

OPTIMAL SPECTRAL DECOMPOSITION (OSD) FOR REMOTELY SENSED OCEAN DATA ASSIMILATION

Peter C. Chu

Naval Postgraduate School, Monterey, CA93943, USA
(pcchu@nps.edu, tel: 831-656-3688, fax: 831-656-3686)

ABSTRACT

Assimilation of remotely sensed ocean data (velocity, temperature, and salinity) into numerical model is of great importance in oceanic and climatic research. However, the data should be reconstructed (onto grids) before assimilation since the original datasets are usually noisy and sparse. This paper describes a recently developed optimal spectral decomposition (OSD) method for mapping and noise filtration with examples of reconstructing the data from the Argo profiling and trajectories, Ocean Surface Current Analyses – Real time (OSCAR), shore-based high-frequency (HF) Doppler radar (CODAR) and Global Temperature-Salinity Profile Program (GTSP).

Index Terms—Optimal spectral decomposition (OSD), Argo profiling and trajectory data, OSCAR, CODAR, GTSP

1. INTRODUCTION

Remotely observed ocean data from Argo profiling (T , S) and trajectories (velocity), Ocean Surface Current Analyses – Real time (OSCAR), Global Temperature-Salinity Profile Program (GTSP), as well as shore-based high-frequency (HF) Doppler radars such as CODAR are usually acquired for various spatial and temporal resolutions with errors and noises. These observations enhance considerably our understanding of oceanic phenomena and processes and are useful for numerous practical applications in climatic, ecology, and marine biology. Direct assimilation of the noisy and sparse data into numerical ocean models may not be feasible. An important question arises: Can we reconstruct these noisy and sparse data into grids before being assimilated into numerical ocean models?

From the theoretical point of view this problem is not simple by the following reasons. First, remote sensed data contain measurement errors of different features, which need to be preliminarily filtrated out. Because a priori real (“true”) state of ocean states is always unknown, it is very difficult to divide noises and signal and to calculate the

error correlation functions. This makes the traditional objective analysis methods not feasible to reconstruct these data.

Base on the fact that any field (temperature, salinity, or velocity) can be decomposed into generalized Fourier series, an optimal spectral decomposition (OSD) method has been developed to fulfill the task. The three dimensional field is then represented by linear combination of the products of basis functions (or called modes) and corresponding Fourier coefficients. If a rectangular closed ocean basin is considered, the basis functions are sinusoidal functions. If a realistic ocean basin is considered, the basis functions are the eigen-values of the three-dimensional Laplace operator with real topography. After the Fourier coefficients are determined from observational data, the gridded field (temperature, salinity, or velocity) is calculated from the generalized Fourier series. Major benefits of using the OSD method are (1) satisfaction of the boundary conditions for the ocean variables, (2) no *a-prior* needed, and (3) noise filtration.

2. OSD

Let (\mathbf{x}, t) be the spatial coordinates and time, and Ω the domain. A scalar variable $c(\mathbf{x}, t)$ is decomposed using the generalized Fourier series^{[1]-[5]},

$$c(\mathbf{x}, t) = D_0(t) + \sum_{i=1}^I D_i(t) \Xi_i(\mathbf{x}), \quad (1)$$

where I is the truncated mode number, $\Xi_i(\mathbf{x})$ and $A_i(t)$ are the orthogonal basis functions (or called modes). A vector field (such as velocity) is represented by

$$\mathbf{U}(\mathbf{x}, t) = C(t) \nabla \Psi_0(\mathbf{x}) + \sum_{n=1}^N A_n(t) \nabla \times [\mathbf{k} \Psi_n(\mathbf{x})] + \sum_{m=1}^M B_m(t) \nabla \cdot \Phi_m(\mathbf{x}), \quad (2)$$

where $\{\Psi_n(\mathbf{x}), \Phi_m(\mathbf{x})\}$ are the basis functions; $\{A_n, B_m\}$ are the spectral coefficients; \mathbf{k} is the unit vector in the vertical

(upward positive) direction; ∇ is the horizontal gradient operator; $[\Psi_0(\mathbf{x}), C]$ are the harmonic function and its coefficient for area of Ω with three boundary Γ . If a rectangular closed ocean basin is considered, the basis functions are sinusoidal functions. If a realistic ocean basin is considered, the basis functions are the eigen-values of the three-dimensional Laplace operator with real topography. One major benefit of using the OSD method is that the boundary conditions for the ocean variables (temperature, salinity, velocity) are always satisfied.

