called a subsurface layer, 50–300 m was called a top layer and greater 300 m was called a bottom layer respectively. It was clearly that the simulated temperature at the subsurface layer without assimilation was relatively higher than the observation and the deviations at east of the region could be up to 3°C. The temperature deviation was positive at the top layer in the whole region and the larger value mainly distributed in southeast of the region with a maximum could be up to 2.7°C. As for the bottom layer, the whole area temperature deviation was very small, all within 0.5°C. The negative deviation was mainly located in the western region while the positive deviation was mainly distributed in the eastern region. After assimilation, the temperature deviation at the subsurface and top layers decreased significantly, especially in the region where the deviation was large before assimilation. The absolute deviation at the subsurface and top layers in most of the region could be restricted in 0.5°C. While at the bottom layer, the deviation changed little after assimilation. It was because the deviation was so small before assimilation that the degree of improvement was also small after assimilation. What we should pay attention to was that the bottom temperature deviation with assimilation was somewhat larger than that without assimilation in the east tropical Pacific. It was probable because the deviation before assimilation in the small region where the deviation was larger after assimilation was so small (less than 0.25°C) that the assimilation effect was not clear to the simulation results. If the assimilation was carried out in these regions, the analysis field after assimilation may drift to some extent. In short, the temperature error was reduced in different degrees at all three levels, but the degree of improvement was relatively small below 300 m.

The horizontal distribution of time-averaged D20 before and after assimilation are shown in Fig. 5. The observed D20 is gradually shallow from southwest to northeast. The minimum value that located in the east of the region is 40 m and the maximum that located in the southwest of the region could be up to 240 m.
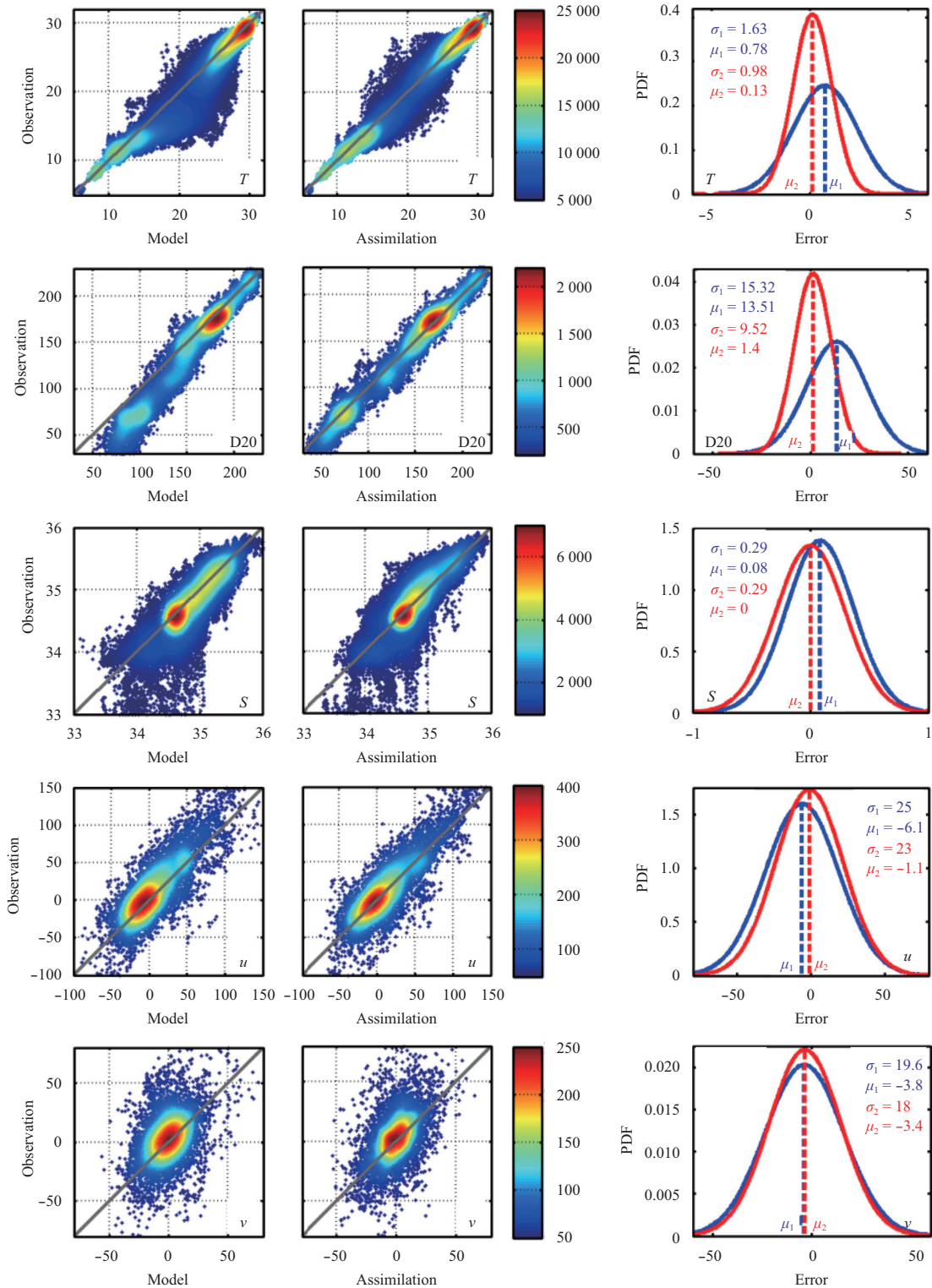
Fig. 2. Scatter diagrams of model simulations versus observations and corresponding probability density distributions of temperature (°C), D20 (m), salinity and two horizontal components (*u* and *v*, cm/s) of current from top to bottom. In the scatter diagrams, the left column represents the case without the data assimilation and the right column represents the data assimilation case. The color bar represents the number of record of every variable listed in the corresponding scatter diagrams. In the probability density distribution, the blue and red lines indicate the case without and with the data assimilation, respectively.

