

Articles

Home > ARTICLES >

ANALYZING TROPICAL WAVES USING THE PARALLEL ENSEMBLE EMPIRICAL MODEL DECOMPOSITION METHOD: PRELIMINARY RESULTS FROM HURRICANE SANDY

Analyzing Tropical Waves Using the Parallel Ensemble Empirical Model Decomposition Method: Preliminary Results from Hurricane Sandy

Earthzine / December 2, 2013 / [ESTO Showcase 2013](#), [Original](#)

Bo-Wen Shen^{1,2}, Samson Chueng³, Jui-Lin F. Li⁴ and Yu-ling Wu⁵

¹UMCP/ESSIC, ²NASA/GSFC, ³NASA/ARC, ⁴Caltech/JPL

¹Earth System Science Interdisciplinary Center

Abstract

In this study, we discuss the performance of the parallel ensemble empirical mode decomposition (EMD) in the analysis of tropical waves that are associated with tropical cyclone (TC) formation. To efficiently analyze high-resolution, global, multiple-dimensional data sets, we first implement multi-level parallelism into the ensemble EMD (EEMD) and obtain a parallel speedup of 720 using 200 eight-core processors. We then apply the parallel EEMD (PEEMD) to extract the intrinsic mode functions (IMFs) from preselected data sets that represent (1) idealized tropical waves and (2) large-scale environmental flows associated with Hurricane Sandy (2012). Results indicate that the PEEMD is efficient and effective in revealing the major wave characteristics of the data, such as wavelengths and periods, by sifting out the dominant (wave) components. This approach has a potential for hurricane climate study by examining the statistical relationship between tropical waves and TC formation.

1 Introduction

Recent studies using high-resolution model simulations and satellite data have shown a potential for improving our understanding of TC formation and thus extending the lead time of TC genesis prediction. In a series of papers with a high-resolution global model, Shen et al. (2010a,b; 2012a) have conducted case studies to examine the relationship between TC formation and different kinds of tropical waves, including (i) TC Nargis (2008) and its association with an Equatorial Rossby (ER) wave; (ii) Hurricane Helene (2006) and its association with an

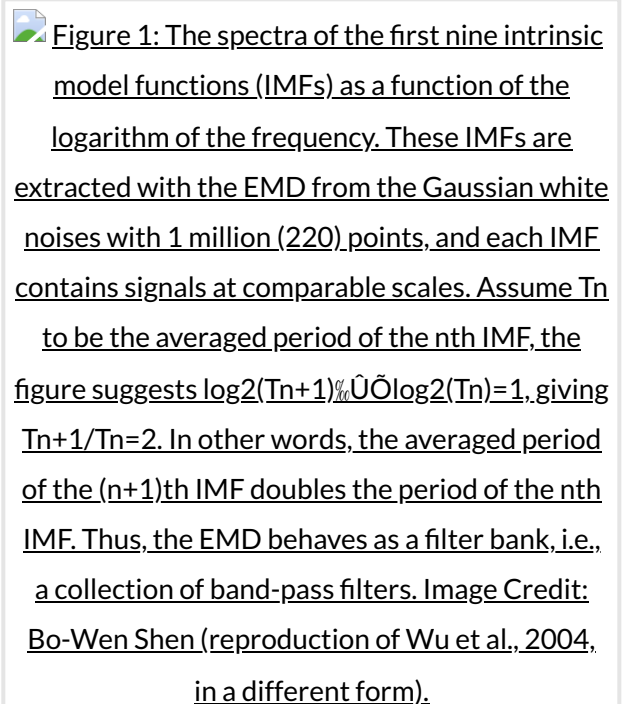
Figure 1: The spectra of the first nine intrinsic model functions (IMFs) as a function of the logarithm of the frequency. These IMFs are extracted with the EMD from the Gaussian white noises with 1 million (220) points, and each IMF contains signals at comparable scales. Assume T_n to be the averaged period of the n th IMF, the figure suggests $\log_2(T_{n+1})/\log_2(T_n)=1$, giving $T_{n+1}/T_n=2$. In other words, the averaged period of the $(n+1)$ th IMF doubles the period of the n th IMF. Thus, the EMD behaves as a filter bank, i.e., a collection of band-pass filters. Image Credit: Bo-Wen Shen (reproduction of Wu et al., 2004, in a different form).

