Satellite Images

Monthly composites MODIS Level 3 Chl-*a* concentration sea surface temperature (SST) data from October 2002 to September 2013 were obtained from NASA's Goddard Space Flight Center (GSFC, http://oceancolor.gsfc.nasa.gov/). On a pixel-by-pixel basis, 55.9% of the Chl-*a* data were missing, while SST has more available data. The data interpolating of empirical orthogonal function (DINEOF) method developed by Beckers and Rixen (2003) was then performed for missing data reconstruction. Images with less than 5% data coverage (the data coverage for each image was the ratio of available pixels to the sum of all pixels) and pixels with less than 10% coverage (the pixel coverage was the number of available images for each pixel divided by all valid images) from images were removed prior to processing to ensure an accurate reconstruction (Alvera-Azcárate et al., 2007). Of the initial 132 months, 117 monthly images were retained as a result. In addition, a set of 92 daily images with the same spatial resolution from July 1 to September 30, 2013 were downloaded for the validation of DINEOF method using cruise collected measurements.

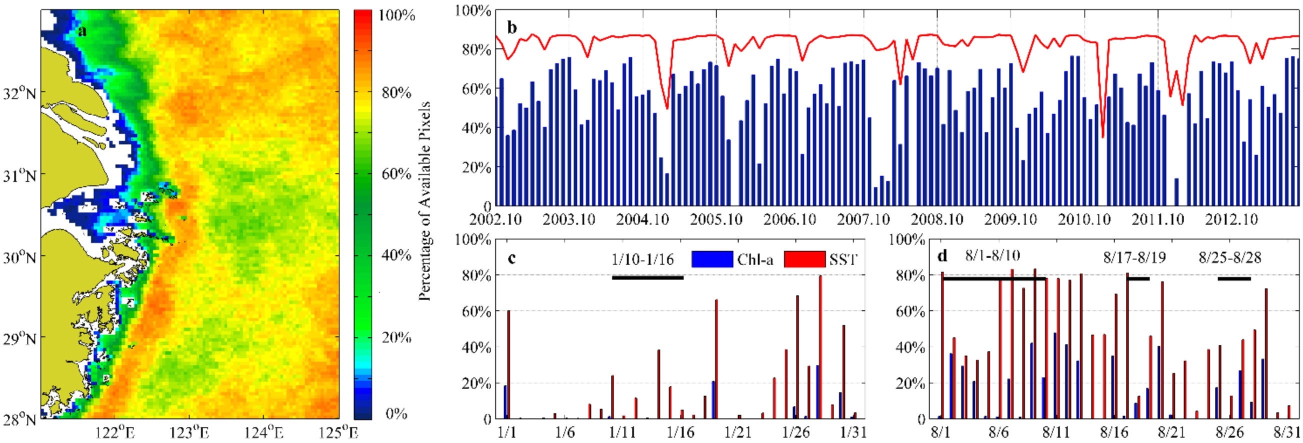

Fig. 1. (a) Spatial distribution of the percentage of available Chl-*a* data in the data set from October 2002 to September 2013, and (b) the availability of monthly Chl-*a* (blue bar) and SST (red line) composites during this period. Also shown are the coverage of available Chl-*a* (blue bar) and SST (red bar) pixels in January (c) and August (d), 2013 during two cruises. The black line represents each period of field measurement.

DINEOF Method

DINEOF is a self-consistent, parameter-free technique for the reconstruction of missing data in oceanographic data sets (Beckers and Rixen, 2003). It is about 30 times faster than the widely used local OI package provided by the Harvard Ocean Prediction System (Carter and Robinson, 1987), and has been proven to be reliable when compared to *in situ* observations (e.g., Alvera-Azcárate et al., 2005).

