A REVIEW ON HILBERT-HUANG TRANSFORM: METHOD AND ITS APPLICATIONS TO GEOPHYSICAL STUDIES

Norden E. Huang[1] and Zhaohua Wu[2]

Received 10 May 2007; accepted 22 October 2007; published 6 June 2008.

[1] Data analysis has been one of the core activities in scientific research, but limited by the availability of analysis methods in the past, data analysis was often relegated to data processing. To accommodate the variety of data generated by nonlinear and nonstationary processes in nature, the analysis method would have to be adaptive. Hilbert-Huang transform, consisting of empirical mode decomposition and Hilbert spectral analysis, is a newly developed adaptive data analysis method, which has been used extensively in geophysical research. In this review, we will briefly introduce the method, list some recent developments, demonstrate the usefulness of the method, summarize some applications in various geophysical research areas, and finally, discuss the outstanding open problems. We hope this review will serve as an introduction of the method for those new to the concepts, as well as a summary of the present frontiers of its applications for experienced research scientists.

Citation: Huang, N. E., and Z. Wu (2008), A review on Hilbert-Huang transform: Method and its applications to geophysical studies, *Rev. Geophys.*, *46*, RG2006, doi:10.1029/2007RG000228.

1. INTRODUCTION

[2] Data are the only link we have with the unexplained reality; therefore, data analysis is the only way through which we can find out the underlying processes of any given phenomenon. As the most important goal of scientific research is to understand nature, data analysis is a critical link in the scientific research cycle of observation, analysis, synthesizing, and theorizing. Because of the limitations of available methodologies for analyzing data, the crucial phase of data analysis has in the past been relegated to "data processing," where data are routinely put through certain well-established algorithms to extract some standard parameters. Most traditional data processing methodologies are developed under rigorous mathematic rules; and we pay a price for this strict adherence to mathematical rigor, as described by *Einstein* [1983, p. 28]: "As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality." As a result, data processing has never received the deserved attention as data analysis should, and data processing has never fulfilled its full potential of revealing the truth and extracting the coherence hidden in scattered numbers.

[3] This is precisely the dilemma we face: in order not to deviate from mathematical rigor, we are forced to live in a pseudoreality world, in which every process is either linear or stationary and for most cases both linear and stationary. Traditional data analysis needs these conditions to work; most of the statistical measures have validity only under these restrictive conditions. For example, spectral analysis is synonymous with Fourier-based analysis. As the Fourier spectra can only give meaningful interpretation to linear and stationary processes, its application to data from nonlinear and nonstationary processes is often problematical. And probability distributions can only represent global properties, which imply homogeneity (or stationarity) in the population. The real world is neither linear nor stationary; thus the inadequacy of the linear and stationary data analysis methods that strictly adhere to mathematical rigor is becoming glaringly obvious.

[4] To better understand the physical mechanisms hidden in data, the dual complications of nonstationarity and nonlinearity should be properly dealt with. A more suitable approach to revealing nonlinearity and nonstationarity in data is to let the data speak for themselves and not to let the analyzer impose irrelevant mathematical rules; that is, the method of analysis should be adaptive to the nature of the data. For the methods that involve decomposition of data, the adaptive requirement calls for an adaptive basis. Here adaptivity means that the definition of the basis has to be based on and derived from the data. Unfortunately, most

[1]Research Center for Adaptive Data Analysis, National Central University, Chungli, Taiwan.
[2]Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland, USA.

8755-1209/08/2007RG000228$15.00
Paper number 2007RG000228

currently available data decomposition methods have an a priori basis (such as trigonometric functions in Fourier analysis), and they are not adaptive. Once the basis is determined, the analysis is reduced to a convolution computation. This well-established paradigm is mathematically sound and rigorous. However, the ultimate goal for data analysis is not to find the mathematical properties of data; rather, it is to unearth the physical insights and implications hidden in the data. There is no a priori reason to believe that a basis arbitrarily selected is able to represent the variety of underlying physical processes. Therefore, the results produced, though mathematically correct, might not be informative.

[5] A few adaptive methods are available for signal analysis, as summarized by *Widrow and Stearns* [1985], which relies mostly on feedback loops but not on an adaptive basis. Such methods are, however, designed for stationary processes. For nonstationary and nonlinear data, adaptivity is absolutely necessary; unfortunately, few effective methods are available. How should the general topic of an adaptive method for data analysis be approached? How do we define adaptive bases? What are the mathematical properties and problems of these basis functions? An a posteriori adaptive basis provides a totally different approach from the established mathematical paradigm, and it may also present a great challenge to the mathematical community.

[6] The combination of the well-known Hilbert spectral analysis (HAS) and the recently developed empirical mode decomposition (EMD) [*Huang et al.*, 1996, 1998, 1999], designated as the Hilbert-Huang transform (HHT) by NASA, indeed, represents such a paradigm shift of data analysis methodology. The HHT is designed specifically for analyzing nonlinear and nonstationary data. The key part of HHT is EMD with which any complicated data set can be decomposed into a finite and often small number of intrinsic mode functions (IMFs). The instantaneous frequency defined using the Hilbert transform denotes the physical meaning of local phase change better for IMFs than for any other non-IMF time series. This decomposition method is adaptive and therefore highly efficient. As the decomposition is based on the local characteristics of the data, it is applicable to nonlinear and nonstationary processes.

[7] This new empirical method offers a potentially viable method for nonlinear and nonstationary data analysis, especially for time-frequency-energy representations. It has been tested and validated exhaustively [*Huang and Attoh-Okine*, 2005; *Huang and Shen*, 2005], though only empirically. In almost all the cases studied, HHT gives results much sharper than most of the traditional analysis methods. Additionally, it has been demonstrated to have unprecedented prowess in revealing hidden physical meanings in data. In this article, we will review the method and its recent developments and applications to geophysical sciences, as well as the directions for future developments.

2. A BRIEF DESCRIPTION OF HHT

[8] The HHT consists of empirical mode decomposition and Hilbert spectral analysis [*Huang et al.*, 1996, 1998, 1999]. In this section, we will introduce briefly both components of HHT and present some properties of HHT. It will be shown that the Hilbert transform (HT) can lead to an apparent time-frequency-energy description of a time series; however, this description may not be consistent with physically meaningful definitions of instantaneous frequency and instantaneous amplitude. The EMD can generate components of the time series whose Hilbert transform can lead to physically meaningful definitions of these two instantaneous quantities, and hence the combination of HT and EMD provides a more physically meaningful time-frequency-energy description of a time series.

2.1. Hilbert Spectral Analysis

[9] As emphasized in section 1, the purpose of the development of HHT is to provide an alternative view of the time-frequency-energy paradigm of data. In this approach, the nonlinearity and nonstationarity can be dealt with better than by using the traditional paradigm of constant frequency and amplitude. One way to express the nonstationarity is to find instantaneous frequency and instantaneous amplitude. This was the reason why Hilbert spectrum analysis was included as a part of HHT.

[10] For any function $x(t)$ of L^p class, its Hilbert transform $y(t)$ is

$$y(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau, \tag{1}$$

where P is the Cauchy principal value of the singular integral. With the Hilbert transform $y(t)$ of the function $x(t)$, we obtain the analytic function,

$$z(t) = x(t) + iy(t) = a(t)e^{i\theta(t)}, \tag{2}$$

where $i = \sqrt{-1}$,

$$a(t) = \left(x^2 + y^2\right)^{1/2}, \qquad \theta(t) = \tan^{-1}\frac{y}{x}. \tag{3}$$

Here a is the instantaneous amplitude, and θ is the instantaneous phase function. The instantaneous frequency is simply

$$\omega = \frac{d\theta}{dt}. \tag{4}$$

[11] With both amplitude and frequency being a function of time, we can express the amplitude (or energy, the square of amplitude) in terms of a function of time and frequency, $H(\omega, t)$. The marginal spectrum can then be defined as

$$h(\omega) = \int_0^T H(\omega, t)dt, \tag{5}$$

where $[0, T]$ is the temporal domain within which the data is defined. The marginal spectrum represents the accumulated amplitude (energy) over the entire data span in a

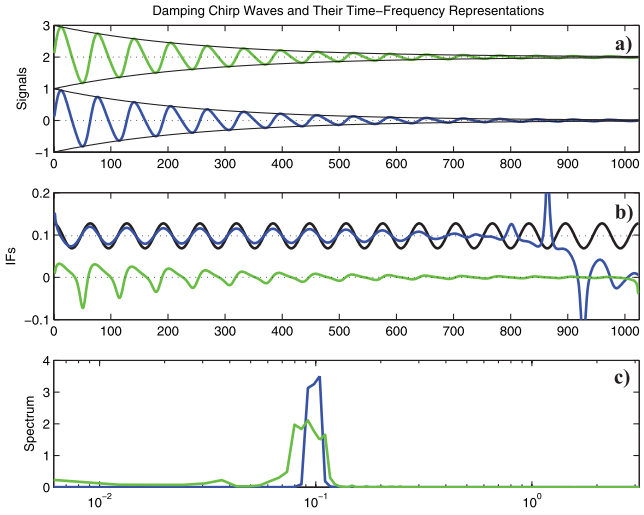

Figure 1. Instantaneous frequencies, marginal and Fourier spectra of a simple oscillatory function with a zero mean or a nonzero mean. (a) Bold blue and green lines show a simple oscillatory function specified in equation (6) riding on a zero mean and a nonzero mean with a value of 2, respectively. The four thin black lines are the upper and lower envelopes, respectively. (b) True instantaneous frequency specified in equation (6) is given by the black line. The Hilbert transform–based instantaneous frequencies of the simple oscillatory function riding on a zero mean and a nonzero mean are displayed by the blue and green lines, respectively. (c) Fourier spectrum (blue line) and the marginal spectrum (green line) of the simple oscillatory function are displayed.

probabilistic sense and offers a measure of the total amplitude (or energy) contribution from each frequency value, serving as an alternative spectrum expression of the data to the traditional Fourier spectrum. A perfect IMF,

$$x(t) = e^{-t/256}\sin\left[\frac{\pi t}{32} + 0.3\sin\left(\frac{\pi t}{32}\right)\right], \tag{6}$$

is displayed in Figure 1, as well as its true and Hilbert transform–based instantaneous frequency and its Fourier spectrum and marginal spectrum.

[12] The term on the right-hand side of equation (2) provides an apparent time-frequency-energy expression for function $x(t)$, such as the one given by equation (6). However, for an arbitrary $x(t)$, the instantaneous frequency obtained using the above method is not necessarily physically meaningful. For example, the instantaneous frequency method applied to a sinusoidal function of a constant frequency riding on a nonzero reference level (e.g., $\cos\omega t + C$, where C is a constant) does not yield a constant frequency of ω; rather, the obtained frequency bears unexpected fluctuations [*Huang et al.*, 1998]. An example of such distortion associated with Hilbert transform–based instantaneous frequency is also displayed in Figure 1, in which the signal is specified by equation (6) but with a constant mean of 2.0. Clearly, the Hilbert transform–based instantaneous frequency contains unphysical negative values and is qualitatively different from that of the riding chirp wave at any temporal location, indicating the limited applicability of the Hilbert transform. To explore the applicability of the Hilbert transform, *Huang et al.* [1998] showed that a purely oscillatory function (or a monocomponent) with a zero reference level is a necessary condition for the above instantaneous frequency calculation method to work appropriately [*Huang et al.*, 1998]. Indeed, searching for the expression of an arbitrary $x(t)$ in terms of a sum of a small number of purely oscillatory functions of which Hilbert transform–based instantaneous frequencies are physically meaningful was the exact motivation for the early development of EMD [*Huang et al.*, 1998].

2.2. Empirical Mode Decomposition

[13] The EMD, in contrast to almost all the previous methods, works in temporal space directly rather than in the corresponding frequency space; it is intuitive, direct, and adaptive, with an a posteriori defined basis derived from the data. The decomposition has implicitly a simple assumption that, at any given time, the data may have many coexisting simple oscillatory modes of significantly different frequencies, one superimposed on the other. Each component is defined as an intrinsic mode function (IMF) satisfying the following conditions: (1) In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one. (2) At any data point, the mean value of the envelope defined using the local maxima and the envelope defined using the local minima is zero.

[14] With the above definition of an IMF in mind, one can then decompose any function through a sifting process. An example of obtaining an IMF from an arbitrarily given time series is displayed in Figure 2. The input, $x(t)$, is displayed as the bold solid line in Figures 2a and 2b. The sifting starts with identifying all the local extrema (see Figure 2b, with maxima marked with diamonds and minima marked with circles) and then connecting all the local maxima (minima) by a cubic spline line to form the upper (lower) envelope as shown by the thin solid lines in Figure 2c. The upper and lower envelopes usually encompass all the data between them as shown in Figure 2c. Their mean is designated as m_1, the dashed line in Figure 2c. The difference between the input and m_1 is the first protomode, h_1, shown in Figure 2d, i.e.,

$$h_1 = x(t) - m_1. \tag{7}$$

[15] By construction, h_1 is expected to satisfy the definition of an IMF. However, that is usually not the case since changing a local zero from a rectangular to a curvilinear coordinate system may introduce new extrema, and further adjustments are needed. Therefore, a repeat of the above procedure, the sifting, is necessary. This sifting process serves two purposes: (1) to eliminate background waves on which the IMF is riding and (2) to make the wave profiles more symmetric. The sifting process has to be repeated as many times as is required to make the extracted signal satisfy the definition of an IMF. In the iterating

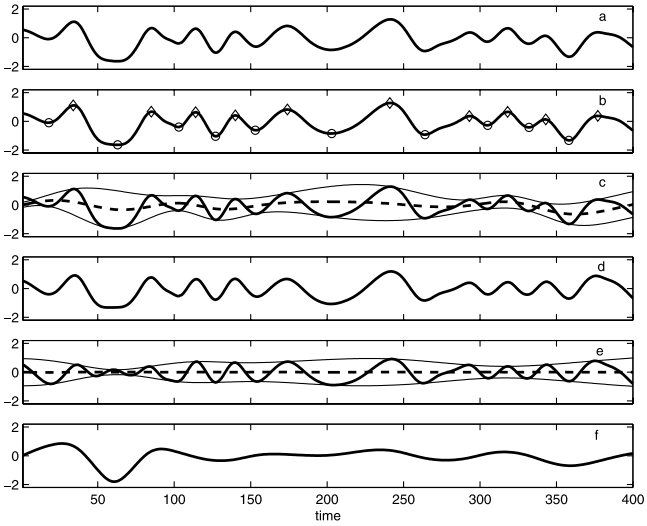

Figure 2. Sifting process of the empirical mode decomposition: (a) an arbitrary input; (b) identified maxima (diamonds) and minima (circles) superimposed on the input; (c) upper envelope and lower envelope (thin solid lines) and their mean (dashed line); (d) prototype intrinsic mode function (IMF) (the difference between the bold solid line and the dashed line in Figure 2c) that is to be refined; (e) upper envelope and lower envelope (thin solid lines) and their mean (dashed line) of a refined IMF; and (f) remainder after an IMF is subtracted from the input.

processes, h_1 can only be treated as a proto-IMF, which is treated as the data in the next iteration:

$$h_1 - m_{11} = h_{11}. \tag{8}$$

After k times of iterations,

$$h_{1(k-1)} - m_{1k} = h_{1k}; \tag{9}$$

the approximate local envelope symmetry condition is satisfied, and h_{1k} becomes the IMF c_1, i.e.,

$$c_1 = h_{1k}, \tag{10}$$

which is the first IMF component shown in Figure 2e.

[16] The approximate local envelope symmetry condition in the sifting process is called the stoppage (of sifting) criterion. In the past, several different types of stoppage criteria were adopted: the most widely used type, which originated from *Huang et al.* [1998], is given by a Cauchy type of convergence test, the normalized squared difference between two successive sifting operations defined as

$$\mathrm{SD}_k = \frac{\sum_{t=0}^{T} |h_{k-1}(t) - h_k(t)|^2}{\sum_{t=0}^{T} h_{k-1}^2(t)}, \tag{11a}$$

must be smaller than a predetermined value. This definition is slightly different from the one given by *Huang et al.* [1998] with the summation signs operating for the numerator and denominator separately in order to prevent the SD_k from becoming too dependent on local small-amplitude values of the sifting time series. An alternative stoppage criterion of this type is

$$\mathrm{SD}_k = \frac{\sum_{t=0}^{T} |m_{1k}(t)|^2}{\sum_{t=0}^{T} |h_{1k}(t)|^2}, \tag{11b}$$

which is smaller than a predetermined value.