FIGURE 1: THE SPECTRA OF THE FIRST NINE INTRINSIC MODEL FUNCTIONS (IMFS) AS A FUNCTION OF THE LOGARITHM OF THE FREQUENCY. THESE IMFS ARE EXTRACTED WITH THE EMD FROM THE GAUSSIAN WHITE NOISES WITH 1 MILLION (220) POINTS, AND EACH IMF CONTAINS SIGNALS AT COMPARABLE SCALES. ASSUME T_N TO BE THE AVERAGED PERIOD OF THE N TH IMF, THE FIGURE SUGGESTS $\log_2(T_{N+1})/\log_2(T_N)=1$, GIVING $T_{N+1}/T_N=2$. IN OTHER WORDS, THE AVERAGED PERIOD OF THE $(N+1)$ TH IMF DOUBLES THE PERIOD OF THE N TH IMF. THUS, THE EMD BEHAVES AS A FILTER BANK, I.E., A COLLECTION OF BAND-PASS FILTERS. IMAGE CREDIT: BO-WEN SHEN (REPRODUCTION OF WU ET AL., 2004, IN A DIFFERENT FORM).

intensifying African easterly wave; and (iii) Twin TCs (2002) and their association with a mixed Rossby-gravity (MRG) wave embedded in organized large-scale convective systems. Multiscale processes and their interactions that could lead to the formation of the aforementioned TCs are illustrated with advanced scientific 4-D (X-Y-Z-time) visualizations by Shen et al. (2013a). Based on these studies, it has been hypothesized that improved predictability of TC formation can be achieved by refining hierarchical scale interactions of a TC and its environmental flows, such as different types of tropical waves.

To verify the above hypothesis comprehensively, it is crucial to perform multiscale analysis to identify the spatial and/or temporal distributions of tropical waves and their correlation with TC activities. The former is illustrated with the case of Hurricane Sandy in this study. The original EMD (Huang et al., 1998) or EEMD (Wu et al., 2009) was selected to perform the multiscale analysis as the involved processes are non-stationary and nonlinear. The EMD decomposes one set of observational data into the so-called intrinsic mode functions (IMFs). The EEMD was originally developed to overcome the scale (or mode) mixing problem that may appear in the original EMD. The EEMD deals with an ensemble of data sets, each of which contains the original observational data and finite amplitude white noise; one set of IMFs is obtained for each ensemble member with the EMD, and an ensemble average is applied to obtain the final IMFs. It has been shown that the decomposed mean IMFs using EEMD stay within the natural filter period windows, significantly reducing the chance of scale mixing while still preserving a filter bank property (e.g., Flandrin et al., 2003; We et al., 2004). Depending on the required level of accuracy in the decomposed IMFs, typical number of ensemble members used in EEMD is in the range of 200-400 or higher, tremendously increasing computational demands. Therefore, to reduce wall-time to solution, it becomes important to improve the performance of the EEMD algorithm.

Recently, a hybrid multi-level parallelism has been implemented in the EEMD (Shen et al., 2012b; Cheung et al., 2013). In section 2.1, we briefly discuss the design of the parallelism in the parallel EEMD (PEEMD). In section 2.2, we first present the hierarchical filtering properties of the EMD, and then illustrate the performance of the PEEMD in the analysis of idealized tropical wave solutions. The latter quantitatively examines the differences between the specific IMF(s) and the corresponding analytical component(s) in the original data. In section 2.3, we apply the PEEMD to analyze the multiscale environmental flows that were associated with the formation and initial movement of Hurricane Sandy (2012).

