

Reconstruction and analysis of long-term satellite-derived sea surface temperature for the South China Sea

Hong-Ngu T. Huynh¹ · Aida Alvera-Azcárate¹ · Alexander Barth¹ · Jean-Marie Beckers¹

Received: 3 June 2015 / Revised: 13 March 2016 / Accepted: 15 March 2016 / Published online: 6 April 2016
© The Oceanographic Society of Japan and Springer Japan 2016

Abstract Sea surface temperature (SST) is one of the key variables often used to investigate ocean dynamics, ocean-atmosphere interaction, and climate change. Unfortunately, the SST data sources in the South China Sea (SCS) are not abundant due to sparse measurements of in situ SST and a high percentage of missing data in the satellite-derived SST. Therefore, SST data sets with low resolution and/or a short-term period have often been used in previous researches. Here we used Data INterpolating Empirical Orthogonal Functions, a self-consistent and parameter-free method for filling in missing data, to reconstruct the daily nighttime 4-km AVHRR Pathfinder SST for the long-term period spanning from 1989 to 2009. In addition to the reconstructed field, we also estimated the local error map for each reconstructed image. Comparisons between the reconstructed and other data sets (satellite-derived microwave and in situ SSTs) show that the results are reliable for use in many different researches, such as validating numerical models, or identifying and tracking meso-scale oceanic features. Moreover, the Empirical Orthogonal Function (EOF) analysis of the reconstructed SST and the reconstructed SST anomalies clearly shows the subseasonal, seasonal, and interannual variability of SST under the influence of monsoon and El Niño-Southern Oscillation (ENSO), as well as reveals some oceanic features that could not be captured well in previous EOF analyses. The SCS SST often lags ENSO by about half a year. However,

in this study, we see that the time lag changes with the frequencies of the SST variability, from 1 to 6 months.

Keywords South China Sea · AVHRR Pathfinder SST · Subseasonal · Seasonal and interannual variability · Monsoon · ENSO · DINEOF

1 Introduction

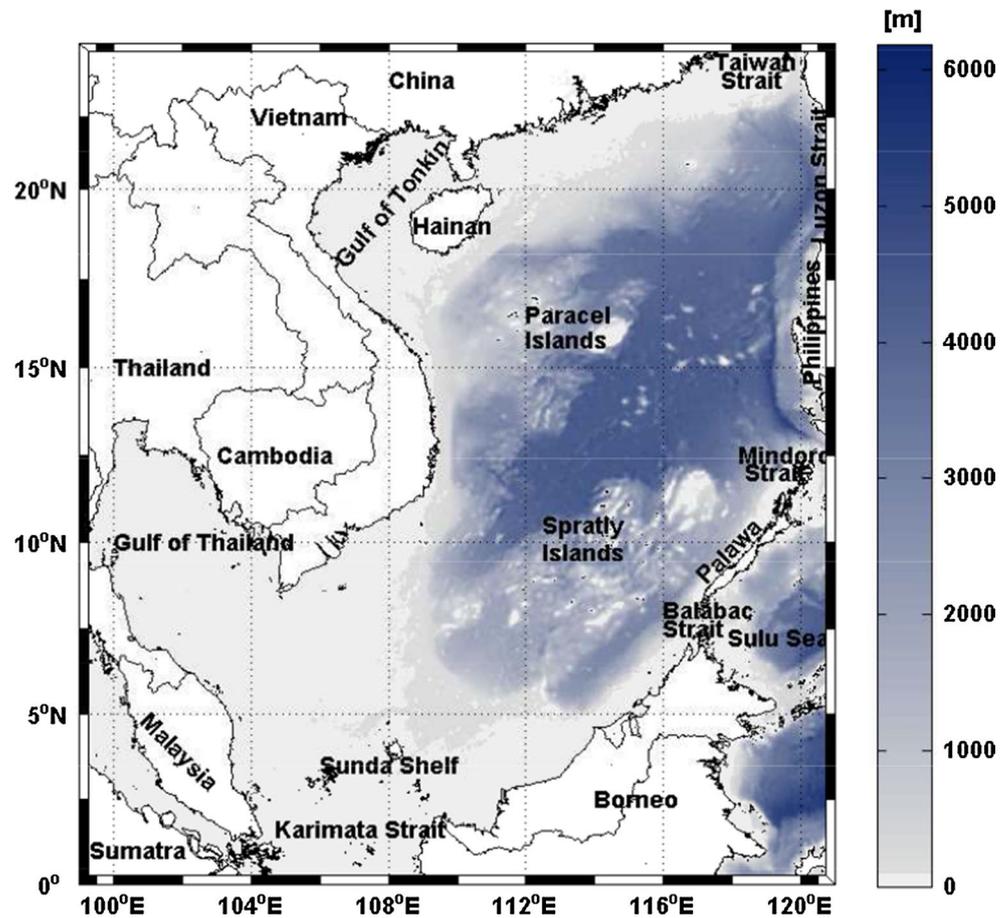
The South China Sea (SCS), a semi-closed marginal sea, covers a large area (3.5×10^6 km²), from 0–23°N and 99–121°E, with an average depth over 2000 m (Fig. 1). It is almost surrounded by high mountain ranges. The sea is connected to the East China Sea, the Sulu Sea, the Java Sea, and the Indian Ocean through narrow and shallow straits, except for the western Pacific Ocean through the deep and wide Luzon Strait, which is the major pathway of water entering into the SCS. The topography of the sea is rather complex: wide and shallow continental shelves in the northwest and southwest, steep slopes without shelves in the east, a deep basin with a maximum depth reaching more than 5000 m in the northeast, and numerous islands over the sea.

Located in southeastern Asia, the SCS is strongly influenced by the East Asian monsoon (Wyrтки 1961; Chao et al. 1996; Chu et al. 1997). The cold and strong northeast monsoon lasts from November to April. The warm and weaker southwest monsoon lasts from mid-May to mid-September. In the two periods of monsoon transition, the prevailing direction of the surface wind is not clear. Affected by the orographic features of the surrounding land, the surface wind is intensive along the northeast-southwest diagonal of the basin, and a wind jet is generated off the southeastern coast of Vietnam in winter as well as in summer (Xie et al.

✉ Hong-Ngu T. Huynh
hongnguhuyhthi@gmail.com; HN.Huynh@student.ulg.ac.be

¹ GeoHydrodynamics and Environment Research, AGO, University of Liège, Quartier Agora, Allée du 6 Août 19, B5a, Sart-Tilman, 4000 Liège, Belgium

Fig. 1 Bathymetric map of the South China Sea



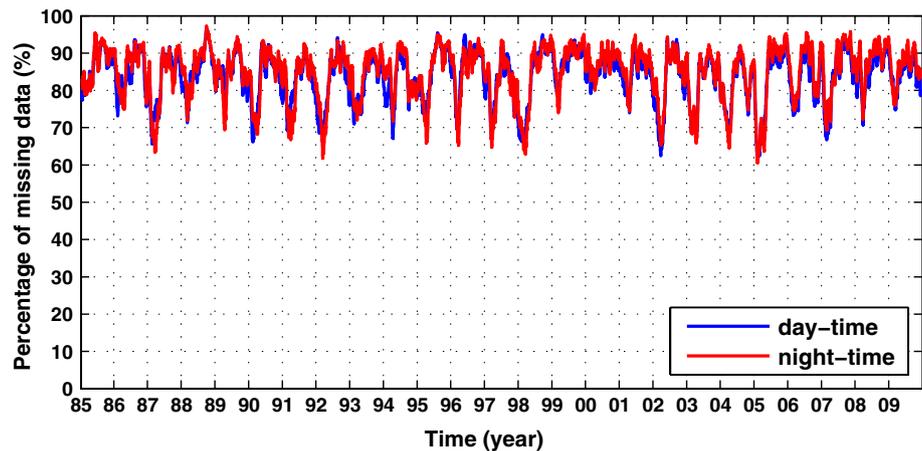
2003, 2007; Liu et al. 2004). In response to the seasonally reversing monsoon associated with the intrusion of the Kuroshio current through the Luzon Strait, there is a basin-scale cyclonic circulation in winter, while there are a weak northern cyclonic circulation and a southern anticyclonic circulation in summer (Wyrтки 1961; Chao et al. 1996; Chu et al. 1999; Qu et al. 2000). Associated with the circulation, a strong current flows along the western coast of the sea. Under the impact of the orographic wind jet and the shallow shelf topography, the southward and northeastward coastal jet separation occurs from the coast of central Vietnam in winter and summer, respectively (Chu et al. 1999; Liu et al. 2004; Xie et al. 2007; Gan and Qu 2008). The seasonal reversal in the monsoon-induced circulation also leads to the variability of sea surface temperature (SST) in the SCS. In winter, there are water exchanges with the East China Sea and the Pacific Ocean through the Taiwan and Luzon Straits, respectively. The western boundary current and cyclonic circulation transport the waters from the north to the south and cause a distinct cold tongue in the southern Vietnam coast (Liu et al. 2004; Gan and Qu 2008). In summer, SST is rather uniform, around 28.5–29.5°C. In addition, under the effect of the monsoon and coastal features, upwelling occurs along the coasts. The upwelling

appears in the southeastern Vietnam coast (Wyrтки 1961; Kuo et al. 2000; Xie et al. 2003; Dippner et al. 2007), the Taiwan Strait (Tang et al. 2002, 2004; Hong et al. 2009), and Hainan Island (Lü et al. 2008; Su and Pohlmann 2009; Li et al. 2012). In contrast, it occurs in the northeast off the Luzon Strait (Wang et al. 2010) and the northwest Borneo Island (Yan et al. 2015) in winter.

The influence of El Niño-Southern Oscillation (ENSO) on the SCS is also well documented (e.g., Chao et al. 1996; Klein et al. 1999; Wang et al. 2000; Liu et al. 2004; Qu et al. 2004; Fang et al. 2006; Wang et al. 2006; Xie et al. 2007; Lin et al. 2011; Wu et al. 2014), and is explained by the fact that the effect of ENSO is located in the western Pacific Ocean (e.g., Klein et al. 1999; Wang et al. 2000; Wang 2002; Qu et al. 2004). The interannual variability of the SCS in response to ENSO is due to the tropical atmospheric bridge (Klein et al. 1999; Wang et al. 2000; Wang 2002) and the water transport through the Luzon Strait (Qu et al. 2004). The response of the SCS lags moderate and strong ENSO events by about half a year (Klein et al. 1999; Fang et al. 2006; Wang et al. 2006).

In general, under the impact of the complex topography, monsoon, and ENSO, the temporal and spatial variability of the dynamics in the SCS is very complex. Therefore,

Fig. 2 Percentage of missing data of daily daytime and night-time AVHRR Pathfinder SST in the SCS in 1985–2009. The lines are smoothed by a 30-day low-pass filter



in order to better understand the SCS's dynamics as well as its relationship with the monsoon and ENSO, the need for long-term and high-resolution data sources of oceanographic variables is essential for researchers. This has been proved through taking advantage of the long-term and high-resolution satellite data sets to investigate the seasonal, intraseasonal and interannual variations of the SCS and the overlying atmosphere in recent researches (e.g., Xie et al. 2003, 2007; Liu et al. 2004).

SST is one of the key variables often used to investigate ocean dynamics, ocean-atmosphere interaction, and climate change. However, in situ measurements of SST in the seas and oceans are sparse. For recent decades, based on the development of remote sensing systems, SST measured from sensors on board of satellites is a valuable source providing data for researchers. Among the satellite-derived data sources, the Advanced Very High Resolution Radiometer (AVHRR) Pathfinder SST (Kilpatrick et al. 2001; Casey et al. 2010), measured by infrared sensors, has been widely used because of its high resolution and long time-series. The disadvantage of the AVHRR SST as any infrared sensor is a high percentage of missing data due to cloud coverage. This becomes more serious in the SCS because it is located in the tropical region, frequently covered by clouds (Guan and Kawamura 2003). The average percentage of missing data of daily AVHRR Pathfinder SST in the SCS is often more than 80 % (Fig. 2). In order to overcome this problem, in this paper, we applied Data Interpolating Empirical Orthogonal Functions (DINEOF), a self-consistent and parameter-free method used for accurately reconstructing incomplete geophysical data sets with low computational cost (Beckers and Rixen 2003; Alvera-Azcárate et al. 2005), to create a daily 4-km cloud-free SST field spanning from 1989 to 2009 for the whole SCS.

There have been previous researches using observation data to investigate the variability of SST in the SCS. For example, Chu et al. (1997) studied the temporal and

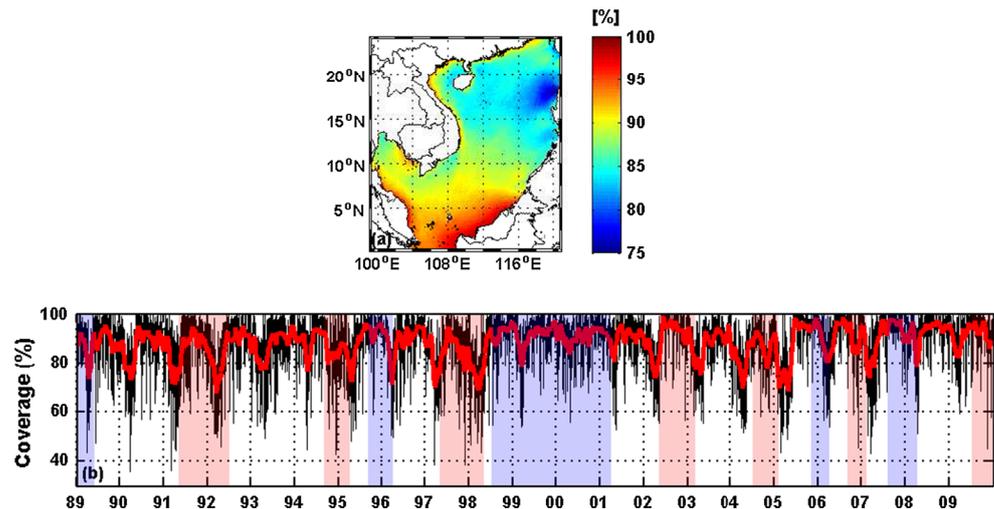
spatial variability of SST under the influence of the monsoon, Fang et al. (2006) investigated the interannual variability of SST in relation to sea surface height and surface wind, Wang et al. (2006) investigated the interannual variability of the SCS associated with El Niño, Lin et al. (2011) studied the variability of SST in relation to the western Pacific warm pool, and Wu and Chen (2015) investigated the intraseasonal SST during boreal winter. However, the data sets used in their researches have low resolution in time and space, and/or short time-series. In this paper, analysing the EOFs in high temporal and spatial resolution of the reconstructed SST and the reconstructed SST anomalies for the period 1989–2009, we can show more details of the SST variability in relation to the monsoon and ENSO, as well as reveal some oceanic features that could not be captured well in previous EOF analyses.