The simulated D20 before assimilation is deeper than the observation except a small mid-west region. The spatial difference of the deviation is significant. The distribution of large deviation corresponds well to the location of small D20. After assimilation, the deviation of the whole region significantly reduced in general especially in the region where the deviation was large before as-
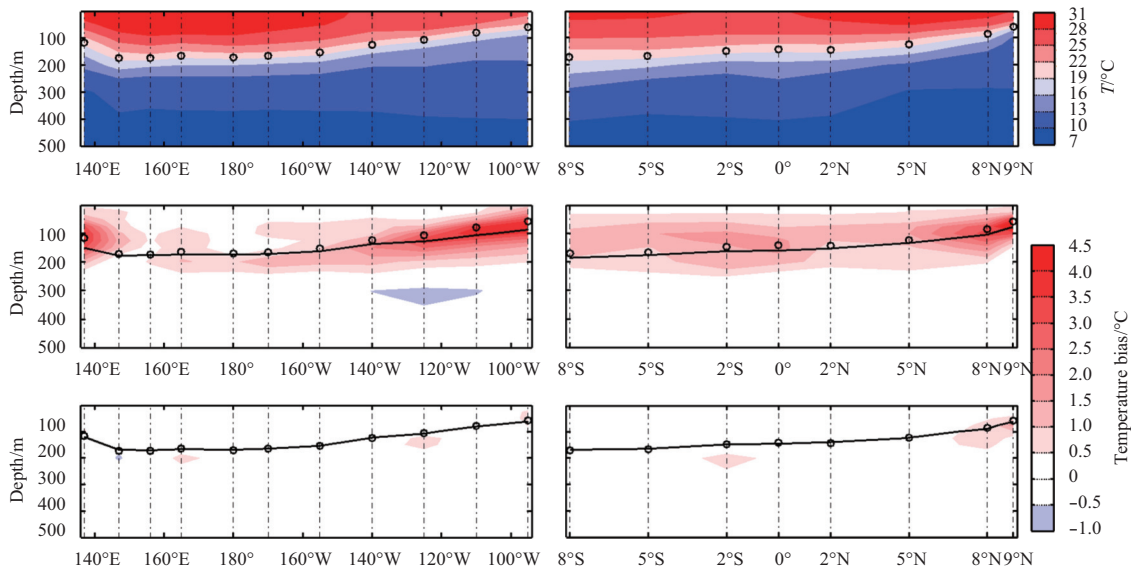
Fig. 3. The vertical distributions of the temperature (top), the simulated temperature bias without the data assimilation (middle) and with the data assimilation (bottom). The left column represents meridional-time-averaged results while the right column zonal-time-averaged. The hollow circles are the observed D20 and the black solid lines are the simulated results of D20.
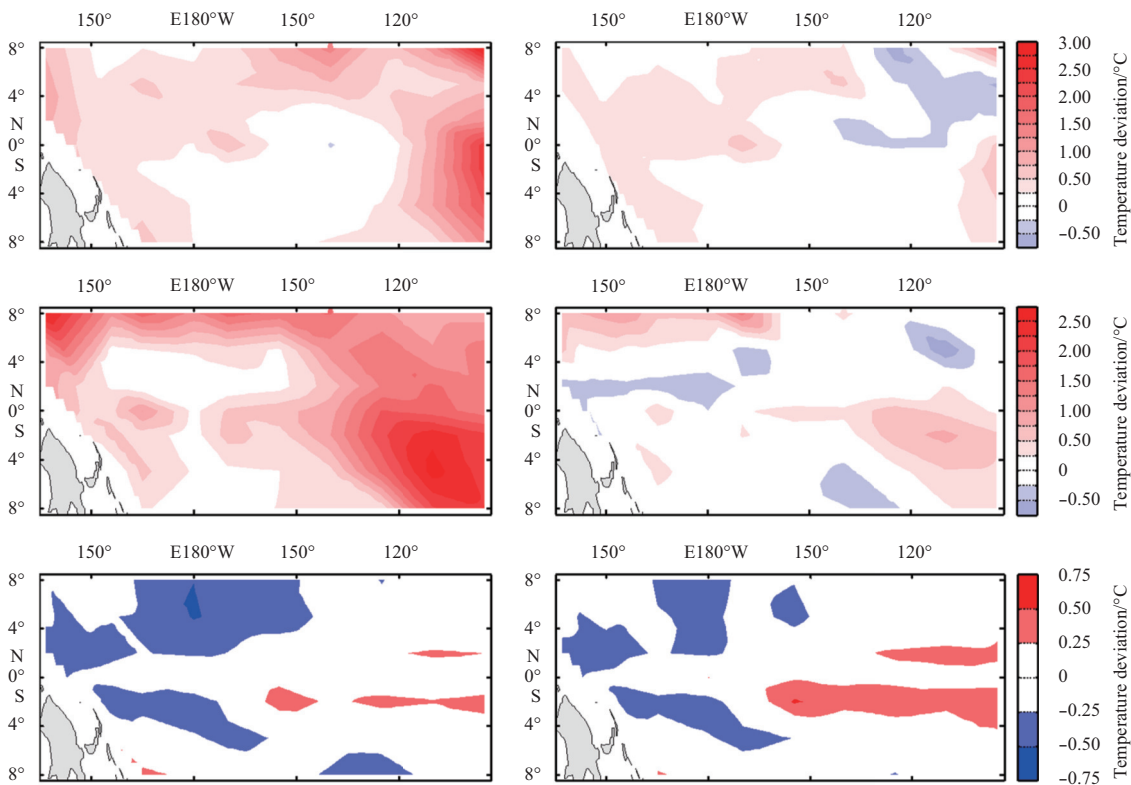

Fig. 4. The horizontal distributions of the temperature deviation from observation at 0–50 m averaged subsurface layer (top), 51–300 m averaged top layer (middle) and bottom layer below 300 m (bottom). The left column represents without the data assimilation and the right column with the data assimilation.

simulation. The bias of D20 could be reduced by 73% averagely and restricted in ±15 m after assimilation.

Similarly, the horizontal distribution of the time-averaged salinity at 1 m underwater and its deviation from the observation are shown in Fig. 6. The observed salinity increased gradually from north to south. The small value of the salinity was mainly

distributed in the northeast of the region with minimal 32.5 and the large value of the salinity was mainly located in the south of the equator with maximal 35.75. The simulated salinity was larger than the observation before assimilation especially in the north of the region where the observed salinity was small. After assimilation, the salinity deviation in the whole region was gener-
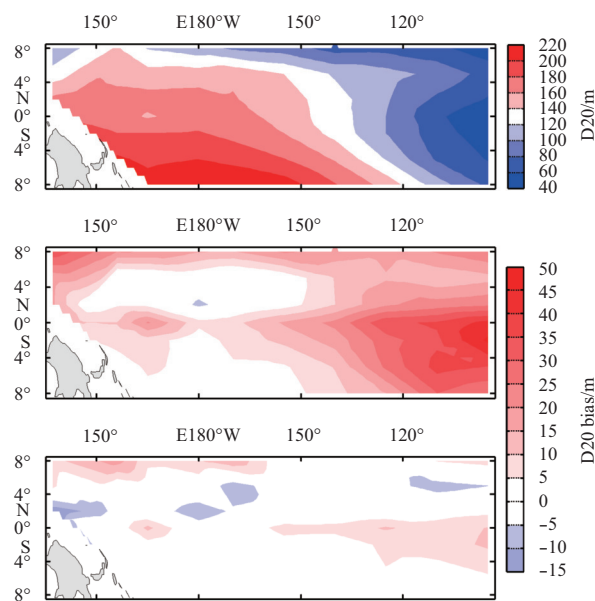
Fig. 5. The horizontal distributions of time-averaged observed D20 (top), simulated D20 bias without the data assimilation (middle) and with the data assimilation (bottom).
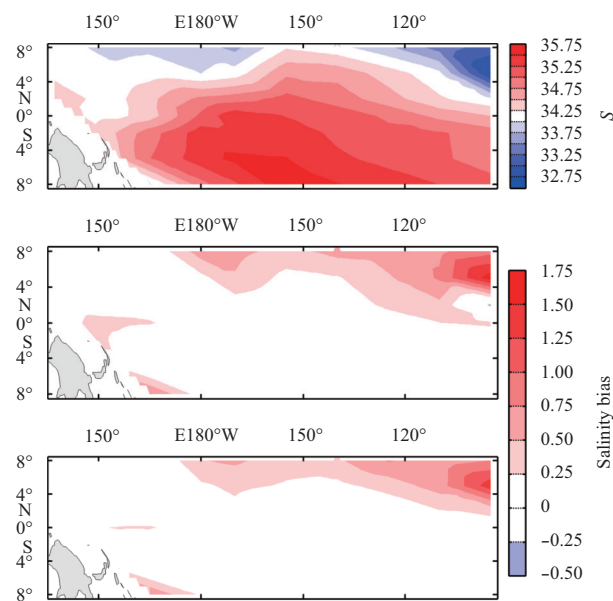