[17] These Cauchy types of stoppage criteria are seemingly rigorous mathematically. However, it is difficult to implement this criterion for the following reasons: First, how small is small enough begs an answer. Second, this criterion does not depend on the definition of the IMFs for the squared difference might be small, but there is no guarantee that the function will have the same numbers of zero crossings and extrema. To remedy these shortcomings, *Huang et al.* [1999, 2003] proposed the second type of criterion, termed the S stoppage. With this type of stoppage criterion, the sifting process stops only after the numbers of zero crossings and extrema are (1) equal or at most differ by one and (2) stay the same for S consecutive times. Extensive tests by *Huang et al.* [2003] suggest that the optimal range for S should be between 3 and 8, but the lower number is favored. Obviously, any selection is ad hoc, and a rigorous justification is needed.

[18] This first IMF should contain the finest scale or the shortest-period oscillation in the signal, which can be extracted from the data by

$$x(t) - c_1 = r_1. \tag{12}$$

[19] The residue, r_1, still contains longer-period variations, as shown in Figure 2d. This residual is then treated as the new data and subjected to the same sifting process as described above to obtain an IMF of lower frequency. The procedure can be repeatedly applied to all subsequent r_j, and the result is

$$\begin{aligned} r_1 - c_2 &= r_2 \\ &\cdots \\ r_{n-1} - c_n &= r_n. \end{aligned} \tag{13}$$

[20] The decomposition process finally stops when the residue, r_n, becomes a monotonic function or a function with only one extremum from which no more IMF can be extracted. By summing up equations (12) and (13), we have

$$x(t) = \sum_{j=1}^{n} c_j + r_n. \tag{14}$$

Thus, the original data are decomposed into n IMFs and a residue obtained, r_n, which can be either the adaptive trend or a constant. An example is given in Figure 3 in which the decomposition of Remote Sensing Systems (RSS) T2, the channel 2 tropospheric temperature of the microwave sounding unit [*Mears et al.*, 2003], is presented.

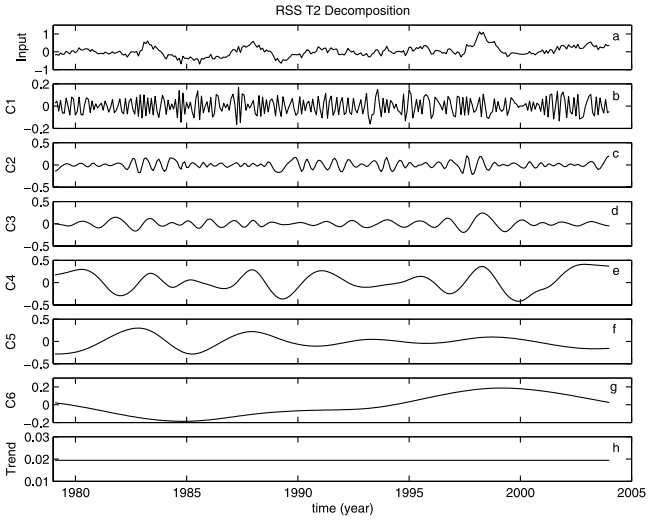

Figure 3. RSS T2 decomposed using the empirical mode decomposition. (a) RSS T2 data are shown. (b–h) IMFs of high to low frequency, respectively, are displayed.

[21] In the EMD method, a constant mean or zero reference is not required since the EMD technique only uses information related to local extrema; and the zero reference for each IMF is generated automatically by the sifting process. Without the need of the zero reference, EMD avoids the troublesome step of removing the trend, which could cause low-frequency terms in the resulting spectra. Therefore, the detrending operation is automatically achieved, an unexpected benefit.

2.3. Some Properties of HHT

[22] The IMFs obtained by sifting processes constitute an adaptive basis. This basis usually satisfies empirically all the major mathematical requirements for a time series decomposition method, including convergence, completeness, orthogonality, and uniqueness, as discussed by *Huang et al.* [1998].

[23] For an arbitrary time series of length N, $x(t)$, if EMD is used and instantaneous frequencies and instantaneous amplitudes of IMFs are obtained, $x(t)$ can be expressed as

$$x(t) = \mathrm{Re}\left[\sum_{j=1}^{n} a_j(t) e^{i\omega_j(t)dt}\right], \tag{15}$$

where Re[ ] represents the real part of terms within brackets. Here the residue, r_n, is not expressed in terns of a simple oscillatory form on purpose for it is either a monotonic function or a function with only one extrema not containing enough information to confirm whether it is an oscillatory component whose frequency is physical meaningful.

[24] Equation (15) gives both amplitude and frequency of each component as functions of time. The same data expanded in a Fourier representation would be

$$x(t) = \mathrm{Re}\sum_{j=1}^{\infty} a_j e^{i\omega_j t}, \tag{16}$$

where both a_j and ω_j are constants. The contrast between equations (15) and (16) is clear: the IMF represents, to a large degree, a generalized Fourier expansion. The variable amplitude and the instantaneous frequency not only improve the efficiency of the expansion but also enable the expansion to accommodate nonlinear and nonstationary variations in data. The IMF expansion lifts the restriction of constant amplitude and fixed frequency in the Fourier expansion, allowing a variable amplitude and frequency representation along the time axis.

[25] As demonstrated by *Huang et al.* [1998, 1999], the distribution of amplitude (energy) in time-frequency domain, $H(\omega, t)$, in HHT can be regarded as a skeleton form of that in continuous wavelet analysis, which has been widely pursued in the wavelet community. For example, *Bacry et al.* [1991] and *Carmona et al.* [1998] tried to extract the wavelet skeleton as the local maximum of the continuous wavelet coefficient. While that approach reduces significantly the "blurredness" in the energy distribution in time-frequency domain caused by redundancy and subharmonics in ideal cases, it is difficult to apply it to complex data, for it is still encumbered by the harmonics, a side effect always associated with an a priori basis. In contrast, the time-frequency representation of data in HHT does not involve spurious harmonics and hence can present more natural and quantitative results.

[26] The HHT provides an alternative, and possibly more physically meaningful, representation of data. The rational for this statement can be justified by its ability to enable an implementation of a novel concept: using instantaneous frequency to describe intrawave frequency modulation. Frequency is traditionally defined as

$$\omega = \frac{1}{T}, \tag{17}$$

where T is the period of an oscillation. Although this definition is almost the standard for frequency, it is very crude for it fails to account for the possible intrawave frequency modulation, a hallmark for nonlinear oscillators. This failure can be demonstrated using the Duffing equation

$$\frac{d^2x}{dt^2} + x + \varepsilon x^3 = \gamma \cos \alpha t, \tag{18}$$

where ε, γ, and α are prescribed constants. The cubic term in equation (18) is the cause of nonlinearity. If we rewrite equation (18) as

$$\frac{d^2x}{dt^2} + \left(1 + \varepsilon x^2\right)x = \gamma \cos \alpha t, \tag{19}$$

clearly, the term within the parentheses could be interpreted as a variable spring constant or varying pendulum length. Either way, the frequency is ever changing even within a single period. This intrawave frequency modulation was traditionally represented by harmonics. As discussed by *Huang et al.* [1998, 2003] and *Huang* [2005a, 2005b], such a harmonic representation is a mathematical artifact with

TABLE 1. Comparison Between Fourier, Wavelet, and HHT Analysis

	Fourier	Wavelet	HHT
Basis	a priori	a priori	a posteriori adaptive
Frequency	convolution over global domain, uncertainty	convolution over global domain, uncertainty	differentiation over local domain, certainty
Presentation	energy in frequency space	energy in time-frequency space	energy in time-frequency space
Nonlinearity	no	no	yes
Nonstationarity	no	yes	yes
Feature extraction	no	discrete, no; continuous, yes	yes
Theoretical base	complete mathematical theory	complete mathematical theory	empirical

little physical meaning; its existence depends on the basis selected. The instantaneous frequency introduced here is physical and depends on the differentiation of the phase function, which is fully capable of describing not only interwave frequency changes due to nonstationarity but also the intrawave frequency modulation due to nonlinearity.

[27] Numerous tests demonstrated empirically that HHT is a powerful tool for time-frequency analysis of nonlinear and nonstationary data. It is based on an adaptive basis, and the frequency is defined through the Hilbert transform. Consequently, there is no need for the spurious harmonics to represent nonlinear waveform deformations as in any methods with an a priori basis, and there is no uncertainty principle limitation on time or frequency resolution from the convolution pairs based on a priori bases. A summary of comparison between Fourier, wavelet, and HHT analyses is given in Table 1.

3. RECENT DEVELOPMENTS

[28] Since the development of the basics of HHT, some new advances have been made in the following areas: (1) improvements in instantaneous frequency calculation methods; (2) determination of the confidence limits of IMFs and improvement in the robustness of the EMD algorithm; (3) creation of statistical significance test of IMFs; (4) development of the ensemble EMD, a noise-assisted data analysis method; and (5) development of two-dimensional EMD methods. The advances in each area are discussed in sections 3.1–3.5, respectively.

3.1. Normalized Hilbert Transform and the Direct Quadrature

[29] For any IMF $x(t)$, one obtains envelope function $a(t)$ and phase function $\theta(t)$ through Hilbert transform of $x(t)$, as expressed by equation (3), i.e., $x(t) = a(t)\cos\theta(t)$. On the basis of the definition of the IMF, it is always true that $a(t)$ contains fluctuations of significantly lower frequency than $\theta(t)$ does at any temporal location. For such a function, the physically meaningful instantaneous frequency should only be determined by $\theta(t)$, which is equivalent to requiring the Hilbert transform to satisfy

$$\mathrm{H}[a(t)\cos\theta(t)] = a(t)\mathrm{H}[\cos\theta(t)], \tag{20}$$

where H[ ] is the Hilbert transform of "terms within brackets." Unfortunately, equation (20) is usually not true since it cannot satisfy the Bedrosian theorem [*Bedrosian*, 1963]: the Hilbert transform for the product of two functions, $f(t)$ and $h(t)$, can be written as

$$\mathrm{H}[f(t)h(t)] = f(t)\mathrm{H}[h(t)] \tag{21}$$

only if the Fourier spectra for $f(t)$ and $h(t)$ are totally disjoint in frequency space and the frequency range of the spectrum for $h(t)$ is higher than that of $f(t)$. According to the Bedrosian theorem, equation (20) can only be true if the amplitude is varying so slowly that the frequency spectra of the envelope and the carrier waves are disjoint, a condition seldom satisfied for any practical data.

[30] To circumvent this difficulty, *Huang* [2005a] and *Huang et al.* [2008] have proposed a normalization scheme, which is essentially an empirical method to separate the IMF into amplitude modulation (AM) and frequency modulation (FM) parts uniquely. With this separation, we can conduct the Hilbert transform on the FM part alone and avoid the difficulty stated in the Bedrosian theorem.

[31] The normalization scheme is given as follows: for an IMF, $x(t)$, as given in Figure 1 and equation (15), we first find absolute values of the IMF and then identify all the maxima of these absolute values, followed by defining the envelope by a spline through all these maxima and designating it as $e_1(t)$. For any IMF, the envelope so defined is unique if the type (order) of spline is given. The normalization is given by

$$f_1(t) = \frac{x(t)}{e_1(t)}. \tag{22}$$

If $e_1(t)$ based on extrema fitting is identical to $a(t)$, $f_1(t)$ should be $\cos\theta(t)$ with all the values of $f_1(t)$ either equal to or less than unity. However, this is often not the case, for the fitted envelope $e_1(t)$ through the extrema often has cuts through $x(t)$, especially when $a(t)$ (the AM part of $x(t)$) undergoes large changes within a relatively small temporal duration. In such a case, $e_1(t)$ could have values smaller than the data $x(t)$ at some temporal locations and cause the normalized function to have amplitude greater than unity at some temporal locations. These conditions may be rare but

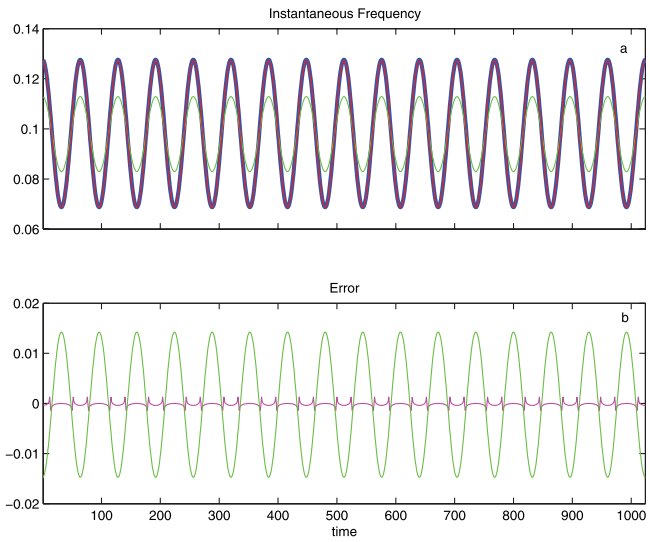

Figure 4. Instantaneous frequencies (IFs) of the idealized IMF specified by equation (6) and their errors, defined as the difference of IF calculated using a particular method and the truth. (a) Blue line is the true IF specified in equation (6), the magenta line is the IF based on the "direct quadrature" method, and the green line is the IF calculated based on the normalized Hilbert transform method. (b) Magenta line is the error associated with the "direct quadrature" method, and the green line is the error associated with the normalized Hilbert transform method.

are certainly not impossible. Therefore, the normalization is again an iterative process, as follows:

$$f_2(t) = \frac{f_1(t)}{e_2(t)}; \ \ldots \ f_n(t) = \frac{f_{n-1}(t)}{e_n(t)}. \tag{23}$$

The iteration stops at the nth step, when the normalized maximum values are all unity. Then the empirical FM and AM components of $x(t)$ are defined as

$$F(t) = f_n(t) \tag{24a}$$

$$A(t) = \frac{x(t)}{F(t)}, \tag{24b}$$

respectively. This iteration will gradually yield the empirical AM and FM parts of the IMF. The combination of this normalizing iteration process and application of the Hilbert transform to the empirical AM signal is called the normalized Hilbert transform method, NHT for short.

[32] The instantaneous frequency calculation method based on NHT improves the result significantly, as demonstrated in Figure 4. The instantaneous frequency obtained using NHT corrects, to a great degree, the irregular oscillations in the instantaneous frequency calculated using the Hilbert transform shown directly in Figure 1b. However, it is also evident from Figure 4 that the calculated instantaneous frequency using NHT still underestimates the true frequency fluctuation although its mean and periodicity are correctly determined. The cause of this underestimation is summarized in the Nuttall theorem [*Nuttall*, 1966].

[33] The Nuttall theorem states that for a complicated phase function with its time derivative (the frequency) not constant the Hilbert transform of the cosine of this phase function (FM part of the IMF) is not necessarily a simple 90° phase shift. The error bound, ΔE, defined as the difference between C_h, the Hilbert transform, and C_q, the true quadrature (with phase shift of exactly 90°) of the function can be expressed as

$$\Delta E = \int_{t=0}^{T} \left| C_q(t) - C_h(t) \right|^2 dt = \int_{-\infty}^{0} S_q(\omega)\, d\omega, \tag{25}$$

in which S_q is the Fourier spectrum of the quadrature function. Though the theorem is elegant and the proof rigorous, the result is hardly useful: first, it is expressed in terms of the Fourier spectrum of a still unknown quadrature, and, second, it gives a constant error bound over the whole data range. For a nonstationary time series, such a constant bound is not informative. Finally, the global bound will not reveal the location of the error on the time axis.

