The Application of Hilbert-Huang Transforms to Meteorological Datasets

Dean G. Duffy

Code 912
NASA/Goddard Space Flight Center
Greenbelt, MD 20771

April 2003

Submitted to Journal of Atmospheric and Oceanic Technology

Email address: duffy@agnes.gsfc.nasa.gov
Abstract

Recently a new spectral technique has been developed for the analysis of aperiodic and nonlinear signals – the Hilbert-Huang transform. This paper shows how these transforms can be used to discover synoptic and climatic features: For sea level data, the transforms capture the oceanic tides as well as large, aperiodic river outflows. In the case of solar radiation, we observe variations in the diurnal and seasonal cycles. Finally, from barographic data, the Hilbert-Huang transform reveals the passage of extratropical cyclones, fronts, and troughs. Thus, this technique can flag significant weather events such as a flood or the passage of a squall line.
1. Introduction.

For generations researchers have used Fourier analysis to analyze signals. Because both the signal and its Fourier transform are important in understanding most processes, a combined temporal-frequency analysis has the potential of painting a more revealing picture. The purpose of this paper is to illustrate the advantage of one of these schemes — the Hilbert-Huang transform — on meteorological data sets.

The earliest time-frequency representation (TFR) was the short-time Fourier transform (STFT). See Allen and Rabiner (1977). This scheme divides the temporal signal \( f(t) \) into a series of small overlapping pieces. Each piece is then windowed and individually Fourier transformed. The STFT of a function \( f(t) \) is defined by

\[
F_{ST}(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) h(t - \tau) e^{-i\omega\tau} d\tau,
\]

where \( h(\tau) \) is the window function. Contour plots of the energy density function \( |F_{ST}(t, \omega)| \) are typically presented. This scheme is most useful when the physical process is linear so that the superposition of sinusoidal solutions is valid and the time series is stationary so that the Fourier coefficients are constant.

One of the drawbacks of STFT is the presence of a fixed window although Czerwinski and Jones (1997) have developed a short-time Fourier analysis with an adaptively adjusting window. Wavelet analysis (see Daubechies 1992) seeks to address this defect by decomposing the time series into local, time-dilated and time-translated wavelet components using time-frequency atoms or wavelets \( \psi \). The wavelet transform of the signal \( f(t) \) is then

\[
F_{WT}(t, \omega) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi \left( \frac{t - b}{a} \right) dt,
\]
where $a$ is the scale and $b$ is the time shift. The wavelet transform represents the energy in the signal of temporal scale $a$ at $t = b$. Wavelet analysis is attractive because (1) it is local, although higher frequencies are more localized, (2) it has uniform temporal resolution for all frequency scales, and (3) it is useful for characterizing gradual frequency changes. However, it is non-adaptive because the same basic wavelet is used for all data.

Finally empirical orthogonal function (EOF) analysis or its Fourier transform version, the singular spectral analysis (SSA), decomposes a time series using eigenfunctions of the covariance matrix. See Ghil et al. (2002). This analysis is quite different from the short-time Fourier transform or wavelet analyses because the EOFs are derived from data. However, its distribution of eigenvalues does not yield characteristic time or frequency scales. Furthermore, the eigenfunctions themselves are not necessarily linear or stationary and therefore are not easily analyzed by spectral modes.

From our experience with short-time Fourier transforms, wavelets, and EOF analyses, an ideal scheme for the spectral analysis of signals would be complete, orthogonal, local and adaptive. This method would also allow us to extract local time and frequency scales. The Hilbert-Huang transform is the next step toward this goal.

The Hilbert-Huang transform differs from the conventional concept of transforms. Although a Hilbert transform is applied to signals, it is first decomposed into a set of intrinsic mode functions (IMFs) which are band-limited, oscillatory, and derived from the data. Once we have a IMF and its Hilbert transform, we can form an analytic signal $A(t)e^{i\theta(t)}$, where $A(t)$ and $\theta(t)$ are the time dependent amplitude and phase, respectively. This amplitude can then be plotted as a function of time and (instantaneous) frequency.
Huang-Hilbert transforms have been successfully applied to a wide variety of signals. Some of these signals include pulmonary blood pressure (Huang et al. 1998), earthquake data (Huang et al. 2001), and rotational residuals from the solar convection zone (Komm et al. 2001). In the atmospheric sciences, these transforms have been applied to climatic signals (Pan et al. 2003, Salisbury and Wimbush 2002, Wu et al. 1999, Xie et al. 2002). In this paper we extend these applications to signals that include synoptic as well as climatic signals.