Fig. 6. The horizontal distributions of time-averaged observed salinity at 1 m (top), simulated Salinity bias without the data assimilation (middle) and with the data assimilation (bottom).

ally less than that before assimilation, but the amplitude of reduction was much smaller than the temperature and D20 case.

3.3 *Reconstruction of physical process analysis*

In this part, time-related physical processes of variables were analyzed. First, we discussed the variation of the temperature amplitude at different depths by the vertical distribution of zonal- and meridional-averaged standard deviation of the temperature. In a meridional-averaged case (Fig. 7, left column), The observation results show that the amplitude variation of the temperature is the biggest near the thermocline. Farther away from the thermocline, the smaller amplitude of the temperature changes. Before assimilation, the simulated upper mixed layer was deeper

than the observation especially on the east and west sides. Compared with the observation, the amplitude variation of the temperature was generally small. But it was noteworthy that on the east side, there existed a region under 20°C isotherm where the simulated amplitude was bigger than the observation. It may be because the simulated upper mixed layer without assimilation is deeper than the observation which results in an illusory amplitude. After assimilation, the depth of the upper mixed layer was well agreeable with the observations and there was a significant improvement of the amplitude although it was also smaller than the observed values. The phenomenon of the illusory amplitude before assimilation was disappeared. Similarly, in the zonal-averaged case (Fig. 7, right column), according to the observation, the
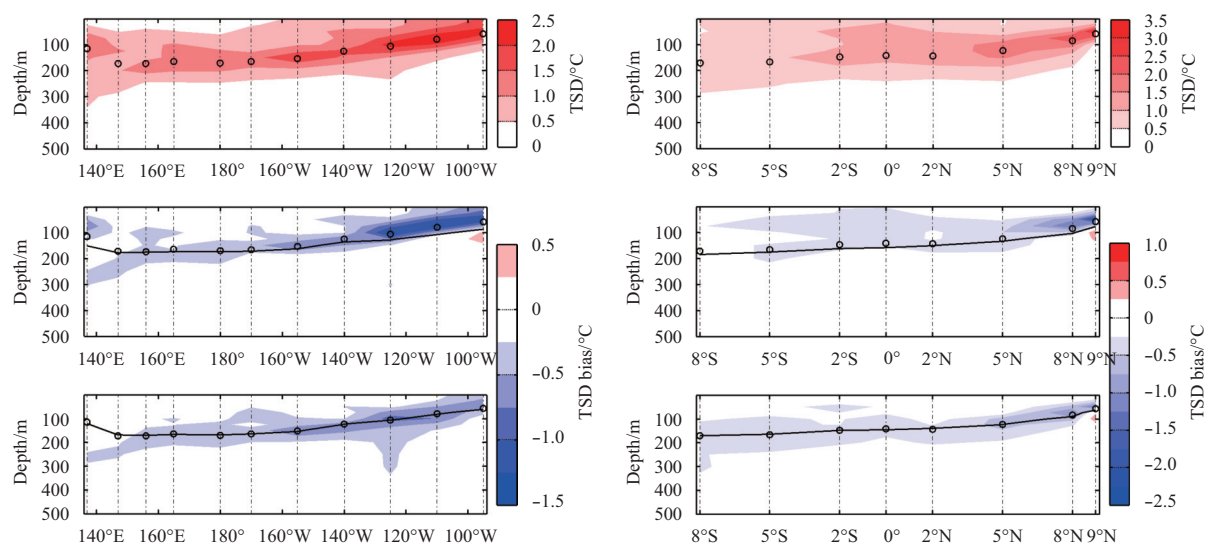

Fig. 7. The vertical distributions of the temperature standard deviation (TSD) (top), the simulated TSD bias without the data assimilation (middle) and with the data assimilation (bottom). The left column represents meridional-time-averaged results while the right column zonal-time-averaged. The hollow circles are the observed D20 and the black solid lines are the simulated results of D20.
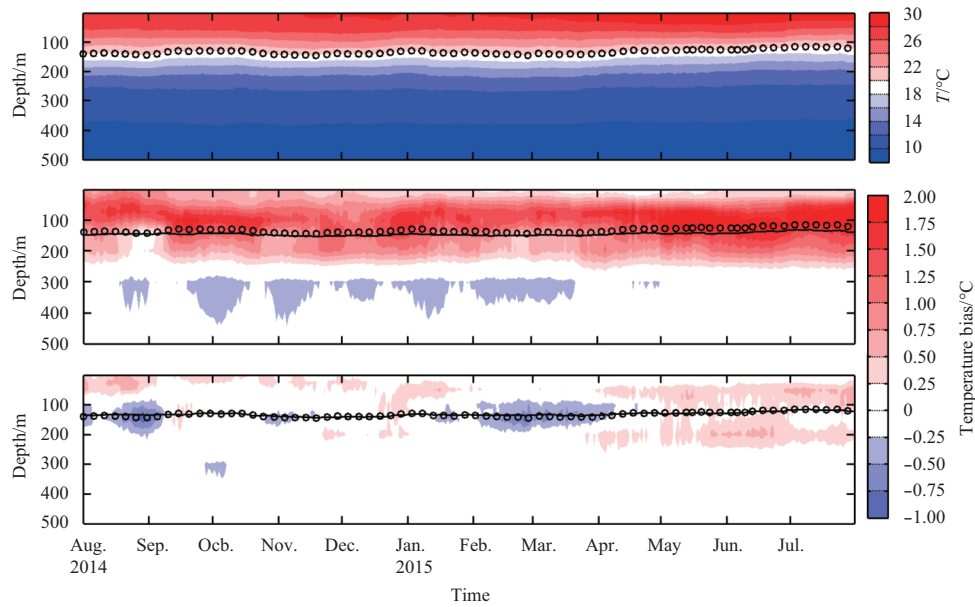
Fig. 8. The time-depth distribution of the horizontally averaged temperature (top), simulated temperature bias without the data assimilation (middle) and with the data assimilation (bottom).
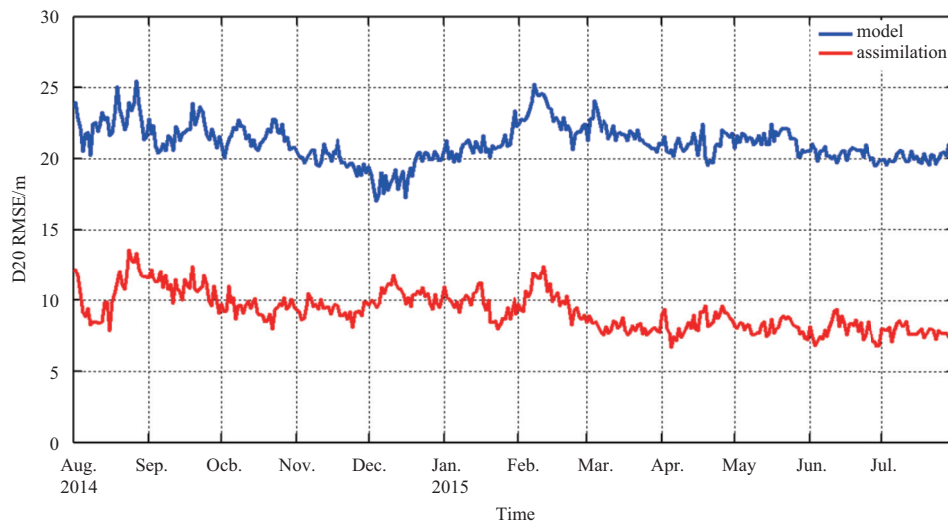