Similarly to the methods of Beckers and Rixen (2003) and Alvera-Azcárate et al. (2005), the DINEOF interpolation is performed as follows: (1) X is the initial matrix of dimensions $m \times n$

calculated from the Chl-*a* temporal anomaly data (i.e., temporal means are removed from the original matrix), where *m* is the spatial dimension and *n* is the temporal dimension. The missing data points are initialized to zero to prevent a biased first guess. In eqn., *k* is a given number of EOFs. (2) The singular value decomposition (SVD) method is used to compute the first EOF values, and missing data are then replaced by the value obtained with the EOF decomposition:

$$X_{i,j} = \sum_{p=1}^{k} \rho_p (U_p)_i (V_p^T)_j$$

where *i* and *j* are the spatial and temporal indexes of the missing data; U_p and V_p are the p^{th} column of the spatial and temporal EOF, respectively; and ρ_p (where p = 1…*k*) is the corresponding singular value. With the new values for the missing data, the SVD is performed again. (3) The last two steps are repeated until convergence is reached, and then the number of EOFs is increased for the subsequent iteration to obtain *k* EOFs. (4) Cross validation is used to decide the optimal number of EOFs retained for the reconstruction. About 1% of the initial valid data are set aside. The optimal number of EOFs is obtained when the initial data that has been set aside and the reconstructed value reach a minimum level of error. (5) Once step 4 has been completed, the whole process is restarted. At this time, the data set aside for cross-validation are also included. Final values for the missing data are then computed.

Validation of DINEOF Method

To examine the efficiency of the DINEOF method, we reconstructed monthly and daily MODIS Chl-*a* and SST data for a comparison with *in-situ* measurements. Although the daily reconstructed Chl-*a* (3.23 mg/m^3) was slightly overestimated compared to the measured data (2.81 mg/m^3), the concentrations were still statistically significant ($p<0.01$), with a correlation of 0.81 and a root mean square error (RMSE) of 1.23. Similar results were detected after checking the difference between monthly (correlation coefficient, 0.42–0.85; RMSE, 1.41–1.62) and daily (correlation coefficient, 0.65–0.91; RMSE, 1.17–1.22) SST data sets. The DINEOF interpolated Chl-*a* satellite data were clearly overestimated in our study area, due to the highly turbid coastal waters and presence of complicated constituents such as colored dissolved organic matter. We believe that the results derived from daily interpolated Chl-*a* and SST are more accurate for evaluating the ability of the DINEOF method when compared to *in-situ* data. Therefore, the good correlation between our *in-situ* measurements and the remotely-sensed daily data provides further confidence that data reconstructed by DINEOF is able to reflect the general temporal and spatial pattern of Chl-*a* and SST in the study area.