The paper begins with a detailed description of the scheme in Section 2. In Section 3, this scheme is then applied to data sets of sea level heights, incoming solar radiation, and barographic data. We present our conclusions in Section 4.

2. Procedure

The process of developing time-frequency diagrams using Hilbert-Huang transforms consists of three steps. The first step decomposes or "sifts" the signal into its intrinsic mode functions. Using a cubic spline, two envelope curves are generated; one of these curves connects the maxima of the signal while the other, the minima. From these curves the mean is computed. The difference between the signal and this mean constitutes a first-guess of the first IMF. If the IMF were sinusoidal, then the number of extrema would equal the number of zero crossings, or differ by one. This is usually not the case and suggests that our first guess, while good, needs further refinement. This is done by repeating our sifting process on this IMF. Originally Huang et al. (1998) repeated this sifting process until a Cauchy-like integral condition was satisfied. See their equation (5.5). In later
papers they adopted a stopping criterion based on the number extrema and zero-crossings. When these quantities were equal, or differed by one, for three consecutive times, the IMF is set to the values found in the final iteration.

To illustrate the decomposition process, Figure 1 presents various steps in computing the first IMF. In frame (a) we present a small portion of the original dataset – sea level observations taken at the mouth of the Chesapeake Bay during the 1980’s. In addition to the oscillations due to the oceanic tides, there is a large peak at 133 h from the passage of Hurricane Gloria during the early hours of 27 September 1985.

Frames (b) and (c) show the first IMF after it has satisfied the convergence criteria for the first and last (third) time, respectively. We have also included the top and bottom envelopes (the dashed lines) as well as the mean (the dotted line). Note how the sifting process has generated a mode that is symmetry with respect to the abscissa.

Shortly we will show that this first IMF represents the semidiurnal tides. This mode varies smoothly except in the interval 125 to 150 hours. If we examine the original data record in frame (a), we can associate this behavior with several “kinks” in the data record. It is unclear whether these kinks are real or a failure of the instrument.

To examine the effects that such kinks have on our construction of the IMFs, let us smooth the original data with a single pass of a simple Shapiro filter that eliminates waves with a period of $2\Delta t$. After sifting, frame (d) gives the first IMF of the original (solid line) and smoothed (dashed line) data. This improved behavior in the structure of the IMF might argue for smoothing the data before applying Hilbert-Huang transforms. Unfortunately, filtering is a two-edged sword: It modifies both the true signal as well as
eliminates noise. Because Huang-Hilbert transforms were designed specifically to analyze aperiodic and nonlinear signals, the use of a linear filter could alter the signal in way that might compromise the usefulness of the technique. For this reason, we will simply endure the presence of the noise.

Once the first IMF is found, it is subtracted from the signal and the process begins anew with the residual to obtain the second IMF. Frame (e) shows the residual after the first IMF has been removed. As we might expect, this residual is much smoother and has a lower frequency than the original signal. Further modes are found in the same manner and the sifting process concludes when there is no longer any maxima or minima in the residual.

One drawback in using a cubic spline to obtain the envelope curve is the possibility of large swings near the endpoints. We have adopted the method by Komm et al. (2001) of adding buffer zones on each end that equal the length of the original data. The signal was extrapolated into these buffer zones by fitting a sine wave using the closest extrema and zero crossing to define the frequency and amplitude of the wave.

Having determined the IMFs, the second step consists of computing the Hilbert transform of each mode \( \tilde{f}(t) \) and then the corresponding analytic signal \( z(t) = A(t)e^{i\theta(t)} = f(t) + i\tilde{f}(t) \), where \( f(t) \) denotes the IMF. Because we only have numerical values, the conventional method of computing \( \tilde{f}(t) \) is to take the fast Fourier transform of the data, multiply the transform by \( i \text{sgn}(\omega) \), and then take the inverse Fourier transform of this product. See Čižek (1970) or Henrici (1986, p. 203). Here \( \text{sgn}(\cdot) \) denotes the sign function.