The OSD method has three components: (1) determination of the basis functions, (2) optimal mode truncation, and (3) determination of the Fourier coefficients through solving a set of ill-posed algebraic equations. Determination of basis functions is to solve the eigen-value problem. Chu et al. [2][3] (2003a, b) developed a theory to obtain the basis functions with open boundaries. The basis functions are only dependent on the geometry of the ocean basin, not dependent on the oceanic variables. This is to say, no matter which variable (temperature, salinity, or velocity) is concerned, the basis functions are the same, and can be pre-determined before the data analysis. For data without error, the more the modes, the more the accuracy of the processed field. For data with error, this rule of the thumb is no longer true. Inclusion of high-order modes leads to increasing error. The Vapnik variational principal [6] is used to determine the optimal mode truncation. After the mode truncation, optimal field estimation is to solve a set of a linear algebraic equation of the Fourier coefficients. This algebraic equation is usually ill-posed. The rotation method [7] is developed to change the matrix of the algebraic equation from ill-posed to well-posed such that a realistic set of the Fourier coefficients are obtained.

3. OSD FOR ARGO TRAJECTORY DATA ASSIMILATION

Since its inception in 1999 over 30 nations have committed support towards building the Argo array, and a number of other countries have assisted in deploying floats. Deployment of Argo floats started in 2000 and more than 3,000 floats (one float per $3^\circ \times 3^\circ$ box) surveying the global ocean. Every 10 days, the float pumps fluid into an external bladder and rises to the surface with measuring temperature and salinity. At the surface, the data are transmitted to satellite. The bladder then deflates and float sinks to the depth to drift until the cycle is repeated.

Between November 2003 and January 2005, over 56000 float days (cumulative) of data were collected in the North Atlantic (10°N - 60°N) in general at three parking depths: 1000 m, 1500 m and 2000 m. The floats parking at 2000 m, depths shallower than 1000 m, and unknown depths are excluded from the analysis. Temperature at 950 m and trajectories at 1000 m and 1500 m are extracted from all the

existing Argo floats. The data from 1000 m and 1500 m were grouped together to represent the mid-depth.

Current velocities computed along the original (non-smoothed) Argo tracks are shown in Fig. 1 as red arrows. After using the OSD method, the velocity vectors are shown in the blue arrows. A strong contribution from intensive eddies, such as that shown in inset B, narrow jets and measurement errors is clearly identified here. Visually, the velocity pattern corresponding to the original data looks quite chaotic, and there are 500-600 km spatial gaps in observation coverage. To understand the reconstruction skill for such data we applied three criteria: (1) the formal mean square error (the reconstruction error) computed by the ‘‘laminar ensemble’’ technique, (2) statistics of angle (α) between the reconstructed and observed velocity at float locations, and (3) stability degree of the reconstructed snapshot on observation sampling.