[34] To construct a calculable local difference (error) between the Hilbert transform of the cosine of a complicated phase function and its true quadrature, *Huang* [2005b, 2006] and *Huang et al.* [2008] have proposed an alternative variable error bound based on the normalization scheme described in this section. The error is defined as the difference between unity and the squared sum of the FM component and its Hilbert transform. The rationale of this concept is simple: if the Hilbert transform is exactly the quadrature, the squared sum of the FM component and its Hilbert transform should be unity, and the difference (the alternative error bound) should be zero. If the squared sum is not exactly unity, then the Hilbert transform cannot be exactly the quadrature. The errors have to come from a highly complicated phase function as discussed by *Huang et al.* [1998], whenever the phase plane formed from the Hilbert transform is not a perfect circle.

[35] Although the error bounds defined by the *Nuttall* [1966] theorem and, especially, by *Huang* [2005b] provide insights into the possible causes of the systematic underestimation of the fluctuations of the instantaneous frequency using the Hilbert transform, as shown in Figure 4a, there was no solution being proposed in those studies. It seems that the problem stems from the Hilbert transform itself, for the Hilbert transform uses a global domain integral and implicitly a global domain weighted average, while instantaneous frequency is a temporal local quantity. For this reason, to improve the calculation of the instantaneous frequency, an approach without using Hilbert transform seems to be a necessary choice.

[36] There are indeed methods other than Hilbert transform that can be used to calculate instantaneous frequency, such as the Wigner-Ville distribution [*Cohen*, 1995], the Teager energy operator [*Kaiser*, 1990; *Maragos et al.*, 1993a, 1993b], and wavelet analysis. Unfortunately, the

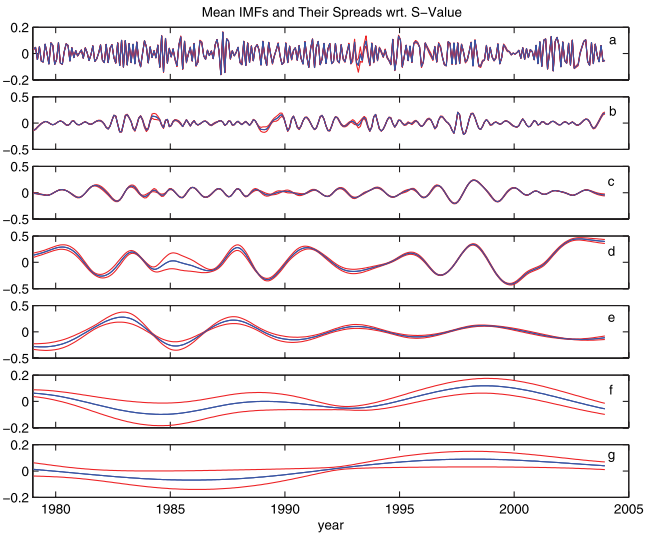

Figure 5. Confidence limit of IMFs of RSS T2 data. The mean (blue line) and its 1-standard-deviation confidence limit (red line) for each IMF are plotted for eight different decompositions with S numbers ranging from 5 to 12.

Wigner-Ville distribution is based on a global temporal domain integration of the lagged autocorrelation so that the instantaneous frequency can hardly be temporally local. The Teager energy operator appears to be a nonlinear operator, but its instantaneous frequency definition only has physical meaning when the signal has no frequency and amplitude modulation over the global domain. The wavelet method, as mentioned in section 1, suffered greatly from its subharmonics. It seems that modifying these methods cannot lead to noticeable improvement of the calculation of instantaneous frequencies of a complicated time series.

[37] One alternative is to abandon the Hilbert transform and other above mentioned methods and use the good properties of IMF. As we mentioned earlier, an IMF is approximately an AM/FM separated component. Therefore, instantaneous frequencies of an IMF can be determined from the FM part. Although such an alternative has existed ever since the development of EMD, the actual implementation was not feasible earlier because the determination of the FM part necessitates using the envelope (AM part) and IMF itself, and the empirical envelope of an IMF may have cut into the IMF. That problem has been solved with the normalization scheme described earlier. After normalization, the FM part of the signal is obtained, and we can then compute the phase function using the arc-cosine of the normalized data. This method is called the "direct quadrature" (DQ) method, and its details are given by *Huang et al.* [2008].

[38] As the DQ method is direct and simple, the results of the calculated instantaneous frequency using DQ show a dramatically improved version. An example is displayed in Figure 4. The magenta line in Figure 4a is almost identical to the true instantaneous frequency specified in equation (6). Its error (the difference between the calculated values using DQ and the true frequency) is an order smaller than the corresponding error when the Hilbert transform is applied to the same FM part. It is also clear that the error associated with DQ has a special pattern: the error is larger in the temporal locations where IMF has its extrema. Such systematic error can be easily corrected; see *Huang et al.* [2008] for more details.

3.2. A Confidence Limit

[39] A confidence limit is always desirable in any statistical analysis, for it provides a measure of reliability of the results. For Fourier spectral analysis, the confidence limit is routinely computed based on the ergodic assumption and is determined from the statistical spread of subdivided N sections of the original data. When all the conditions of the ergodic theory are satisfied, this temporal average is treated as the ensemble average. Unfortunately, ergodic conditions can be satisfied only if the processes are stationary; otherwise, the averaging operation itself will not make sense.

[40] Since the EMD is an empirical algorithm and involves a prescribed stoppage criterion to carry out the sifting moves, we have to know the degree of sensitivity in the decomposition of an input to the sifting process, so the reliability of a particular decomposition can further be determined. Therefore, a confidence limit of the EMD is a desirable quantity.

[41] Since the ergodic approach in Fourier spectrum analysis cannot be applied to obtain a confidence limit of EMD, which is a time domain decomposition, *Huang et al.* [2003] proposed a different approach to examining the confidence limit. In this approach, the fact that there are infinitely many ways to decompose any given function into different components using various methods is utilized. To find the confidence limit, the coherence of the subset of the decompositions must be found. For EMD, many different sets of IMFs may be generated by only changing the stoppage criterion, the S number. From these different sets of IMFs, we can calculate the mean and the spread of any corresponding IMFs, thereby determining the confidence limit quantitatively. This approach does not depend on the ergodic assumption, and, by using the same data length, there is no downgrading of the spectral resolution in frequency space by dividing the data into sections.

[42] An example of such a defined confidence limit of the decomposition of RSS T2 data (see Figure 2a) is displayed in Figure 5, in which eight sets of decompositions with different S numbers, from 5 to 12, are used. For the first four components, the means of the corresponding IMFs from different decompositions are almost identical to the corresponding IMFs displayed in Figure 3, for which the S number is selected to be 5. The spreads of decompositions are also quite small for these components. For the last two components, the means resemble the corresponding IMFs in Figure 3, but the spreads are relatively larger.

[43] From the confidence limit study, an unexpected result is the determination of the optimal S number in the range of 4 to 8. The result is consistent with logic: the S number should not be too high that it would lead to

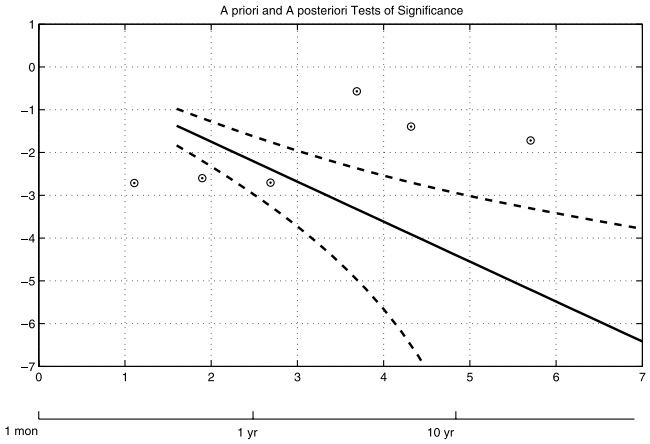

Figure 6. Statistical significance test of IMFs of RSS T2 against white noise null hypotheses. Each "target" sign represents the energy of an IMF as a function of mean period of the IMF, ranging from the IMF 1 to IMF 6. The solid line is the expected energy distribution of IMFs of white noise; and the upper and lower dashed lines provide the first and 99th percentiles of noise IMF energy distribution as a function of mean period, respectively. Anything that stays above the upper dashed line is considered statistically significant at the 99th percentile.

excessive sifting, which drains all physical meanings out of the IMF, nor too low that it would lead to undersifting, which leaves some riding waves remaining in the resulting IMFs. An alternative confidence limit for EMD is given by the ensemble EMD [*Wu and Huang*, 2008], which will be discussed in section 3.4.

3.3. Statistical Significance of IMFs

[44] The EMD is a method of separating data into different components by their scales. Since the data almost always contain noise, a natural question to ask is whether a component (an IMF for EMD) contains a true signal or is only a component of noise. To answer this question, the characteristics of noise in EMD should be understood first.

[45] This task was recently carried out by a few research groups. *Flandrin et al.* [2004, 2005] and *Flandrin and Gonçalves* [2004] studied the Fourier spectra of IMFs of fractional Gaussian noise, which are widely used in the signal processing community and in financial data simulation. They found that the spectra of all IMFs of any fractional Gaussian noise except the first one collapse to a single shape along the axis of the logarithm of frequency (or period) with appropriate amplitude scaling of the spectra. The center frequencies (periods) of the spectra of the neighboring IMFs are approximately halved (and hence doubled); therefore, the EMD is essentially a dyadic filter bank. Independently, *Wu and Huang* [2004, 2005a] found the same result for white noise (which is a special case of fractional Gaussian noise).

[46] From the characteristics they obtained, *Wu and Huang* [2004, 2005a] further derived the expected energy distribution of IMFs of white noise. However, the expected energy distribution of IMFs is not enough to tell whether an IMF of real data contains only noise components or has signal components, because the energy distribution of IMFs of a noise series can deviate significantly from the expected energy distribution. For this reason, the spread of the energy distribution of IMFs of noise must be determined. *Wu and Huang* [2004, 2005a] argued using the central limit theorem that each IMF of Gaussian noise is approximately Gaussian distributed, and therefore the energy of each IMF must be a X^2 distribution. By determining the number of degrees of freedom of the X^2 distribution for each IMF of noise, they derived the analytic form of the spread function of the energy of the IMF. From these results, one would be able to discriminate an IMF of data containing signals from that of only white noise with any arbitrary statistical significance level. They verified their analytic results with those from the Monte Carlo test and found consistency. An example of such a test for the IMFs of RSS T2 as displayed in Figure 3 is presented in Figure 6.

[47] Another way of testing the statistical significance of an IMF of data is purely based on the Monte Carlo method, as proposed by *Coughlin and Tung* [2004a, 2004b] and *Flandrin et al.* [2005]. In this method, the special characteristics of the data, such as lagged autocorrelation, are used to make a null hypothesis for the type of noise process for the data. By generating a large number of samples of the noise series with the same length as that of the data and decomposing them into a large number of sets of IMFs, one obtains numerically the distribution of a certain metric (such as energy) of the corresponding IMFs. By comparing the location of the same metric of IMF data with this distribution, one can tell whether an IMF contains signals with any given confidence level. A minor drawback of this approach is the demand of large computational resources to obtain an accurate distribution of a metric of an IMF when the size of the data to be analyzed is very large.

3.4. Ensemble Empirical Mode Decomposition

[48] One of the major drawbacks of EMD is mode mixing, which is defined as a single IMF either consisting of signals of widely disparate scales or a signal of a similar scale residing in different IMF components. Mode mixing is a consequence of signal intermittency. As discussed by *Huang et al.* [1998, 1999], the intermittency could not only cause serious aliasing in the time-frequency distribution but could also make the individual IMF lose its physical meaning. Another side effect of mode mixing is the lack of physical uniqueness. Suppose that two observations of the same oscillation are made simultaneously: one contains a low level of random noise and the other does not. The EMD decompositions for the corresponding two records are significantly different, as shown by *Wu and Huang* [2005b, 2008]. An example of the mode mixing and its consequences is illustrated in Figure 7, in which the decomposition of the University of Alabama at Huntsville (UAH) T2, an alternative analysis of the same channel 2 tropospheric temperature of the microwave sounding unit [*Christy et al.*, 2000], as well as the RSS T2, is presented. The decomposition of UAH T2 has obvious scale mixing: local

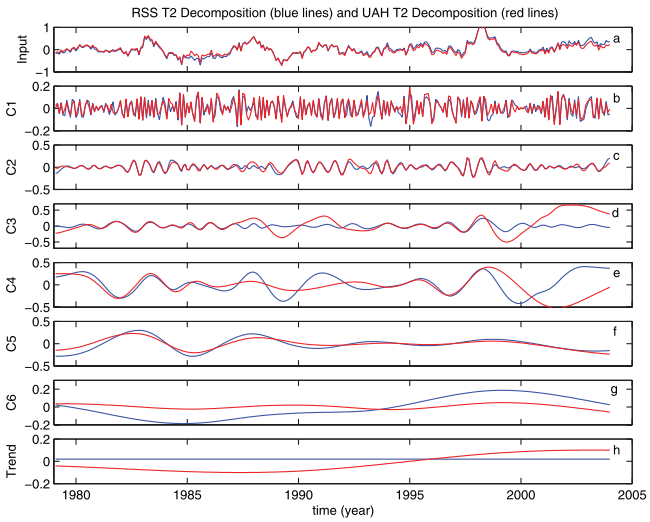

Figure 7. Mode mixing and the sensitivity of decomposition of the empirical mode decomposition to low-level noise. (a–h) Blue lines correspond to RSS T2 as displayed in Figure 2, and red lines correspond to UAH T2 (see the discussion in the text). RSS T2 and UAH T2, from data analysis view, can mutually be considered a noise-perturbed copy of the other.

periods around 1990 and 2002 of the third component are much larger than local periods elsewhere and are also much larger than the local periods of the corresponding IMF of RSS T2 around 1990 and 2002. The significant difference in decompositions but only minor difference in the original inputs raises a question: Which decomposition is reliable?

[49] The answer to this question can never be definite. However, since the cause of the problem is due to mode mixing, one expects that the decomposition would be reliable if the mode mixing problem is alleviated or eliminated. To achieve the latter goal, *Huang et al.* [1999] proposed an intermittency test. However, the approach itself has its own problems: First, the test is based on a subjectively selected scale, which makes EMD not totally adaptive. Second, the subjective selection of scales may not work if the scales are not clearly separable. To overcome the scale mixing problem, a new noise-assisted data analysis method was proposed, the ensemble EMD (EEMD) [*Wu and Huang*, 2005b, 2008], which defines the true IMF components as the mean of an ensemble of trials, each consisting of the signal plus a white noise of finite amplitude.

[50] Ensemble EMD is also an algorithm, which contains the following steps: (1) add a white noise series to the targeted data; (2) decompose the data with added white noise into IMFs; (3) repeat steps 1 and 2 again and again but with different white noise series each time; and (4) obtain the (ensemble) means of corresponding IMFs of the decompositions as the final result.

[51] The principle of EEMD is simple: the added white noise would populate the whole time-frequency space uniformly with the constituting components at different scales. When the signal is added to this uniform background, the bits of signals of different scales are automatically projected onto proper scales of reference established by the white noise. Although each individual trial may produce very noisy results, the noise in each trial is canceled out in the ensemble mean of enough trials; the ensemble mean is treated as the true answer.

[52] The critical concepts advanced in EEMD are based on the following observations:

[53] 1. A collection of white noise cancels each other out in a time-space ensemble mean. Therefore, only the true components of the input time series can survive and persist in the final ensemble mean.

[54] 2. Finite, not infinitesimal, amplitude white noise is necessary to force the ensemble to exhaust all possible solutions; the finite magnitude noise makes the different scale signals reside in the corresponding IMFs, dictated by the dyadic filter banks, and renders the resulting ensemble mean more physically meaningful.

[55] 3. The physically meaningful result of the EMD of data is not from the data without noise; it is designated to be the ensemble mean of a large number of EMD trials of the input time series with added noise.

[56] As an analog to a physical experiment that could be repeated many times, the added white noise is treated as the possible random noise that would be encountered in the measurement process. Under such conditions, the *i*th "artificial" observation will be

$$x_i(t) = x(t) + w_i(t), \tag{26}$$

where $w_i(t)$ is the *i*th realization of the white noise series. In this way, multiple artificial observations are mimicked.

