

# A new mapping method for propagating data

Peng Yu, Jorge Zavala-Hidalgo, Steven L. Morey, and James J. O'Brien

Center for Ocean-Atmospheric Prediction Studies

The Florida State University

Tallahassee, FL 32306-2840, U.S.A.

peng@coaps.fsu.edu

**Abstract**-Satellite altimetry data have been very useful in the study of oceanic eddies. However, eddies are only detected while they cross the satellite tracks but are missed when they are between tracks. A natural question is: should we use information of both time and space to do the data mapping? A new satellite data processing method combines propagating empirical orthogonal functions (EOF) analysis and an interpolation algorithm based on variance decomposition and maps the data into new time and/or space grids objectively. Here, as an example, the method is applied to the Topex/Poseidon (T/P) along-track sea surface height (SSH) anomaly data. For each EOF mode, space and time information is obtained, and later mapped onto a finer grid. Finally, the field is reconstructed onto the new grid by summing the top several significant modes. The method was tested by performing three simple experiments and another experiment using synthetic SSH from a realistic high resolution model. Then a reconstruction of SSH from satellite data was applied to the Gulf of Mexico (GoM).

## I. INTRODUCTION

In oceanographic research, because of the relatively high cost of observations, it is often impossible to collect observations in time or space of sufficient regularity and spacing to resolve the phenomenon of interest. Efforts are usually made to space measurements as evenly as possible but, for a variety of reasons, station spacing is often problematic, and almost invariably leads to unwanted gaps in the data set. This problem was partly solved with the introduction of many satellites, and since then much more data than ever before could be obtained. However the difficulty induced by the data gaps is still not completely solved.

Data gaps or holes are a problem fundamental to many geophysical data records. Gappy data are frequently the consequence of uneven or irregular sampling (in time and/or space), or they may result from the removal of erroneous values during editing and from sporadic recording system failures [1]. For example, in the study of eddies in the Gulf of Mexico (GoM) with Topex/Poseidon (T/P) data, major difficulties arise when eddies are located between tracks (Fig. 1), *i.e.*, how could these eddies be detected? This problem cannot be solved unless the information from both time and space is taken into account. An option to do this is using the propagating Empirical Orthogonal Functions (EOF) analysis.

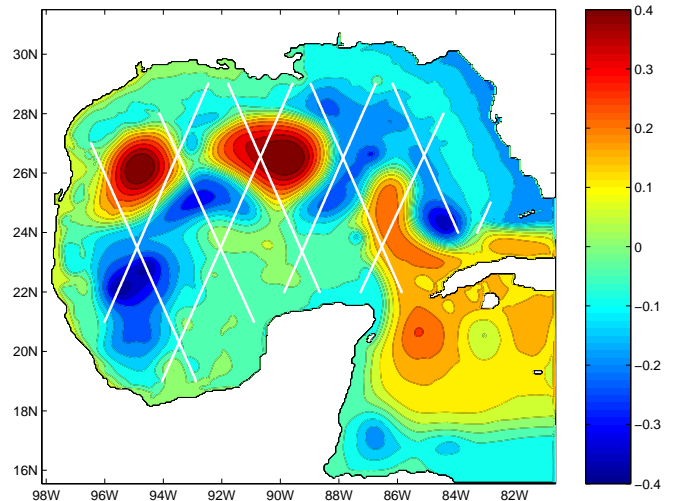


Fig. 1. The sea surface height (SSH) from the COAPS/FSU Gulf of Mexico numerical simulation using the Navy Coastal Ocean Model (NCOM). The white lines are simulated Topex/Poseidon (T/P) tracks.

While some useful computational techniques have been developed for unevenly spaced data as in [2] and [3] and even some advantages to having a range of Nyquist frequencies within a given data set have been shown as in [4], most analysis methods require data values that are regularly spaced in time or space. As a consequence, it is generally necessary to use an interpolation procedure to create the required regular set of data values as part of the data processing. In this paper an alternative to the conventional interpolation method is discussed.