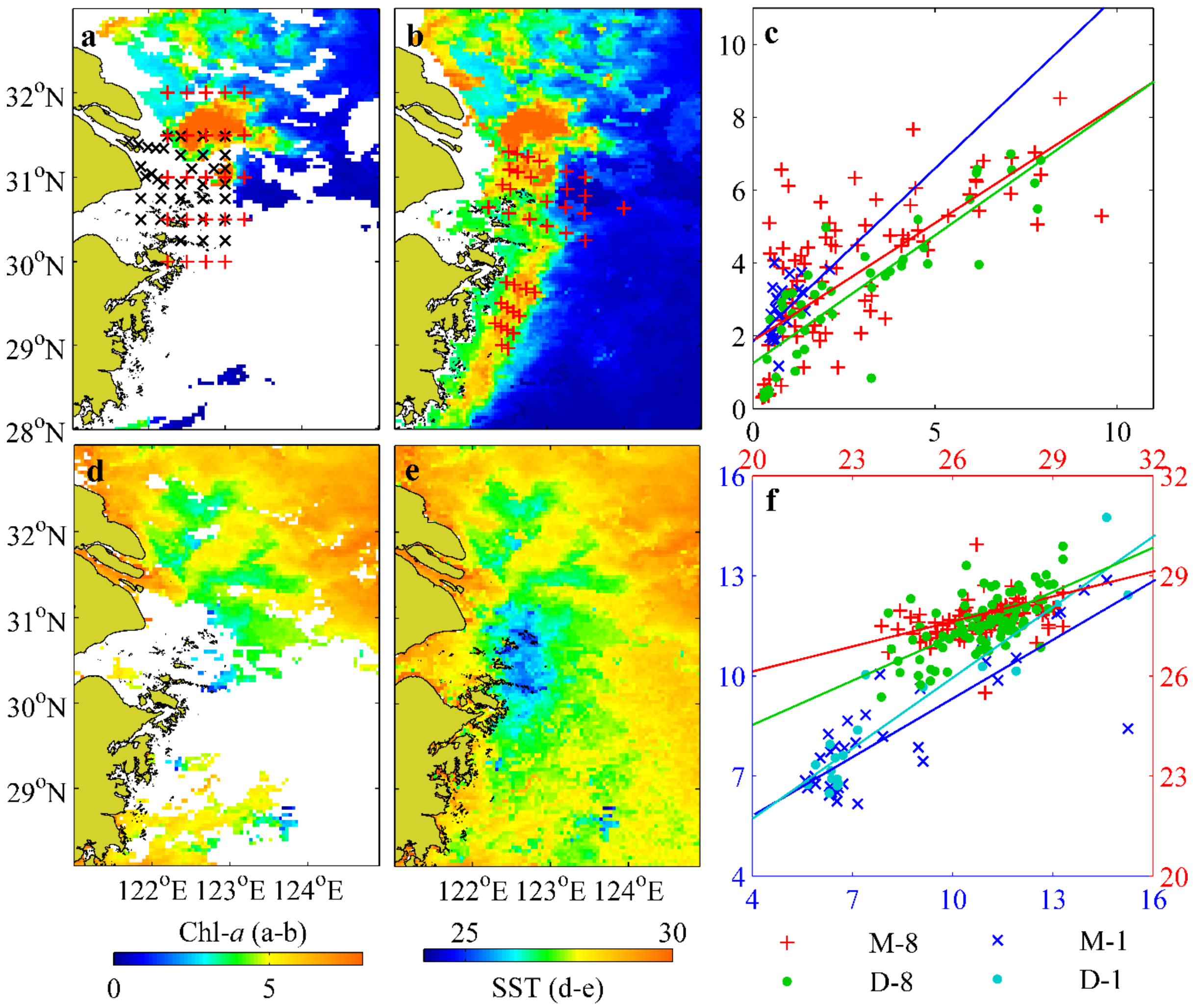

Fig. 2. Original Chl-*a* (a, mg/m^3) and SST (d, C) images for August 27th, 2013, and examples of the corresponding reconstructed images (b and e, respectively). Red plus signs are sampling stations during the cruise, black multiplication signs are sampling stations during the winter cruise. (c) and (f) show the correlation between measured Chl-*a* and SST data (x-axis) and DINEOF reconstructed satellite data (y-axis). Monthly (M-1, blue cross; M-8, red plus) and daily (D-1, green dot; D-8, light blue dot) interpolated data are compared.