[57] As the ensemble number approaches infinity, the truth, $c_j(t)$, as defined by EEMD, is

$$c_j(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \left[c_{jk}(t)\right], \tag{27}$$

in which

$$c_{jk}(t) = c_j(t) + r_{jk}(t), \tag{28}$$

where $r_{jk}(t)$ is the contribution to the *j*th IMF from the added white noise of the *k*th trial of the *j*th IMF in the noise-added signal. The amplitude of noise $w_i(t)$ is not necessarily small. But, the ensemble number of the trials, N, has to be large. The difference between the truth and the result of the ensemble is governed by the well-known statistical rule: it decreases as 1 over the square root of N [*Wu and Huang*, 2005b, 2008].

[58] With EEMD, the mode mixing is largely eliminated, and the consistency of the decompositions of slightly different pairs of data, such as RSS T2 and UAH T2, is greatly improved, as illustrated in Figure 8. Indeed, EEMD represents a major improvement over the original EMD. As the level of added noise is not of critical importance, as long as it is of finite amplitude while allowing for a fair ensemble

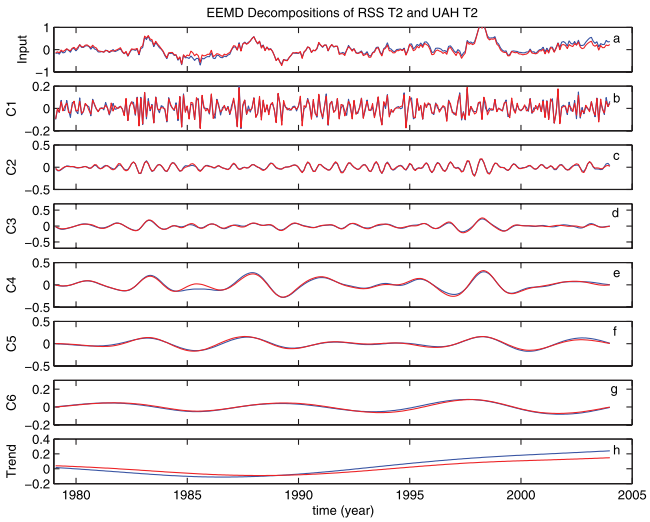

Figure 8. Ensemble empirical mode decomposition of low-noise perturbed data. (a–h) Blue lines correspond to RSS T2 as displayed in Figure 2, and red lines correspond to UAH T2 (see the discussion in text). RSS T2 and UAH T2, from data analysis view, can mutually be considered a noise-perturbed copy of the other. For both decompositions, the ensemble number is 100, and the added noise has an amplitude of 0.2 of that of the standard deviation of the corresponding data.

of all the possibilities, EEMD can be used without any significant subjective intervention; thus, it provides a truly adaptive data analysis method. By eliminating the problem of mode mixing, it also produces a set of IMFs that may bear the full physical meaning and a time-frequency distribution without transitional gaps. The EMD, with the ensemble approach, has become a more mature tool for nonlinear and nonstationary time series (and other one-dimensional data) analysis.

[59] The EEMD resolves, to a large degree, the problem of mode mixing [*Huang et al.*, 1999, 2003]. It might have resolved the physical uniqueness problem too, for the finite magnitude perturbations introduced by the added noises have produced the mean in the neighborhood of all possibilities.

[60] While the EEMD offers great improvement over the original EMD, there are still some unsettled problems. The EEMD-produced results might not satisfy the strict definition of IMF. One possible solution is to conduct another round of sifting on the IMFs produced by EEMD. As the IMFs results from EEMD are of comparable scales, mode mixing would not be a critical problem here, and a simple sift could separate the riding waves without any problem.

3.5. Two-Dimensional EMD

[61] Another important development is to generalize EMD to a two-dimensional analysis tool. The first attempt was initiated by *Huang* [2001], in which the two-dimensional image was treated as a collection of one-dimensional slices. Each slice was treated as one-dimensional data. Such an approach is called the pseudo-two-dimensional EMD. This method was later used by *Long* [2005] on wave data and produced excellent patterns and statistics of surface ripple riding on underlying long waves. The pseudo-two-dimensional EMD has shortcomings, and one of them is the interslice discontinuity. Recently, *Wu et al.* [2007b] used this pseudo-two-dimensional EMD approach but with EEMD replacing EMD and found that interslice discontinuity can be greatly reduced. However, in such an approach, the spatial structure is essentially determined by timescales. If spatial structures of different timescales are easily distinguishable, this approach would be appropriate, as demonstrated by *Wu et al.* [2007b] using the North Atlantic sea surface temperature (SST). If it is not, the applicability of this approach is significantly reduced.

[62] To overcome the shortcomings of the pseudo-two-dimensional EMD, one would have to use genuine two-dimensional EMD. For the development of a genuine two-dimensional EMD, the first difficulty comes from the definition of extrema. For example, should the ridge of a saddle be considered a series of maxima? The second difficulty comes from generating a smooth fitting surface to the identified maxima and minima. Currently, there are several versions of two-dimensional EMD, each containing a fitting surface determined by different methods. *Nunes et al.* [2003a, 2003b, 2005] used a radial basis function for surface interpretation and the Riesz transform rather than the Hilbert transform for computing the local wave number. *Linderhed* [2004, 2005] used the spline for surface interpretation to develop two-dimensional EMD data for an image compression scheme, which has been demonstrated to retain a much higher degree of fidelity than any of the data compression schemes using various wavelet bases. *Song and Zhang* [2001], *Damerval et al.* [2005] and *Yuan et al.* [2008] used a third way based on Delaunay triangulation and on piecewise cubic polynomial interpretation to obtain an upper surface and a lower surface. *Xu et al.* [2006] provided the fourth approach by using a mesh fitting method based on finite elements. All these two-dimensional approaches are computationally expensive. Though the spline-fitted surface serves the purpose well, the fittings offer only an approximation and could not go through all the actual data points. These points need further attention before the two-dimensional decomposition can become practically operational. Nevertheless, the 2-D method has been applied already. Examples of geophysical applications have been reported by *Han et al.* [2002], where they used the method to reduce speckle in synthetic aperture radar images, and by *Sinclair and Pegram* [2005], where they used the EMD for temporal-spatial rainfall data.

4. A DEMONSTRATION OF HHT AS A POWERFUL ANALYSIS METHOD

[63] In sections 2 and 3, we introduced the HHT method and its recent developments. We emphasized that HHT is an adaptive and temporally local analysis method that is capable of identifying nonlinear nonstationary processes

hidden in data. In this section, we will use examples to demonstrate that HHT is, indeed, a powerful tool for analyzing data.

[64] To achieve this goal, time series with rich characteristics and well-understood physics behind the data are ideal, so that the usefulness of the results analyzed using HHT can be easily verified. In the geophysical sciences, there are a few such time series, such as the Southern Oscillation Index (SOI) [*Ropelewski and Jones*, 1987], Vostok temperature (VT) derived from ice core [*Petit et al.*, 1997, 1999], and length-of-day data (LOD) [*Gross*, 2001]. SOI has already been decomposed using EEMD by previous studies [e.g., *Wu and Huang*, 2005b, 2008], showing that HHT can lead to physically meaningful decomposition. Therefore, in the following, we will only present the decompositions of LOD and VT.

[65] The LOD was previously analyzed using HHT and was studied extensively by *Huang et al.* [2003]. In that study, an intermittency test was performed to properly separate oscillations of different timescale when EMD is used to decompose the data. Here we decompose it using EEMD instead of EMD. The LOD being decomposed here are from 1 January 1964 to 31 December 1999. LOD (or part of it) has previously been studied by many researchers [e.g., *Barnes et al.*, 1983; *Rosen and Salstein*, 1983; *Chao*, 1989; *Höpfner*, 1998; *Razmi*, 2005]. A common approach in previous studies was to use Fourier spectrum analysis to identify the spectrum peaks. With the consideration of possible quasiperiodic response to periodic forcing (e.g., the tidal effect caused by revolution of the Moon around the Earth), one designs band-pass filters to isolate various physically meaningful quasiperiodic components [e.g., *Höpfner*, 1998]. However, the filtered results are usually sensitive to the structures of band-pass filters, and often, a posteriori calibration is needed to obtain ''neat'' results with ''neatness'' judged by the a priori knowledge a researcher has on the possible mechanisms involved in the change of LOD. An even more serious problem is that a filtered component for any earlier period may be changed if later data are added for the calculation of the spectrum of the data and the same filter is reapplied to obtain the component. Furthermore, most frequency domain filters are linear. If the waveforms are distorted by nonlinear processes, the band-pass filter could only pick up the fundamental mode and could leave out all the needed harmonics to reconstitute the full waveforms. These drawbacks raise concerns about the reliability of the filtered components and, consequently, the corresponding scientific interpretations.

[66] The locality and adaptivity of HHT overcome the drawbacks described above. The a priori knowledge of the physical mechanisms behind LOD is no longer needed; the filtered components of their local periods smaller than data length would not be affected by adding data later on; and the need for multiple filters to isolate components of different timescales is no longer necessary.

[67] The LOD data and its EEMD decomposition are displayed in Figure 9 and in Figure 10, which is an enlargement of Figure 9 from January 1982 to December 1983 for the first three components. In this decomposition, some standard postprocessing schemes, such as combination of neighboring components and additional EMD decomposition of EEMD outputs, as discussed in more detail by *Wu and Huang* [2005b, 2008], are applied. In Figure 9, eight IMF-like components (C1 and C5) and IMFs (C2–4 and C6–8) are presented, as well as the low-frequency component displayed over the LOD data.

[68] The first component, C1, has an averaged amplitude 1 order smaller than any other components. It has quasi-regular spikes with an average period around 14 d superimposed on random high-frequency oscillations. These random high-frequency oscillations may be related to weather storms [*Huang et al.*, 2003]. The second component, C2, has an average period of about 14 d, which was linked to semimonthly tides [*Huang et al.*, 2003]. The amplitude of C2 has small semiannual modulation superimposed on a 19-year modulation. The 19-year modulation is believed to be related to the Metonic cycle. C3 has an average period of about 28 d, which was linked to monthly tides. Its amplitude appears to be relatively small in El Niño years, as indicated by the red arrows in Figure 9.

[69] The fourth component, C4, is a component with periods between a month and one-half year, with a relatively smaller averaged amplitude that is only a fraction of those of other components excluding C1. C5 and C6 are semiannual and annual components, respectively. The causes of these cycles in length of day have been attributed to both the semiannual and annual cycles of the atmospheric circulation and to other factors, such as tidal strength change related to the revolution of the Earth around the Sun [*Höpfner*, 1998].

[70] The next two components are the variations between interannual timescales. C7 is quasi-biannual, and C8 has an average period slightly longer than 4 years. In previous studies [e.g., *Chao*, 1989; *Gross et al.*, 1996], it was documented that in El Niño years, the length of day tended to be longer. Here we present a systematic phase locking of C8 to the El Niño phenomenon, which is represented by the sign-flipped Southern Oscillation Index (with amplitude reduced to 1/10).

[71] Two more extra features of the decomposition of LOD are worthy of special attention: First, the modulation of the annual cycle (C8) reported here is much larger than the corresponding modulation of the annual cycle extracted using a predesigned band-pass filter [*Höpfner*, 1998]. The reason for that is understandable: The use of Fourier analysis–based filtering implicitly includes the global domain averaging, and therefore it tends to smooth amplitude change. This argument, indeed, can be verified by applying the same filter to subsections of LOD data; that leads to larger-amplitude modulations of the annual cycle than the case of applying it to the full length of LOD data.

[72] The second feature is the nonstationarity of some components. It appears from C1, C2, C7, and C8 that some characteristics of LOD may be changed. The semiannual modulation of amplitudes of high-frequency components is more obvious before the early 1980s than after. C7 has

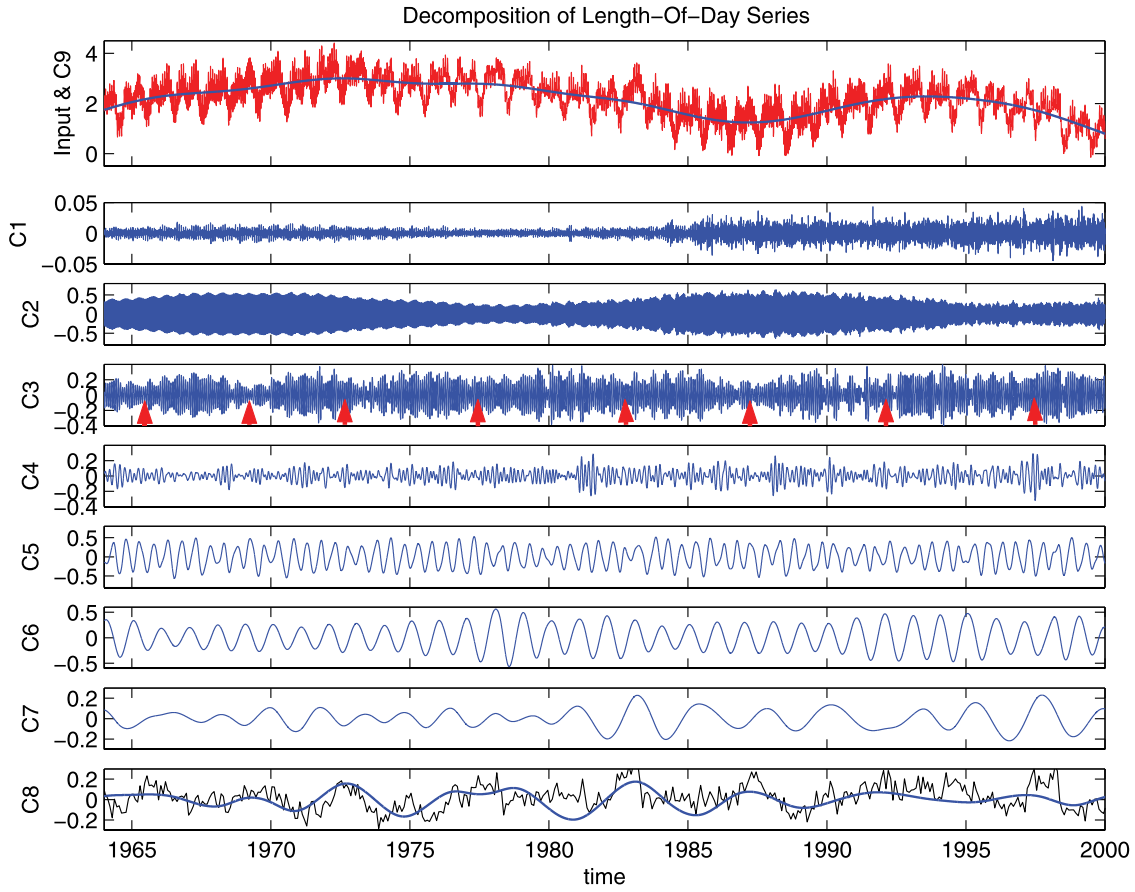

Figure 9. Ensemble empirical mode decomposition of the length-of-day data (the red line in the top plot) from 1 January 1964 to 31 December 1999. In the decomposition, noise of standard deviation 0.2 (absolute value not relative as in the case displayed in Figure 8) is added for the ensemble calculation, and the ensemble number is 800. The direct output (D1−D13, not shown here) of the decomposition has been reprocessed with combination of its components and additional EMD calculation. C1 is D1; C2 is the first mode of the combination of D2 and D3 subjected to additional EMD; the difference (D2 + D3 − C2) is added to D4; and the sum is subject to an additional EMD to obtain C3. The left over in this decomposition is added to D5 and D6. This latter sum is decomposed using an additional EMD to obtain C4 and C5. The sum of D7, D8, and D9 is decomposed using an addition EMD to obtain C6, C7, and C8. D10−1D3 are combined and displayed as the blue line in the top plot.

relatively smaller amplitude before the early 1980s than after, while the amplitude of C8 appears to be opposite. Part of these characteristic changes may be attributed to the change in the density of data [*Huang*, 2005b]. However, it remains to be investigated as to what causes such changes.