The paper is structured as follows. The data used for this study are briefly described in Sect. 2. In Sect. 3, we introduce the main features of DINEOF. The results and their validation are presented in Sect. 4. The analysis of EOFs is implemented in Sect. 5. The conclusions of our work are presented in Sect. 6.

2 Data and data processing

In order to reconstruct a long-term and high-resolution data set of SST, we chose the daily nighttime 4-km AVHRR Pathfinder SST version 5.0, developed by the University of Miami's Rosenstiel School of Marine and Atmospheric Science, and the NOAA National Oceanographic Data Center in partnership with NASA's Physical Oceanography Distributed Active Archive Center, using the Pathfinder algorithm (Kilpatrick et al. 2001; Casey et al. 2010). We extracted the SST data covering the SCS, 0–24°N and 99–121°E (548 × 502 pixels), spanning from 1989 to 2009 (21 years), including 7670 images.

Fig. 3 Percentage of cloud-covered AVHRR Pathfinder SST in the SCS: **a** spatial variation of cloud-covered SST, **b** temporal variation of cloud-covered SST (black line) with a 30-day low-pass filter (red line). The ENSO events are marked with coloured stripes: El Niño in pink and La Niña in blue. The percentage of missing data decreases during El Niño due to the reduction in cloud coverage over the SCS, except for the 2002–2003 El Niño event, which may be related to the volcanic eruption in this area at that time. In contrast, missing data increases during La Niña



Only nighttime images were used for the reconstruction to avoid the skin-temperature problem. Quality levels are assigned to each pixel in this data set, ranging from 0 (the poorest quality) to 7 (the highest quality) (Kilpatrick et al. 2001). To have the initial data set with enough accuracy for the reconstruction, we only kept pixels assigned with the quality flags of 5, 6 and 7. Statistically, the original data set has a high percentage of missing data (Fig. 3), approximately 88 % of the total data set. Since the Intertropical Convergence Zone (ITCZ) is often positioned over the southern SCS, more clouds cover this region than the others (Guan and Kawamura 2003), resulting in more than 90 % of missing data. Although the northeastern SCS (15–20°N, 116–120°E) is covered by fewer clouds than the other regions, its missing data percentage is still up to 75 %. Moreover, the coverage of clouds also changes with seasons due to the migration of the ITCZ (Guan and Kawamura 2003). The percentage of missing data is higher during the southwest monsoon and lower during the northeast monsoon. Note that the percentage of missing data decreases during El Niño due to the reduction in cloud coverage over the SCS, whereas it increases during La Niña (Klein et al. 1999).

Before applying DINEOF to reconstruct the SST field, some procedures were done to process data. Only images containing at least 5 % of data were retained. Pixels missing more than 95 % of the time were not reconstructed and were considered as “land”. We have applied these procedures because these images and pixels do not contain enough statistical information for the reconstruction and could affect the quality of the overall result (Alvera-Azcárate et al. 2005, 2009). After these steps were carried out, a new data set with 5119 images, 67 % of the initial data set, was kept. Its percentage of missing data was about 83 % of the total data set. A large number of pixels near the coast were marked as land, especially near the Borneo

Island. The total number of pixels for the reconstruction was about 7.19×10^8 pixels, excluding land.

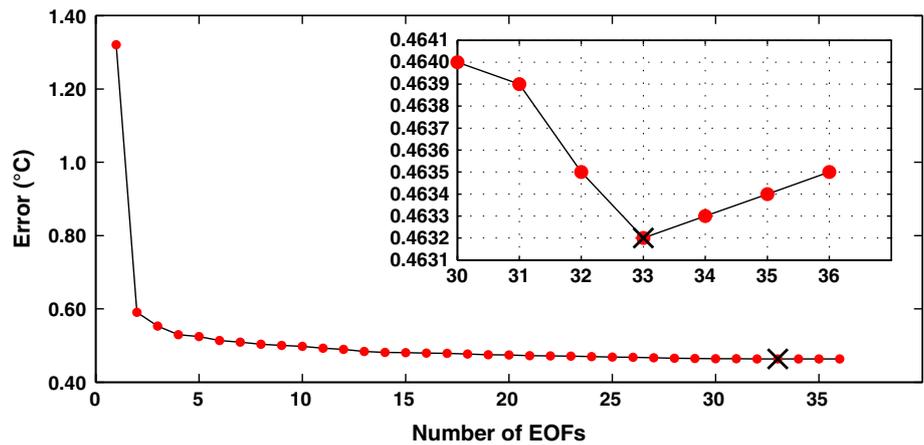
3 Method

The original SST data set we wanted to reconstruct was thus very large and highly cloud-covered. Therefore, we chose DINEOF, a self-consistent and parameter-free method used for accurately reconstructing incomplete geophysical data sets with a reasonable computational cost, even at high resolution (Beckers and Rixen 2003; Alvera-Azcárate et al. 2005). Moreover, we could analyse the by-product of DINEOF, the temporal and spatial EOFs, to understand the variability of SST.

Here we will briefly describe the reconstruction method of DINEOF. For more detailed descriptions, please refer to Beckers and Rixen (2003) and Alvera-Azcárate et al. (2005). We consider the missing data set as a matrix \mathbf{X} of dimensions $[m \times n]$, with the condition $m > n$. Here $m = 140,540$ is the number of sea pixels (both containing present and missing data) of each image and $n = 5119$ is the number of images (in temporal order). Firstly, the mean value over time and space is subtracted from the matrix. To avoid a biased initial estimation, the missing data within the matrix are set to zero; and they are also marked to differentiate them from those existing pixels on the mean. Secondly, a Singular Value Decomposition (SVD) is applied to compute the dominant modes of variability. The EOF decomposition is based on the efficient Lanczos method presented by Toumazou and Cretaux (2001). The procedure to compute EOFs in DINEOF is implemented as follows:

1. An EOF decomposition based on the Lanczos method is realized to obtain the first estimation of the singular values and singular vectors

Fig. 4 Expected errors obtained from cross-validation. The embedded figure on the upper-right corner zooms in the error with the optimal number of EOFs



$$USV^T = X \tag{1}$$

where **U** are the spatial modes, **V** are the temporal modes (or principal components) and **S** are the singular values.

- These values are then used to infer the flagged missing data, initially present data are unchanged

$$X_{i,j} = \sum_{p=1}^k \rho_p(\mathbf{u}_p)_i(\mathbf{v}_p^T)_j \tag{2}$$

if *i, j* correspond to a missing data point

where *k* is the retained number of EOF modes and ρ are the singular values. These two steps are repeated until the ratio between the root mean square of successive missing data reconstruction and the standard deviation of existing data lower than a threshold value, which is the precision criterion stopping the DINEOF iteration, is reached. Here we set the threshold value to 10^{-3} . Another data set extracted by artificially covering 3 % of valid data on the 63 cleanest images, including about 3.8×10^6 data points, is set aside for cross-validation (e.g., Brankart and Brasseur 1996). The optimal number of EOFs will be obtained when the global error between the reconstructed and cross-validation data points is minimal. Finally, the optimal number of EOFs will be used to reconstruct the whole data set, reintegrating the cross-validation data set.

There are some methodologies that improve the standard DINEOF described above (e.g., Ding et al. 2008; Alvera-Azcárate et al. 2009; Sorjamaa et al. 2010). In this paper, in order to reduce spurious time variability in the DINEOF reconstruction for a long-term data set, the filter technique described in Alvera-Azcárate et al. (2009) was applied to the temporal covariance matrix before the SVD decomposition. Here we used the filter with three iterations and its strength value of 0.01 day^2 .

When we applied the filter with these parameters, the frequencies higher than 0.91 per day were filtered out. Moreover, applying this filter within DINEOF increases the retained number of EOFs, which better presents small-scale oceanic features for the reconstruction, as well as decreases the cross-validation error.

In addition to the reconstructed SST field, we also estimated the local error map for each reconstructed image according to Beckers et al. (2006) (see “Appendix”).

4 Results

4.1 Validation

After the application, DINEOF retained 33 EOFs, accounting for 99.4 % of the total variance, to reconstruct the SST field for the SCS in 1989–2009. With this number of optimal modes, the reconstructed field can present from large- to small-scale oceanic features. The expected error calculated by the cross-validation reaches 0.46°C (Fig. 4). In addition to the cross-validation error estimation carried out within DINEOF, we also compared the results to in situ and satellite-derived microwave SSTs.

4.1.1 Comparison between reconstructed and in situ SSTs

An in situ data set extracted from the World Ocean Database 2009 (WOD09) (Locarnini et al. 2010) is used to validate the quality of the reconstructed SST. This data set was collected from measurements at a depth of 1–5 m. Only nighttime data (20:00–08:00 LT) were kept in order to avoid the diurnal variation of SST. The quality flags of the data source were used to extract in situ SSTs. Because WOD09 was collected from different instruments, to select reliable SSTs for the validation, we discarded the values that deviate more than \pm three standard deviations from the

Fig. 5 Nighttime in situ observations for validation: **a** spatial distribution of observations, **b** number of observations for each month from 1989 to 2009

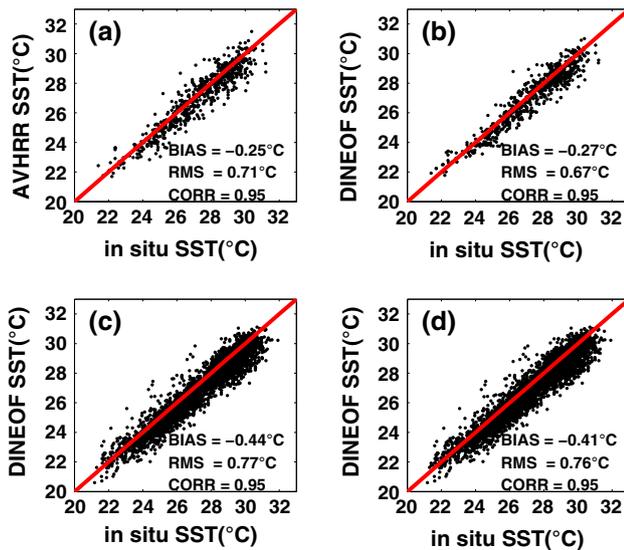
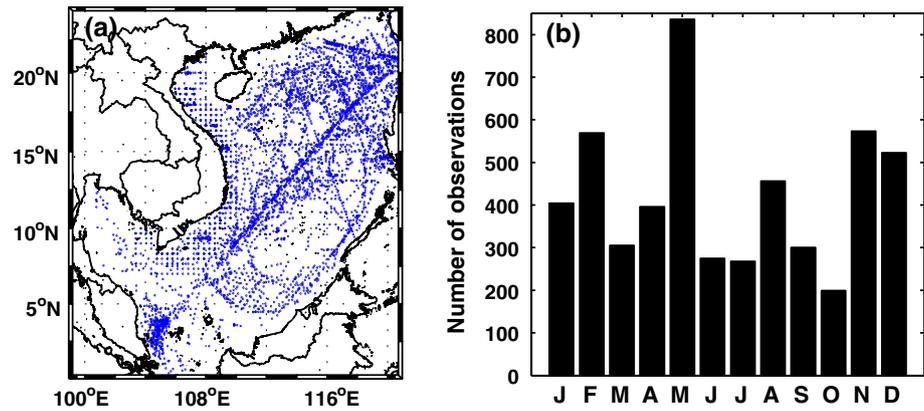


Fig. 6 Comparisons between in situ and reconstructed SSTs: **a** the original SST with in situ data, **b** the reconstructed SST at cloud-free positions with in situ data, **c** the reconstructed SST at cloud-covered positions with in situ data, and **d** the reconstructed SST at cloud-free and cloud-covered positions with in situ data

mean of the data set (Emery and Thomson 2001). Finally, 5100 data points ranging from 21 to 32 °C were retained for the period from 1989 to 2009. Figure 5 shows the spatial and temporal distribution of in situ data for the validation. More data were collected in the northern than southern SCS. The temporal distribution of data also shows that there were many observations measured during the northeast monsoon and the transitional period from the northeast monsoon to the southwest monsoon. A linear interpolation had been applied to extract satellite data at the positions of in situ data on the same date when the comparison was implemented.