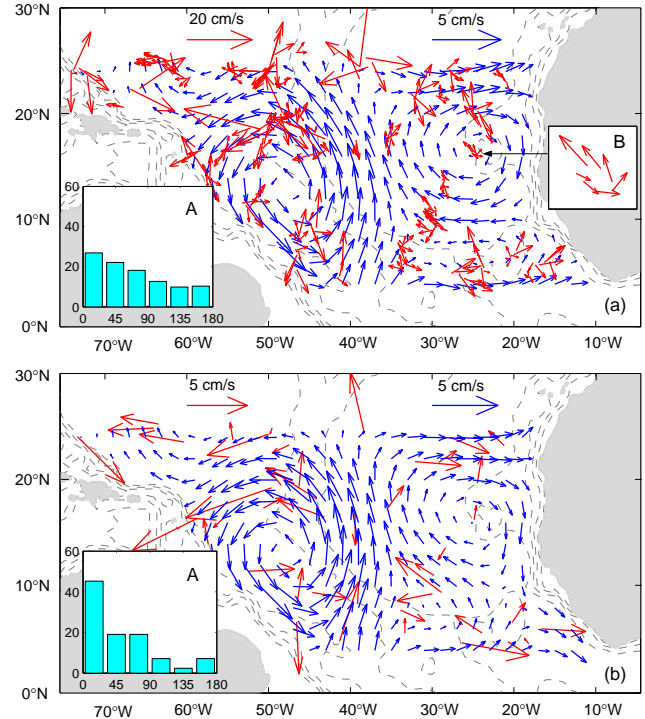


Figure 1. Sensitivity of the reconstructed circulation patterns to filtration of the original data: OSD is applied to (a) the original data (November-December, 04); and (b) the original data (November-December, 04) filtered with a 2-month window. Blue and red arrows correspond to the reconstructed circulation and original/filtered data. For α -histograms (inserts A) the x-axes is the angle α and the y-axes is the number of comparisons (%) (from [8]).

4. OSD FOR CODAR DATA ASSIMILATION

Use of CODAR to probe the ocean surface currents is based on the concept that radio waves are backscattered from the moving ocean surface by resonant surface waves of one-half the incident radar wavelength. This Bragg scattering effect

results in two discrete peaks in the Doppler spectrum. In the absence of a surface current, the spectral peaks are symmetric and their frequency (σ) is offset from the origin by an amount proportional to $2c_0\lambda^{-1}$, where c_0 represents the linear phase speed of the surface wave and λ is the radar wavelength. If there is an underlying surface current, the Bragg peaks in the Doppler spectrum are displaced by an amount of $\Delta\sigma = 2V\lambda^{-1}$, where V is the radial component of current along the direction of the radar. Using two radar stations, separated in space by a baseline distance of 20–30 km, the two-dimensional velocity vector is resolved. Capability of the OSD method is demonstrated by reconstructing the raw noisy CODAR data (Fig. 2a) at 1700 UT on December 1, 1999 into the gridded data (Fig. 2b). After using the OSD method, the reconstructed surface velocity field is ready for assimilation into coastal numerical ocean models.

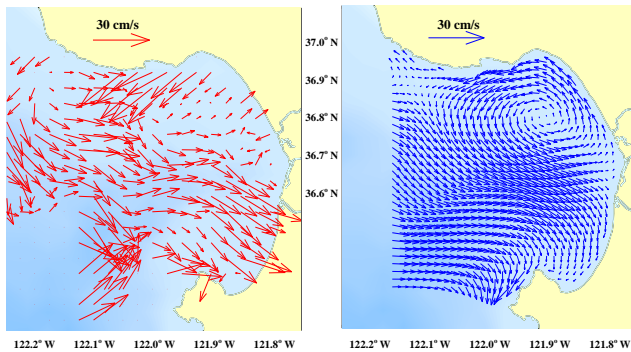


Figure 2. CODAR derived surface currents for the Monterey Bay at 17:00 UT December 1, 1999: (a) before and (b) after using the OSD method.

5. OSD FOR OSCAR DATA ASSIMILATION

Near-real time ocean surface currents derived from satellite altimeter (JASON-1, GFO, ENVISAT) and scatterometer data on $1^\circ \times 1^\circ$ resolution for world oceans (59.5° S to 59.5° N) posted online as “Ocean Surface Current Analyses – Real Time (OSCAR)”, provide invaluable resources online (<http://www.oscar.noaa.gov/index.html>) for various uses include large scale climate diagnostics and prediction, fisheries management and recruitment, monitoring debris drift, larvae drift, oil spills, fronts and eddies, plus opportunities for search and rescue, naval and maritime operations. The methodology for OSCAR combines geostrophic, Ekman and Stommel shear dynamics, and a complementary term from the surface buoyancy gradient.

A major weakness of the OSCAR dataset is its inability to represent the currents near the lateral boundary since it uses idealized dynamics (geostrophic, Ekman, and Stommel shear dynamics). The most evident western boundary currents such as the Gulf Stream and Kuroshio are missing (Fig. 3a).