[73] The second example to show the power of HHT is the decomposition of Vostok temperature derived from ice cores. The most well-known science hidden in this time series is the relation between Milankovitch cycles and glacier-interglacier changes. It will be shown in this example that HHT can, indeed, reveal the role of three Milankovitch cycles related to the Earth's eccentricity (about 100 ka), axial tilt (about 41 ka), and precession (about 23 ka).

[74] The decomposition of Vostok temperature using the EEMD is displayed in Figure 11. From Figure 11, it is seen that the precession component is a mixture of large oscillations on precession timescale and higher-frequency oscillations of small amplitude. The oscillations on precession

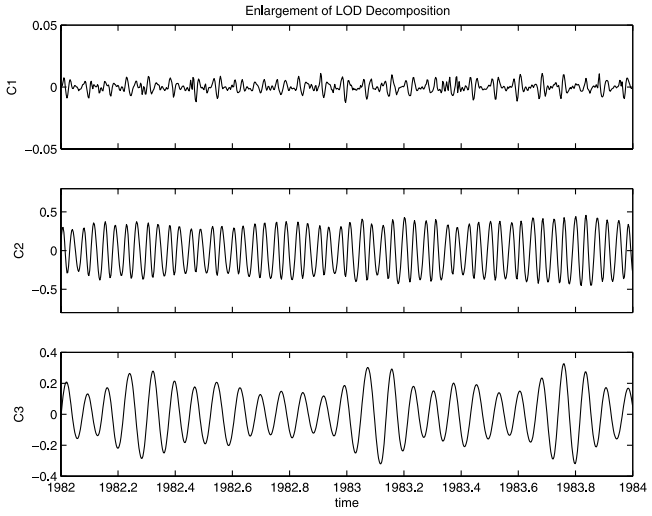

Figure 10. Enlargements of the first three components displayed in Figure 9 for the period from 1 January 1982 to 31 December 1983.

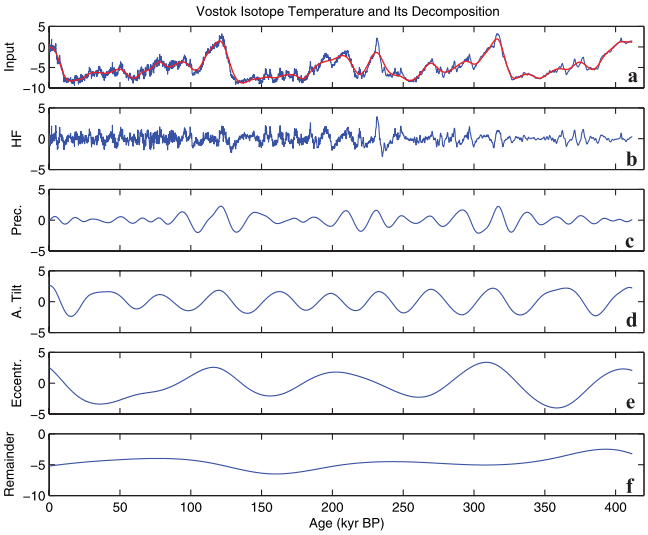

Figure 11. Ensemble empirical mode decomposition of Vostok temperature data based on ice core. (a) Data, (b) the high-frequency (HF) component, (c–e) the three components corresponding to three Milankovitch cycles, and (f) the low-frequency component are displayed. The difference of the input data and its high-frequency component is also shown (the red line) in Figure 11a.

timescale only appear in the temporal locations in interglacier periods and glacier-interglacier transition periods. The axial tilt component has relatively regular oscillations in both period and frequency. The eccentricity component has relatively flat amplitude but contains relatively larger frequency modulation. These results imply that the response of the Earth's climate system to Milankovitch cycles is highly nonlinear. From these results, it is seen that the response to precession cycles is larger around interglacier peaks than at other times. Also, for each interglacier peak, the response to the axial tilt cycles and the eccentricity cycles seems phase locked.

[75] An advantage of the HHT analysis of Vostok temperature (or other paleoclimate proxy data) is that the physical interpretation based on the decomposition will not be affected qualitatively by the dating error. It has been argued that the dating error of the ice core could reach to an order of a thousand years [*EPICA Community Members*, 2004]. However, since the EMD/EEMD decomposition is based on extrema, the dating error would only cause the peaks of the components displayed to shift slightly for components related to Milankovitch cycles. As pointed out by J. Chiang of University of California, Berkeley (personal communication, 2006), such an advantage can, indeed, be used to correct dating errors in Vostok temperature derived from the ice core if other well-dated individual paleoclimate events are used to calibrate the dating.

5. NEW FINDINGS FACILITATED BY HHT IN GEOSCIENCES

[76] Over the last few years, there have been many applications of HHT in scientific research and engineering. Two books based on papers presented in symposia were published: *Huang and Shen* [2005] discussed applications in various scientific fields, and *Huang and Attoh-Okine* [2005] discussed applications in engineering. The papers contained in these two books cover only selected topics of the theory and applications. In this section, we will provide a brief and incomplete survey on new discoveries in the geosciences facilitated by HHT.

5.1. Geophysical Studies

[77] The applications of HHT to geophysical data started with the introduction of the method by *Huang et al.* [1998], where they suggested that the Hilbert spectral representation for an earthquake can reveal the nonstationary and nonlinear nature of the phenomenon. Subsequently, *Huang et al.* [2001] applied HHT to the devastating earthquake in Chi-Chi, Taiwan, in 1999. Their study demonstrated that the Fourier-based representations (including the response spectrum) seriously underrepresented low-frequency energy because the linear characteristics of the Fourier transform generate artificial high-frequency harmonics. This misrepresentation is particularly severe when the signal is highly nonstationary. The HHT-based spectral representation was also used by *Loh et al.* [2000, 2001] and *Zhang et al.* [2003a, 2003b] to study ground motions and structure responses.

[78] As seismic waves are definitely nonstationary and the near-field motion is highly nonlinear, it is expected that the application of HHT to the study of seismic waves will be potentially fruitful. Indeed, the HHT has been used to investigate seismic wave propagation and explore source characteristics. *Vasudevan and Cook* [2000] used EMD to extract the scaling behavior of reflected waves. *Zhang et al.* [2003a, 2003b] and *Zhang* [2006] reported their efforts in tracing seismic wave propagation. They concluded that certain IMF components (higher-frequency ones) could be identified as being generated near the hypocenter, with the high-frequency content related to a large stress drop associated with the initiation of earthquake events. The lower-frequency components represent a general migration of the source region away from the hypocenter associated with longer-period signals from the rupture propagations.

5.2. Atmospheric and Climate Studies

[79] Some of the applications of HHT in the studies of meteorological phenomena of spatial scales from local to global have been summarized by *Duffy* [2004]. As atmospheric and climatic phenomena are highly nonstationary and nonlinear, HHT should potentially catch more insights into these phenomena. Specific applications are discussed as follows.

[80] First, let us survey wind field investigations. Among the first applications of HHT in atmospheric science was the study by *Lundquist* [2003], where she studied the intermittent and elliptical inertial oscillations in the atmospheric boundary layer, which has long been hypothesized to be associated with the evening transition phenomena. Her analysis using HHT revealed the intermittent nature of the

wind field and confirmed significant correlations of inertial motions with frontal passage. *Li et al.* [2005] also found strong intermittency of the turbulence content. In fact, the whole wind field structure is highly nonuniform [*Pan et al.*, 2003; *Xu and Chen*, 2004; *Chen and Xu*, 2004]. The inhomogeneity of the small-scale wind field has caused some difficulties in measuring the local wind velocity using lidar, which requires multiscanning and averaging. Recently, *Wu et al.* [2006], using HHT as a filter, succeeded in removing local small-scale fluctuations and obtained a stable mean. HHT was also applied to study wind field–related pollutant dispersion [*Janosi and Muller*, 2005], regional wind variations over Turkey [*Tatli et al.*, 2005], and multidecadal variability of wind over France [*Abarca-Del-Rio and Mestre*, 2006].

[81] Another area that takes advantage of the power of HHT is the study of rainfall that is highly intermittent. *Baines* [2005] found that the long-term rainfall variations of southwest Western Australia were directly related to the African monsoon. *El-Askary et al.* [2004] found that the 3- to 5-year cycles of rainfall over Virginia correlated well with the Southern Oscillation Index with a correlation coefficient of 0.68 at a confidence level of 95%, which indicated that the El Niño–Southern Oscillation could be strongly teleconnected to the eastern coast of the United States. *Molla et al.* [2006] and *Peel and McMahon* [2006] reported that global warming, indeed, has caused flood-drought cycles. There were also attempts to use HHT as a tool for long-term rainfall predictions. The basic principle behind this is quite simple: although the total rainfall record could be highly nonstationary, a narrowband IMF could make it amenable to prediction. Such data-based predictions were tried by *Mwale et al.* [2004] for central southern Africa, *Iyengar and Kanth* [2005, 2006] for the seasonal monsoon in India, and *Wan et al.* [2005] for climate patterns.

[82] HHT could provide better trend/cycle information than any other method because of its adaptivity and temporal locality [*Wu et al.*, 2007a]. *Li and Davis* [2006] found decadal variation of Antarctic ice sheet elevation using satellite altimeter data. *Bindschadler et al.* [2005] identified significant changes in the new snow thickness over a large area in the data collected with a vertically polarized passive radiometer at 85 GHz. *Hu and Wu* [2004] used HHT to study the impact of global warming on the variability of the North Atlantic Oscillation (NAO), and they identified shifts of the centers of action of the NAO associated with large-scale atmospheric wave pattern. *Xie et al.* [2002] revealed the interannual and interdecadal variations of landfalling tropical cyclones in the southeastern United States. *Slayback et al.* [2003] and *Pinzón et al.* [2005] detected a trend of longer growth season after they had painstakingly removed the satellite drift with HHT and constructed a uniformly valid advanced very high resolution radiometer (AVHRR) normalized difference vegetation index.

[83] The climate on the Earth is connected to the Sun at almost all timescales. In fact, the long-term variations (of the order of 10 to 100 ka) of Milankovitch scales have been a popular subject and even a conference topic (e.g., "Astronomical (Milankovitch) Calibration of the Geological Time-scale," a workshop organized by the Royal Society of London in 1998). *Lin and Wang* [2004, 2006] reported their analysis of Pleistocene (1 Ma B.P. to 20 ka B.P.) climate and found that the eccentricity band signal at that time is much larger than earlier estimations. On a much shorter timescale of around 10 years, the clear influence of solar activities on the Earth's climate had not been convincingly identified until the studies of *Coughlin and Tung* [2004a, 2004b, 2005], where they extracted clear 11-year cycles in stratosphere temperature that is related to the 11-year cycle of solar activity and its downward propagation to the lower troposphere.

[84] Another area of the application of HHT to climate study is in the examination of the climate variability of anomaly with respect to an alternative reference frame of amplitude-frequency modulated annual cycle (MAC) instead of the traditional repeatable annual cycle [*Wu et al.*, 2007b, 2008]. As is expected, when a reference frame is changed, the corresponding anomaly is changed. Consequently, the physical explanations for an anomaly may change as well.

[85] Recently, *Wu et al.* [2007b, 2008] developed a method based on HHT to extract MAC in climate data. With MAC, they defined an alternative copy of the anomaly. On the basis of that alternative anomaly, they found that the "reemergence" mechanism may be better interpreted as a mechanism for explaining the change of annual cycle rather than for explaining the interannual to interdecadal persistence of SST anomaly. They also found that the apparent ENSO phase locking is largely due to the residual annual cycle (the difference of the MAC and the corresponding traditional annual cycle) contained in the traditional anomaly and hence can be interpreted as a scenario of a part of annual cycle phase locked to annual cycle itself. They illustrated the problems of concepts such as "decadal variability of summer (winter) climate" in the climate study and suggested more logically consistent concepts of interannual or/and decadal variability of climate. Furthermore, they also demonstrated the drawbacks related to the stationary assumption in previous studies of extreme weather and climate and proposed a nonstationary framework to study extreme weather and climate.

[86] HHT has also been applied to the study of upper atmosphere and even extraterrestrial atmosphere. *Chen et al.* [2001] found that the vertical velocity fluctuations of plasma induced by the electric field followed exactly what is predicted by the Boltzmann relation down to the length scale between 50 and 300 m. *Komm et al.* [2001] studied the rotation residuals of the solar convection and found that the torsional oscillation pattern disappeared with increasing depth.

5.3. Oceanographic Studies

[87] HHT was initially motivated by the study of nonlinear wave evolution as reported by *Huang et al.* [1998, 1999]. As a result, the applications of HHT to water wave problems still occupy a highly visible position. As is known

to all the wave theorists, wave study started with the definition of a gradually changing phase function, $\theta(x, t)$, whose wave number and frequency are defined as

$$k = \frac{\partial \theta}{\partial x} \quad \omega = -\frac{\partial \theta}{\partial t}, \tag{29}$$

and therefore

$$\frac{\partial k}{\partial t} + \frac{\partial \omega}{\partial x} = 0. \tag{30}$$

Equation (30) is the kinematical conservation law. On the basis of this definition, the frequency should be a continuous and differentiable function of time and space. However, traditional wave studies had always used the Fourier-based frequency that is constant along the temporal axis. As a result, the very foundation of nonlinear wave governing equations was built on the assumption that the carrier wave frequency is constant; therefore, the wave envelope is governed by the nonlinear Schrödinger equation [*Infeld and Rowland*, 1990]. The introduction of HHT provided a way to define the instantaneous frequency empirically. *Huang et al.* [1998, 1999] have shown that nonlinear wave evolution processes are discrete, local, and abrupt. The observed phenomena are quite different from the idealized theoretical model. The results of their studies have the potential to drastically change wave studies, when a full theoretical foundation of HHT is established. We desperately need a new theory and additional new data to advance the study of waves.

[88] Observation of water waves either in the laboratory or in the field indicated that Fourier-based analysis is insufficient. For example, there is critical inconsistency in the traditional Fourier-based wave spectral representation. As waves are nonlinear, Fourier analysis needs harmonics to simulate the distorted wave profiles. However, harmonics are mathematical artifacts generated by imposing the linear Fourier representation on a fully nonlinear wave system. As a result, in the wave spectral representation we could not separate the true energy content from the fictitious harmonic components. Most field studies such as those of *Hwang et al.* [2003], *Schlurmann* [2002], *Datig and Schlurmann* [2004], *Veltcheva* [2002], *Veltcheva et al.* [2003], and *Veltcheva and Soares* [2004] confirmed this observation. Additionally, *Hwang et al.* [2006] also observed the drastic change of wavefield energy density across the Gulf Stream fronts, a nonlinear interaction yet to be fully explored.

[89] Water waves are nonlinear, but the effects of nonlinearity are shown most clearly only near the point of wave breaking. HHT has been used to study such phenomena from freak waves [*Wu and Yao*, 2004] to breaking, bubble generation, and turbulence energy dissipation [*Banner et al.*, 2002; *Gemmrich and Farmer*, 2004; *Vagle et al.*, 2005]. Recently, HHT was also employed by *Wang et al.* [2006] to study underwater acoustic signals.

[90] When we examined waves of longer scale, we also found HHT extremely useful. *Huang et al.* [2000] used HHT to study seiches on the coast of Puerto Rico. Traditionally, these waves were thought to be generated thousands of miles away, but using dispersive properties, *Huang et al.* [2000] proved that the seiches were, in fact, generated locally on the continental shelf. *Lai and Huang* [2005] and *Zheng et al.* [2006] also studied the inertia waves generated by passing hurricanes. They found the hurricanes could generate inertia waves in the upper layer. In the Gulf region, the deep layer waves (depth below the seasonal thermocline) are actually topographic Rossby waves.