We carried out four cases of comparison between satellite and in situ data (Fig. 6). Firstly, the original SST was compared with in situ data in order to evaluate the bias,

correlation and error between original satellite data and in situ measurements (Fig. 6a). The total root mean square (RMS) error of this case is 0.71 °C. The comparison also shows that the original satellite data are colder than in situ data, with a bias of -0.25 °C. Secondly, we compared the reconstructed SST at cloud-free positions with in situ data (Fig. 6b). The RMS error is reduced to 0.67 °C and the increase in bias is negligible. The small reduction of the RMS error appears because DINEOF eliminates the noise of the original satellite data through the use of a truncated EOF basis (Beckers and Rixen 2003; Alvera-Azcárate et al. 2005). Thirdly, the comparison between the reconstructed SST at positions of missing data with in situ data was implemented (Fig. 6c). Although the missing data percentage is very high, the total RMS error for this case is only 0.77 °C but the bias nearly doubles, -0.44 °C. The high increase in bias may be partially related to the seasonal and regional distribution of validation data. As mentioned above, there are large seasonal and regional variations in the availability of daily AVHRR Pathfinder SST in the SCS (Fig. 3). Therefore, 79 % of the data in the first and second cases are highly concentrated in the northern SCS, especially in the northeastern SCS. In addition, 62 % of the data in these cases were also collected during the northeast monsoon. Meanwhile, the distribution of data in the third case is less heterogeneous in space (62 % of the data in the northern SCS) and more homogeneous in time (51 % of the data during the northeast monsoon) than that in the first and second cases. Note that there are significant regional biases of the AVHRR SST in the SCS (Qiu et al. 2009). Moreover, calculating the monthly ERA-Interim wind climatology of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al. 2011) in the SCS from 1989 to 2009 (figure not shown), we see that the surface wind speed in the northeastern SCS is always greater than 6 m s^{-1} during the northeast monsoon. At wind speeds greater than 6 m s^{-1} , the difference between satellite-derived and in situ SSTs decreases (Donlon et al. 2002). The above reasons may lead to the large differences of bias among the

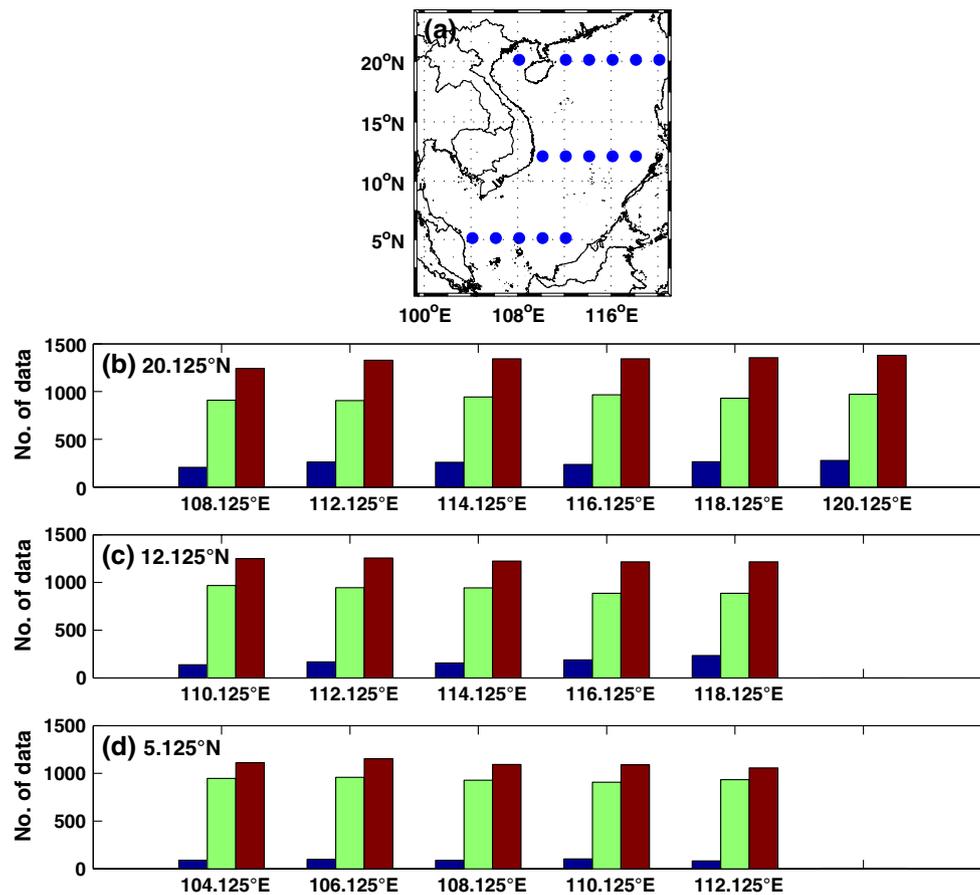


Fig. 7 Distribution of the locations where the nighttime TMI data were extracted for validation and the number of data at each location: **a** spatial distribution of the locations; **b–d** number of data at each location from 1998 to 2009: data used for validating the original

AVHRR and reconstructed SSTs in the cloud-free case (*blue*), data used for validating the reconstructed SST in the cloud-covered case (*green*), and data used for validating the reconstructed SST in the cloud-free and cloud-covered case (*red*)

validation cases. Finally, we computed the total RMS error of the reconstructed and in situ SSTs at cloud-free as well as cloud-covered positions (Fig. 6d). Its value is 0.76 °C. In general, the correlation coefficients including the seasonal cycle are high for all the cases, about 0.95.

4.1.2 Comparison between the reconstructed and satellite-derived microwave SSTs

To enhance the reliability of the reconstructed data set, we compared the reconstructed field with satellite-derived microwave SST that is not affected by clouds. Here we used the 25-km Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) version 7. We extracted a data set from the nighttime TMI (20:00–8:00 LT) SST at 16 locations (Fig. 7a) from 1998 to 2009. The linear interpolation had been applied to extract the reconstructed SST at 16 locations, where the TMI data had been extracted, on the same date when the comparison was implemented. Figure 7b–d show the number of data used for comparison

at each location. The results of the comparison are presented in Fig. 8.

The comparison between the original AVHRR and TMI SSTs shows that the AVHRR SST is a bit colder than the TMI SST (Fig. 8a, d, g), about -0.04 to -0.2 °C, except for the locations near the coast, 5.125°N, 104.125°E and 12.125°N, 110.125°E, with a bias of -0.5 °C, which might be attributed to the land contamination of the TMI SST (Qiu et al. 2009); therefore we will exclude these two locations from our later discussion. The RMS errors between the AVHRR and TMI SSTs are about 0.64–1.0 °C (Fig. 8b, e, h), and the correlation coefficients including the seasonal cycle are greater than 0.8 (Fig. 8c, f, i). The results indicate that we can use the TMI SST as a data source to validate the reconstructed SST.

In Fig. 8, the results of comparison between the reconstructed and TMI SSTs in the cloud-free case also show that the RMS errors are slightly reduced. In the cloud-covered case, the biases range from -0.05 to -0.26 °C; the RMS errors are about 0.7–1.1 °C; and the correlation

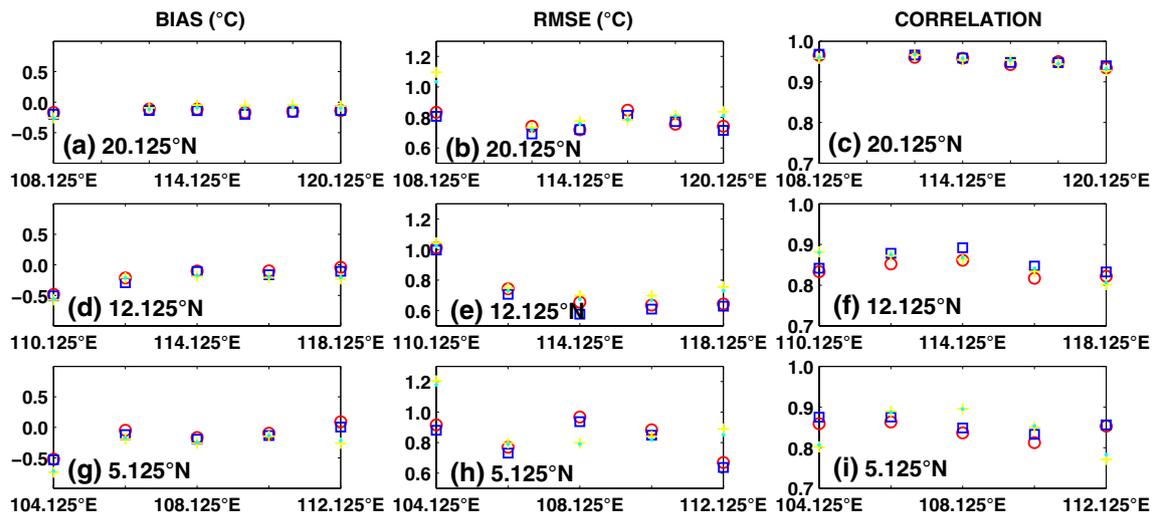


Fig. 8 Comparisons between the reconstructed and TMI SSTs: **a–c** biases, RMS errors, and correlations of the data respectively, at 20.125°N, **d–f** at 12.125°N, and **g–i** at 5.125°N. There are four cases of comparison: the original AVHRR and TMI SSTs in the cloud-free

case (*red circle*); the reconstructed and TMI SSTs in the cloud-free case (*blue square*); the reconstructed and TMI SSTs in the cloud-covered case (*yellow plus sign*); and the reconstructed and TMI SSTs in the cloud-free and cloud-covered case (*cyan point*)

coefficients including the seasonal cycle are high, 0.77–0.97. In the cloud-free and cloud-covered case, the comparison between the DINEOF and TMI SSTs has the similar results.

In summary, the results of validation with the in situ and TMI SSTs give us confidence in the reliability of the reconstructed data set.

4.2 An example of the daily reconstructed images

As an example of the quality as well as the application of the reconstructed SST, we will use the filled images to monitor the spatial and temporal variability of the coastal upwelling regions in summer. Figure 9 shows the original and reconstructed SST images in the summer of 1993. Although the original images have high percentage of missing data (Fig. 9a, c, e), from the reconstructed ones (Fig. 9b, d, f) we can clearly see the low SST regions, which have been reported as the upwellings in previous researches, including three regions: (1) in the southern Taiwan Strait (TW) (Tang et al. 2002, 2004; Hong et al. 2009); (2) around Hainan Island (HN) (Lü et al. 2008; Su and Pohlmann 2009; Li et al. 2012); and (3) along the southeastern Vietnam (VN) (Wyrski 1961; Kuo et al. 2000, 2004; Xie et al. 2003, 2007; Dippner et al. 2007).

For an upwelling region, we can directly measure the area and SST from the AVHRR images (Kuo et al. 2000). Therefore, in order to estimate the size of upwelling, from the daily reconstructed SST climatology we extracted pixels that are defined by an SST difference greater than 1 °C between the upwelling and non-upwelling areas, and

then multiplied the total number of pixels by $4 \times 4 \text{ km}^2$. The total size of the TW upwelling varies from about $1.5 \times 10^4 \text{ km}^2$ to $2.5 \times 10^4 \text{ km}^2$. The upwelling area often shrinks and a sharp front is formed between the upwelling and non-upwelling areas in July and August. The TW SST is lower than 24 °C in late May and June, and about 25–26.5 °C from July to September, similar to Tang et al. (2002)'s observations. The TW upwelling clearly splits into two small regions from early July to August: (1) along the coast of China, with an area of about 1.0×10^4 – $1.5 \times 10^4 \text{ km}^2$; (2) off the western coast of Taiwan Island, with an area of about 3000–4000 km^2 . Two small upwellings merge in September.

There have been a few previous studies about the spatial and temporal variability of upwellings around the Hainan Island. The areas of the HN upwellings are quite small and narrow. Therefore, it is difficult to monitor their signals, even if the high-resolution $1/8^\circ$ daily SST images from the Modular Ocean Data Analysis System (MODAS) are used (Su and Pohlmann 2009). However, we can see two regions with low SST around the coast of Hainan Island (Fig. 9b): (1) the upwelling induced by tidal mixing front is off the western coast (Lü et al. 2008), with an area of about 2000 km^2 and a cold centre of about 27–28 °C; (2) The coastal upwelling off the eastern coast is induced by the effect of wind, topography, and large-scale circulation (Su and Pohlmann 2009; Li et al. 2012), with an area of about $1.0 \times 10^4 \text{ km}^2$ and a cold centre of about 26.5–27 °C. The SSTs of the upwellings are quite similar to those in previous studies using cruise observations and numerical modeling (e.g., Lü et al. 2008; Su and Pohlmann 2009; Li et al.

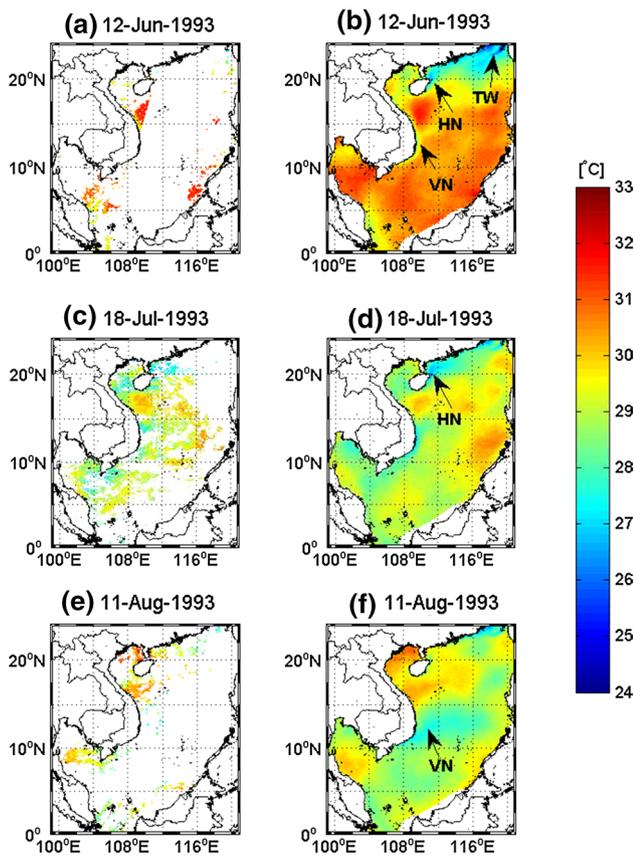


Fig. 9 Examples of the reconstructed SST in the summer of 1993. **a**, **c**, **e** are the original images with high percentage of missing data. **b**, **d**, **f** are the reconstructed images by DINEOF. The *arrows* and *letters* in **b** show three upwelling regions: (1) in the southern Taiwan Strait (TW); (2) around the Hainan Island (HN); and (3) along the southeastern Vietnam (VN). The *arrow* in **d** indicates a cold filament in the northeastern Hainan Island. The *arrow* in **f** shows a broad and strong cold filament spreading from the VN upwelling

2012). The HN upwellings often appear in June, strengthen from mid-July to mid-August, and weaken in late August. The strength of the upwelling in the eastern coast is stronger than that in the western coast. Also, note that there is a cold filament often spreading northeastward from the north of the island (20°N, 111°E) (Fig. 9d).