Spatial and Temporal Pattern of Reconstructed Data

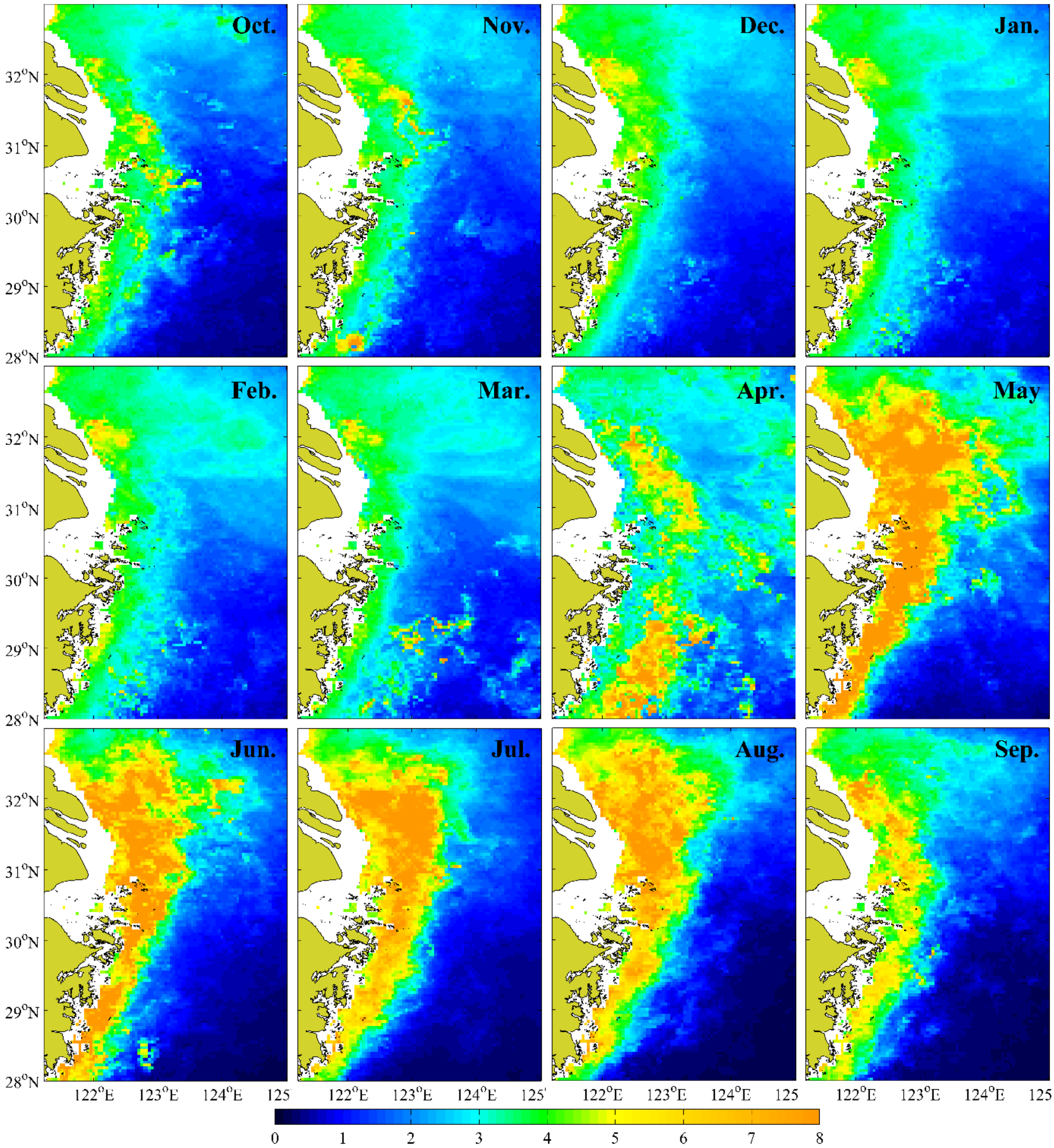

Fig. 3. Monthly Chl-*a* concentration (mg/m^3) images obtained from DINEOF reconstructions for the study area from October 2002 to September 2013.

Wavelet Analysis

Wavelet analysis was used to identify the main signals of the first three temporal modes of Chl-*a* variability. Wavelet analysis decomposes a time series into time-frequency signals, thus the dominant modes of variability and how they vary over time are easy to detect. In this study, Normalized Wavelet Power Spectrum (NWPS) and global NWPS were estimated using a Morlet Wavelet as the mother wavelet, so as to indicate the frequency of variability. The 95% confidence levels were determined from a χ^2 test using the red noise background spectrum. For a detailed description, see Torrence and Compo (1998) and Goubanova et al. (2013).