2. Parallel Ensemble EMD and its Applications

2.1. Design of Parallelism in the PEEMD

The EMD, first proposed by Huang et al. (1998) at NASA/GFSC in late 1990s, extracts a finite number of the so-called IMFs from the local extrema of the data. Considering the IMFs as basis functions that represent the original data, the EMD is adaptive. In contrast, the Fourier transform and wavelet analysis as well are not adaptive because they are based on an a priori basis. Due to the complexity of real world data that reside at a wide range of time (or spatial) scales, it has been challenging for a basis, which is selected a priori, to capture the essential components that reflect the underlying physical processes in the data. As a result of the adaptive and local features, the EMD has the advantages of dealing with non-stationary and nonlinear data.

The ensemble EMD (EEMD) method was developed to overcome the mode-mixing issues of the EMD (Wu et al., 2009). This method requires large computational resources that are linearly proportional to the number of ensemble trials. It becomes even more challenging to extend the one-dimensional EEMD to multidimensional EEMD for multiple fields at different vertical levels. Thus, a parallel version is proposed to reduce the time to obtain solutions. In our parallel EEMD (PEEMD), a three-level parallelism has been implemented with MPI tasks for the first two levels and with OpenMP threads in the third level. The first level parallelism decomposes a global (or limited-area) domain into M different sub-domains, running with M processes in parallel, while the second level decomposes the total number of ensemble trials into N groups, running with N processes in parallel. The third level parallelism is achieved with the fine-grain OpenMP inside each of the N processes in each of the M tasks. Major features of the newly-developed PEEMD include: (1) promising scalability that yields a speedup of up to 720 using 200 eight-core processors; (2) bit-by-bit consistency that assures the binary identical results with different CPU layouts; and (3) sustainability enabled by a well-designed interface, making it easy to plug in new versions of the EMD. More details in the parallel implementation are given by Cheung et al (2013). In the following sections, we discuss the performance of the PEEMD in the analysis of idealized wave solutions and environmental flows associated with Hurricane Sandy.

2.2 Hierarchical Scale Decompositions of Idealized Tropical Wave Solutions

Following the discussions of Flandrin et al. (2003) and We et al. (2004), we first illustrate the unique feature of the EMD that behaves as a filter bank, i.e., a collection of band-pass filters. To achieve this, we apply the EMD to decompose Gaussian white noises with 1 million (2^{20}) points into 9 IMFs, and perform Fourier analysis to obtain the spectrum for each of the IMFs (Figure 1). The spectrum of each IMF displays a Gaussian distribution, containing signals within a range of frequencies. In general, a higher order IMF has lower frequencies. Assume f_n (T_n) to be the averaged frequency (period) of the n th IMF, Figure 1 suggests an approximate relation: $\log_2(f_n) - \log_2(f_{n+1}) = 1$, where $n=2-8$. This relation is equivalent to the following form: $\log_2(T_{n+1}) - \log_2(T_n) = 1$, which gives $T_{n+1}/T_n = 2$. In other words, the averaged period of the $(n+1)$ th IMF doubles the period of n th IMF. This feature is similar to that of a dyadic filter. Numerical experiments suggest that such a feature can be clearly shown with sufficient data points. However, a typical set of observation or model data may not have so many points to clearly display the feature of the dyadic filter (with doubling periods). Nonetheless, without loss of generality and accuracy, it can be said that the EMD performs a special kind of hierarchical filtering. This feature is ideal for multiscale analysis.