[91] Another area of using HHT is in the determination of the ocean-climate relationship. *Yan et al.* [2002] found length of day related to the Western Pacific Warm Pool variations. *Yan et al.* [2006] also used HHT to track the movement of the middle-scale eddies in the Mediterranean Sea in remote sensing data. Oceanic data, when analyzed with HHT, revealed a rich mix of timescales from Rossby waves discussed above and by *Susanto et al.* [2000] to the interannual timescale of the Southern Oscillation [*Salisbury and Wimbush*, 2002]. Using sea surface temperature data collected from Nimbus 7 scanning multichannel microwave radiometer (SMMR) for 8 years and NOAA AVHRR for 13 years, *Gloersen and Huang* [2004] also found that the El Niño–Southern Oscillation is not confined only to the equatorial region but may have impact on both the North and the South Pacific with even larger amplitudes than at the equator. Additional work on the climate and ocean relationship is underway at this writing.

[92] The sea ice variation has been gaining increasing attention. *Gloersen and Huang* [1999, 2003], using HHT, found a pattern of Antarctic circumpolar waves in the perimeter of the ice pack. This eastward propagating wave had a quasi-quadrennial period. The geophysical implication is still unknown. Recently, *Zwally et al.* [2002] also found a puzzling result using HHT: from SMMR data covering the period from 1979 to 1998, they found an increase in the Antarctic sea ice but a decrease in Arctic ice. With the recent collapse of the Antarctic ice bank, their observation added complexity to the global climate change scenarios.

[93] From the successful applications mentioned above, HHT is demonstrated to be a powerful tool in our future research tool kit. It is expected that more successful applications will be made in the geophysical sciences.

6. SEARCHING FOR A THEORETICAL FOUNDATION

[94] In sections 2–5, we have introduced HHT, its recent methodological developments, and applications in the geosciences. The power of the method has been demonstrated, and the advantages of HHT over many previously widely used methods are also explained. However, HHT and its most recent developments are all empirical. A theoretical basis for HHT, or adaptive data analysis in general, is highly desired.

[95] By far, most of the theoretical studies on adaptive time-frequency analysis can fit into one of three major

categories: the first category of the studies concentrates on finding the necessary or sufficient or both conditions for positive definite instantaneous frequency defined as the derivative of the phase function defined using the Hilbert transform (see equation (11)). *Sharpley and Vatchev* [2006] found that a weak IMF, defined as any function having the same number of extrema and zero crossings, should satisfy a simple ordinary self-adjoint differential equation. However, such a week IMF failed to satisfy the limitation imposed by the Bedrosian theorem. *Xu and Yan* [2006] investigated the necessary and sufficient conditions to ensure the validity of Bedrosian theory. They found a stronger sufficient condition for Hilbert transform of a product of two functions in temporal domain, which treats the classical Bedrosian theorem in frequency domain as a special case. *Qian* [2006a, 2006b] pursued other conditions for positive definite instantaneous frequency defined by equation (11). *Qian and Chen* [2006] also defined the condition for positive definite instantaneous frequency for special types of equations.

[96] The above mentioned studies provide new insights on the validity of using the Hilbert transform to define physically meaningful instantaneous frequency. Unfortunately, with EMD components (IMFs) defined in the current way, these new insights are not going to overcome the limitation imposed by the Bedrosian theorem. However, the Bedrosian limitation does not vindicate the fact that the Hilbert transform is not capable of defining correct instantaneous frequency; rather, it can be circumvented by using the normalized Hilbert transform as described in section 3.1 and in greater detail by *Huang et al.* [2008].

[97] The second category of study uses alternative decomposition methods to EMD to find monocomponents of the data. A representative study of this category is that of *Olhede and Walden* [2004]. They used wavelet-based projections and obtained a set of monocomponent like functions, from which smooth instantaneous frequency could be found through Hilbert transform. The problem of this approach is that the decomposition is based on a priori defined wavelet basis; therefore, the decomposition is linear with respect to the selected basis. The fact that the instantaneous frequency values obtained through Hilbert transform are smooth is a consequence of eliminating the intrawave frequency modulation; therefore, this type of study lacks the unique ability to represent the nonlinear waveforms as HHT does.

[98] The third category of study is through modifying EMD algorithm: using B spline local medium fitting to replace the cubic spline envelope mean in traditional EMD. With such modification, *Chen et al.* [2006] made an effort to give EMD an analytic expression. The algorithm modification leads to noticeably different decompositions. However, the overall characteristics of the decompositions remained the same. It is hoped that the B spline with the analytic expression can lead to a proof of convergence, for the varying diminishing property guarantees that no straight line intersects a B spline curve more times than it intersects the curve's control polyline. Therefore, the B spline should have less variation and contain fewer extrema than the original collection of extrema. A formal proof of convergence is still wanting at this time.

7. OUTSTANDING OPEN PROBLEMS

[99] Since the introduction of HHT, it has gained some following and recognition. Up to this time, most of the progress in HHT has been in the application areas, while the underlying mathematical problems are mostly left untreated. The status of HHT is similar to wavelet analysis in the early 1980s. The future of the work depends on the arrival of someone to lay the mathematical foundation for HHT, as *Daubechies* [1992] did for wavelets. The following discussions of some outstanding mathematical problems are essentially based on the work of *Huang* [2005b] with some updating. The problems are set forth in sections 7.1–7.6.

7.1. Adaptive Data Analysis

[100] One of the most powerful methods in data analysis is to decompose a complicated data set into information-concentrated components. The established approach for data decomposition is to define a basis (such as trigonometric functions in Fourier analysis, for example); the analysis is then reduced to a convolution computation. This well-established paradigm is specious, for there is no a priori reason to believe that the basis selected truly represents the underlying processes for all the data all the time. Consequently, most of the results produced are not informative. It does, however, provide a definitive quantification with respect to a known metric for certain properties of the data based on the basis selected.

[101] If we give up this paradigm, there is no solid foundation to tread on. Yet data analysis methods need to be adaptive for most of the natural processes are both nonlinear and nonstationary. Therefore, there is no universal basis that could fit all the data all the time. Furthermore, the goal of data analysis is to discover the underlying physical processes; only the adaptive method can let the data reveal their underlying processes without any undue influence from the basis. Unfortunately, there is no mathematical model or precedent for such an approach. Recently, adaptive data processing has gained some attention especially in the adaptive wavelet analysis for filtering and denoising [*Chang et al.*, 2000]. Adaptation in such applications is not universal but is restricted to utilizing the difference between the characteristics of noise and data. The other adaptive approach depends on feedback [*Widrow and Stearns*, 1985], which will depend on the stationarity assumption. To generalize these available methods to nonlinear and nonstationary conditions is not easy; to generalize the adaptation as a methodology requires a totally new paradigm in mathematics. Related issues such as the uniqueness and convergence, for example, are tasks yet to be established.

7.2. Nonlinear System Identification

[102] Traditional system identification methods are based on the input-output relationship. Rigorous as the traditional

approach is, it is impossible to apply such methods to natural systems because in most of the cases the luxury of having such data is not available. All that might be available is a set of data describing a phenomenon; it is usually not known what the input actually is or the identity of the system itself. The only additional information that might be available is some general knowledge of the underlying controlling processes connected with the data. For example, the atmosphere and ocean are all controlled by the generalized equations for fluid dynamics and thermodynamics, which are nonlinear. Such a priori knowledge could guide the search for the characteristics or the signatures of nonlinearity. The question is whether it is possible to identify the nonlinear characteristics from the data. This might well be an ill-posed problem. Whether it is possible or not to identify the system through data only is an open question.

[103] So far, most of the available tests for nonlinearity from any data are only necessary conditions, for example, various probability distributions, higher-order spectral analysis, harmonic analysis, instantaneous frequency, etc. [see, e.g., *Bendat*, 1990; *Priestley*, 1988; *Tong*, 1990; *Kantz and Schreiber*, 1997; *Diks*, 1999]. There are many difficult problems in making such identifications from observed data only. This difficulty has made some scientists talk about only nonlinear systems rather than nonlinear data. Such reservations are reasonable and understandable, but this choice of terms still does not resolve the basic problem. The problem is made even more difficult when the process is also stochastic and nonstationary. For a nonstationary process, the various probabilities and the Fourier-based spectral analyses are all problematic, for those methods are based on global properties, with linear and stationary assumptions.

[104] Through the study of instantaneous frequency, intrawave frequency modulation is proposed as an indicator for nonlinearity. The advantage of the new approach is that the nonlinearity is revealed at the frequency near the fundamental mode rather appearing in the higher-frequency harmonics, which are usually buried in or mingled with the noise, making the identification even more difficult. More recently, *Huang* [2005b] has also identified the Teager energy operator [*Kaiser*, 1990; *Maragos et al.*, 1993a, 1993b] as an extremely local and sharp test for harmonic distortions within any IMF derived from data. The combination of these local methods offers some hope for system identification, but the problem is not solved, for this approach is based on the assumption that the input is linear. Furthermore, all these local methods also depend on local harmonic distortion; they cannot distinguish a quasi-linear system from a truly nonlinear system. A test or definition for nonlinear system identification based on observed output only is still needed.

7.3. End Effects of EMD

[105] End effects have unavoidably plagued all known data analysis methods from the beginning. The accepted, albeit timid, way to deal with it is using various kinds of windowing, as is done routinely in Fourier spectral analysis. Although sound in theory, such practices inevitably sacrifice some precious data near the ends, a serious hindrance when the data are few. In the HHT approach, we opt to extend the data beyond the existing range; otherwise, we have to stop the spline at the last extremum. Consequently, a method is needed to determine the spline curve between the last available extremum and the end of the data range. *Huang et al.* [1998] introduced the idea of using a "window frame," a way to extend the data beyond the existing range in order to tame the spline envelopes and extract some information from all the data available.

[106] The extension of data, or data prediction, is a difficult procedure even for linear and stationary processes. The problem that must be faced is how to make predictions for nonlinear and nonstationary stochastic processes. Here the cozy shelter of the linear, stationary, low-dimensional, and deterministic assumptions must be abandoned, and the complicated real world must be faced. The data are mostly from high-dimensional nonlinear and nonstationary stochastic systems. Are these systems predictable? What are the conditions needed to make them predictable? How well can the goodness of a prediction be quantified? In principle, data prediction cannot be made based on past data alone; the underlying processes have to be involved. Can the available data be used to extract enough information to make a prediction? These are open questions at present.

[107] In practice, however, our problem is not as daunting as discussed above, for we do not have to make the prediction on the whole data but only on the IMF, which has a much narrower bandwidth, for all the IMF should have the same number of extrema and zero crossings. Furthermore, all that is needed is the value and location of the next extrema not all the data. Such a limited goal, notwithstanding, the task is still challenging.

7.4. Spline Problems

[108] EMD is a "Reynolds type" decomposition: to extract variations from the data by separating the mean, in this case the local mean, from the fluctuations using spline fits. Although this approach is totally adaptive, several unresolved problems arise from its implementation.

[109] First, among all the spline methods, which one is the best? This is critical for it can be easily shown that all the IMFs other than the first are a summation of spline functions. This is because, from equations (7) to (9), we obtain

$$c_1 = X(t) - \left(m_{1k} + m_{1(k-1)} + \cdots + m_{11} + m_1\right), \qquad (31)$$

in which all the intermittent mean functions are generated by splines. Therefore, from equation (6),

$$r_1 = X(t) - c_1 = \left(m_{1k} + m_{1(k-1)} + \cdots + m_{11} + m_1\right) \qquad (32)$$

is totally determined by splines. As a result, all the rest of the IMFs are also totally determined by spline functions. What kind of spline is the best fit for the EMD? How can

one quantify the selection of one spline versus another? On the basis of our experience, it was found that the higher-order spline functions needed additional subjectively determined parameters, which violates the adaptive spirit of the approach. Furthermore, higher-order spline functions could also introduce additional length scales, which could lead to slow convergence and even nonconvergence. Furthermore, higher-order splines are also more time-consuming in computation. Such shortcomings dictate that only the cubic spline be selected. However, even with the cubic spline, should we use the B spline as proposed by *Chen et al.* [2006] or the natural cubic spline as used here?

[110] Finally, there is also the critical question of convergence of the EMD method: Is there a guarantee that in finite steps, a function can always be reduced into a finite number of IMFs? All intuitive reasoning and experience suggest that the procedure is converging. Under rather restrictive assumptions, the convergence can even be proved rigorously. The restricted and simplified case studied had sifting with middle points only. With further restriction of the middle point sifting to linearly connected extrema, the convergence proof can be established by reductio ad absurdum, and the number of extrema of the residue function has to be less than or equal to that in the original function. The case of equality only exists when the oscillation amplitudes in the data are either monotonically increasing or decreasing. In this case, the sifting may never converge and may forever have the same number in the original data and the IMF extracted. The proof is not complete in another aspect: Can one prove the convergence once the linear connection is replaced by the cubic spline? Therefore, this approach to establishing a proof is not complete.

[111] Recently, *Chen et al.* [2006] have used a B spline to implement the sifting. If one uses B spline as the base for sifting, then one can invoke the variation-diminishing property of the B spline and show that the spline curve will have less extrema. Details of this proof remain to be established.

7.5. Best IMF Selection and Uniqueness

[112] Does EMD generate a unique set of IMFs? By varying the parameters of sifting, the EMD method is a tool to generate infinite sets of IMFs. How are these different sets of IMF related? What is a criterion, or are there criteria to guide the sifting? What is the statistical distribution and significance of the different IMF sets? Therefore, a critical question is the following: How do we optimize the sifting procedure to produce the best IMF set? The difficulty is not to sift too many times, which would drain all the physical meanings out of each IMF component, and, at the same time, not to sift too few times and fail to get clean IMFs. Recently, *Huang et al.* [2003] has studied the problem of different sifting parameters and established a confidence limit for the resulting IMFs and Hilbert spectrum. However, the study was empirical and limited to cubic splines only. Optimization of the sifting process is still an open question.

[113] This question of uniqueness of the IMF can be traced to this more fundamental one: How do we define the IMF more rigorously? The definition given by *Huang et al.* [1998, 1999] is hard to quantify. Fortunately, the results are quite forgiving: even with the vague definition, the results produced are similar enough. Is it possible to give a rigorous mathematical definition and also find an algorithm that can be implemented automatically? Is EEMD the solution?

7.6. Hilbert Transform and Quadrature

[114] Traditionally, the Hilbert transform has been considered unusable or imprecise in defining the instantaneous frequency by two well-known theorems: the Bedrosian theorem [*Bedrosian*, 1963] and the Nuttall theorem [*Nuttall*, 1966]. The Bedrosian theorem states that the Hilbert transform for product functions can only be expressed in terms of the product of the low-frequency function with the Hilbert transform of the high-frequency one if the spectra of the two functions are disjointed. This guarantees that the Hilbert transform of $a(t)\cos\theta(t)$ is given by $a(t)\sin\theta(t)$. And, the Nuttall theorem [*Nuttall*, 1966] further stipulates that the Hilbert transform of $\cos\theta(t)$ is not necessarily $\sin\theta(t)$ for an arbitrary function $\theta(t)$. In other words, there is a discrepancy between the Hilbert transform and the perfect quadrature of an arbitrary function $\theta(t)$. Unfortunately, the error bound given by *Nuttall* [1966] is expressed in terms of the integral of the spectrum of the quadrature, an unknown quantity. Therefore, the single-valued error bound cannot be evaluated.

[115] Through research, the restriction of the Bedrosian theorem has been overcome through the EMD process and normalization of the resulting IMFs [*Huang et al.*, 2003]. With this new approach, the error bound given by *Nuttall* [1966] has been improved by expressing the error bound as a function of time in terms of instantaneous energy. However, the influence of the normalization procedure must be quantified. As the normalization procedure depends on a nonlinear amplification of the data, what is the influence of this amplification on the final results? Even if the normalization is accepted for an arbitrary $\theta(t)$ function, the instantaneous frequency is only an approximation. How can this approximation be improved?

[116] One possible alternative is to abandon the Hilbert transform and to compute the phase function using the arccosine of the normalized data. Two complications arise from this approach: the first one is the high precision needed for computing the phase function when its value is near $n\pi/2$. The second one is that the normalization scheme is only an approximation; therefore, the normalized functional value can occasionally exceed unity. Either way, some approximations are needed.