The next section discusses the oceanic and synthetic data used in the research. Section 3 first discusses propagating EOF analysis and then explains how it is used to map the data. Results and discussion will be in section 4 followed by the summary and conclusions in the last section.

## II. MODEL AND DATA

### A. Synthetic data

Two experiments were conducted based on the synthetic data. In experiment one, a sinusoidal wave propagates to the right with a constant speed. The propagating wave is sampled at each 1/4 cycle of itself (Fig. 2). In experiment two, a synthetic eddy propagates westward with a constant speed. The whole field is sampled at the simulated satellite track observation points.

### B. Model

The Navy Coastal Ocean Model (NCOM) is used to simulate the GoM. The model domain includes the whole GoM and the western Caribbean Sea, from 98.15°W to 80.60°W and from 15.55°N to 31.50°N, with model equations discretized on a 1/20-degree grid, in latitude and longitude [5]. The model uses 20 sigma layers uniformly distributed in the upper 100 meters and 40 unevenly distributed z-level layers below 100m. The NCOM is a three-dimensional primitive equation ocean model with the commonly used hydrostatic, incompressible, and Boussinesq approximations; it is forced with climatological monthly surface fluxes of heat and momentum derived from the Comprehensive Ocean Atmosphere Data Set [6]. The discharge of 30 major rivers is prescribed using monthly means for those in the United States and annual means for those in Mexico. A surface salinity flux uniform in space and time is applied to balance the river input of fresh water. Climatological volume flux and temperature and salinity fields are prescribed along the open boundary in the Caribbean Sea, generating a similar current structure in the Yucatan Strait to that obtained from recent measurements, as in [7] and [8]. The last seven years of a 10-year model simulation are analyzed. Detailed information of the NCOM and the GoM simulation are in [9] and [10], respectively.

### C. Topex/Poseidon data

Satellite altimeters measure sea surface height (SSH) relative to an Earth-centered coordinate system, providing an absolute reference frame for studies of sea level rise. Satellite altimeters also provide frequent near-global data coverage of the world's oceans. This allows altimeter-derived estimates of sea surface height to be based on data collected over much of the ocean surface, in contrast to the geographically sparse coastal data provided by tide gauges.

Before the T/P mission, large errors in satellite altimeter SSH measurements prevented accurate SSH estimates from being obtained with this technology. Since its launch in August of 1992, the T/P satellite has been providing SSH measurements with an accuracy of better than 5 cm RMS [11].

T/P SSH data were extracted from the geophysical data records (AVISO 1992) and analyzed for the period January 1993 through February 2001. Corrections described in AVISO (1998) were applied to the data.

## III. METHODOLOGY

### A. Complex empirical orthogonal function analysis

The goal of this technique is to extract physical information on propagating features from a two-dimensional data array,

$$D = (h_{nm})_{N \times M} \quad (3.1)$$

where  $D$  is an  $N \times M$  matrix with elements  $h_{nm}$  and  $N$  and  $M$  being the number of space and time points, respectively. Through the use of CEOF analysis the component eigenmodes of  $D$  are obtained. The theoretical development and other details of CEOF are omitted here, see [12].

The component eigenmodes of  $D$  consist of  $M$  pairs of complex vectors  $T_i$  and  $S_i$  where

$$T_i = (T_m)_{1 \times M} \quad (3.2)$$

$$S_i = (S_n)_{N \times 1} \quad (3.3)$$

The vector  $S$  will be referred to as the spatial function (SF) and  $T$  as the temporal function (TF). As will be shown, it is possible to obtain phase and amplitude information from the spatial and temporal components, respectively, of a given eigenmode, yielding some very useful physical information.