The VN upwelling, from 11°N to 17°N, has been well documented in previous studies (e.g., Wyrski 1961; Kuo et al. 2000, 2004; Xie et al. 2003, 2007; Dippner et al. 2007). It is generated in late May or early June, and strengthens in August, with a lowest temperature of about 27 °C. This upwelling region is rather special due to its deformation. In July and August, the VN upwelling spreads northeastward offshore as a strong and broad cold filament at about 12°N (Kuo et al. 2000; Xie et al. 2003, 2007; Dippner et al. 2007). We can clearly see this striking feature in the reconstructed image (Fig. 9f).

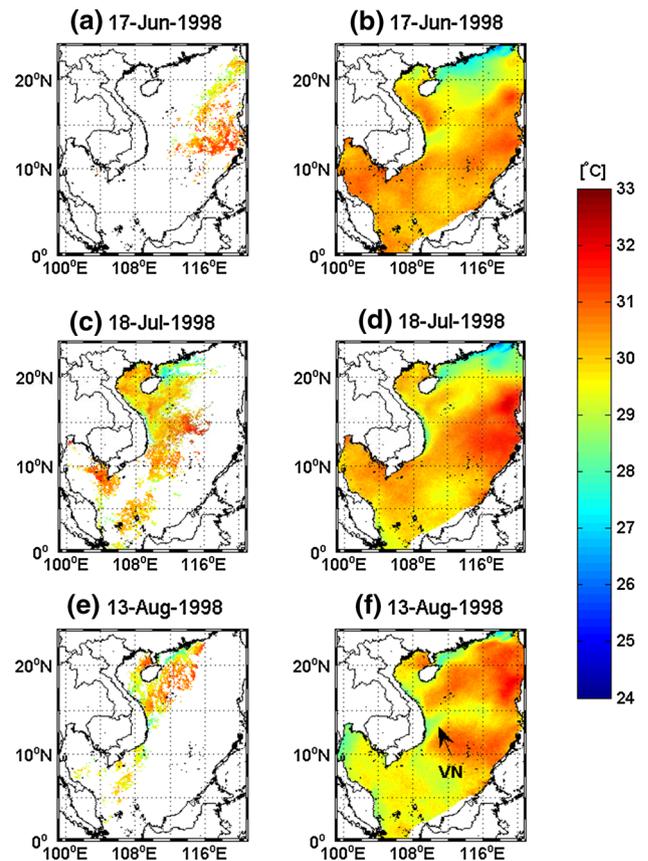


Fig. 10 As in Fig. 9, except for the summer of 1998. The reconstructed images show the upwelling regions under the influence of El Niño. The *arrow* in **f** indicates the weakening of cold filament spreading from the VN upwelling

Figures 10 and 11 are the original and filled SST images in the summers of 1998 and 2000. The reconstructed images show the variability of SST in the upwelling regions under the influence of the ENSO events. In the summer of 1998, post-El Niño event, the SCS SST increased. Xie et al. (2003) used the monthly data of the TMI SST, sea surface height (SSH), SeaWiFS chlorophyll, and surface wind to investigate the stretch of the VN upwelling, but they did not see its presence from 10°N to 14°N. Using the reconstructed images, we observe that this phenomenon did not totally disappear (Fig. 10f). It still occurred over short periods of time in July and August; however, the strength of the cold filament was very weak. It coincides with findings of Kuo et al. (2004). In that work, they used the 3-day composite 1.1 km AVHRR SST, surface wind, TOPEX/POSEIDON and ERS-2 altimeter SSHs to investigate the response of the VN upwelling to the 1997–1998 ENSO; the result showed that the stretch of the VN upwelling in the summer of 1998 was weak due to the fact that it was confined in two anticyclonic

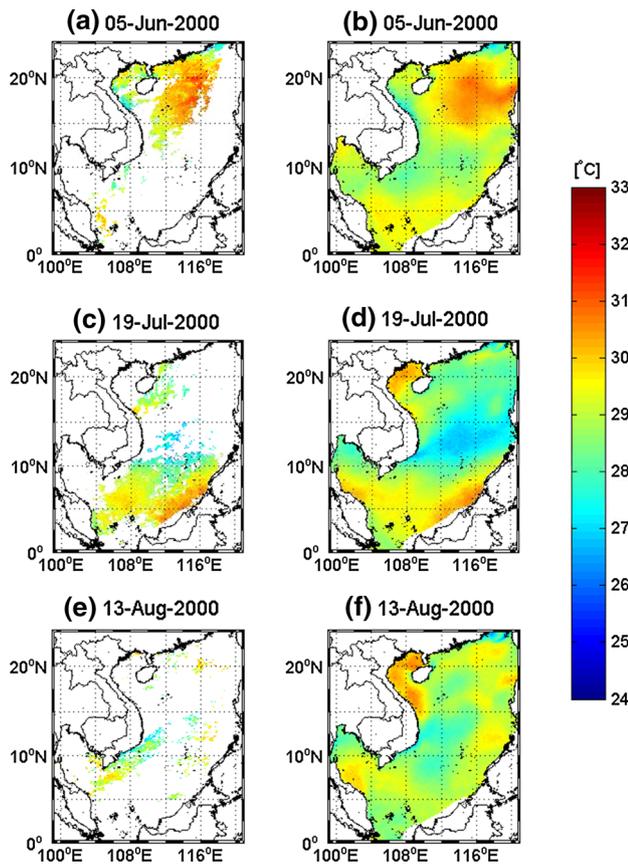


Fig. 11 As in Fig. 9, except for the summer of 2000. The reconstructed images show the upwelling regions under the influence of La Niña. The cold filament of the VN upwelling is stronger in **d**

circulations. Figure 11 shows the effect of La Niña on SST of the upwelling regions in the summer of 2000. The cold filament of the VN upwelling was very strong (Fig. 11d). SST of the western HN upwelling was also warmer in the summer of 1998 and colder in the summer of 2000. In contrast, SSTs of the TW upwelling and the eastern HN upwelling in the summer of 1998 were colder than those in the summer of 2000.

5 EOF analysis

As mentioned in the introduction section, the SST variability in the SCS is affected by the monsoon. We thus analysed the EOFs of SST associated with those of surface wind (SW) to explain more clearly the temporal and spatial variability of SST. The SW EOFs were computed from the daily 10-m ERA-Interim wind data of ECMWF (Dee et al. 2011) with a resolution of $1.5^\circ \times 1.5^\circ$. Here we analyse the first three EOFs of the reconstructed SST and the reconstructed SST anomalies.

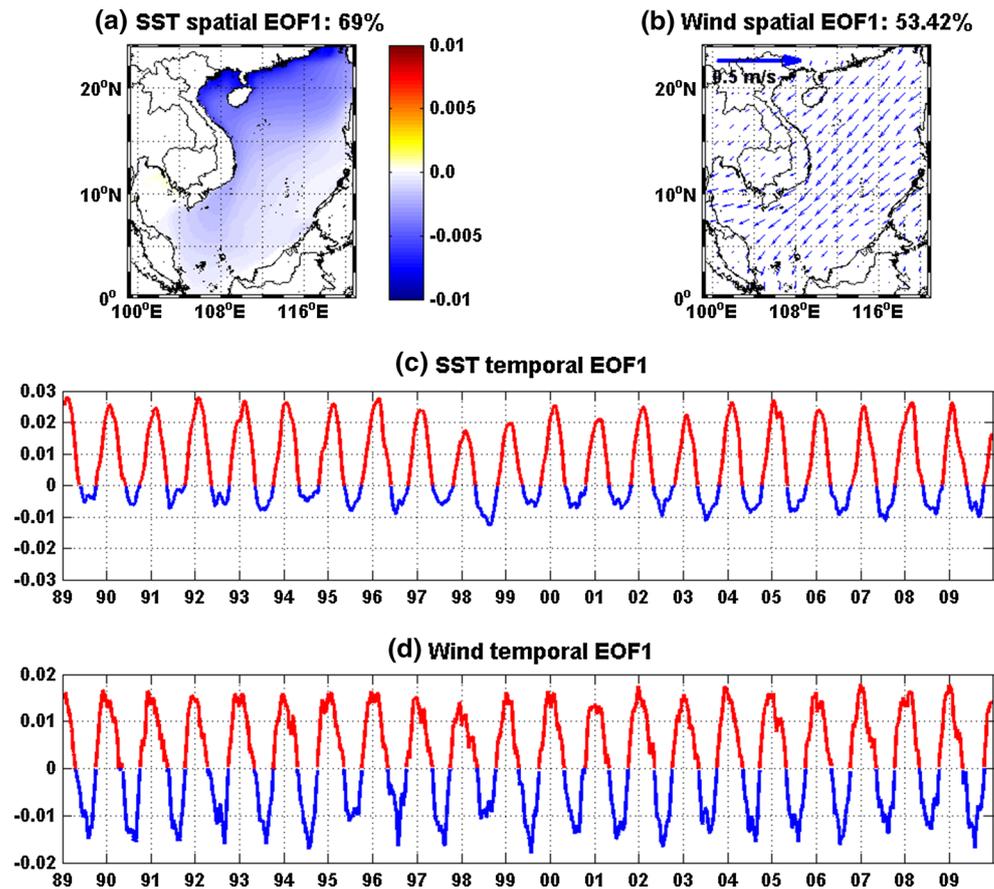
5.1 EOFs of the reconstructed SST

5.1.1 EOF1

The first SW mode (Fig. 12b, d), accounting for 53.42 % of the total variance, shows the seasonal variability. It indicates that the SW blows northeasterly in winter and southwesterly in summer over the SCS, except for the Gulf of Thailand where the SW blows easterly in winter and westerly in summer. The SW is strong along the northeast–southwest diagonal of the basin, the southeastern Vietnam coast, and in the northern SCS. It has the maximum value near the southeastern coast of Vietnam, at $10\text{--}12^\circ\text{N}$. The SW is weaker along the coast of the Borneo and Palawa Islands and the Karimata Strait. This situation is due to the effect of the orographic features. The first temporal mode (Fig. 12d) indicates that the peaks (northeast monsoon) and troughs (southwest monsoon) often occur in November–February and June–August, respectively. The first temporal mode of SW also shows a significant interannual variability.

Corresponding to the first SW mode, the first SST mode (Fig. 12a, c), contributing up to 69 % of the total variance, clearly shows the seasonal variability. The first SST spatial mode exhibits negative values in the whole basin (Fig. 12a), except in the Gulf of Thailand. The spatial variability of SST is high in the northwestern part of the basin and the southeastern Vietnam coast, and low along the western coast of the Borneo and Palawa Islands, the Karimata Strait, and in the Gulf of Thailand. The first temporal mode (Fig. 12c) often reaches peaks in December–February and troughs in June–August, responding to the cold northeast monsoon and the warm southwest monsoon, respectively. The amplitude of the peaks is about three times higher than that of the troughs. It means that this mode strongly develops in winter. Considering the first SST spatial mode associated with its temporal mode (Fig. 12a, c), we can see that the first SST mode exhibits a cooling in the whole basin in winter and a warming mainly in the northwest of the sea in summer. The solar insolation through seasons, monsoon, water exchange, and topography are the major factors influencing the first SST mode. In winter, the northeast monsoon causes an Ekman pumping in the southern half of the basin. This Ekman pumping spins up a basin-scale cyclonic circulation over the SCS, with a western boundary current flowing southward on the southern coast of Vietnam and the eastern continental slope of the Sunda Shelf (Liu et al. 2004). In Fig. 12a, we can clearly see the signal of the strong western boundary current flowing southward in winter, resulting as a distinct cold tongue in the southeastern coast of Vietnam and the Sunda slope at $105\text{--}110^\circ\text{E}$. In summer, along with the increase in solar radiation and the

Fig. 12 The first SST ($\rho = 0.4885 \times 10^5$) and SW ($\rho = 0.5218 \times 10^4$) EOFs. Spatial EOF1s: **a** SST ($^{\circ}\text{C}$) and **b** SW (m/s); temporal EOF1s with a 30-day low-pass filter: **c** SST and **d** SW



prevailing southwest monsoon, this pattern becomes warm and mainly occurs in the northwest of the basin. The first SST mode also shows that it has a significant interannual variability.