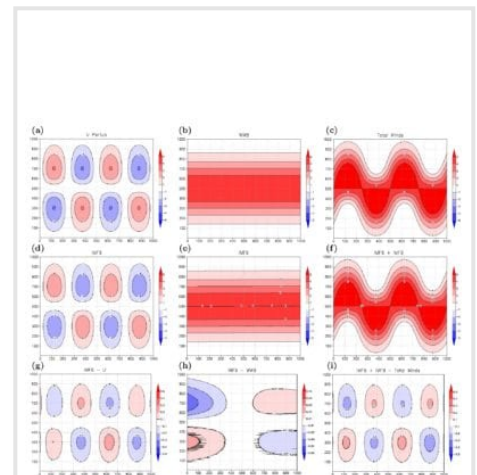


FIGURE 2: DECOMPOSITIONS OF IDEALIZED TROPICAL WAVE SOLUTIONS. (A, B, C) THE ZONAL WIND PERTURBATIONS (U'), THE ZONAL BASIC WINDS, AND THE TOTAL WINDS, RESPECTIVELY (REPRODUCED FROM SHEN ET AL., 2012A). THE FIRST FIELD IS FROM AN ANALYTICAL MIXED ROSSBY-GRAVITY WAVE SOLUTION, WHILE THE SECOND FIELD IS THE SOLUTION OF AN IDEALIZED WEST WIND BELT (WWB). THEIR SUMMATION ($U'+WWB$) REPRESENTS THE TOTAL WINDS. (D, E) THE SIXTH AND NINTH IMFS EXTRACTED FROM THE TOTAL WINDS, SEQUENTIALLY. (F) THE SUMMATION OF THE SIXTH AND NINTH IMFS (I.E., IMF_6+IMF_9). (G, H, I) THE DIFFERENCES BETWEEN IMFS AND CORRESPONDING IDEALIZED COMPONENTS, WHICH ARE (IMF_6-U'), (IMF_9-WWB) AND ($IMF_6+IMF_9-U'+WWB$), RESPECTIVELY. NOTE THAT THE

To verify the performance of the PEEMD in decomposing meteorological data into different wave modes, we begin with analyzing idealized wave solutions, which consist of an analytical solution of a mixed Rossby-gravity (MRG) wave as wind perturbations (U'), and an analytical form of a west wind belt (WWB) as basic winds. The

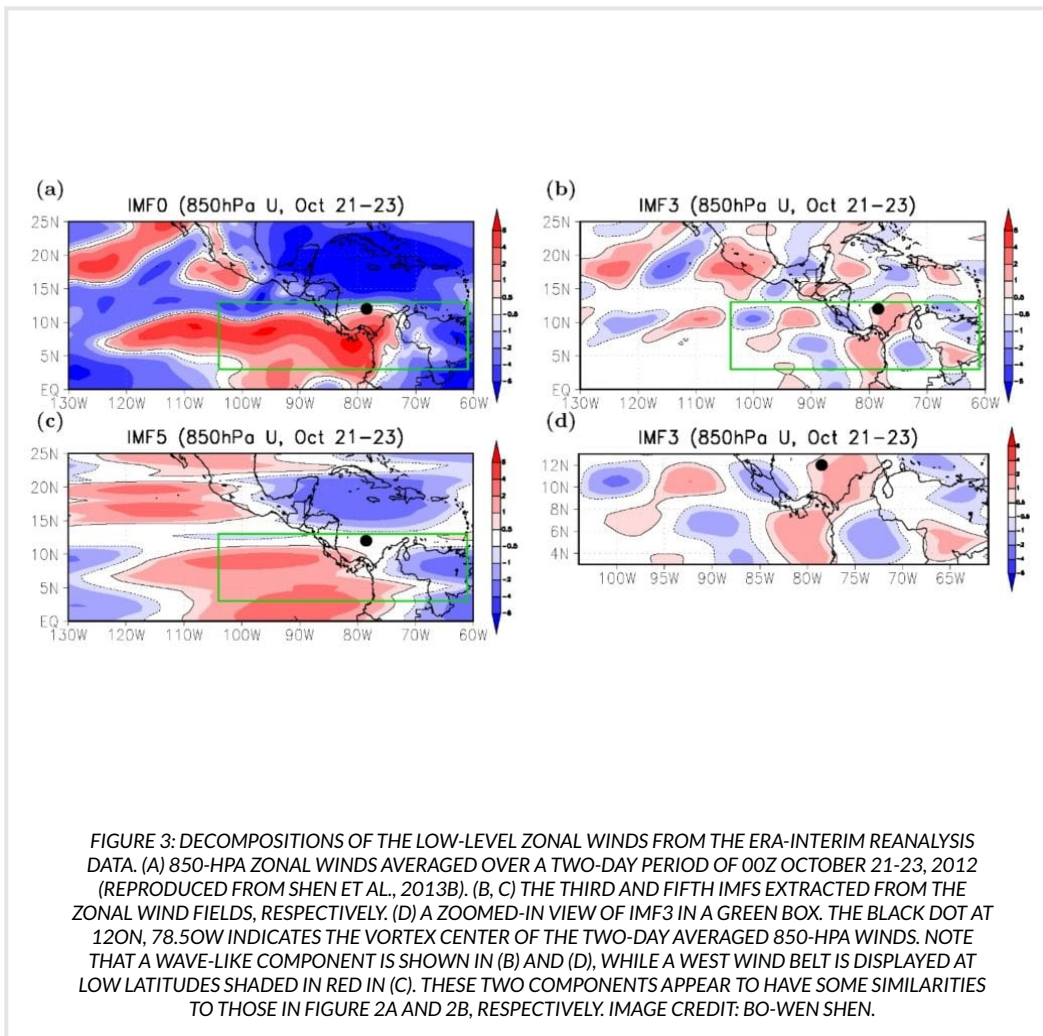
COLOR SCALES IN THE BOTTOM PANELS ARE DIFFERENT FROM THOSE IN THE TOP AND MIDDLE PANELS. THE SIXTH IMF REPRESENTS MORE THAN 90 PERCENT OF THE ZONAL WIND PERTURBATIONS (A,G), WHILE THE NINTH IMF REPRESENTS MORE THAN 99 PERCENT OF THE ZONAL BASIC WINDS (WWB) (B, H). IMAGE CREDIT: BO-WEN SHEN.