By definition a complex vector  $C(x)$  can be represented as

$$C(x) = A(x) \exp[i\Theta(x)] \quad (3.4)$$

where  $A(x)$  is an amplitude and  $\Theta(x)$  is a phase. We can use the vector identity shown above to represent  $S$  and  $T$  as

$$S(x) = E(x) \exp[i\Theta(x)] \quad (3.5a)$$

$$T(t) = R(t) \exp[i\phi(t)] \quad (3.5b)$$

Their product (real part) is

$$ST = \text{Re} \{ E(x)R(t) \exp[i(\Theta(x) + \phi(t))] \} \\ = E(x)R(t) \cos(\Theta(x) + \phi(t)) \quad (3.6)$$

$E(x)$  (spatial amplitude function) shows the variability of amplitude in space associated with a given eigenmode.  $R(t)$  (temporal amplitude function) provides information on the variability of amplitude. For a line of constant phase (letting  $\xi(x, t) = \Theta(x) + \phi(t)$ ),

$$d\xi(x, t) = \frac{\partial \Theta(x)}{\partial x} \partial x + \frac{\partial \phi(t)}{\partial t} \partial t = 0 \quad (3.7)$$

or rearranging,

$$\frac{\partial x}{\partial t} = c = - \frac{\partial \phi(t) / \partial t}{\partial \Theta(x) / \partial x} \quad (3.8)$$

since the phase speed  $c = \omega / k$ , then

$$k = \frac{\partial \Theta(x)}{\partial x} \quad (3.9)$$

$$\omega = - \frac{\partial \phi(t)}{\partial t} \quad (3.10)$$

### B. The objective space-time interpolation

Propagating EOF analysis could decompose propagating data into orthogonal modes and identify the propagating information from each mode. After the

application of propagating EOF analysis, the spatial and temporal information is obtained. The top several most significant modes (modes with bigger eigenvalues compared to others) are picked among all modes. After the manipulation of the complex temporal function (TF) and spatial function (SF) associated with each mode, the phase and amplitude information in each mode is mapped onto a regular finer grid in time and/or space. With the new information, the new TFs and SFs are rebuilt and finally the reconstruction of data is done by summing the product of each TF and SF. Further, this method could be used for data interpolation in both time and space at the same time.

#### IV. RESULTS AND DISCUSSION

Before the study of real satellite data, three examples are given based on the synthetic data and model simulation output data.

Experiment one is a very simple case with only time interpolation (Fig. 2). In this case a sinusoidal wave propagates to the right. The wave is sampled at each 1/2 of its period, *i.e.*, there are two samples during each cycle. The propagating EOF analysis was conducted for the sampled data and the top several significant modes were picked. Then the phase and amplitude of the time series for each mode was interpolated over time to a finer time grid. After completion of interpolation, the new time series for each mode was reconstructed and the product of it and the spatial pattern for each mode were summed. Fig. 2 clearly shows that the new method does identify the propagating information in comparison to the linear interpolation method.

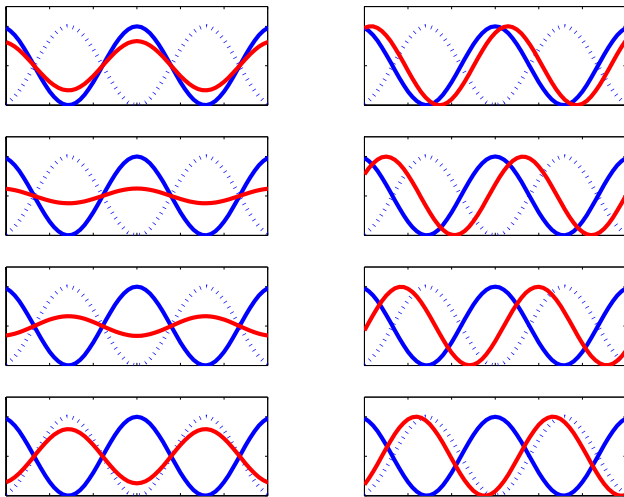


Fig. 2. A sinusoidal wave propagates to the right, the blue solid lines are the initial position of the wave and the red ones are the position of the wave at next sampling time. The left column shows the result using linear interpolation, and the right one shows the result with the new method.