Figure 13 shows the correlation between the first temporal EOFs and the Niño 3 SST region (5°S – 5°N , 90°W – 150°W). To find the correlation between the temporal EOFs and ENSO, we calculated the monthly anomalies of each time-series, smoothed them with a 5-month running mean to filter out the intraseasonal variations (Trenberth 1997), and removed the linear trend over 21 years. The maximum correlation between the first SST temporal EOF and the Niño 3 SST is -0.56 with a 6-month lag (Fig. 13a). During the period from 1989 to 2009, El Niño occurred in 1991–1992, 1994–1995, 1997–1998, 2002–2003, 2004–2005, 2006–2007, and 2009, while La Niña occurred in 1989, 1995–1996, 1998–2001, 2005–2006, and 2007–2008. However, we can see that the interannual variability of the first SST EOF is clearly influenced by ENSO only in 1989, 1995–1996, 1997–1998, 2002–2003, and 2007–2008. It means that this pattern strengthens (weakens) during or after the moderate and strong La Niña (El Niño) events. Meanwhile, the maximum correlation

between the first SW temporal EOF and the Niño 3 SST is negligible, -0.24 without a lag (Fig. 13b).

5.1.2 EOF2

The second SW mode (Fig. 14b, d), accounting for 13.5 % of the total variance, also has an annual component. It exhibits part of an anticyclone (cyclone). The anticyclone often appears from winter to early summer and the cyclone exists for the rest of the year. The anticyclone (cyclone) fully develops in the transition of the monsoon. Looking at the first and second SW modes (Figs. 12b, d, 14b, d), we can see that the anticyclone (cyclone) weakens the north-east (southwest) monsoon in the north of the basin. This pattern may be the southwesternmost part of the western North Pacific anticyclonic (cyclonic) circulation that plays an important role in the Pacific-East Asia teleconnection leading to the influence of ENSO on the East Asian monsoon (Klein et al. 1999; Wang et al. 2000). The maximum correlation between the second SW temporal EOF and the Niño 3 SST is -0.68 with a 3-month lag (Fig. 15b). We can see that the anticyclone often strongly develops in the 1991–1992, 1994–1995, 1997–1998, 2002–2003, and

Fig. 13 Correlation between the monthly anomalies of the first SST and SW temporal EOFs and the Niño 3 SST region (5°S–5°N, 90°W–150°W). For each time-series, data are smoothed by a 5-month running mean and the linear trend over 21 years is removed. Also note that each time-series is normalized. The coloured areas indicate El Niño (pink) and La Niña (blue): **a** SST and Niño 3 with a correlation coefficient of -0.56 and a 6-month lag; **b** SW and Niño 3 with a correlation coefficient of -0.24 and without a lag. All correlation coefficients are significant at the 95 % confidence level

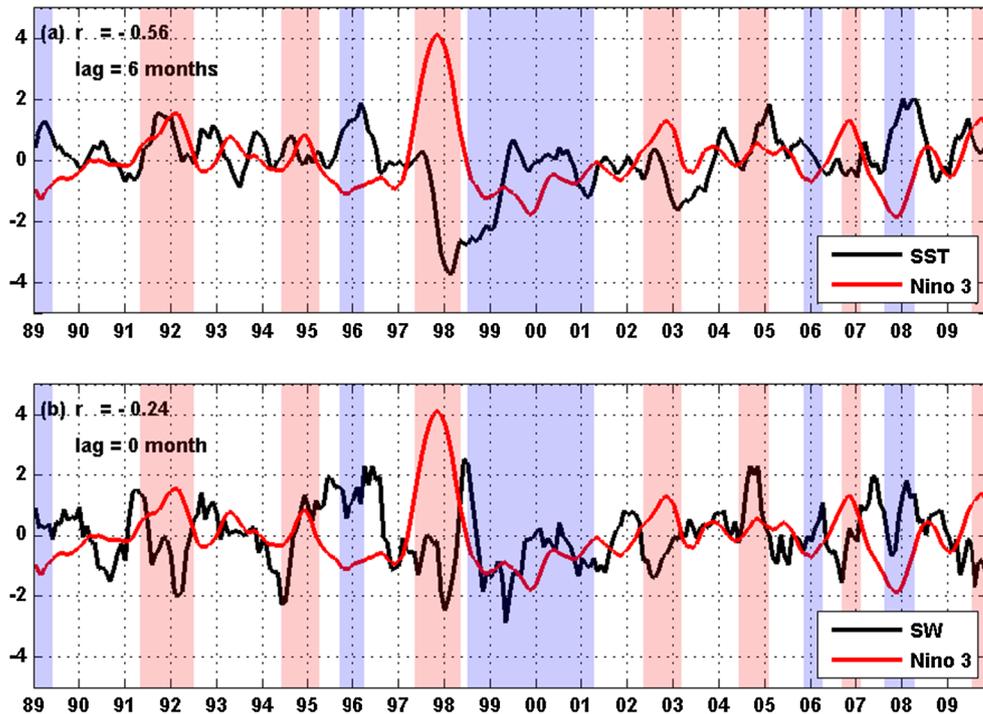


Fig. 14 As in Fig. 12 except for the second SST ($\rho = 0.2926 \times 10^5$) and SW ($\rho = 0.2624 \times 10^4$) EOFs

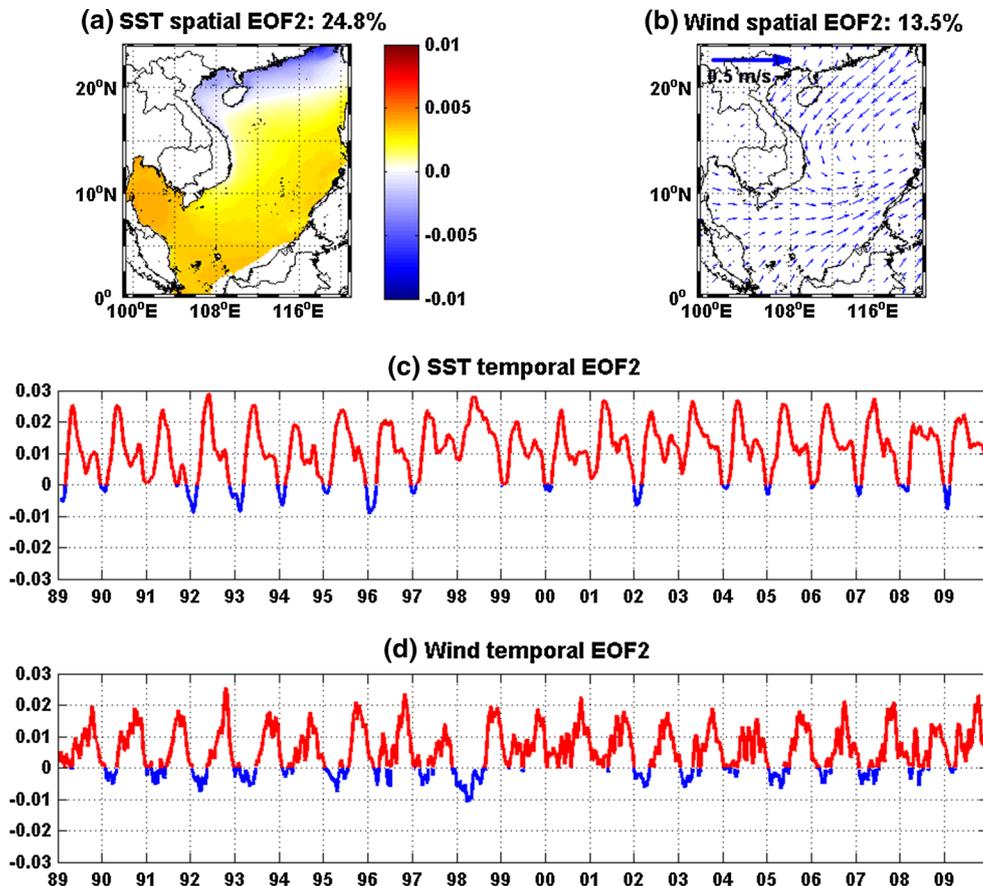
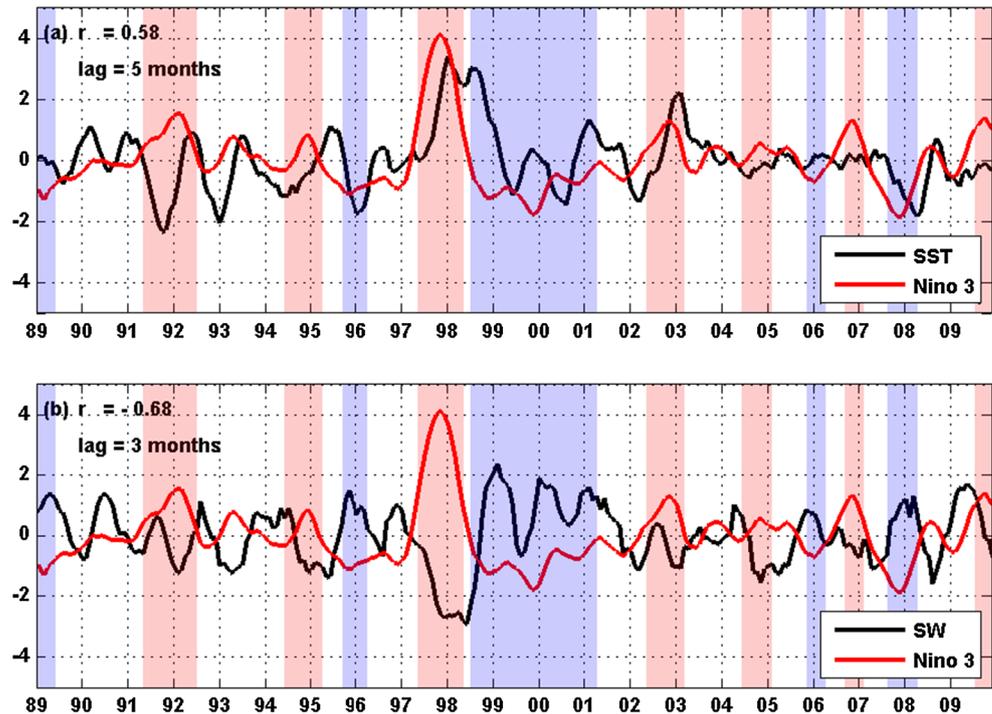


Fig. 15 As in Fig. 13 except for the second SST and SW EOFs: **a** SST and Niño 3 with a correlation coefficient of 0.58 and a 5-month lag; **b** SW and Niño 3 with a correlation coefficient of -0.68 and a 3-month lag. All correlation coefficients are significant at the 95 % confidence level



2004–2005 El Niño events, and weakens/disappears in the 1989, 1995–1996, 1998–2001, and 2007–2008 La Niña events (Figs. 14b, 15b).

The second SST mode (Fig. 14a, c), explaining 24.8 % of the total variance, presents the annual variability of the thermal advection along the northeast–southwest diagonal of the basin from two opposite directions. The temporal and spatial variability of the second SST mode is partially influenced by the atmospheric anticyclone (cyclone). In late April, the strong development of the anticyclone weakens the northeast monsoon (Figs. 12b, d, 14b, d). At that time, the loss of latent heat flux from the ocean, the amount of cloud and rainfall over the SCS decrease; and more solar radiation can enter into the ocean (Wang and Wang 2006). As a result, the SCS SST increases and reaches its peak in May (Fig. 14c). The increase in SST provides heat and vapor for the onset of the southwest monsoon (Yan 1997). The changes of SW lead to those of the surface circulation in the basin. There is an anticyclone in the south and a weaker cyclone in the north. These circulations advect warm water from the south to the northeast and cold water from the north to the southwest. In addition, under the effect of the southwest monsoon, the mixed layer depth deepens and the vertical entrainment pumps cold water into the mixed layer, cooling the surface temperature (Qu 2001; Wang and Wang 2006). During that time, the surface heat flux also decreases over the basin due to the increase in cloud cover and rainfall, and strong wind speed. Therefore, SST decreases after May. Looking at the second SST temporal EOF (Fig. 14c), we can see that the SCS SST often

reaches its trough in July–August and peaks again in September–October. The second spatial mode (Fig. 14a) shows two regions with the highest values. The first one is located in the southwest of the Philippines and the Palawa Island, which may be the result of the influence of the atmospheric anticyclone (cyclone). The second one is in the Gulf of Thailand, and is the anticyclonic circulation often occurring in this region during the southwest monsoon (Wyrtki 1961). These features could not be captured well in previous EOF analyses (e.g., Chu et al. 1997; Fang et al. 2006; Kuo et al. 2009).

Under the influence of the anticyclone (cyclone), the second SST mode presents an interannual variability in response to ENSO. The maximum correlation between the second SST temporal EOF and the Niño 3 SST is 0.58 with a lag of 5 months (Fig. 15a). SST in the second mode is warmer during or after the 1991–1992, 1994–1995, 1997–1998, and 2002–2003 El Niño events. The opposite occurs in the La Niña events. SST in this pattern is colder during or after the 1995–1996, 1998–2000, and 2007–2008 La Niña events.