summation of the two fields produces the total winds, i.e., $U'+WWB$. These wind fields are shown in the top panels of Figure 2 (see also Shen et al., 2012a). The total winds are then decomposed into nine IMFs. Only the sixth IMF, the ninth IMF and their summation are displayed in the middle panels of Figure 2. It is clear that the sixth IMF (Figure 2d) and ninth IMF (Figure 2e) resemble the MRG wave mode (Figure 2a) and WWB (Figure 2b), respectively. The deviations of a specific IMF from the original component (shown in the bottom panels) suggest that the sixth IMF can capture more than 90 percent of the original MRG wave mode, while the ninth IMF can represent more than 99 percent of the original WWB component. Encouraged by these promising decompositions, we use the PEEMD to analyze the environmental flows associated with Hurricane Sandy.

2.3 Multiscale Processes associated with Hurricane Sandy

In late October 2012, Storm Sandy made landfall near Brigantine, New Jersey, devastating surrounding areas and causing tremendous economic loss and hundreds of fatalities (e.g., Shen et al. 2013b and references therein). An estimated damage of \$50 billion made Sandy the second costliest TC in US history, surpassed only by Hurricane Katrina (2005). A central question to be addressed is to what extent the lead time of severe storm prediction such as Sandy can be extended (e.g., Emanuel 2012; Kerr 2012). In the study of Shen et al., (2013b), we demonstrated our model's capability to realistically predict Sandy's genesis with a lead time of up to six days and subsequent evolution for the next two-day period of October 22-24. The multiscale processes that may have contributed to the extended lead time are analyzed with the PEEMD below.

Figure 3 shows the ECMWF-Interim zonal winds (Dee et al., 2011) averaged over a two-day period of 00Z October 21-23 (Figure 3a), and the selected IMFs extracted from the zonal winds. While the third IMF reveals a dominant wave mode (Figure 3b), the fifth IMF represents a WWB at low latitudes (Figure 3c). Note that the spatial distributions of the zoomed-in third IMF (Figure 3d) appears to have some similarities to those in Figure 2a. Figure 4 shows the time longitude diagram of the meridional winds at 200 hPa and the selected IMFs, which are used to determine the wavelentgh and phase speed of the wave-like disturbance. Among the IMFs, the fourth IMF (Figure 4c) displays a more visible wave mode than the original data (Figure 4a). This wave mode has an estimated wavelentgh of 45 degrees (~4535.1 km). Plug this wavelength into theoretical wave dispersion relations (e.g., Shen et al. 2013b and references therein), we obtain an estimated period of 4.914 and 7.407 days for an MRG and ER wave, respectively Therefore, analyses in Figures 3-4 provide the insightful understanding of the following multiscale processes associated with the formation and initial evolution of Sandy: (1) evolution of the (low-level) WWB and its interaction with the easterly wave (e.g., Figure 3c); and (2) the location of an upper-level trough, which is indicated by the white areas between 85 and 80 degrees west longitude and appeared in association with the MRG and ER waves (Figure 4c).