Experiment two is a case to show spatial interpolation (Fig. 3). In this case, a synthetic eddy propagates westward in a square domain. The whole field is sampled on the simulated tracks shown in Fig. 3. First the propagating EOF was applied to the data and the top six most significant modes were picked. Then the phase and amplitude of the spatial pattern for each mode were interpolated onto a new finer regular grid and the new spatial pattern for each mode was reconstructed using the new phase and amplitude information. Finally, we added up the product of new spatial pattern and the original time series for each mode. This experiment clearly shows that the new method does very well to recover the eddy and its propagation.

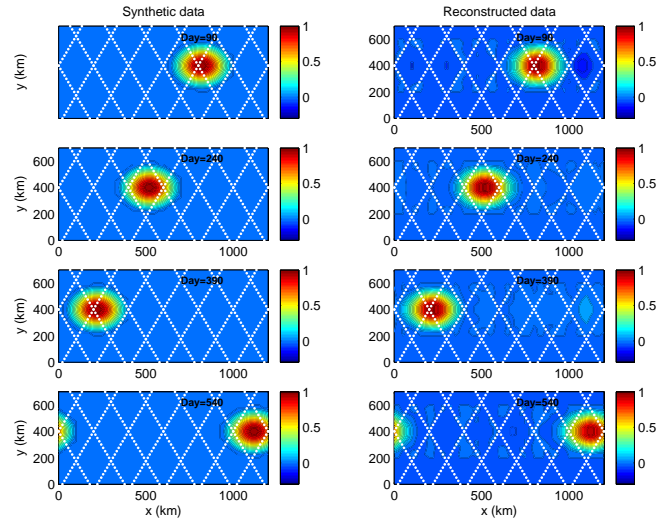


Fig. 3. The left column is the synthetic data with an eddy moving to the west. The white dots are the simulated satellite tracks, *i.e.*, the positions where “observations” are obtained. The right column shows the reconstructed data based on the simulated satellite observations.

Experiment three is a study of the sea surface height from the COAPS/FSU Gulf of Mexico numerical simulation using the NCOM. In this case, nine tracks were simulated as in the GoM and data were only sampled at the simulated satellite track observation positions. Based on the “observations”, the same method was applied. The reconstruction is again successful in this case.

After the above three experiments, the method was applied to the T/P altimetry SSH anomaly data. Synchronization of all tracks has to be done before the application of this method, since in each cycle the sampling time for each track position is different. Because the satellite passes by the GoM very fast each time, an assumption was taken that the observation for each track in the same cycle are simultaneous. At first the data for each track were processed using the method depicted in experiment one and then all the data for different tracks were interpolated to the same time grid to synchronize all the tracks. Next, the

procedures explained in experiment two were applied and finally the reconstructed data were obtained. Further, more could be done like the

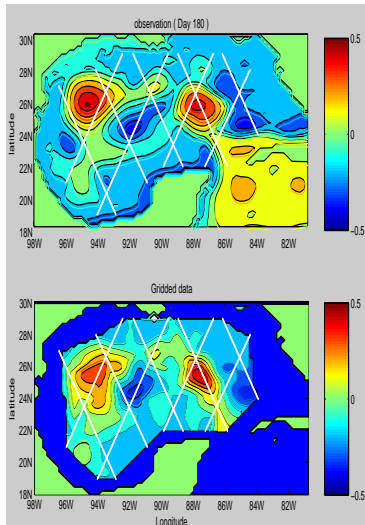


Fig. 4. The upper panel is the SSH from the NCOM GoM simulation; the lower panel shows the reconstructed data. The white dots are the simulated satellite tracks.

interpolation in both time and space at the same time, which will not be explained here. In Fig. 5, a zonal section along 26°N in the GoM from the reconstructed data shows that the life cycles for eddies, especially those big events like the one between 1998 and 1999, are clear and continuous.