5.1.3 EOF3

The third SW mode accounts for 8.94 % of the total variance. The spatial and temporal EOFs (Fig. 16b, d) show that the SW mainly blows southeasterly in the northern SCS. This pattern strengthens in April–May, and weakens in October–December. The maximum correlation between the third SW temporal EOF and the Niño 3 SST is $-$

Fig. 16 As in Fig. 12 except for the third SST ($\rho = 0.6514 \times 10^4$) and SW ($\rho = 0.2135 \times 10^4$) EOFs

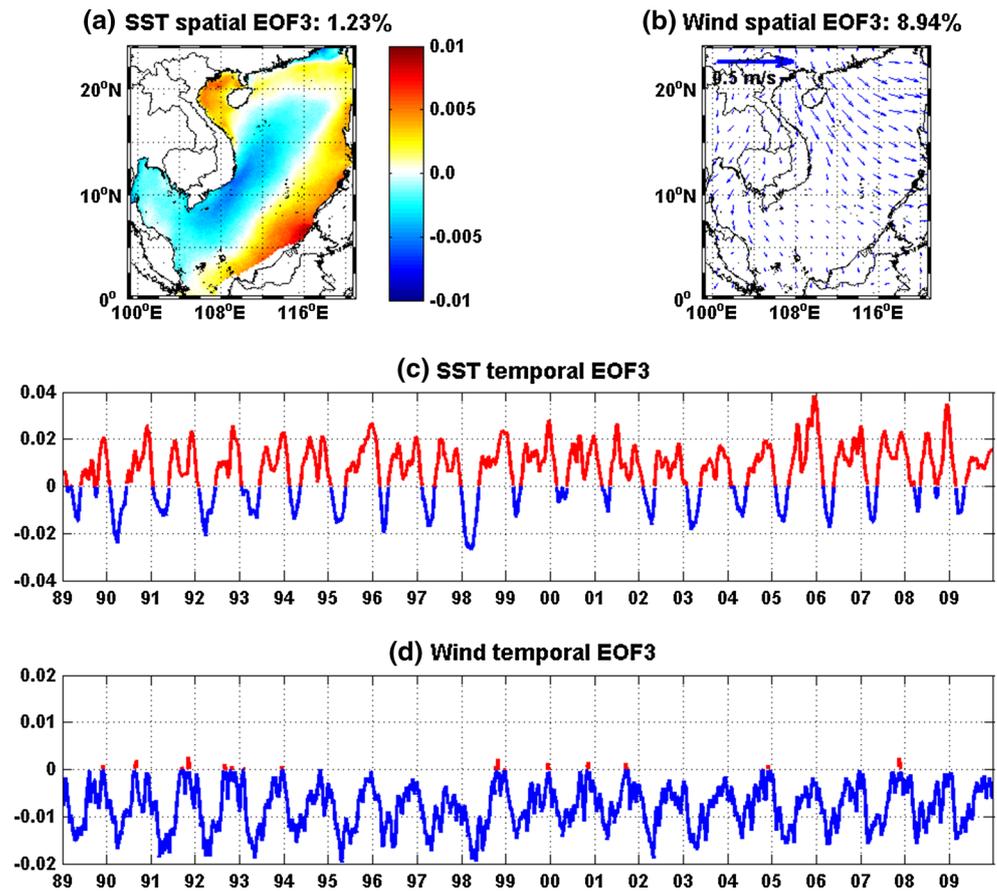
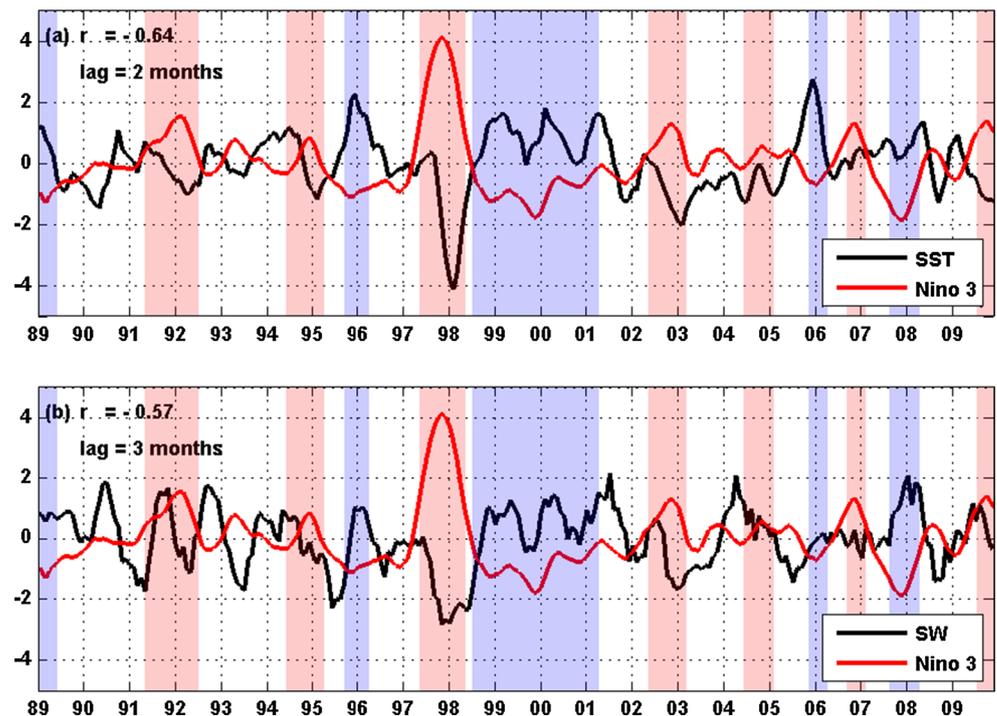


Fig. 17 As in Fig. 13 except for the third SST and SW EOFs: **a** SST and Niño 3 with a correlation coefficient of -0.64 and a 2-month lag; **b** SW and Niño 3 with a correlation coefficient of -0.57 and a 3-month lag. All correlation coefficients are significant at the 95 % confidence level



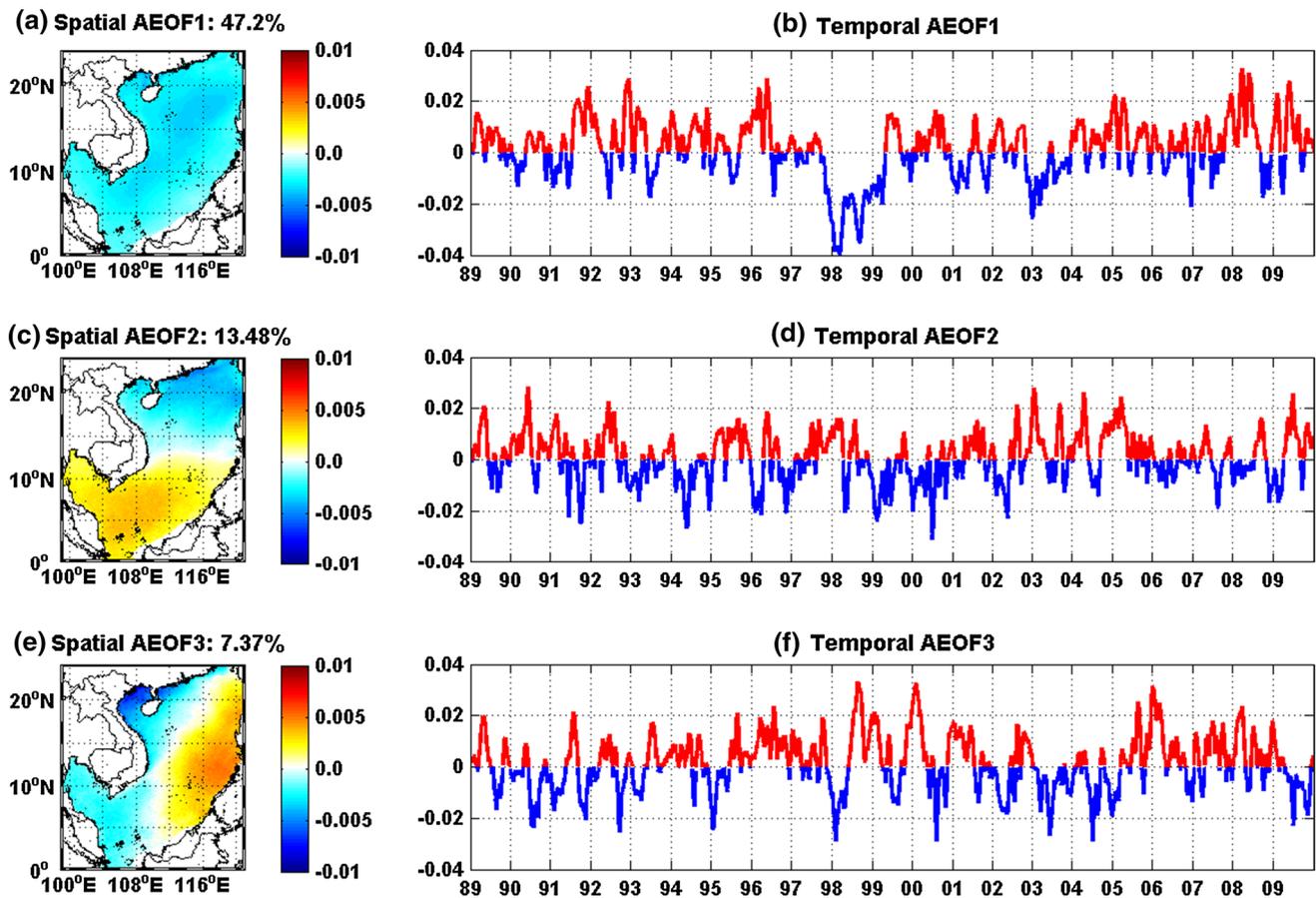


Fig. 18 The first three SST AEOFs: **a, b** spatial and temporal AEOF1s ($\rho = 0.1240 \times 10^5$); **c, d** spatial and temporal AEOF2s ($\rho = 0.6626 \times 10^4$); **e, f** spatial and temporal AEOF3s ($\rho = 0.4899 \times 10^4$). Temporal AEOFs are smoothed with a 30-day low-pass filter

0.57 with a 3-month lag (Fig. 17b). Looking at Figs. 16d and 17b, we can see that this mode strengthens in the 1994–1995, 1997–1998, and 2002–2003 El Niño years; and it weakens in the 1998–2001 La Niña years. However, it appears that there is insignificant correlation between this pattern and the third SST mode, which will be presented below.

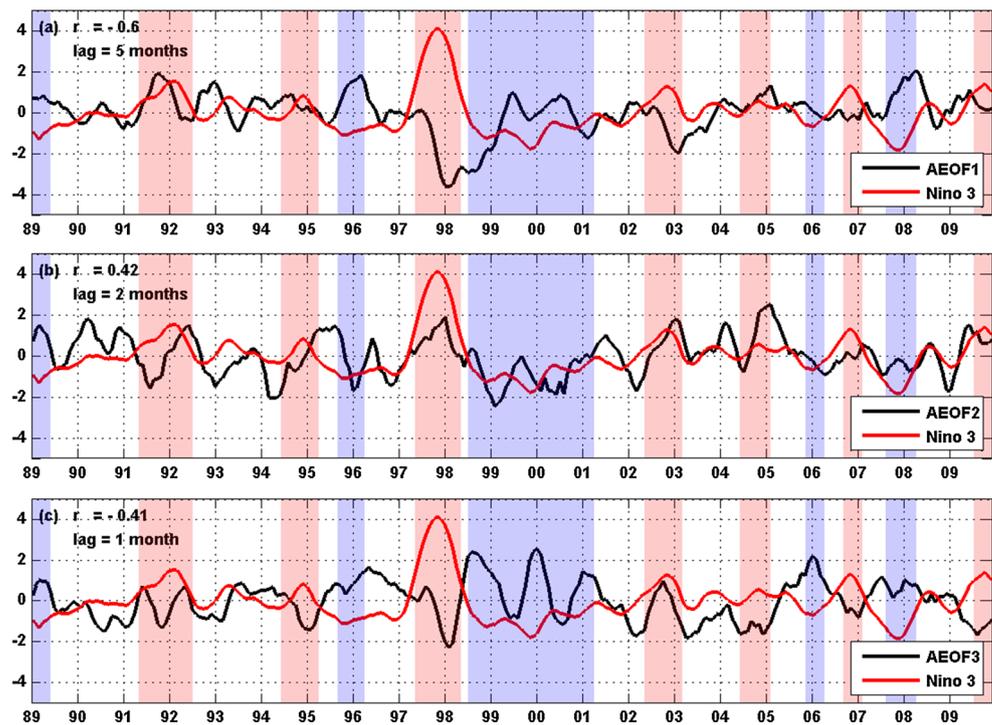
The third SST mode (Fig. 16a, c), accounting for 1.23 % of the total variance, has not been presented in previous EOF analyses (e.g., Chu et al. 1997; Fang et al. 2006; Kuo et al. 2009). It shows a contrast of SST between the west and the east of the basin. In general, the SST in the west is warm (cold) in January–May (June–December), while the SST in the east is cold (warm) in January–May (June–December). From the spatial and temporal EOFs, we can see the cooling in the west in summer, related to the upwelling in the Taiwan Strait and the offshore spread of cold water from the southeastern Vietnam coast, which are presented in Sect. 4.2. Looking at the temporal EOF (Fig. 16c), we can see that the mid-summer cooling due to the offshore spread of cold water from the southeastern

Vietnam coast weakens (strengthens) in the El Niño (La Niña) years. Figure 16c also indicates that under the influence of the strong El Niño event in 1998, the cooling in mid-summer was weak, but did not totally disappear as Xie et al. (2003). The anomalous cooling in the northwest Borneo Island (5–7°N, 115–117°E) in January–May coincides with the upwelling investigated recently by Yan et al. (2015). The maximum correlation between the third SST mode and the Niño 3 index is -0.64 , with a 2-month lag (Fig. 17a).

5.2 EOFs of the reconstructed SST anomalies

As presented above, the first three modes of the reconstructed SST are dominated by the annual cycles. We thus analysed the EOFs of the daily reconstructed SST anomalies (AEOFs) to reveal transient variations. Before calculating the AEOFs, we removed the seasonal cycle and trend from the reconstructed SST. The first three spatial EOFs of the daily ERA-Interim SW anomalies (figures not shown) are quite similar to those of the daily ERA-Interim SW. It

Fig. 19 As in Fig. 13 except for the first three daily temporal SST AEOFs: **a** AEOF1 and Niño 3 with a correlation coefficient of -0.6 and a 5-month lag; **b** AEOF2 and Niño 3 with a correlation coefficient of 0.42 and a 2-month lag; **c** AEOF3 and Niño 3 with a correlation coefficient of -0.41 and a 1-month lag. All correlation coefficients are significant at the 95 % confidence level



is due to the fact that the low-resolution ERA-Interim data do not properly resolve high-frequency, small-scale SW features. Therefore, we analysed only the SST AEOFs in this case.