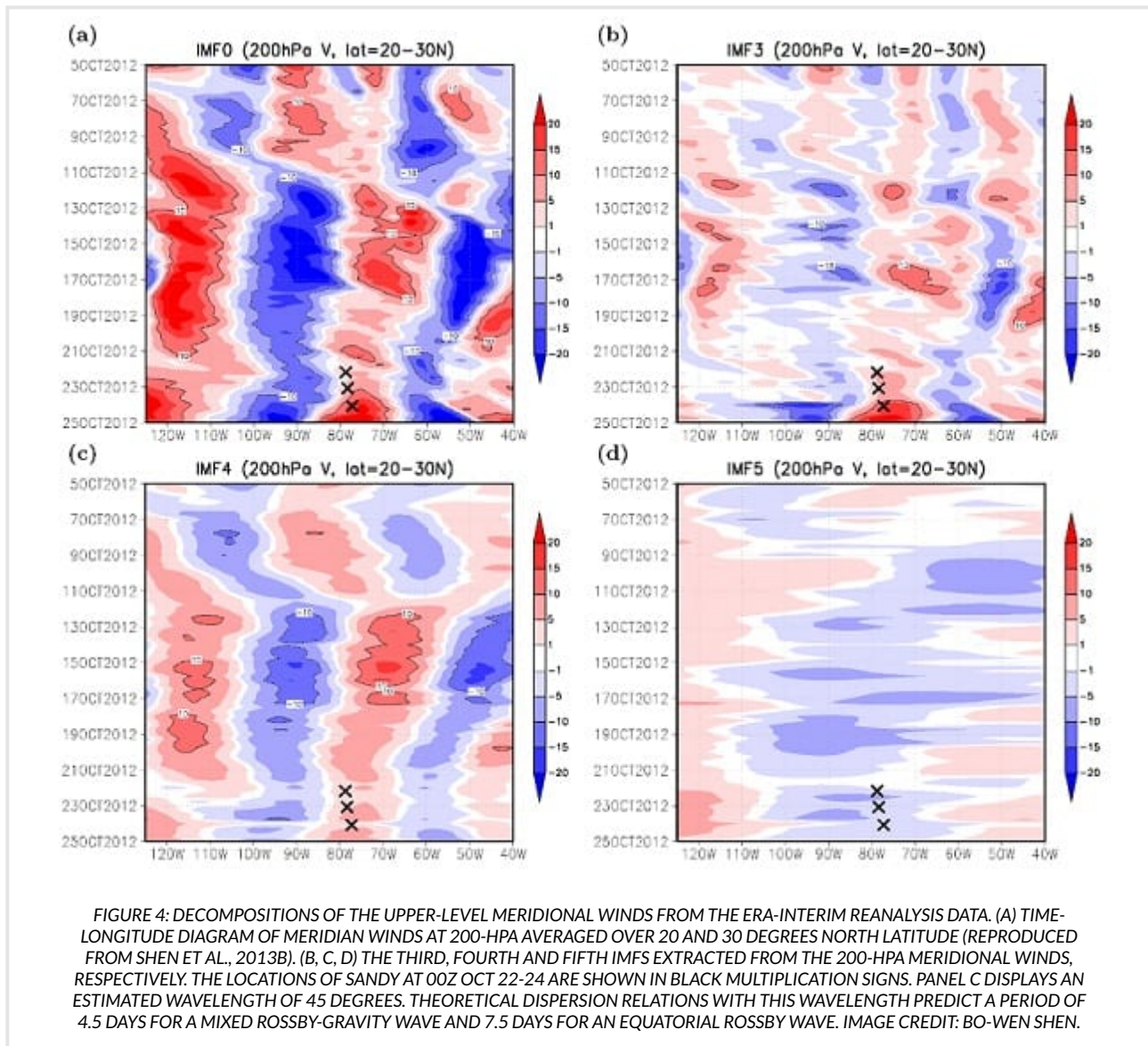


3. Concluding Remarks

To efficiently and effectively reveal multiscale processes from high-resolution, global, multi-dimensional Earth Science data, we have developed the parallel version of the ensemble empirical model decomposition (PEEMD) method with a three-level parallelism. In this study, we illustrated the performance of the PEEMD in extracting tropical wave components from two different data sets: (1) idealized tropical wave solutions and (2) large-scale environmental flows associated with Sandy. Compared to conventional methods, our analyses with the PEEMD provide better representations of the dominant modes (e.g., MRG wave and WWB) associated with Sandy's formation. Spatial and temporal correlation between tropical waves and hurricane formation are still under investigation. Our ultimate goal is to integrate the PEEMD with other modules (e.g., schemes of spectral analysis, readers for different satellite data, and numerical models) into the multiscale analysis package for hurricane climate study.

Acknowledgments:

We are grateful to the following organizations for their support: the NASA Earth Science Technology Office, the Advanced Information Systems Technology Program, and the NASA Computational Modeling Algorithms and Cyberinfrastructure program. *Resources supporting this work were provided by the NASA High-End Computing*



References:

Cheung, S., B.-W. Shen, P. Mehrotra, J.-L. F. Li, 2013: Parallelization of the Ensemble Empirical Model Decomposition (PEEMD) Method on Multi- and Many-core Processors. AGU 2013 Fall Meeting, San Francisco, December 9-13.

Dee, D.P., et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, 137: 553–597.

Flandrin, P.; Rilling, G.; Gonçalves, P., 2003: "Empirical Mode Decomposition as a Filterbank" (pdf). *IEEE Signal Processing Letters* 11 (2): 112-114. doi:10.1109/LSP.2003.821662.