The OSD method is used to reconstruct the OSCAR data. After the OSD analysis, the reconstructed OSCAR data show realistic surface circulations including western boundary currents such as Gulf stream, Kuroshio, Brazilian Currents, Somali Currents, and eastern boundary currents such as California Currents, Peru Currents, etc. (Fig. 3b).

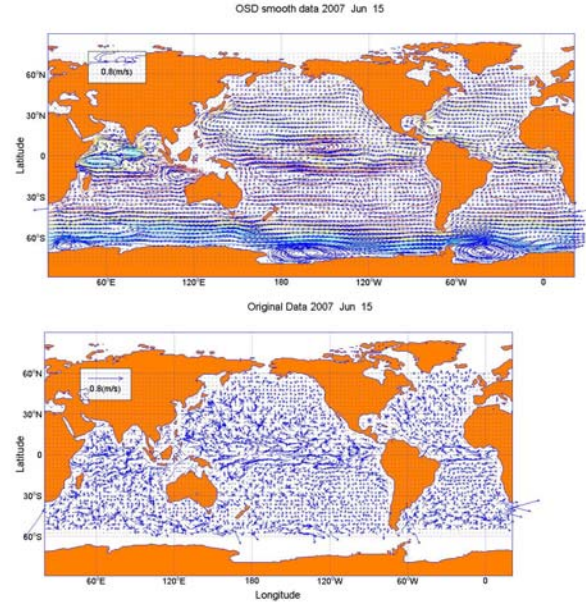


Figure 3. Ocean surface currents on 15 June 2007 from the OSCAR data: (a) after and (b) before the OSD analysis.

6. OSD FOR GTSP-ARGO (T, S) PROFILING DATA ASSIMILATION

Complexity of GTSP-Argo profiling (T, S) data causes difficulty in retrieving useful information. For example, the following sources of uncertainty in Argo float data lead to computational errors in the velocity estimate: (a) the trajectories of Argo floats are not continuous; (b) the vertical shear causes an increase or decrease of the real distance between the points of ascending from and descending to the parking depth; (c) the sequence of float trajectory segments (tracks) only approximate the real Lagrangian paths. We may fix float positions after their ascent to the ocean surface and before their descent to the parking depth only; (d) preliminary computations^[8] demonstrated that high resolution elements of mid-depth circulation in the North Atlantic, such as the northern recirculation gyre in the western North Atlantic, Deep Western Boundary Current (DWBC), the systems of eastward and westward zonal flows in the equatorial Atlantic are also detectable using the Argo float data. However, such a resolution is not available for the whole North Atlantic and high-energetic mesoscale eddies as well as narrow boundary currents should be classified as “noise” and removed from the analysis or parameterized; (e) there are large spatial gaps (from 200 km to 600 km) in Argo

float observation coverage. Using the OSD method, we are computing global monthly (T, S) fields for oceanographic and climatic studies, e.g., temperature at the surface (Fig. 4) and 30 m depth (Fig. 5) for June 2006. After the OSD analysis, the reconstructed GTSP-Argo data are ready to be assimilated into global ocean models.

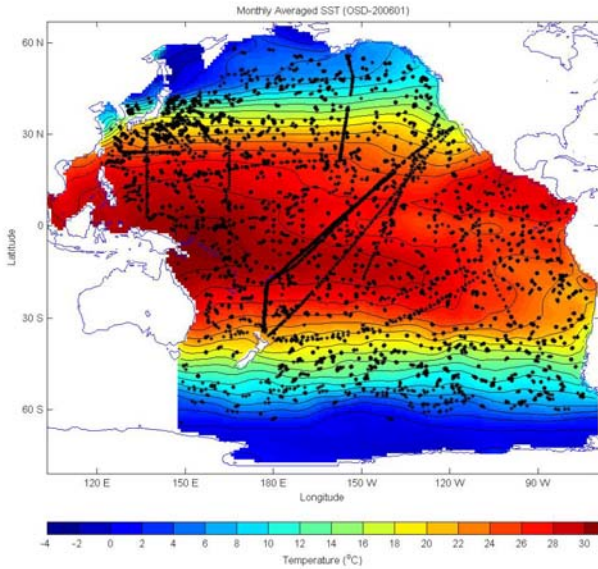


Fig. 4. Sea surface temperature of the Pacific Ocean for June 2006 calculated from GTSP-Argo data using the OSD method.