Fig. 9. The time series of the RMSEs with (red) and without (blue) the data assimilation for horizontally averaged D20.

farther away from the thermocline, the smaller amplitude of the temperature changes. The amplitude variation of the temperature was generally smaller than the observation besides east of the region below the thermocline before assimilation. After assimilation, the results were well agreeable with the observation and the physical evolution process of the temperature through the whole vertical especially in the upper ocean was more reasonable.

Then we inspected the process of the temperature change along time at different depths. As was shown in Fig. 8, the simulated upper mixed layer before assimilation was generally deeper during the comparison period than the observation especially from September to November 2014 and from April to July 2015 which led to a great temperature positive deviation near 20°C isotherm and the maximum deviation could be up to 2°C. In addition, the simulated temperature was somewhat lower than the observation below 300 m in most of the time, which was probably associated with the larger simulated thermocline layer

depth. The larger simulated thermocline layer depth probably means that the temperature at the top of the thermocline layer is lower and the temperature on the bottom of the thermocline layer is also lower after a sharp cooling from the top to bottom of the thermocline layer. After assimilation, we could clearly see that there was a significant improvement of the depth of mixed layer which was more consistent with the observation. The temperature deviation throughout the water column was significantly reduced by 76.4% averagely especially near the 20°C isotherm which suggested that the complex variation process in the upper mixed layer had greatly improved after assimilation.

Figure 9 shows the time series of the RMSE of D20. In most of the time, the RMSE was less than 10 m after assimilation while the RMSE was more than 20 m before assimilation. The RMSE of D20 after assimilation was sharply reduced by 56.4% compared with the case without assimilation. This also indicates that the variation process of the temperature near the thermocline was

more realistic after assimilation.

It is noteworthy that the 20°C isotherm in the tropical Pacific is usually regarded as the measurement of an effective heat content in the upper ocean. So, it could be speculated that there might be a great improvement for the variation process of the heat content in the upper ocean in the tropical Pacific after assimilation. What is more, the location of the thermocline in the tropical ocean which is typically represented by D20 plays a key role in an ENSO cycle (Zhang and Levitus, 1996, 1997). So we can infer that the assimilation data set can play an important role in studying the process of the ENSO in the future research due to its high precision for D20.

Next, we inspected the time evolution characteristics of the salinity profiles S1, S2, SC1 and SC2. As shown in Figs 10–13,

these four simulated salinity profiles both had some degree of improvement on the whole after assimilation, but the effect was not obvious. There might be two reasons, one was attributed to fewer amounts of salinity data to be assimilated and the other was probably resulted from the imperfect parameterization process in model such as evaporation, precipitation and so on.

Finally, we focus on discussing the dynamic process of the time evolution characteristics of the EUC recorded by Stes SC1 and SC2 at the background of the strong 2014/2016 El Niño event.

Although the research of the strong El Niño in 2014/2016 is limited and consensus has not been reached in many respects, the process is generally divided into three stages. The first stage is the process of weak central pattern El Niño from the autumn of
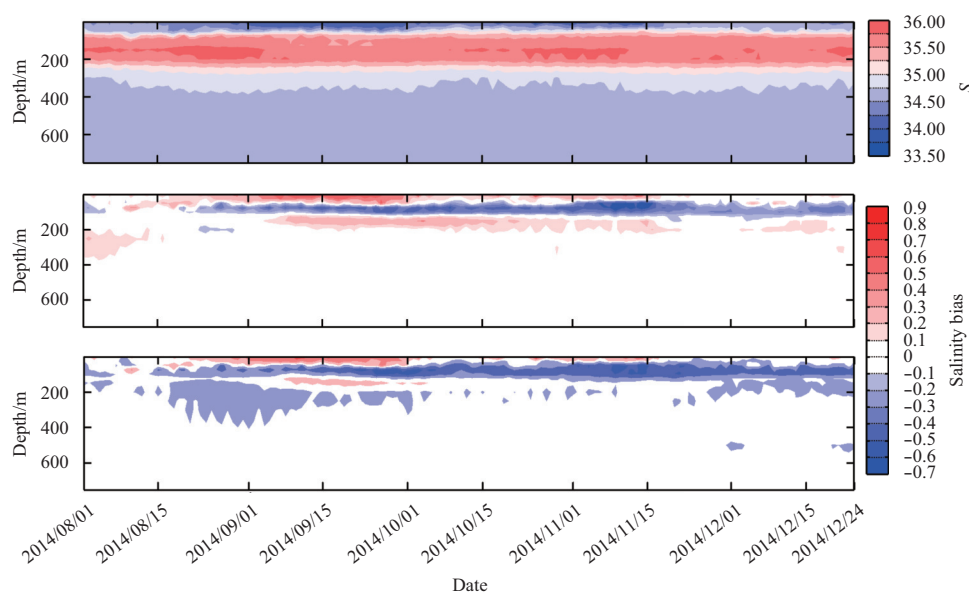