The first AEOF (Fig. 18a, b) accounts for 47.2 % of the total variance. It presents an anomalous cooling (warming) in the whole basin. The high variations are located along the northeast-southwest diagonal of the basin, in the deep basin ($15\text{--}20^{\circ}\text{N}$, $112\text{--}120^{\circ}\text{E}$) and the southeast Vietnam coast ($9\text{--}12^{\circ}\text{N}$, $107\text{--}110^{\circ}\text{E}$), where the SW is strongest due to the orographic forcing by the mountains; therefore, the two highly anomalous regions may be the effect of the wind stress anomalies. This spatial pattern is quite similar to the ones in Chu et al. (1997) and Fang et al. (2006). The temporal AEOF1 (Fig. 18b) clearly shows the influence of ENSO on this pattern, especially with an abnormal warming in the whole basin in the strong 1997–1998 El Niño event; however, a closer look at the temporal AEOF1 indicates that the decrease in SST still occurred in the summer of 1998. The maximum correlation between the first temporal AEOF and the Niño 3 index is -0.6 with a 5-month lag (Fig. 19a). This pattern becomes colder in the 1989, 1995–1996, 1998–2001, and 2007–2008 La Niña years, and warmer in the 1997–1998 and 2002–2003 El Niño years.

The second AEOF (Fig. 18c, d) accounts for 13.48 % of the total variance. It shows an anomalous SST contrast between the north and the south of the basin across 11°N where the western boundary current leaves the southeast Vietnam coast forming the summer southeast Vietnam

offshore current (Shaw and Chao 1994; Xie et al. 2003, 2007; Fang et al. 2012). From the spatial pattern, we can see two highly cold (warm) anomalies: the first one is in the southern SCS ($3\text{--}10^{\circ}\text{N}$, $104\text{--}112^{\circ}\text{E}$) that coincides with the position of the anticyclonic gyre investigated by Fang et al. (2002) and Xie et al. (2003); the second one is in the northern SCS, which shows the waters from the Pacific intruding into the SCS through the Luzon Strait. This pattern is quite similar to the second EOF in Chu et al. (1997); however, the intrusion of the Pacific waters could not be captured well in their result. The reason is probably due to the low resolution of their data set. The second temporal AEOF positively correlates with the Niño 3 index, with a maximum correlation coefficient of 0.42 and a 2-month lag (Fig. 19b).

The third AEOF (Fig. 18e, f) explains 7.37 % of the total variance. This pattern presents an anomalous SST contrast between the shelves and the deep basin. On the shelves, the highest anomaly is located in the Gulf of Tonkin, which is due to the bathymetric effect (Liu et al. 2004). On the deep basin, the highest anomaly is located along the western Palawa and Philippine Islands, which may be influenced by the anomalous atmospheric cyclone (anticyclone). This mode has the high variations in the 1997–1998, 1998–2000, and 2005–2006 ENSO events. The maximum correlation between the temporal AEOF3 and the Niño 3 index is -0.41 , with a lag of 1 month.

To enhance the reliability of our work, we calculated the SST EOFs and AEOFs of the data set extracted from

the 17-year daily $1/8^\circ$ MODAS SST (1993–2009) (Barron and Kara 2006) (figures not shown). The results are quite similar to our work. However, from the MODAS EOFs and AEOFs, the SST variability related to the upwelling phenomena in the Taiwan Strait and the northwest Borneo Island, and the intrusion of the Pacific waters is not clear.

6 Conclusions

In this study, we used DINEOF to reconstruct 21 years of daily nighttime 4-km AVHRR Pathfinder SST in the SCS. The 33 optimal modes, accounting for 99.4 % of the total variance, were retained for the SST reconstruction of the whole SCS, except for a small region near the western coast of Borneo Island that was excluded from this study because the percentage of missing data in this region is higher than 95 %, not containing enough information for the reconstruction. With this number of optimal modes, the reconstructed field can present from large- to small-scale oceanic features. This was proved by applying the reconstructed images to monitor the spatial and temporal variability of the coastal upwelling regions in summer, and analysing the EOFs. Besides the cross-validation implemented within DINEOF with an RMS error of 0.46°C , we also compared the reconstructed data with the nighttime in situ and TMI SSTs. The RMS errors are 0.76°C for the comparison with in situ data, and $0.7\text{--}1.1^\circ\text{C}$ for the comparison with TMI data. Therefore, the accuracy of the reconstructed SST field is reliable for use in many different researches, such as validating oceanic numerical models, or identifying and tracking meso-scale oceanic features.

The analysis of the first three modes of the reconstructed SST, associated with the first three SW modes computed from the ERA-Interim product, clearly exhibits the seasonal and interannual variability of the SCS SST under the influence of monsoon and ENSO. The first mode, accounting for 69 % of the total variance, presents a cooling in the whole basin during the northeast monsoon and a warming mainly in the northwest of the sea during the southwest monsoon. The first SST mode is affected by ENSO with a 6-month lag. The second mode, accounting for 24.8 % of the total variance, shows the thermal advection along the northeast–southwest diagonal of the basin from two opposite directions. Under the impact of the atmospheric anticyclone (cyclone), which is the atmospheric bridge of ENSO in the tropics, the second SST mode presents an interannual variability in response to ENSO with a 5-month lag. The third mode, contributing to 1.23 % of the total variance, has not been analysed in previous researches. It highlights an SST contrast between the west and the east of the basin. From this mode, we can see that the cooling of SST in summer, due to the offshore spread of cold water from the

southeastern Vietnam coast, was weak in 1998, but did not totally disappear as in Xie et al.'s (2003) work. The third SST mode lags ENSO by 2 months.

Due to the fact that the first three modes of the reconstructed SST are dominated by the annual cycles, we analysed the AEOFs to reveal transient variations. The first AEOF, explaining 47.2 % of the total variance, presents two cool (warm) SST anomalies located along the northeast-southwest diagonal of the basin, in the deep basin and the southeastern Vietnam coast. The temporal AEOF1 also shows that the decrease in SST still occurred in the summer of 1998. This mode lags ENSO by 5 months. The second AEOF, accounting for 13.48 % of the total variance, shows a contrast of the SST anomalies between the north and the south of the basin across 11°N , with the strong anomalies related to the anticyclonic gyre in the south and the intrusion of the Pacific waters in the north. The third AEOF, accounting for 7.37 % of the total variance, shows a contrast of the SST anomalies between the deep basin and the shelves. This mode also indicates the high variation of SST located along the Palawa and Philippine Islands. The second and third AEOFs have significant correlations with the Niño 3 index, with a 1- or 2-month lag.

The analysis of the SST EOFs and AEOFs in high spatial and temporal resolution clearly exhibits the variability of some oceanic features that could not be captured well in previous EOF analyses for the whole SCS: (1) the signal of the influence of the atmospheric anticyclone (cyclone) on the SCS SST is located along the Philippine and the Palawa Islands. (2) The anticyclonic circulation is in the Gulf of Thailand. (3) The cooling of SST related to the upwelling appears in the northwest Borneo Island in winter.

As well documented in previous researches, the SCS SST often lags ENSO by half a year; however, comparing the temporal EOFs and AEOFs with the Niño 3 index, we see that the time lag changes with the frequencies of the SST variability, from 1 to 6 months.

Acknowledgments Daily nighttime 4-km AVHRR Pathfinder SST data used in this study were obtained from the Physical Oceanography Distributed Active Archive Center (PODAAC) at the NASA Jet Propulsion Laboratory (<ftp://podaac-ftp.jpl.nasa.gov>). TMI data produced by Remote Sensing Systems and sponsored by the NASA Earth Science MEaSUREs DISCOVER Project are available at <http://www.remss.com>. In situ data were downloaded from NOAA National Oceanographic Data Center (NODC), World Ocean Database 2009 (WOD09) (<http://www.nodc.noaa.gov>). Wind data were provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (<http://data-portal.ecmwf.int>). The Niño 3 SST was downloaded from <ftp://ftp.cpc.ncep.noaa.gov/wd52dg/data/indices>. Calculations were run on the super-computer NIC3 of the University of Liège, and HMEM of the Université catholique de Louvain (CISM/UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de la Recherche Scientifique de Belgique (FRS-FNRS). We are grateful to the anonymous reviewers for valuable comments on the manuscript.

Suggestions by Prof. Joji Ishizaka are appreciated. This study was carried out within the context of the BESST (SR/12/158) project funded by the Belgian Science Policy (BELSPO) within the framework of the Research Program for Earth Observation STEREO II. The Vietnam Ministry of Education and Training is gratefully acknowledged for funding H.-N.T. Huynh's Ph.D. scholarship.

Appendix

When using the reconstructed SST, perhaps the users would also like to know the local error map associated with each reconstructed image. Therefore, we will concisely present the method and parameters that were used to compute the error field. For more detailed descriptions, please refer to Beckers et al. (2006).

Beckers et al. (2006) proved that a very efficient least-square fit of EOF amplitudes to an observed subset of data is equivalent to optimal interpolation (OI) if the filtered covariance matrix of DINEOF is used as the ad hoc covariance of OI. Hence the error estimates of OI can be used as a proxy for the error fields of DINEOF. To calculate the error field, a method similar to the least-square fit, instead of an equivalent standard OI error calculation, is applied due to its reduced calculation cost.

In standard OI, the covariance matrix between data points is the sum of the observational error-covariance matrix \mathbf{R} and the target field-covariance matrix \mathbf{B} . The observational errors and target field are assumed to be uncorrelated. We can use the covariance matrix based on SVD decomposition from DINEOF, with the N -retained significant EOFs ($N = 33$), in an OI approach. If we define scaled spatial EOFs

$$\mathbf{L} = \frac{1}{\sqrt{n}} \mathbf{U}^N \mathbf{S}^N \quad (3)$$

where \mathbf{L} is a matrix with N columns, each of which is the scaled spatial EOF, this leads to the field covariance

$$\mathbf{B} = \frac{1}{n} \mathbf{X}^N \mathbf{X}^{N^T} = \mathbf{L} \mathbf{L}^T \quad (4)$$

DINEOF cannot for certain separate noise from signal. It can only say via the cross-validation technique used, that with the amount of data available, no more useful information can be extracted and kept in the retained EOFs. Therefore, if we consider that the N -retained EOFs contain signals and the remaining EOFs contain noise, then the noise can be estimated as the difference between the original values x and the reconstructed ones x^r

$$\mu^2 = \frac{1}{m_p} \sum_{x_{ij} \text{ present}} (x_{ij}^2 - x_{ij}^{r2}) \quad (5)$$

with m_p being present pixels. We can use this equation to estimate the error variance from the variance filtered by the EOF reconstruction. Assuming the observational error uncorrelated, the error-covariance matrix \mathbf{R} has the diagonal form

$$\mathbf{R} = \mu^2 \mathbf{I} \quad (6)$$

where \mathbf{I} is the identity matrix.

The covariance matrix can be calculated for a given image

$$\mathbf{C} = \mu^2 (\mathbf{L}_p^T \mathbf{L}_p + \mu^2 \mathbf{I}_N)^{-1} \quad (7)$$

where \mathbf{L}_p are the EOF values of points for which data are available.

For satellite data, atmospheric corrections and associated errors are likely to contain spatial correlations. Therefore, μ_{eff}^2 is used instead of μ^2

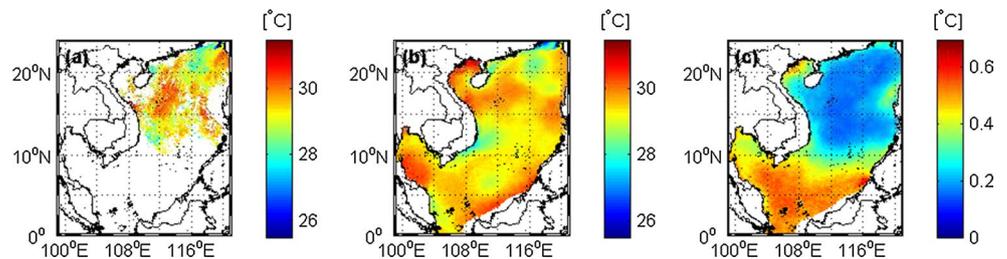
$$\mu_{\text{eff}}^2 = \mu^2 L^2 / (\Delta x \Delta y) \quad (8)$$

where L is the correlation length of the observational error, Δx and Δy are the zonal and meridional resolution, respectively. Different values L were used until the error fields from the analysis gave on average a value of 0.46°C under the clouded regions. Here we chose $L = 88$ km.

The error variance in each grid point as the quadratic form is then calculated

$$\epsilon^2 = \mathbf{i}^T \mathbf{C} \mathbf{i} \quad (9)$$

Fig. 20 An example of error map of the reconstructed SST on 30 June 2009: **a** the original AVHRR SST, **b** the reconstructed AVHRR SST, and **c** the error map of the reconstructed SST



where \mathbf{i} is column array of dimension $N \times 1$ containing the values of the N scaled EOFs at grid point i (irrespective of whether or not the data are missing).

A sample of error map is shown in Fig. 20.