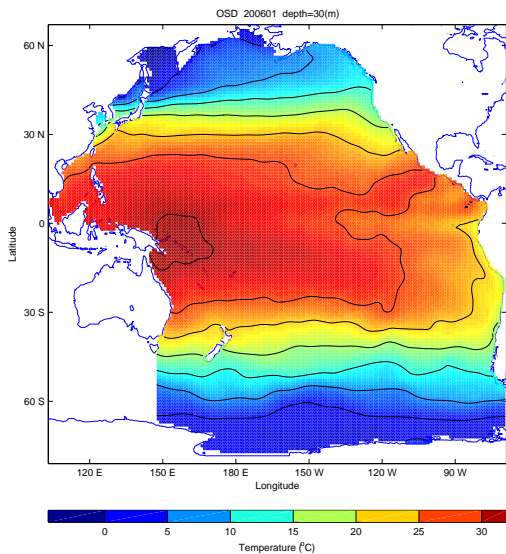


Fig. 5. Temperature (30 m depth) of the Pacific Ocean for June 2006 calculated from GTSP-Argo data using the OSD method.

7. CONCLUSIONS

The OSD is an effective method to reconstruct sparse and noisy ocean data for ocean data assimilation. This scheme includes three components: (1) determination of the basis functions, (2) optimal mode truncation, and (3)

determination of the Fourier coefficients through solving a set of ill-posed algebraic equations. This study demonstrates its capability to process remotely sensed ocean data such as Argo trajectories, Argo-GTSPP profiling (T, S), OSCAR, and CODAR data. With installation of the OSD algorithm into numerical ocean models, the remotely sensed data are easily assimilated into models.

ACKNOWLEDGMENTS

This research was supported by the NOAA/NODC, Office of Naval Research, Naval Oceanographic Office, and Naval Postgraduate School.

REFERENCES

- [1] P.C. Chu, "Fundamental circulation functions for determination of open boundary conditions," Third Conference on Coastal Atmospheric and Oceanic Prediction and Processes, New Orleans, LA, American Meteorological Society, 389–394, 1999.
- [2] P.C. Chu, L.M. Ivanov, T.M. Margolina, T.P. Korzhova, and O.V. Melnichenko, "Analysis of sparse and noisy ocean current data using flow decomposition. I. Theory," *Journal of Atmospheric and Oceanic Technology*, 20, 478 – 491, 2003a.
- [3] P.C. Chu, L.M. Ivanov, T.M., Margolina, T.P. Korzhova, and O.V. Melnichenko, "Analysis of sparse and noisy ocean current data using flow decomposition. Part 2. Applications to Eulerian and Lagrangian data," *Journal of Atmospheric Oceanic Technology*, 20, 492–512, 2003b.
- [4] P.C. Chu, L.M. Ivanov, and O.M. Melnichenko, "Fall-winter current reversals on the Texas-Louisiana continental shelf," *Journal of Physical Oceanography*, 35, 902-910, 2005a.
- [5] P.C. Chu, L.M. Ivanov, and T.M. Margolina, "Seasonal variability of the Black Sea Chlorophyll-a concentration," *Journal of Marine Systems*, 56, 243-261, 2005b.
- [6] V.H. Vapnik, "Reconstruction of Empirical Laws from Observations (in Russian)," Nauka, 447 pp, 1983.
- [7] P.C. Chu, L.M. Ivanov, and T.M. Margolina, "Rotation method for reconstructing process and fields from imperfect data.," *International Journal of Bifurcation and Chaos*, 14 (8), 2991-2997, 2004.
- [8] P.C. Chu, L.M. Ivanov, O.V. Melnichenko, and N.C. Wells, "On long baroclinic Rossby waves in the tropical North Atlantic observed from profiling floats," *Journal of Geophysical Research*, 112, C05032, doi:10.1029/2006JC003698, 2007.