Fig. 10. The time-depth distribution of the salinity profile at Site S1 (top), the simulated salinity bias without the data assimilation (middle) and with the data assimilation (bottom).
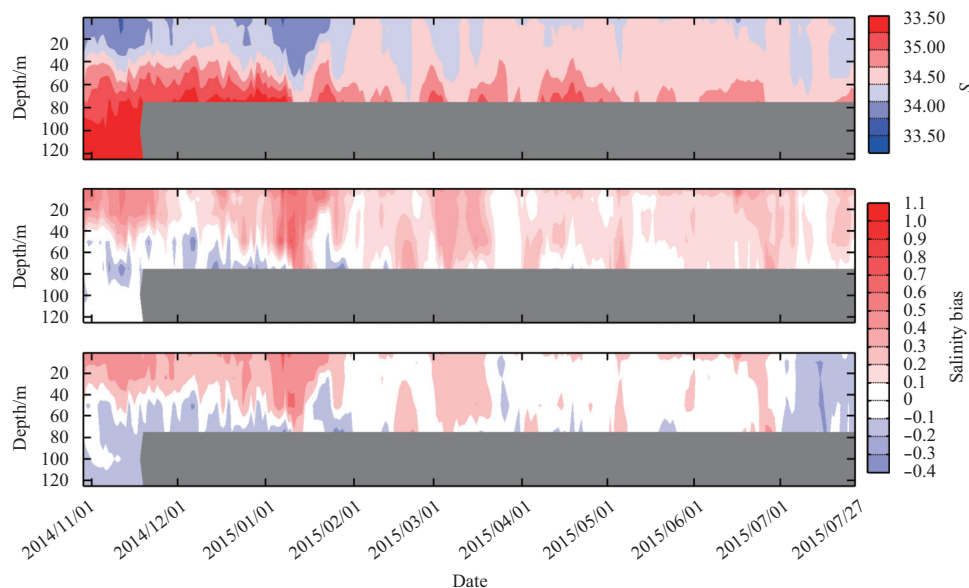

Fig. 11. The time-depth distribution of the salinity profile at Site S2 (top), the simulated salinity bias without the data assimilation (middle) and with the data assimilation (bottom).
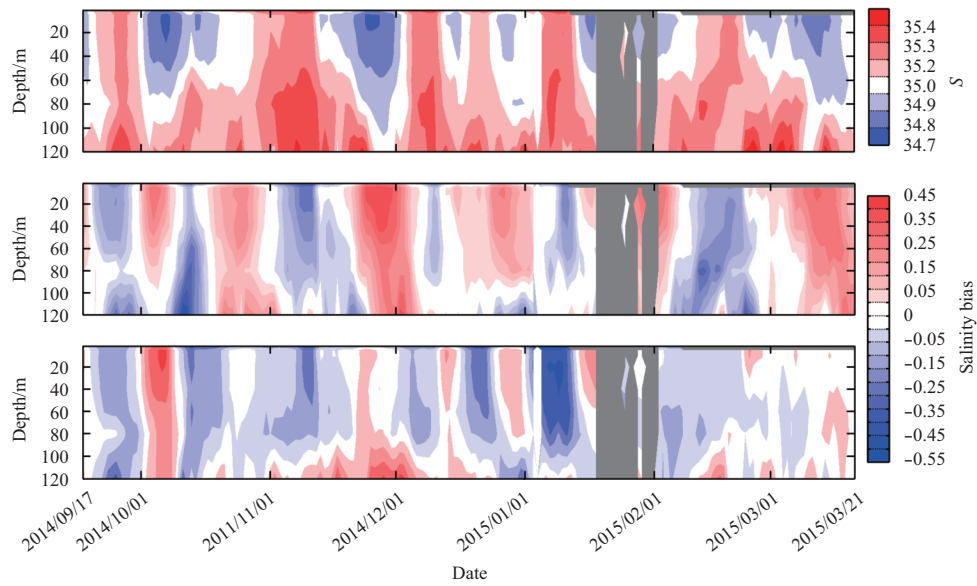
Fig. 12. The time-depth distribution of the salinity profile at Site SC1 (top), the simulated salinity bias without the data assimilation (middle) and with the data assimilation (bottom).
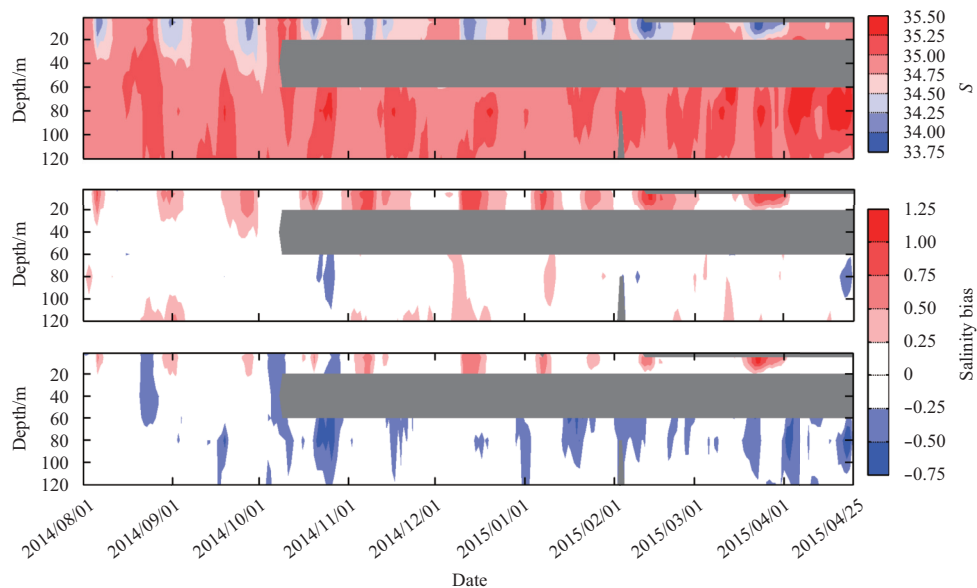

Fig. 13. The time-depth distribution of the salinity profile at Site SC2 (top), the simulated salinity bias without the data assimilation (middle) and with the data assimilation (bottom).

2014 to the spring of 2015. The second stage is the process of strong eastern pattern El Niño between the spring of 2015 and the end of 2015. The third stage is the process of recession in 2016(Yuan et al., 2016; Ren et al., 2016; Zhai et al., 2016).

　　The time-depth distribution of the zonal current at Site SC1 is shown in Fig. 14. Generally, the zonal current in the Assim data set is more close to the observation in most of the time and depth than that in the No_assim data set though the velocity was not assimilated. The mean absolute error of the zonal current in the Assim data set was reduced by 26.5% compared with that in the No_assim data set. The observation shows that the zonal current at SC1 is controlled by the South Equatorial Current (SEC) and the EUC alternately between September and December 2014. During this period, the EUC increased gradually and then decreased. The strength of the EUC reached the maximum in Octo-

ber when the SEC disappeared on the sea surface. The strength of the simulated EUC before assimilation, by contrast, was almost stable before December 2014 which was very different from the observation case. Well, the EUC extracted from the Assim data set could starkly reproduce the characteristic of the time-depth distribution that depicted by the observation during the period of 2014 except September.