8. CONCLUSION

[117] HHT offers a potentially viable method for nonlinear and nonstationary data analysis, especially for time-frequency-energy representations. It has been tested widely in various applications other than geophysical research but only empirically. In all the cases studied, HHT gives results

much sharper than most of the traditional analysis methods. And in most cases, it reveals true physical meanings. In order to make the method more robust, rigorous, and friendlier in application, an analytic mathematical foundation is needed. The reasons are more than aesthetic; there are important practical justifications:

[118] 1. Only with a mathematical foundation can we make a unified general conclusion on the validity of the empirical results and deduce principles or laws. In the history of science, there are plenty of examples of great discoveries based on the empirical approach, but only with theoretical foundation can we unify the diverse observations. A prominent example is in electromagnetism: in the 1820s, Hans Oersted discovered that a wire carrying an electric current can generate a magnetic field. In the 1830s, Michael Faraday discovered that a moving magnet can produce an electric current in a moving wire passing through the magnetic field. Both of these observations are important and have practical applications, but it was Maxwell's equations that unified electromagnetism, which covers all of the particular phenomena.

[119] 2. Only with a mathematical foundation can we extend the validity to phenomena too subtle to be observed readily.

[120] Over the past few years, the HHT method has gained some following and recognition. Up to this time, most of the progress in HHT was in the application areas, while the underlying mathematical problems were mostly left untreated. Therefore, all of the results generated thus far are case-by-case comparisons conducted empirically. The work with HHT is presently at the stage corresponding historically to wavelet analysis in the early 1980s: producing great results but waiting for a unifying mathematical foundation on which to rest its case. We sorely need some one like *Daubechies* [1992], who laid the theoretical foundation of wavelet analysis in her famous 10 lectures. I hope this review will also draw the attention of mathematicians to HHT.

[121] In our view, the most likely solutions to the outstanding problems associated with HHT can only be formulated in terms of optimization: the selection of spline, the extension of the end, etc. can only be optimized for the time being perhaps. This may be the area of future research.

[122] **ACKNOWLEDGMENTS.** Throughout the years in developing the HHT method, we have received guidance and encouragement from many colleagues. We would like to single out Theodore T. Y. Wu of the California Institute of Technology and Owen M. Phillips of Johns Hopkins University, for without their guidance, HHT would never be what it is today. N. E. H. is supported in part by a Chair at NCU endowed by Taiwan Semiconductor Manufacturing Company, Ltd. and a grant, NSC 95-2119-M-008-031-MY3, from the National Research Council, Taiwan. Z. W. is supported by the National Science Foundation of the United States under grants ATM-0342104 and ATM-0653123.

[123] The Editor responsible for this paper was Gerald North. He thanks two technical reviewers and one anonymous cross-disciplinary reviewer.

REFERENCES

Abarca-Del-Rio, R., and O. Mestre (2006), Decadal to secular time scales variability in temperature measurements over France, *Geophys. Res. Lett.*, *33*, L13705, doi:10.1029/2006GL026019.

Bacry, E., A. Arnéodo, U. Frisch, Y. Gagne, and E. Hopfinger (1991), Wavelet analysis of fully developed turbulence data and measurement of scaling exponents, in *Proceedings of Turbulence 89: Organized Structures and Turbulence in Fluid Mechanics, Grenoble, Sept. 1989*, edited by M. Lesieur and O. Métais, pp. 203–215, Kluwer, Dordrecht, Netherlands.

Baines, P. G. (2005), Long-term variations in winter rainfall of southwest Australia and the African monsoon, *Aust. Meteorol. Mag.*, *54*, 91–102.

Banner, M. L., J. R. Gemmrich, and D. M. Farmer (2002), Multiscale measurements of ocean wave breaking probability, *J. Phys. Oceanogr.*, *32*, 3364–3375, doi:10.1175/1520-0485(2002)032<3364:MMOOWB>2.0.CO;2.

Barnes, R. T. H., R. Hide, A. A. White, and C. A. Wilson (1983), Atmospheric angular momentum fluctuations, length-of-day changes and polar motion, *Proc. R. Soc. London, Ser. A*, *387*, 31–73.

Bedrosian, E. (1963), On the quadrature approximation to the Hilbert transform of modulated signals, *Proc. IEEE*, *51*, 868–869.

Bendat, J. S. (1990), *Nonlinear System Analysis and Identification From Random Data*, 267 pp., Wiley Interscience, New York.

Bindschadler, R., H. Choi, C. Schuman, and T. Markus (2005), Detecting and measuring new snow accumulation on ice by satellite remote sensing, *Remote Sens. Environ.*, *98*, 388–402, doi:10.1016/j.rse.2005.07.014.

Carmona, R., W. L. Hwang, and B. Torresani (1998), *Practical Time-Frequency Analysis: Gabor and Wavelet Transform With an Implementation in S*, 490 pp., Academic, San Diego, Calif.

Chang, S. G., B. Yu, and M. Vetterli (2000), Adaptive wavelet thresholding for image denoising and compression, *IEEE Trans. Image Process.*, *9*, 1532–1546, doi:10.1109/83.862633.

Chao, B. F. (1989), Length-of-day variations caused by El Niño-Southern Oscillation and quasi-biennial oscillation, *Science*, *243*, 923–925, doi:10.1126/science.243.4893.923.

Chen, J., and X. L. Xu (2004), On modeling of typhoon-induced non-stationary wind speed for tall buildings, *Struct. Design Tall Spec. Build.*, *13*, 145–163, doi:10.1002/tal.247.

Chen, K. Y., H. C. Yeh, S. Y. Su, C. H. Liu, and N. E. Huang (2001), Anatomy of plasma structures in an equatorial spread *F* event, *Geophys. Res. Lett.*, *28*, 3107–3110, doi:10.1029/2000GL012805.

Chen, Q., N. E. Huang, S. Riemenschneider, and Y. Xu (2006), A B-spline approach for empirical mode decomposition, *Adv. Comput. Math.*, *24*, 171–195, doi:10.1007/s10444-004-7614-3.

Christy, J. R., R. W. Spencer, and W. D. Braswell (2000), MSU tropospheric temperatures: Dataset construction and radiosonde comparisons, *J. Atmos. Oceanic Technol.*, *17*, 1153–1170.

Cohen, L. (1995), *Time Frequency Analysis: Theory and Applications*, 320 pp., Prentice-Hall, Englewood Cliffs, N. J.

Coughlin, K., and K.-K. Tung (2004a), 11-year solar cycle in the stratosphere extracted by the empirical mode decomposition method, *Adv. Space Res.*, *34*, 323–329.

Coughlin, K., and K.-K. Tung (2004b), Eleven-year solar cycle signal throughout the lower atmosphere, *J. Geophys. Res.*, *109*, D21105, doi:10.1029/2004JD004873.

Coughlin, K., and K.-K. Tung (2005), Empirical mode decomposition and climate variability, in *Hilbert-Huang Transform: Introduction and Applications*, edited by N. E. Huang and S. S. P. Shen, pp. 149–165, World Sci., Singapore.

Damerval, C., S. Meignen, and V. Perrier (2005), A fast algorithm for bidimensional EMD, *IEEE Signal Process. Lett.*, *12*, 701–704, doi:10.1109/LSP.2005.855548.

Datig, M., and T. Schlurmann (2004), Performance and limitations of the Hilbert-Huang transformation (HHT) with an application to irregular water waves, *Ocean Eng.*, *31*, 1783–1834, doi:10.1016/j.oceaneng.2004.03.007.

Daubechies, I. (1992), *Ten Lectures on Wavelets*, 357 pp., Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Diks, C. (1999), *Nonlinear Time Series Analysis*, 209 pp., World Sci., Singapore.

Duffy, D. G. (2004), The application of Hilbert-Huang transforms to meteorological datasets, *J. Atmos. Oceanic Technol.*, *21*, 599–611, doi:10.1175/1520-0426(2004)021<0599:TAOHTT>2.0.CO;2.

Einstein, A. (1983), *Sidelights on Relativity*, 56 pp., Dover, Mineola, N. Y.

El-Askary, H., S. Sarkar, L. Chiu, M. Kafatos, and T. El-Ghazawi (2004), Rain gauge derived precipitation variability over Virginia and its relation with the El Nino Southern Oscillation, *Adv. Space Res.*, *33*, 338–342.

EPICA Community Members, (2004), Eight glacial cycles from an Antarctic ice core, *Nature*, *429*, 623–628, doi:10.1038/nature02599.

Flandrin, P., and P. Gonçalves (2004), Empirical mode decompositions as data-driven wavelet-like expansions, *Int. J. Wavelets Multiresolut. Inf. Process.*, *2*, 477–496, doi:10.1142/S0219691304000561.

Flandrin, P., G. Rilling, and P. Gonçalves (2004), Empirical mode decomposition as a filter bank, *IEEE Signal Process. Lett.*, *11*, 112–114, doi:10.1109/LSP.2003.821662.

Flandrin, P., P. Gonçalves, and G. Rilling (2005), EMD equivalent filter bank, from interpretation to applications, in *Hilbert-Huang Transform and Its Applications*, edited by N. E. Huang and S. Shen, pp. 57–74, World Sci., Singapore.

Gemmrich, J. R., and D. M. Farmer (2004), Near-surface turbulence in the presence of breaking waves, *J. Phys. Oceanogr.*, *34*, 1067–1086, doi:10.1175/1520-0485(2004)034<1067:NTITPO>2.0.CO;2.

Gloersen, P., and N. E. Huang (1999), In search of an elusive Antarctic circumpolar wave in sea ice extents: 1978–1996, *Polar Res.*, *18*, 167–173, doi:10.1111/j.1751-8369.1999.tb00289.x.

Gloersen, P., and N. E. Huang (2003), Comparison of interannual intrinsic modes in hemispheric sea ice covers and other geophysical parameters, *IEEE Trans. Geosci. Remote Sens.*, *41*, 1062–1074, doi:10.1109/TGRS.2003.811814.

Gloersen, P., and N. E. Huang (2004), Interannual waves in the sea surface temperatures of the Pacific Ocean, *Int. J. Remote Sens.*, *25*, 1325–1330, doi:10.1080/01431160310001592247.

Gross, R. S. (2001), Combinations of Earth orientation measurements: SPACE2000, COMB2000, and POLE2000, *JPL Publ.*, *01-2*.

Gross, R. S., S. L. Marcus, T. M. Eubanks, J. O. Dickey, and C. L. Keppenne (1996), Detection of an ENSO signal in seasonal length-of-day variations, *Geophys. Res. Lett.*, *23*, 3373–3377, doi:10.1029/96GL03260.

Han, C. M., D. D. Guo, C. L. Wang, and D. Fan (2002), A novel method to reduce speckle in SAR images, *Int. J. Remote Sens.*, *23*, 5095–5101, doi:10.1080/01431160210153110.

Höpfner, J. (1998), Seasonal variations in length of day and atmospheric angular momentum, *Geophys. J. Int.*, *135*, 407–437, doi:10.1046/j.1365-246X.1998.00648.x.

Hu, Z.-Z., and Z. Wu (2004), The intensification and shift of the annual North Atlantic Oscillation in a global warming scenario simulation, *Tellus, Ser. A*, *56*, 112–124.

Huang, N. E. (2001), Computer implemented empirical mode decomposition apparatus, method and article of manufacture for two-dimensional signals, Patent 6311130, U.S. Patent and Trademark Off., Washington, D. C.

Huang, N. E. (2005a), Empirical mode decomposition for analyzing acoustical signals, Patent 6862558, U.S. Patent and Trademark Off., Washington, D. C.

Huang, N. E. (2005b), Computing instantaneous frequency by normalizing Hilbert transform, Patent 6901353, U.S. Patent and Trademark Off., Washington, D. C.

Huang, N. E. (2006), Computing frequency by using generalized zero-crossing applied to intrinsic mode functions, Patent 6990436, U.S. Patent and Trademark Off., Washington, D. C.

Huang, N. E., and N. O. Attoh-Okine (Eds.) (2005), *Hilbert-Huang Transforms in Engineering*, 313 pp., CRC Press, Boca Raton, Fla.

Huang, N. E., and S. S. P. Shen (Eds.) (2005), *Hilbert-Huang Transform and Its Applications*, 311 pp., World Sci., Singapore.

Huang, N. E., S. R. Long, and Z. Shen (1996), The mechanism for frequency downshift in nonlinear wave evolution, *Adv. Appl. Mech.*, *32*, 59–111.

Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. London, Ser. A*, *454*, 903–993.

Huang, N. E., Z. Shen, and S. R. Long (1999), A new view of nonlinear water waves—The Hilbert spectrum, *Annu. Rev. Fluid Mech.*, *31*, 417–457, doi:10.1146/annurev.fluid.31.1.417.

Huang, N. E., H. H. Shih, Z. Shen, S. R. Long, and K. L. Fan (2000), The ages of large amplitude coastal seiches on the Caribbean Coast of Puerto Rico, *J. Phys. Oceanogr.*, *30*, 2001–2012, doi:10.1175/1520-0485(2000)030<2001:TAOLAC>2.0.CO;2.

Huang, N. E., C. C. Chern, K. Huang, L. W. Salvino, S. R. Long, and K. L. Fan (2001), A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129, Chi-Chi, Taiwan, 21 September 1999, *Bull. Seismol. Soc. Am.*, *91*, 1310–1338, doi:10.1785/0120000735.

Huang, N. E., M. L. Wu, S. R. Long, S. S. Shen, W. D. Qu, P. Gloersen, and K. L. Fan (2003), A confidence limit for the position empirical mode decomposition and Hilbert spectral analysis, *Proc. R. Soc. London, Ser. A*, *459*, 2317–2345.

Huang, N. E., Z. Wu, S. R. Long, K. C. Arnold, K. Blank, and T. W. Liu (2008), On instantaneous frequency, *Adv. Adapt. Data Anal.*, in press.

Hwang, P. A., N. E. Huang, and D. W. Wang (2003), A note on analyzing nonlinear and nonstationary ocean wave data, *Appl. Ocean Res.*, *25*, 187–193, doi:10.1016/j.apor.2003.11.001.

Hwang, P. A., J. V. Toporkov, M. A. Sletten, D. Lamb, and D. Perkovic (2006), An experimental investigation of wave measurements using a dual-beam interferometer: Gulf Stream as a surface wave guide, *J. Geophys. Res.*, *111*, C09014, doi:10.1029/2006JC003482.

Infeld, E., and G. Rowland (1990), *Nonlinear Waves, Solitons and Chaos*, 423 pp., Cambridge Univ. Press, New York.

Iyengar, R. N., and S. T. G. R. Kanth (2005), Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall, *Meteorol. Atmos. Phys.*, *90*, 17–36, doi:10.1007/s00703-004-0089-4.

Iyengar, R. N., and S. T. G. R. Kanth (2006), Forecasting of seasonal monsoon rainfall at subdivision level, *Curr. Sci.*, *91*, 350–356.

Janosi, I. M., and R. Muller (2005), Empirical mode decomposition and correlation properties of long daily ozone records, *Phys. Rev. E*, *71*, 056126, doi:10.1103/PhysRevE.71.056126.

Kaiser, J. F. (1990), On Teager's energy algorithm and its generalization to continuous signals, paper presented at 4th IEEE Signal Processing Workshop, Int. Electr. and Electron. Eng., Mohonk, N. Y.

Kantz, H., and T. Schreiber (1997), *Nonlinear Time Series Analysis*, 320 pp., Cambridge University Press, Cambridge, U. K.

Komm, R. W., F. Hill, and R. Howe (2001), Empirical mode decomposition and Hilbert analysis applied to rotation residuals of the solar convection zone, *Astrophys. J.*, *558*, 428–441, doi:10.1086/322464.

Lai, R. J., and N. E. Huang (2005), Investigation of vertical and horizontal momentum transfer in the Gulf of Mexico using empirical mode decomposition method, *J. Phys. Oceanogr.*, *35*, 1383–1402, doi:10.1175/JPO2755.1.

Li, M., W. M. Jiang, X. Li, and Y. F. Pu (2005), A test of the dynamical nonstationarity in atmospheric boundary layer turbulence (in Chinese), *Chin. J. Geophys.*, *48*, 493–500.