Computing the instantaneous frequency \( \omega(t) \) from the data is difficult because it is
the time derivative of $\theta(t)$. Barnes (1992) tested a number of methods to compute it from $f(t)$ and $\hat{f}(t)$. He found that the best representation of the instantaneous frequency is

$$\omega(t) = \frac{1}{2\Delta t} \tan^{-1} \left[ \frac{f(t - \Delta t)\hat{f}(t + \Delta t) - f(t + \Delta t)\hat{f}(t - \Delta t)}{f(t - \Delta t)f(t + \Delta t) + \hat{f}(t + \Delta t)\hat{f}(t - \Delta t)} \right],$$

(3)

where $\Delta t$ denotes the time between observations. This is the method that we will use.

Figure 2 illustrates the instantaneous amplitude $H(t, \omega)$ and period $P(t)$ corresponding to the first IMF shown in Figure 1. The solid and dashed lines give the results for the original signal and Shapiro filtered signal, respectively. In both cases, the period equals approximately 12.5 h outside of the region from 125 and 150 hours. Within the interval there are considerable differences between the original and smoothed data. Because analytic signals cannot have negative frequencies, the appearance of some suggests that the data is not correct here. For this reason, we will discard any amplitudes and periods when they are negative. This is acceptable because the amplitude when the frequencies are negative is small, as Figure 2 shows.

Once we obtained the analytic signal for each IMF, the final task remains to graphically display the results as a time-frequency plot. Because most geophysical data sets contain a mixture of phenomena with time scales varying from hours to years, it was found convenient to work with the logarithm of the instantaneous period rather than period itself.

One problem with displaying the results is the irregularity of the location of the frequencies associated with each IMF at a given instant. Not only are these locations irregular but they also vary with time. Although most graphical packages can handle such irregularly spaced data, the plots are too noisy. One possible solution would be to construct a smoothed field before plotting.
If we view our instantaneous amplitudes as "data" observed at various instantaneous periods, then statisticians have developed regression techniques for fitting a curve through data (Ryan 1997, Chapter 10). Because we have little a priori knowledge about the shape of the curve, we have chosen to use a nonparametric scheme that includes a kernel smoother. Such a scheme was developed by Herrmann (1996); the Fortran 77 code is available at http://www.unizh.ch/biostat/Software/kernf77.html.

For large data sets, this method was modified to include time averaging. There are two possible methods. One method would group all of the instantaneous amplitudes and periods within a particular time interval and assign them a common time. The second method would find curve fits at each time and then these curves would be averaged over an appropriate time interval. Figure 3 shows the instantaneous amplitudes and periods at four instances during the 24 hour interval from 2330 LST 24 September 1985 to 2330 LST 26 September 1985 using sea level data. The solid gives the grouping of instantaneous amplitudes and phases to form a single time value while the dashed line shows the averaging of the curve fits. The first method retains the character of the higher temporal resolution amplitudes compared to the second method.

As we shall shortly see, time-frequency plots contain a wealth of detail. For that reason, it is useful to integrate the square of the instantaneous amplitude, $H^2(t,\omega)$, over time, say from 0 to $T$. Because this marginal spectrum represents the energy of the signal, it is analogous to the power spectrum in Fourier analysis. Here a monochromatic, linear and periodic signal appears as a sharp peak in the marginal spectrum while a nonlinear and nonperiodic phenomenon yields a broad peak in the spectrum. If the marginal spectrum
is normalized by $T$, we have the average marginal spectrum.