　　The observation in 2015 at SC1 shows that the zonal current in the whole vertical is dominated by the EUC from March to mid-June. Then the SEC appears obviously after mid-June and the maximum can be up to 0.6 m/s. Overall, the observed strength of the EUC gradually weakened after March 2015 which reflected the impact of 2014/2016 El Niño on the EUC in the eastern Pacific Ocean. In history, during the process of El Niño, the EUC appeared obvious weakening phenomenon (Firing et al.,
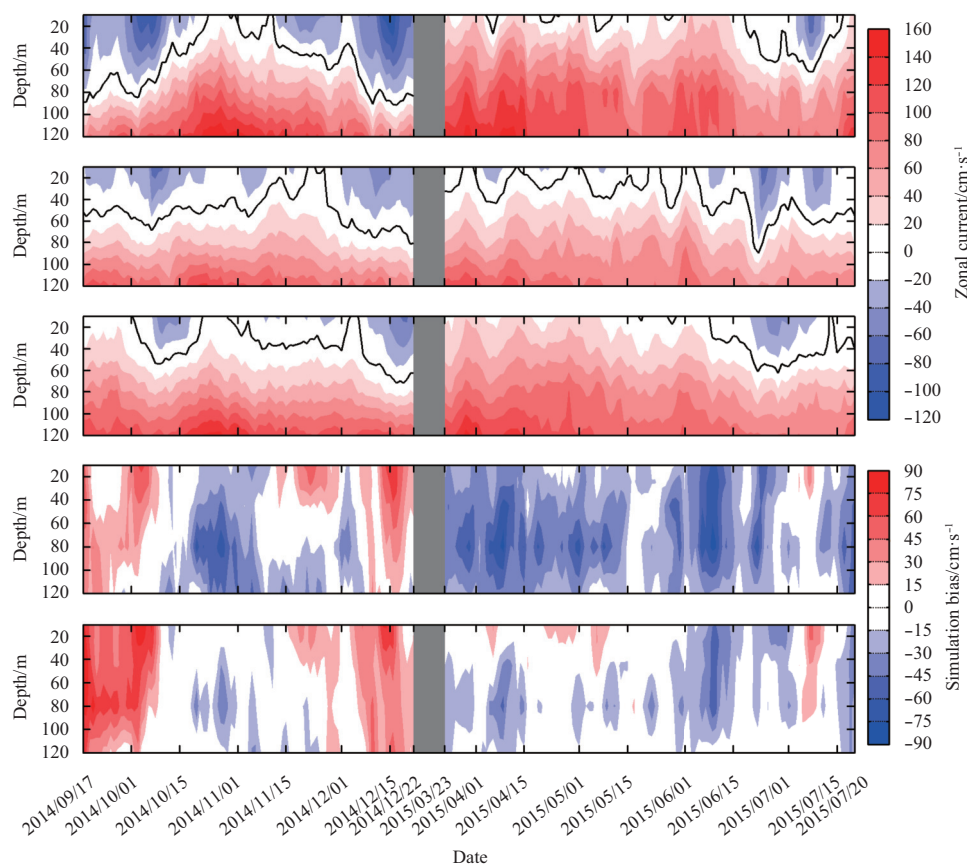
Fig. 14. The time-depth distribution of the zonal current at Site SC1. The observation, the simulation without the data assimilation, the simulation with the data assimilation, the simulation bias without the data assimilation, and the simulation bias with the data assimilation are from top to bottom.

1983; Guan, 1986; Jiang, 1993; Shi et al., 1999). The pattern of the simulated zonal current after assimilation was well consistent with the phenomenon than the case before assimilation which indicated that the zonal current in the Assim data set could catch the signal of El Niño well.

The time-depth distribution of the zonal current at Site SC2 is shown in Fig. 15. The observation shows that the strength of the EUC at SC2 increases gradually from August to the beginning of September 2014. Especially, the strength of the EUC reaches the maximum on September 8, 2014 when the SEC disappears. This phenomenon occurred before the 2014/2016 El Niño event. In fact, the strength of the EUC usually increases gradually before the El Niño event in history (Chen et al., 2005). Therefore, the above phenomenon could be regarded as an important indication of El Niño event. From Fig. 15, though the simulated EUC before assimilation was slightly increased over time between August and the beginning of September 2014, the trend was indeed very small. The assimilation case, by contrast, was more consistent with the observation. Especially, the process that the observed EUC reached to the sea surface at the beginning of September was well reproduced by the assimilation case.

Furthermore, compared with the mean absolute error of the zonal current in the No_assim data set at SC2, the mean absolute error of the zonal current in the Assim data set was reduced by 25.5%. The observed zonal current at SC2 in the whole vertical was dominated by the EUC from the end of April to early June 2015 and the strength of the EUC decreased gradually during this period which was closely related to El Niño (Wang et al., 2009).

Obviously, the strength of the simulated EUC before and after assimilation decreased gradually during this period, but the assimilation results decreased more slowly which was well consistent with the observation. What is more, compared with the observation, the position of the core of EUC in the assimilation case was more precise.

In short, the simulated zonal current in the Assim data set could accurately reproduce the observed variation process characteristics of the EUC during the 2014/2016 El Niño event. Because of the obvious improvement of thermohaline structure in Upper Ocean after assimilation, the current field in the Assim data set tended to be more reasonable.