Li, Y. H., and C. H. Davis (2006), Improved methods for analysis of decadal elevation-change time series over Antarctica, *IEEE Trans. Geosci. Rem. Sens.*, *44*, 2687–2697, doi:10.1109/TGRS.2006.871894.

Lin, Z. S., and S. G. Wang (2004), A study on the problem of 100 ka cycle of astroclimatology (in Chinese), *Chin. J. Geophys.*, *47*, 971–975.

Lin, Z. S., and S. G. Wang (2006), EMD analysis of solar insolation, *Meteorol. Atmos. Phys.*, *93*, 123–128, doi:10.1007/s00703-005-0138-7.

Linderhed, A. (2004), Adaptive image compression with wavelet packets and empirical mode decomposition, Ph.D. dissertation, 226 pp., Dep. of Electr. Eng., Linköping Univ., Linköping, Sweden.

Linderhed, A. (2005), Variable sampling of the empirical mode decomposition of two-dimensional signals, *Int. J. Wavelets Multresolut. Inf. Process.*, *3*, 435–452, doi:10.1142/S0219691305000932.

Loh, C. H., Z. K. Lee, T. C. Wu, and S. Y. Peng (2000), Ground motion characteristics of the Chi-Chi earthquake of 21 September 1999, *Earthquake Eng. Struct. Dyn.*, *29*, 867–897, doi:10.1002/(SICI)1096-9845(200006)29:6<867::AID-EQE943>3.0.CO;2-E.

Loh, C. H., T. C. Wu, and N. E. Huang (2001), Application of the empirical mode decomposition-Hilbert spectrum method to identify near-fault ground-motion characteristics and structural responses, *Bull. Seismol. Soc. Am.*, *91*, 1339–1357, doi:10.1785/0120000715.

Long, S. R. (2005), Applications of HHT in image analysis, in *Hilbert-Huang Transform: Introduction and Applications*, edited by N. E. Huang and S. S. P. Shen, pp. 289–305, World Sci., Singapore.

Lundquist, J. K. (2003), Intermittent and elliptical inertial oscillations in the atmospheric boundary layer, *J. Atmos. Sci.*, *60*, 2661–2673, doi:10.1175/1520-0469(2003)060<2661:IAEIOI>2.0.CO;2.

Maragos, P., J. F. Kaiser, and T. F. Quatieri (1993a), On amplitude and frequency demodulation using energy operators, *IEEE Trans. Signal Process.*, *41*, 1532–1550, doi:10.1109/78.212729.

Maragos, P., J. F. Kaiser, and T. F. Quatieri (1993b), Energy separation in signal modulation with application to speech analysis, *IEEE Trans. Signal Process.*, *41*, 3024–3051, doi:10.1109/78.277799.

Mears, C. A., M. C. Schabel, and F. J. Wentz (2003), A reanalysis of the MSU channel 2 tropospheric temperature record, *J. Clim.*, *16*, 3650–3664.

Molla, M. K. I., M. S. Rahman, A. Sumi, and P. Banik (2006), Empirical mode decomposition analysis of climate changes with special reference to rainfall data, *Discrete Dyn. Nat. Soc.*, *2006*, 45348, doi:10.1155/DDNS/2006/45348.

Mwale, D., T. Y. Gan, and S. S. P. Shen (2004), A new analysis of variability and predictability of seasonal rainfall of central southern Africa for 1950–94, *Int. J. Climatol.*, *24*, 1509–1530, doi:10.1002/joc.1062.

Nunes, J. C., Y. Bouaoune, E. Delechelle, O. Niang, and P. Bunel (2003a), Image analysis by bidimensional empirical mode decomposition, *Image Vis. Comput.*, *21*, 1019–1026, doi:10.1016/S0262-8856(03)00094-5.

Nunes, J. C., O. Niang, Y. Bouaoune, E. Delechelle, and P. Bunel (2003b), Bidimensional empirical mode decomposition modified for texture analysis, in *Image Analysis: 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, June 29 – July 2, 2003 Proceedings*, *Lect. Notes Comput. Sci.*, vol. 2749, pp. 295–296, Springer, New York.

Nunes, J. C., S. Guyot, and E. Delechelle (2005), Texture analysis based on local analysis of the bidimensional empirical mode decomposition, *Mach. Vis. Appl.*, *16*, 177–188, doi:10.1007/s00138-004-0170-5.

Nuttall, A. H. (1966), On the quadrature approximation to the Hilbert transform of modulated signals, *Proc. IEEE*, *54*, 1458–1459.

Olhede, S., and A. T. Walden (2004), The Hilbert spectrum via wavelet projections, *Proc. R. Soc., Ser. A*, *460*, 955–975.

Pan, J. Y., X. H. Yan, Q. Zheng, W. T. Liu, and V. V. Klemas (2003), Interpretation of scatterometer ocean surface wind vector EOFs over the northwestern Pacific, *Remote Sens. Environ.*, *84*, 53–68, doi:10.1016/S0034-4257(02)00073-1.

Peel, M. C., and T. A. McMahon (2006), Recent frequency component changes in interannual climate variability, *Geophys. Res. Lett.*, *33*, L16810, doi:10.1029/2006GL025670.

Petit, J. R., et al. (1997), Four climate cycles in Vostok ice core, *Nature*, *387*, 359–360, doi:10.1038/387359a0.

Petit, J. R., et al. (1999), Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature*, *399*, 429–436, doi:10.1038/20859.

Pinzón, J. E., M. E. Brown, and C. J. Tucker (2005), EMD correction of orbital drift artifacts in satellite data stream, in *Hilbert-Huang Transform: Introduction and Applications*, edited by N. E. Huang and S. S. P. Shen, pp. 167–186, World Sci., Singapore.

Priestley, M. B. (1988), *Nonlinear and Nonstationary Time Series Analysis*, 237 pp., Academic, London.

Qian, T. (2006a), Analytic signals and harmonic measures, *J. Math. Anal. Appl.*, *314*, 526–536, doi:10.1016/j.jmaa.2005.04.003.

Qian, T. (2006b), Mono-components for decomposition of signals, *Math. Methods Appl. Sci.*, *29*, 1187–1198, doi:10.1002/mma.721.

Qian, T., and Q. H. Chen (2006), Characterization of analytic phase signals, *Comput. Math. Appl.*, *51*, 1471–1482, doi:10.1016/j.camwa.2006.01.007.

Razmi, H. (2005), On the tidal force of the Moon on the Earth, *Eur. J. Phys.*, *26*, 927–934, doi:10.1088/0143-0807/26/5/024.

Ropelewski, C. F., and P. D. Jones (1987), An extension of the Tahiti-Darwin Southern Oscillation Index, *Mon. Weather Rev.*, *115*, 2161–2165, doi:10.1175/1520-0493(1987)115<2161:AEOTTS>2.0.CO;2.

Rosen, R. D., and D. A. Salstein (1983), Variations in atmospheric angular momentum on global and regional scales and the length of day, *J. Geophys. Res.*, *88*, 5451–5470, doi:10.1029/JC088iC09p05451.

Salisbury, J. I., and M. Wimbush (2002), Using modern time series analysis techniques to predict ENSO events from the SOI time series, *Nonlinear Processes Geophys.*, *9*, 341–345.

Schlurmann, T. (2002), Spectral analysis of nonlinear water waves based on the Hilbert-Huang transformation, *J. Offshore Mech. Arct. Eng.*, *124*, 22–27, doi:10.1115/1.1423911.

Sharpley, R. C., and V. Vatchev (2006), Analysis of the intrinsic mode functions, *Const. Approximation*, *24*, 17–47, doi:10.1007/s00365-005-0603-z.

Sinclair, S., and G. G. S. Pegram (2005), Empirical mode decomposition in 2-D space and time: A tool for space-time rainfall analysis and nowcasting, *Hydrol. Earth Syst. Sci.*, *9*, 127–137.

Slayback, D. A., J. E. Pinzon, S. O. Los, and C. J. Tucker (2003), Northern Hemisphere photosynthetic trends 1982–99, *Global Change Biol.*, *9*, 1–15, doi:10.1046/j.1365-2486.2003.00507.x.

Song, P., and J. Zhang (2001), On the application of two-dimensional empirical mode decomposition in the information separation of oceanic remote sensing image (in Chinese), *High Technol. Lett.*, *11*, 62–67.

Susanto, R. D., A. L. Gordon, J. Sprintall, and B. Herunadi (2000), Intraseasonal variability and tides in Makassar Strait, *Geophys. Res. Lett.*, *27*, 1499–1502, doi:10.1029/2000GL011414.

Tatli, H., H. N. Dalfes, and S. S. Mentes (2005), Surface air temperature variability over Turkey and its connection to large-scale upper air circulation via multivariate techniques, *Int. J. Climatol.*, *25*, 331–350, doi:10.1002/joc.1133.

Tong, H. (1990), *Nonlinear Time Series Analysis*, 584 pp., Oxford Univ. Press, Oxford, U.K.

Vagle, S., P. Chandler, and D. M. Farmer (2005), On the dense bubble clouds and near bottom turbulence in the surf zone, *J. Geophys. Res.*, *110*, C09018, doi:10.1029/2004JC002603.

Vasudevan, K., and F. A. Cook (2000), Empirical mode skeletonization of deep crustal seismic data: Theory and applications, *J. Geophys. Res.*, *105*, 7845–7856, doi:10.1029/1999JB900445.

Veltcheva, A. D. (2002), Wave and group transformation by a Hilbert spectrum, *Coast. Eng. J.*, *44*, 283–300, doi:10.1142/S057856340200055X.

Veltcheva, A. D., and C. G. Soares (2004), Identification of the components of wave spectra by the Hilbert-Huang transform method, *Appl. Ocean Res.*, *26*, 1–12, doi:10.1016/j.apor.2004.08.004.

Veltcheva, A. D., P. Cavaco, and C. G. Soares (2003), Comparison of methods for calculation of the wave envelope, *Ocean Eng.*, *30*, 937–948, doi:10.1016/S0029-8018(02)00069-0.

Wan, S. Q., G. L. Feng, W. J. Dong, J. P. Li, X. Q. Gao, and W. P. He (2005), On the climate prediction of nonlinear and nonstationary time series with the EMD method, *Chin. Phys.*, *14*, 628–633, doi:10.1088/1009-1963/14/3/036.

Wang, F. T., S. H. Chang, and C. Y. Lee (2006), Signal detection in underwater sound using the empirical mode decomposition, *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, *E89-A*, 2415–2421.

Widrow, B., and S. D. Stearns (1985), *Adaptive Signal Processing*, 474 pp., Prentice Hall, Upper Saddle River, N. J.

Wu, C. H., and A. F. Yao (2004), Laboratory measurements of limiting freak waves on currents, *J. Geophys. Res.*, *109*, C12002, doi:10.1029/2004JC002612.

Wu, S. H., Z. S. Liu, and B. Y. Liu (2006), Enhancement of lidar backscatters signal-to-noise ratio using empirical mode decomposition method, *Opt. Commun.*, *267*, 137–144, doi:10.1016/j.optcom.2006.05.069.

Wu, Z., and N. E. Huang (2004), A study of the characteristics of white noise using the empirical mode decomposition method, *Proc. R. Soc., Ser. A*, *460*, 1597–1611.

Wu, Z., and N. E. Huang (2005a), Statistical significant test of intrinsic mode functions, in *Hilbert-Huang Transform: Introduction and Applications*, edited by N. E. Huang and S. S. P. Shen, pp. 125–148, World Sci., Singapore.

Wu, Z., and N. E. Huang (2005b), Ensemble empirical mode decomposition: A noise-assisted data analysis method, *COLA Tech. Rep. 193*, Cent. for Ocean-Land-Atmos. Stud., Calverton, Md. (Available at ftp://grads.iges.org/pub/ctr/ctr_193.pdf)

Wu, Z., and N. E. Huang (2008), Ensemble empirical mode decomposition: A noise-assisted data analysis method, *Adv. Adapt. Data Anal.*, in press.

Wu, Z., N. E. Huang, S. R. Long, and C.-K. Peng (2007a), On the trend, detrending, and the variability of nonlinear and nonstationary time series, *Proc. Natl. Acad. Sci. U. S. A.*, *104*, 14,889–14,894.

Wu, Z., E. K. Schneider, B. P. Kirtman, E. S. Sarachik, N. E. Huang, and C. J. Tucker (2007b), The modulated annual cycle—An alternative reference frame for climate anomaly, *COLA Tech. Rep. 193*, Cent. For Ocean-Land-Atmos. Stud., Calverton, Md. (Available at ftp://grads.iges.org/pub/ctr/ctr_244.pdf)

Wu, Z., E. K. Schneider, B. P. Kirtman, E. S. Sarachik, N. E. Huang, and C. J. Tucker (2008), Amplitude-frequency modulated annual cycle: An alternative reference frame for climate anomaly, *Clim. Dyn.*, in press.

Xie, L., L. J. Pietrafesa, and K. J. Wu (2002), Interannual and decadal variability of landfalling tropical cyclones in the southeast coastal states of the United States, *Adv. Atmos. Sci.*, *19*, 677–686, doi:10.1007/s00376-002-0007-y.

Xu, Y., and D. Yan (2006), The Bedrosian identity for the Hilbert transform of product functions, *Proc. Am. Math. Soc.*, *134*, 2719–2728, doi:10.1090/S0002-9939-06-08315-8.

Xu, Y., B. Liu, J. Liu, and S. Riemenschneider (2006), Two-dimensional empirical mode decomposition by finite elements, *Proc. R. Soc. Ser.*, *462*, 3081–3096.

Xu, Y. L., and J. Chen (2004), Characterizing nonstationary wind speed using empirical mode decomposition, *J. Struct. Eng.*, *130*, 912–920, doi:10.1061/(ASCE)0733-9445(2004)130:6(912).

Yan, X.-H., Y. Zhou, J. Pan, D. Zheng, M. Fang, X. Liao, M.-X. He, W. T. Liu, and X. Ding (2002), Pacific warm pool excitation, Earth rotation and El Niño southern oscillations, *Geophys. Res. Lett.*, *29*(21), 2031, doi:10.1029/2002GL015685.

Yan, X. H., Y. H. Jo, W. T. Liu, and M. X. He (2006), A new study of the Mediterranean outflow, air-sea interactions, and meddies using multisensor data, *J. Phys. Oceanogr.*, *36*, 691–710, doi:10.1175/JPO2873.1.

Yuan, Y., M. Jin, P. Song, and J. Zhang (2008), Empirical and dynamical detection of the sea bottom topography from synthetic aperture radar image, *Adv. Adapt. Data Anal.*, in press.

Zhang, R. R. (2006), Characterizing and quantifying earthquake-induced site nonlinearity, *Soil Dyn. Earthquake Eng.*, *26*, 799–812, doi:10.1016/j.soildyn.2005.03.004.

Zhang, R. R., S. Ma, and S. Hartzell (2003a), Signatures of the seismic source in END-based characterization of the 1994 Northridge, California, earthquake recordings, *Bull. Seismol. Soc. Am.*, *93*, 501–518, doi:10.1785/0120010285.

Zhang, R. R., S. Ma, E. Safak, and S. Hartzell (2003b), Hilbert-Huang transform analysis of dynamic and earthquake motion recordings, *J. Eng. Mech.*, *129*, 861–875, doi:10.1061/(ASCE)0733-9399(2003)129:8(861).

Zheng, Q., R. J. Lai, N. E. Huang, J. Y. Pan, and L. W. Timothy (2006), Observation of ocean current response to 1998 Hurricane Georges in the Gulf of Mexico, *Acta Oceanol. Sin.*, *25*, 1–14.

Zwally, H. J., J. C. Comiso, C. L. Parkinson, D. J. Cavalieri, and P. Gloersen (2002), Variability of Antarctic sea ice, 1979–1998, *J. Geophys. Res.*, *107*(C5), 3041, doi:10.1029/2000JC000733.

N. E. Huang, Research Center for Adaptive Data Analysis, National Central University, 300 Jhongda Road, Chungli, 32001 Taiwan. (norden@ncu.edu.tw)

Z. Wu, Center for Ocean-Land-Atmosphere Studies, Calverton, MD 20705, USA.