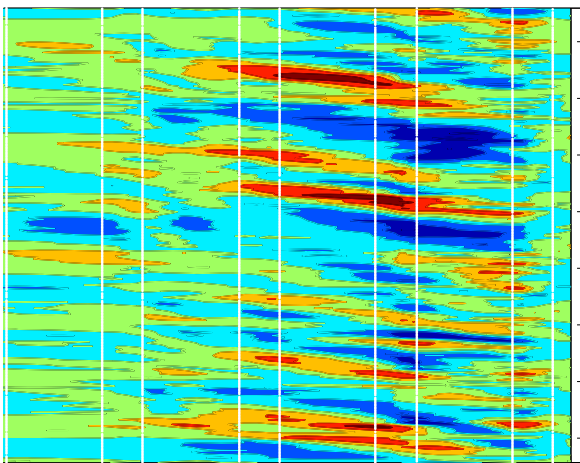


Fig. 5. Reconstructed T/P SSH anomaly of a zonal section (26°N) in the Gulf of Mexico.

## V. Summary and conclusions

The basic idea behind this method is map the reconstructed propagating EOFs onto a finer space/time grid combining information from both time and space at the same time;

This method has been proven to be successful to identify the life cycle of large eddies;

It is a better way to more fully use information from the valuable satellite altimeter data and a good candidate for creating SSH fields for numerical model initialization and data assimilation.

## Acknowledgments

The Center for Ocean-Atmospheric Prediction Studies receives its base funding from the NOAA Office of Global Programs, the NASA Physical Oceanography Program, and the Secretary of Navy Grant from ONR to Dr. James J. O'Brien.

## REFERENCES

- [1] W.J. Emery, and R.E. Thomson, *Data Analysis Methods in Physical Oceanography*, 2<sup>nd</sup> and revised ed., Elsevier, 1997.
- [2] D.D. Meisel, "Fourier transforms of data sampled at unequalled observational intervals". *Astrn.J.*, pp.83,538–545, 1978.
- [3] D.D. Meisel, "Fourier transforms of data sampled in unequally spaced segments", *Astron. J.*, 84,116–126, 1979.
- [4] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes*, 2<sup>nd</sup> ed. Cambridge University Press, Cambridge, 1992.
- [5] J. Zavala-Hidalgo, S. L. Morey, and J. J. O'Brien, "Seasonal Circulation on the Western Shelf of the Gulf of Mexico", unpublished.
- [6] A. DaSilva, A.D. Young, and S. Levitus, *Atlas of surface marine data 1994, Volume 1: Algorithms and procedures*, NOAA Atlas NESDIS 6, U.S. Department of Commerce, Washington, 1994.
- [7] J. Ochoa, J. Sheinbaum, A. Badan, J. Candela, and D. Wilson, "Geostrophy via potential vorticity inversion in the Yucatan Channel", *J. Mar. Res.*, 59, 725–747, 2001.
- [8] J. Sheinbaum, J. Candela, A. Badan, and J. Ochoa, "Flow structure and transport in the Yucatan Channel", *Geophys. Res. Lett.*, 29(3), 1040, doi: 10.1029/2001GL013990, 2002.
- [9] P. Martin, *A description of the Navy Coastal Ocean Model version 1.0 NRL Report NRL/FR /7322-009962*, Naval Research Laboratory, Stennis Space Center, MS, 39 pp., 2000.
- [10] S.L. Morey, P.J. Martin, J.J. O'Brien, A.A. Wallcraft, and J. Zavala-Hidalgo, "Export pathways for river discharge fresh water in the northern Gulf of Mexico", *J. Geophys. Res.*, in press.
- [11] L.L. Fu, E.J. Christensen, C.A. Yamarone Jr., M. Lefebvre, Y. Ménard, M. Dorrer, and P. Escudier, "TOPEX/POSEIDON mission overview", *J. Geophys. Res.*, 99, 23,369–24,381, 1994.
- [12] T.P. Barnett, "Interaction of the monsoon and Pacific trade wind system at interannual time scales, 1; The equatorial zone", *Mon. Weather Rev.*, 111, 756–773, 1983.