4 Discussion and summary

In order to assess the effectiveness of the global high resolution assimilation data set based on the FIOCOM and the EAKF, different variables of data sets before and after assimilation were analyzed in the tropical Pacific from the following three aspects. The overall error statistics of the temperature, D20, the salinity and the velocity showed that the simulated results of all variables after assimilation had been improved with different degrees. The error of the temperature and D20 decreased in the most significant, followed by the zonal current and the salinity. The analysis of the time-averaged physical structure showed that the distribution characteristic of the horizontal and vertical structures of each variable was more reasonable after assimilation. The most obvious improvement was that the simulated structure of the upper mixed layer represented by D20 was significantly close to
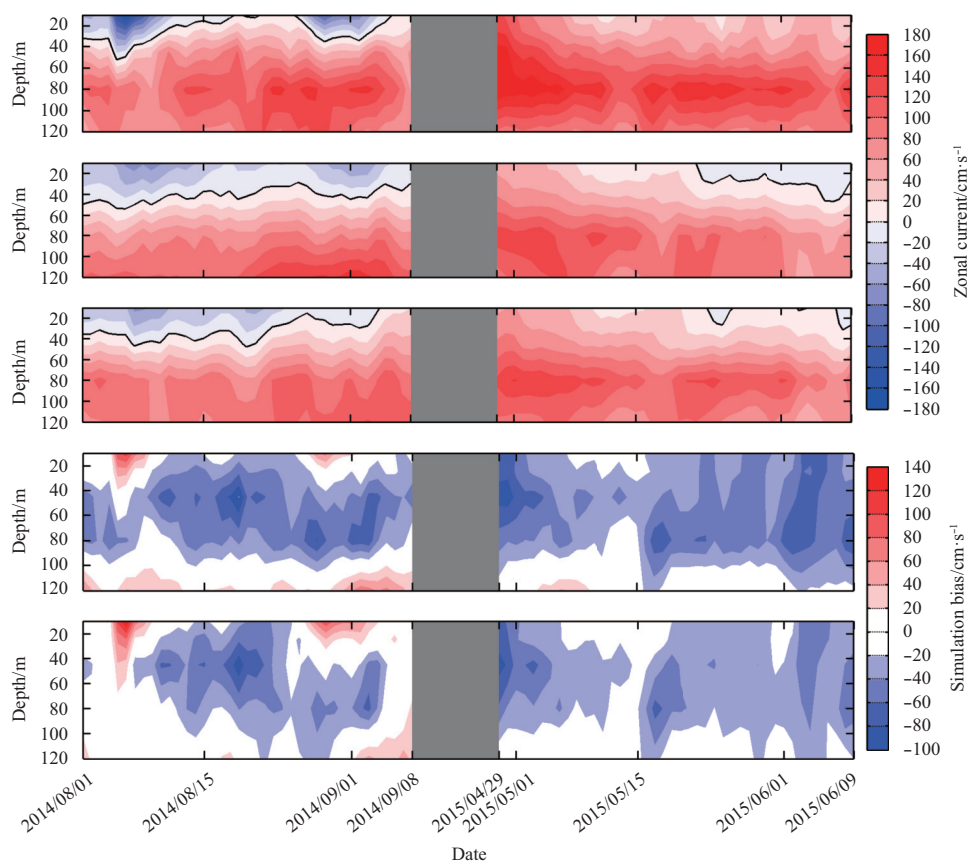
Fig. 15. The time-depth distribution of the zonal current at Site SC2. The observation, the simulation without the data assimilation, the simulation with the data assimilation, the simulation bias without the data assimilation, and the simulation bias with the data assimilation are from top to bottom.

reality after assimilation. Finally the reconstructed physical process was compared with the observation and the results showed that the time evolution characteristics of all variables were more consistent with the observation after assimilation.

It is worth noting that the analysis of dynamic process of the zonal current at two sites shows that the simulated process of the EUC in the data assimilation data set can generally reproduce the observed strong 2014/2016 El Niño event during this period. The current field in the data assimilation data set tends to be more reasonable due to the improvement of the thermohaline structure in the upper ocean after assimilation. Predictably, studying the important events using the high resolution assimilation data set evaluated in this paper will have a positive meaning.

What we should pay attention to is that though the results of *v* component of velocity and the salinity after assimilation have a certain degree of improvement, but the amplitude of improvement is much smaller than other variables. In view of this, some outlook is projected as follows. In terms of a numerical model, a physical processes in model that influences the salinity such as evaporation and precipitation should be better optimized. In terms of assimilation, the salinity and velocity observation data should be collected as much as possible to be absorbed into a model system. In general, the analyzed high resolution assimilation data set was well constructed and could be used for a series of scientific issues efficiently.

References

Anderson J E, Riser S C. 2014. Near-surface variability of temperature and salinity in the near-tropical ocean: observations from profiling floats. Journal of Geophysical Research: Oceans, 119(11): 7433–7448

Balmaseda M, Anderson D. 2009. Impact of initialization strategies and observations on seasonal forecast skill. Geophysical Research Letters, 36(1): L01701

Bennett A F, Chua B S, Harrison D E, et al. 1998. Generalized inversion of tropical atmosphere-ocean data and a coupled model of the tropical Pacific. Journal of Climate, 11(7): 1768–1792

Bhowmick S A, Agarwal N, Ali M M, et al. 2016. Role of ocean heat content in boosting post-monsoon tropical storms over Bay of Bengal during La-Niña events. Climate Dynamics, : doi: 10.1007/s00382-016-3428-5

Chen Jinnian, Lv Xinyan, Hu Dunxin. 2005. Variable properties of the equatorial undercurrent in the pacific and its anomalous warm water eastward propagation. Advances in Water Science (in Chinese), 16(6): 792–798

Chowdary J S, Harsha H S, Gnanaseelan C, et al. 2017. Indian summer monsoon rainfall variability in response to differences in the decay phase of El Niño. Climate Dynamics, 48(7-8): 2707–2727

Firing E, Lukas R, Sadler J, et al. 1983. Equatorial undercurrent disappears during 1982-1983 El Niño. Science, 222(4628): 1121–1123

Fu Weiwei, Zhu Jiang, Yan Changxiang, et al. 2009. Toward a global ocean data assimilation system based on ensemble optimum interpolation: altimetry data assimilation experiment. Ocean Dynamics, 59(4): 587–602

Gao Chuan, Zhang Ronghua. 2017. The roles of atmospheric wind and entrained water temperature (T_e) in the second-year cooling of the 2010-12 La Niña event. Climate Dynamics, 48(1-2): 597–617

Guan Bingxian. 1986. Current structure and its variation in the equat-

orial area of the western north Pacific Ocean. Chinese Journal of Oceanology and Limnology, 4(3): 239–255

Hayes S P, Mangum L J, Picaut J, et al. 1991. TOGA-TAO: a moored array for real-time measurements in the tropical Pacific Ocean. Bulletin of the American Meteorological Society, 72(3): 339–347

Henocq C, Boutin J, Reverdin G, et al. 2010. Vertical variability of near-surface salinity in the tropics: consequences for L-band radiometer calibration and validation. Journal of Atmospheric and Oceanic Technology, 27(1): 192–209

Jiang Jingzhong. 1993. A event of Pacific equatorial undercurrent inversion during El Niño. Donghai Marine Science (in Chinese), 11(1): 1–9

Keppenne C L, Rienecker M M. 2003. Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble Kalman filter. Journal of Marine Systems, 40-41: 363–380

Kimoto M, Yoshikawa I, Ishii M. 1997. An ocean data assimilation system for climate monitoring (gtspecial issueltdata assimilation in meteology and oceanography: theory and practice). Journal of the Meteorological Society of Japan: Series II, 75(1): 471–487

Masuda S, Awaji T, Sugiura N, et al. 2003. Improved estimates of the dynamical state of the north pacific ocean from a 4 dimensional variational data assimilation. Geophysical Research Letters, 30(16): 1868

McPhaden M J. 2002. El Niño and La Niña: causes and global consequences. In: Munn T, ed. Encyclopedia of Global Environmental Change. Chichester, UK: John Wiley and Sons, 353–370

Moore A M. 1991. Data assimilation in a quasi-geostrophic open-ocean model of the gulf stream region using the adjoint method. Journal of Physical Oceanography, 21(3): 398–427