Huang, et al., 1998: "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." *Proc. R. Soc. Lond. A* (1998) 454, 903-995.

Kerr, R., 2012: One Sandy Forecast a Bigger Winner Than Others. *Science* 9 Nov. 2012: 338 pp. 736-737.

Emanuel, K., 2012: Why America Has Fallen Behind the World in Storm Forecasting. *The Wall Street Journal*. October 28 2012.

Shen, B.-W. et al., 2010a: Predicting Tropical Cyclogenesis with a Global Mesoscale Model: Hierarchical Multiscale Interactions During the Formation of Tropical Cyclone Nargis (2008). *J. Geophys. Res.*, 115, D14102.

Shen, B.-W. et al., 2010b: African Easterly Waves in 30-day High-resolution Global Simulations: A Case Study during the 2006 NAMMA Period. *Geophys. Res. Lett.*, L18803.

Shen, B.-W. et al., 2012a: Genesis of Twin Tropical Cyclones Revealed by a Global Mesoscale Model: the Role of Mixed Rossby Gravity Wave. *J. Geophys. Res.*, 117, D13114.

Shen, B.-W., Z. Wu, and S. Cheung, 2012b: Analysis of Tropical Cyclones and Tropical Waves using the Parallel Ensemble Empirical Model Decomposition (EEMD) Method. AGU 2012 Fall Meeting, San Francisco, December 03-07, 2012.

Shen, B.-W. et al., 2013a, "Advanced Visualizations of Scale Interactions of Tropical Cyclone Formation and Tropical Waves," *Computing in Science and Engineering*, vol. 15, no. 2, pp. 47-59, March-April 2013, doi:10.1109/MCSE.2012.64

Shen, B.W., M. DeMaria, J.-L. F. Li, and S. Cheung, 2013b: Genesis of Hurricane Sandy (2012) Simulated with a Global Mesoscale Model. *Geophys. Res. Lett.* 40. DOI: 10.1002/grl.50934.

Wu, Z. and N. E. Huang, 2004: Statistical significance test of intrinsic mode functions. In *Hilbert-Huang Transform: Introduction and Applications*, pp 125-148, Ed. N. E. Huang and S. S. P. Shen, *World Scientific, Singapore*, 311pp.

Wu, Z., and N. E. Huang, 2009: Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method. *Advances in Adaptive Data Analysis*. 1, 1-41.