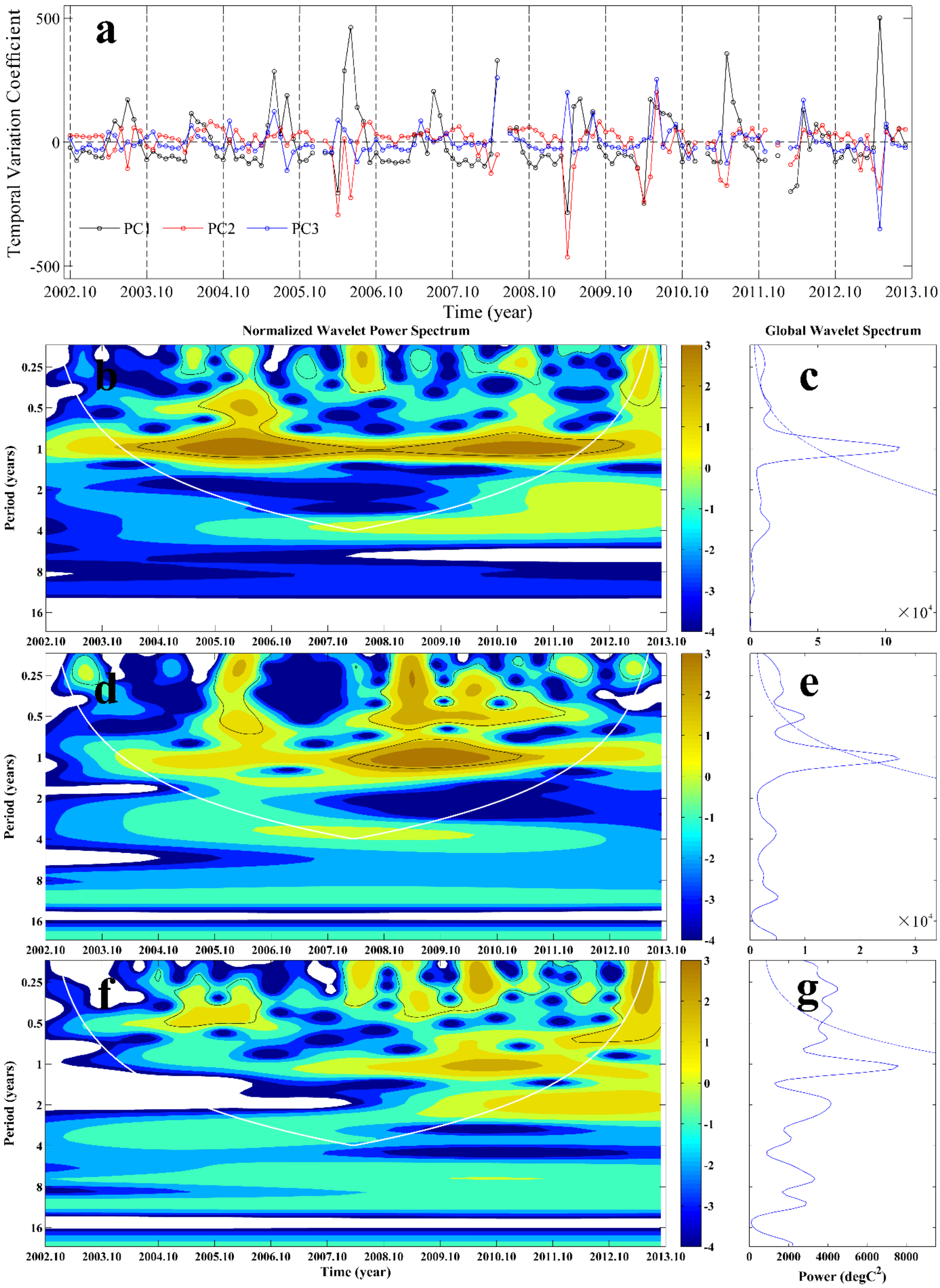

Fig. 4. (a) The temporal evolution functions (i.e., principal components, PCs) for the first three Chl-*a* modes, and the normalized wavelet power spectrum (NWPS; b, d, and f) and global wavelet spectrum (GWS; c, e, and g) for each PC. The colors indicate different degrees of variance (yellow indicates high intensity, blue indicates low intensity). The solid white and black lines in the NWPS depict the cone of influence and 95% confidence level, respectively. The horizontal line in the GWS also shows the 95% confidence level, and the vertical line is the time-averaged intensity at each period.

Beckers, J.M., Rixen, M., 2003. EOF calculations and data filling from incomplete oceanographic datasets. *Journal of Atmospheric & Oceanic Technology*, 20.

Alvera-Azcárate, A., Barth, A., Beckers, J.M., Weisberg, R.H., 2007. Multivariate reconstruction of missing data in sea surface temperature, chlorophyll, and wind satellite fields. *Journal of Geophysical Research*, 112.

Carter, E.F., Robinson, A.R., 1987. Analysis models for the estimation of oceanic fields. *Journal of Atmospheric and Oceanic Technology*, 4, 49-74.

Alvera-Azcárate, A., Barth, A., Rixen, M., Beckers, J.M., 2005. Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic Sea surface temperature. *Ocean Modelling*, 9, 325-346.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. *Bulletion of the American Meteorological Society*, 79, 61-78.

Goubanova, K., Illig, S., Machu, E., Garçon, V., Dewitte, B., 2013. SST subseasonal variability in the central Benguela upwelling system as inferred from satellite observations (1999-2009). *Journal of Geophysical Research: Oceans*, 118, 4092-4110.