Oke P R, Larnicol G, Fujii Y, et al. 2015. Assessing the impact of observations on ocean forecasts and reanalyses: Part 1. Global studies. Journal of Operational Oceanography, 8(S1): S49–S62

Paek H, Yu Jinyi, Qian Chengcheng. 2017. Why were the 2015/2016 and 1997/1998 extreme El Niños different? Geophysical Research Letters, 44(4): 1848–1856

Parent L, Testut C E, Brankart J M, et al. 2003. Comparative assimilation of Topex/Poseidon and ERS altimeter data and of Tao temperature data in the tropical Pacific Ocean during 1994-1998, and the mean sea-surface height issue. Journal of Marine Systems, 40-41: 381–401

Qiao Fangli, Ma Jian, Xia Changshui, et al. 2006. Influences of the surface wave-induced mixing and tidal mixing on the vertical temperature structure of the Yellow and East China seas in summer. Progress in Natural Science, 16(7): 739–746

Qiao Fangli, Yang Yongzeng, Xia Changshui, et al. 2008. The role of surface waves in the ocean mixed layer. Acta Oceanologica Sinica, 27(3): 30–37

Qiao Fangli, Yuan Yeli, Deng Jia, et al. 2016. Wave-turbulence interaction-induced vertical mixing and its effects in ocean and climate models. Philosophical Transactions of the Royal Society: A. Mathematical, Physical and Engineering Sciences, 374(2065): 20150201

Qiao Fangli, Yuan Yeli, Ezer T, et al. 2010. A three-dimensional surface wave-ocean circulation coupled model and its initial testing. Ocean Dynamics, 60(5): 1339–1355

Qiao Fangli, Yuan Yeli, Yang Yongzeng, et al. 2004. Wave-induced mixing in the upper ocean: distribution and application to a global ocean circulation model. Geophysical Research Letters, 31(11): L11303

Ren Hongli, Liu Ying, Zuo Jinqing, et al. 2016. The new generation of ENSO prediction system in Beijing Climate Centre and its predictions for the 2014/2016 super El Niño event. Meteorological Monthly, 42(5): 521–531

Salau O R, Akinyemi S A. 2015. The impacts of El Niño/southern oscillation on changing precipitation over the tropical Pacific. International Journal of Environmental Sciences, 5(5): 995–1010

Shi Qiang, Pu Shuzhen, Su Jie, et al. 1999. Investigation of main current system and equatorial planetary waves in the tropical Pa-

cific during twice untypical El Niño events. Haiyang Xuebao (in Chinese), 21(4): 27–34

Shu Qi, Qiao Fangli, Bao Ying, et al. 2014. Assessment of arctic sea ice simulation by FIO-ESM based on data assimilation experiment. Haiyang Xuebao (in Chinese), 37(11): 33–40

Shu Qi, Qiao Fangli, Song Zhenya, et al. 2011. Improvement of MOM4 by including surface wave-induced vertical mixing. Ocean Modelling, 40(1): 42–51

Stammer D, Köhl A, Awaji T, et al. 2010. Ocean information provided through ensemble ocean syntheses. In: Proceedings of Ocean Obs'09: Sustained Ocean Observations and Information for Society. Venice, Italy: ESA Publication, 920–930

Sun Chaojiao, Rienecker M M, Rosati A, et al. 2007. Comparison and sensitivity of odasi ocean analyses in the tropical pacific. Monthly Weather Review, 135(6): 2242–2264

Torma P, Krámer T. 2017. Modeling the effect of waves on the diurnal temperature stratification of a shallow lake. Periodica Polytechnica Civil Engineering, 61(2): 165–175, doi: 10.3311/PPci.8883

Vidard A, Anderson D L, Balmaseda M. 2007. Impact of ocean observation systems on ocean analysis and seasonal forecasts. Monthly Weather Review, 135(2): 409–429

Wang Hongna, Chen Jinnian, He Yijun. 2009. Variations of Equatorial Undercurrent and its relationship with ENSO cycle. Haiyang Xuebao (in Chinese), 31(3): 1–11

Wang Ou, Fukumori I, Lee T, et al. 2004. Eastern equatorial Pacific Ocean T-S variations with El Niño. Geophysical Research Letters, 31(4): L04305

Wen Na, Liu Zhengyu, Liu Yinghui. 2015. Direct impact of El Niño on East Asian summer precipitation in the observation. Climate Dynamics, 44(11-12): 2979–2987

Wu Lichuan, Rutgersson A, Sahlée E. 2015. Upper-ocean mixing due to surface gravity waves. Journal of Geophysical Research: Oceans, 120(12): 8210–8228, doi: 10.1002/2015JC011329

Xuan Jiliang, Huang Daji, Zhou Feng, et al. 2012. Application of data assimilation to synoptic temperature mapping of the coastal ocean survey. Oceanologia et Limnologia Sinica (in Chinese), 43(1): 17–26

Xue Yan, Wen Caihong, Yang Xiaosong, et al. 2017. Evaluation of tropical Pacific observing systems using NCEP and GFDL ocean data assimilation systems. Climate Dynamics, 49(3): 843–868

Yin Xunqiang, Qiao Fangli, Shu Qi. 2011. Using ensemble adjustment Kalman filter to assimilate Argo profiles in a global OGCM. Ocean Dynamics, 61(7): 1017–1031

Yin Xunqiang, Qiao Fangli, Yang Yongzeng, et al. 2010. An ensemble adjustment Kalman filter study for Argo data. Chinese Journal of Oceanology and Limnology, 28(3): 626–635

Yin Xunqiang, Qiao Fangli, Yang Yongzeng, et al. 2012. Argo data assimilation in ocean general circulation model of northwest Pacific Ocean. Ocean Dynamics, 62(7): 1059–1071

Yuan Yuan, Gao Hui, Jia Xiaolong, et al. 2016. Influences of the 2014-2016 super El Niño event on climate. Meteorological Monthly (in Chinese), 42(5): 532–539

Yuan Yeli, Hua Feng, Pan Zengdi, et al. 1991. LAGFD-WAM numerical wave model-I. Basic physical model. Acta Oceanologica Sinica, 10(4): 483–488

Zhai Panmao, Yu Rong, Guo Yanjun, et al. 2016. The strong El Niño in 2015/2016 and its dominant impacts on global and China's climate. Acta Meteorologica Sinica (in Chinese), 74(3): 309–321

Zhang Ronghua, Levitus S. 1996. Structure and evolution of interannual variability of the tropical Pacific upper ocean temperature. Journal of Geophysical Research: Oceans, 101(C9): 20501–20524

Zhang Ronghua, Levitus S. 1997. Interannual variability of the coupled tropical Pacific ocean-atmosphere system associated with the El Niño-Southern Oscillation. Journal of Climate, 10(6): 1312–1330

Zuo T, Chen Jinnian, Wang Hongna. 2014. Impact of the central Pacific zonal wind divergence and convergence on the central Pacific El Niño event. Acta Oceanologica Sinica, 33(11): 85–89