References

- Alvera-Azcárate A, Barth A, Rixen M, Beckers JM (2005) Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic sea surface temperature. *Ocean Model* 9(4):325–346. doi:[10.1016/j.ocemod.2004.08.001](https://doi.org/10.1016/j.ocemod.2004.08.001)
- Alvera-Azcárate A, Barth A, Sirjacobs D, Beckers JM (2009) Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF. *Ocean Sci Discuss* 6(2):1547–1568
- Barron CN, Kara AB (2006) Satellite-based daily SSTs over the global ocean. *Geophys Res Lett* 33(15). doi:[10.1029/2006GL026356](https://doi.org/10.1029/2006GL026356)
- Beckers JM, Rixen M (2003) EOF calculations and data filling from incomplete oceanographic datasets. *J Atmos Ocean Technol* 20(12):1839–1856
- Beckers JM, Barth A, Alvera-Azcárate A (2006) DINEOF reconstruction of clouded images including error maps—application to the sea surface temperature around Corsican Island. *Ocean Sci* 2(2):183–199. doi:[10.5194/os-2-183-2006](https://doi.org/10.5194/os-2-183-2006)
- Brankart JM, Brasseur P (1996) Optimal analysis of in situ data in the western Mediterranean using statistics and cross-validation. *J Atmos Ocean Technol* 13(2):477–491
- Casey KS, Brandon TB, Cornillon P, Evans R (2010) The past, present, and future of the AVHRR Pathfinder SST program. In: Barale V, Gower J, Alberotanza L (eds) *Oceanography from space: Revisited*, Springer. doi:[10.1007/978-90-481-8681-5_16](https://doi.org/10.1007/978-90-481-8681-5_16)
- Chao SY, Shaw PT, Wu SY (1996) El Niño modulation of the South China Sea circulation. *Prog Oceanogr* 38(1):51–93. doi:[10.1016/S0079-6611\(96\)00010-9](https://doi.org/10.1016/S0079-6611(96)00010-9)
- Chu PC, Lu S, Chen Y (1997) Temporal and spatial variabilities of the South China Sea surface temperature anomaly. *J Geophys Res Oceans* 102(C9):20937–20955. doi:[10.1029/97JC00982](https://doi.org/10.1029/97JC00982)
- Chu PC, Edmons NL, Fan C (1999) Dynamical mechanisms for the South China Sea seasonal circulation and thermohaline variabilities. *J Phys Oceanogr* 29(11):2971–2989
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen I, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavaloto C, Thépaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137(656):553–597. doi:[10.1002/qj.828](https://doi.org/10.1002/qj.828)
- Ding Y, Fu D, Wei Z, Mao Z, Zou J (2008) Reconstruction of incomplete satellite oceanographic data sets based on EOF and Kriging methods. In: *Proceedings of SPIE 7109, image and signal processing for remote sensing XIV*, 710913 (October 10, 2008), vol 7109. doi:[10.1117/12.799894](https://doi.org/10.1117/12.799894)
- Dippner J, Nguyen K, Hein H, Ohde T, Loick N (2007) Monsoon-induced upwelling off the Vietnamese coast. *Ocean Dyn* 57:46–62. doi:[10.1007/s10236-006-0091-0](https://doi.org/10.1007/s10236-006-0091-0)
- Donlon CJ, Minnett PJ, Gentemann C, Nightingale TJ, Barton IJ, Ward B, Murray MJ (2002) Toward improved validation of satellite sea surface skin temperature measurements for climate research. *J Clim* 15(4):353–369
- Emery WJ, Thomson R (2001) *Data analysis methods in physical oceanography*. Elsevier, Amsterdam
- Fang G, Chen H, Wei Z, Wang Y, Wang X, Li C (2006) Trends and interannual variability of the South China Sea surface winds, surface height, and surface temperature in the recent decade. *J Geophys Res* 111:C11S16. doi:[10.1029/2005JC003276](https://doi.org/10.1029/2005JC003276)
- Fang G, Wang G, Fang Y, Fang W (2012) A review on the South China Sea western boundary current. *Acta Oceanol Sin* 31:1–10. doi:[10.1007/s13131-012-0231-y](https://doi.org/10.1007/s13131-012-0231-y)
- Fang W, Fang G, Shi P, Huang Q, Xie Q (2002) Seasonal structures of upper layer circulation in the southern South China Sea from in situ observations. *J Geophys Res Oceans* 107(C11):23-1–23-12. doi:[10.1029/2002JC001343,3202](https://doi.org/10.1029/2002JC001343,3202)
- Gan J, Qu T (2008) Coastal jet separation and associated flow variability in the southwest South China Sea. *Deep Sea Res Part I Oceanogr Res Papers* 55(1):1–19. doi:[10.1016/j.dsr.2007.09.008](https://doi.org/10.1016/j.dsr.2007.09.008)
- Guan L, Kawamura H (2003) SST availabilities of satellite infrared and microwave measurements. *J Oceanogr* 59(2):201–209. doi:[10.1023/A:1025543305658](https://doi.org/10.1023/A:1025543305658)
- Hong H, Zhang C, Shang S, Huang B, Li Y, Li X, Zhang S (2009) Interannual variability of summer coastal upwelling in the Taiwan Strait. *Cont Shelf Res* 29(2):479–484. doi:[10.1016/j.csr.2008.11.007](https://doi.org/10.1016/j.csr.2008.11.007)
- Kilpatrick KA, Podestá GP, Evans R (2001) Overview of the NOAA/NASA advanced very high resolution radiometer pathfinder algorithm for sea surface temperature and associated matchup database. *J Geophys Res Oceans* 106(C5):9179–9197. doi:[10.1029/1999JC000065](https://doi.org/10.1029/1999JC000065)
- Klein SA, Soden BJ, Lau NC (1999) Remote sea surface temperature variations during ENSO: evidence for a tropical atmospheric bridge. *J Clim* 12(4):917–932
- Kuo NJ, Zheng Q, Ho CR (2000) Satellite observation of upwelling along the western coast of the South China Sea. *Remote Sens Environ* 74(3):463–470. doi:[10.1016/S0034-4257\(00\)00138-3](https://doi.org/10.1016/S0034-4257(00)00138-3)
- Kuo NJ, Zheng Q, Ho CR (2004) Response of Vietnam coastal upwelling to the 1997–1998 ENSO event observed by multisensor data. *Remote Sens Environ* 89(1):106–115. doi:[10.1016/j.rse.2003.10.009](https://doi.org/10.1016/j.rse.2003.10.009)
- Kuo NJ, Ho CR, Lo YT, Huang SJ, Chang L (2009) Analysis of chlorophyll-a concentration around the South China Sea from ocean color images. In: *OCEANS 2009—EUROPE*, pp 1–4. doi:[10.1109/OCEANSE.2009.5278318](https://doi.org/10.1109/OCEANSE.2009.5278318)
- Li Y, Peng S, Yang W, Wang D (2012) Numerical simulation of the structure and variation of upwelling off the east coast of Hainan Island using QuikSCAT winds. *Chin J Oceanol Limnol* 30(6):1068–1081. doi:[10.1007/s00343-012-1275-8](https://doi.org/10.1007/s00343-012-1275-8)
- Lin CY, Ho CR, Zheng Q, Huang SJ, Kuo NJ (2011) Variability of sea surface temperature and warm pool area in the South China Sea and its relationship to the western Pacific warm pool. *J Oceanogr* 67:719–724. doi:[10.1007/s10872-011-0072-x](https://doi.org/10.1007/s10872-011-0072-x)
- Liu Q, Jiang X, Xie SP, Liu WT (2004) A gap in the Indo-Pacific warm pool over the South China Sea in boreal winter: Seasonal development and interannual variability. *J Geophys Res Oceans* 109(C7). doi:[10.1029/2003JC002179](https://doi.org/10.1029/2003JC002179)
- Locarnini RA, Mishonov AV, Antonov JJ, Boyer TP, Garcia HE, Baranova OK, Zweng MM, Johnson DR (2010) *World ocean atlas 2009. Temperature*, vol 1. In: Levitus S (ed) *NOAA Atlas NESDIS 68*, Ed., U.S. Government Printing Office, Washington, DC, p 182
- Lü X, Qiao F, Wang G, Xia C, Yuan Y (2008) Upwelling off the west coast of Hainan Island in summer: its detection and mechanisms. *Geophys Res Lett* 35(2). doi:[10.1029/2007GL032440](https://doi.org/10.1029/2007GL032440)
- Qiu C, Wang D, Kawamura H, Guan L, Qin H (2009) Validation of AVHRR and TMI-derived sea surface temperature in the northern South China Sea. *Cont Shelf Res* 29(20):2358–2366. doi:[10.1016/j.csr.2009.10.009](https://doi.org/10.1016/j.csr.2009.10.009)

- Qu T (2001) Role of ocean dynamics in determining the mean seasonal cycle of the South China Sea surface temperature. *J Geophys Res Oceans* 106(C4):6943–6955. doi:[10.1029/2000JC000479](https://doi.org/10.1029/2000JC000479)
- Qu T, Mitsudera H, Yamagata T (2000) Intrusion of the north Pacific waters into the South China Sea. *J Geophys Res Oceans* 105(C3):6415–6424. doi:[10.1029/1999JC900323](https://doi.org/10.1029/1999JC900323)
- Qu T, Kim YY, Yaremchuk M, Tozuka T, Ishida A, Yamagata T (2004) Can Luzon Strait transport play a role in conveying the impact of ENSO to the South China Sea? *J Clim* 17(18):3644–3657
- Shaw PT, Chao SY (1994) Surface circulation in the South China Sea. *Deep Sea Res Part I Oceanogr Res Papers* 41(1112):1663–1683. doi:[10.1016/0967-0637\(94\)90067-1](https://doi.org/10.1016/0967-0637(94)90067-1)
- Sorjamaa A, Lendasse A, Cornet Y, Deleersnijder E (2010) An improved methodology for filling missing values in spatiotemporal climate data set. *Comput Geosci* 14(1):55–64. doi:[10.1007/s10596-009-9132-3](https://doi.org/10.1007/s10596-009-9132-3)
- Su J, Pohlmann T (2009) Wind and topography influence on an upwelling system at the eastern Hainan coast. *J Geophys Res Oceans* 114(C6). doi:[10.1029/2008JC005018](https://doi.org/10.1029/2008JC005018)
- Tang D, Kester DR, Ni IH, Kawamura H, Hong H (2002) Upwelling in the Taiwan Strait during the summer monsoon detected by satellite and shipboard measurements. *Remote Sens Environ* 83(3):457–471. doi:[10.1016/S0034-4257\(02\)00062-7](https://doi.org/10.1016/S0034-4257(02)00062-7)
- Tang D, Kawamura H, Guan L (2004) Long-time observation of annual variation of Taiwan Strait upwelling in summer season. *Adv Space Res* 33(3):307–312. doi:[10.1016/S0273-1177\(03\)00477-0](https://doi.org/10.1016/S0273-1177(03)00477-0)
- Toumazou V, Cretaux JF (2001) Using a Lanczos eigensolver in the computation of empirical orthogonal functions. *Mon Weather Rev* 129(5):1243–1250
- Trenberth KE (1997) The definition of El Niño. *Bull Am Meteorol Soc* 78(12):2771–2777
- Wang B, Wu R, Fu X (2000) Pacific-East Asian teleconnection: how does ENSO affect east Asian climate? *J Clim* 13(9):1517–1536
- Wang C (2002) Atmospheric circulation cells associated with the El Niño-Southern Oscillation. *J Clim* 15(4):399–419
- Wang C, Wang W, Wang D, Wang Q (2006) Interannual variability of the South China Sea associated with El Niño. *J Geophys Res Oceans* 111(C3). doi:[10.1029/2005JC003333](https://doi.org/10.1029/2005JC003333)
- Wang J, Tang D, Sui Y (2010) Winter phytoplankton bloom induced by subsurface upwelling and mixed layer entrainment southwest of Luzon Strait. *J Mar Syst* 83(34):141–149. doi:[10.1016/j.jmarsys.2010.05.006](https://doi.org/10.1016/j.jmarsys.2010.05.006)
- Wang W, Wang C (2006) Formation and decay of the spring warm pool in the South China Sea. *Geophys Res Lett* 33(2). doi:[10.1029/2005GL025097](https://doi.org/10.1029/2005GL025097)
- Wu R, Chen Z (2015) Intraseasonal SST variations in the South China Sea during boreal winter and impacts of the East Asian winter monsoon. *J Geophys Res Atmos* 120(12):5863–5878. doi:[10.1029/2015JD023368](https://doi.org/10.1029/2015JD023368)
- Wu R, Chen W, Wang G, Hu K (2014) Relative contribution of ENSO and East Asian winter monsoon to the South China Sea SST anomalies during ENSO decaying years. *J Geophys Res Atmos* 119(9):5046–5064. doi:[10.1002/2013JD021095](https://doi.org/10.1002/2013JD021095)
- Wyrtki K (1961) *Physical Oceanography of the Southeast Asian waters: scientific results of marine investigations of the South China Sea and the Gulf of Thailand 1959–1961*, vol 2. Naga Report, Scripps Institution of Oceanography, La Jolla
- Xie SP, Xie Q, Wang D, Liu WT (2003) Summer upwelling in the South China Sea and its role in regional climate variations. *J Geophys Res Oceans* 108(C8). doi:[10.1029/2003JC001867](https://doi.org/10.1029/2003JC001867)
- Xie SP, Chang CH, Xie Q, Wang D (2007) Intraseasonal variability in the summer South China Sea: Wind jet, cold filament, and recirculations. *J Geophys Res Oceans* 112(C10). doi:[10.1029/2007JC004238](https://doi.org/10.1029/2007JC004238)
- Yan J (1997) Observational study on the onset of the South China Sea southwest monsoon. *Adv Atmos Sci* 14(2):277–287. doi:[10.1007/s00376-997-0026-9](https://doi.org/10.1007/s00376-997-0026-9)
- Yan Y, Ling Z, Chen C (2015) Winter coastal upwelling off northwest Borneo in the South China Sea. *Acta Oceanol Sin* 34(1):3–10. doi:[10.1007/s13131-015-0590-2](https://doi.org/10.1007/s13131-015-0590-2)