3. Applications

Having presented the procedure for constructing a Hilbert-Huang transform, we now apply this technique to three very different data fields: sea level heights, incoming solar radiation, and barographic observations.

a. Sea level heights

As our first application of the Hilbert-Huang transform to geophysical data, observations were obtained of the sea level heights observed at the Chesapeake Bay Bridge and Tunnel (CBBT) from 0000 LST 8 April 1984 to 2300 LST 28 September 1990. These 48,000 hours of consecutive observations contain a wealth of physical phenomena. A simple Fourier analysis reveals a semidiurnal tide with a period of 12.42 h and a diurnal tide with a period of 23.93 h. Furthermore, because the station lies on the East Coast of the United States, the observations are affected by coastal storm (Northeasters), even hurricanes. Finally, Frison et al. (1999) have estimated that nonlinear effects, such as meteorological forcing, tidal interactions and river outflow, may account for 30% of the measured water level at CBBT.

Figure 4 shows the average marginal spectrum for the sea level heights. Peaks A and B correspond to the semidiurnal and diurnal tides, respectively. This is consistent with a study by Susanto et al. (2000) who showed that Hilbert-Huang transforms capture the diurnal and semidiurnal tides in the Makassar Straits. Peaks C–G arise from the lunar contribution to the tides. Because the Chesapeake Bay is a shallow estuary (average depth
of 8.4 m), the peaks F, E, D, and C are the harmonic overtones of the fundamental located at peak G.

The peaks H, I, and J are the effects of the seasons on the water level. It is noteworthy that these peaks are broader than the peaks associated with the tides. This suggests that the corresponding physical processes are nonstationary and/or nonlinear. As we shall shortly show, these peaks arise due to variable river discharges into the Bay.

Figure 5 shows a contour plot of instantaneous amplitudes as a function of time and period. A persistent maximum is found at approximately 12 h; this is the semidiurnal tides. We note a modulation in the amplitude that has a period of approximately one month. This is the lunar forcing of the tides. There is also a faint green line at 24 h (denoted by the letter A) which is the diurnal tides.

At longer periods there are many local maxima. Two particular large ones have been highlighted by the letters B and C. Maximum B occurs during mid-November 1985 and is associated with the "Election Day Flood", 4–7 November 1985, the second worst river flood in the twentieth century. During this event, heavy rain occurred in the mountainous regions of Virginia and West Virginia and we are seeing the corresponding runoff.

February 1989 was the snowiest month (62 cm) in fifty years (1948–1999) for Norfolk, VA, due in large part from its heaviest 24 h snowfall (36 cm) that occurred during 17–18 February 1989. Point C corresponds to the melt and runoff of this snow pack during late February and early March 1989.

In summary, the Hilbert-Huang transform have discern persistent, periodic features such as the tides as well as episodic events such a snow melt and heavy precipitation events.
b. Solar radiation

To compliment our sea level analysis, Hilbert-Huang transforms were applied to the hourly global horizontal radiation measurements from Norfolk, VA from 1 January 1980 through 13 December 1990. These observations, found at http://rredc.nrel.gov/solar/old_data/nsrdb/hourly, include both the direct and diffuse solar radiation on a horizontal surface.

The average marginal spectrum is shown in Fig. 6. It has two peaks, one corresponding to the diurnal cycle and the other to the season cycle. Unlike the marginal spectrum for the sea level measurements, the peaks are rather broad. This occurs due to the highly nonlinear nature of the signal: During the night it equals zero while during the day it equals a product of trigonometric functions that contain both the diurnal and seasonal cycles.

Figure 7 shows the spectrogram for the solar radiation signal. The maximum values occur along the line corresponding to a period of 24 h, the diurnal cycle, and during the summer months. For periods between $10^2$–$10^3$ h, there are relative minima during the winter months due to the increased cloud cover that reduces the incoming radiation. Finally, for periods near $10^4$ h we clearly see the seasonal cycle.

c. Barographic observations

Our third test involves barographic data averaged to produce observations at 2 minute intervals from a network located in central Illinois (Grivet-Talocia et al. 1999). This data was grouped by station and quarter of the year.
Figure 8 is the average marginal spectrum for the latter half of 1995. The largest peak occurs at 100 h and is due the dominance of baroclinic instability.

Figure 9 is the spectrogram for the final quarter of 1995. Overall the amplitudes increase with time due to the onset of winter. Using the daily weather maps from 12 Z published by NOAA, each maxima was easily correlated with the nearby passage of a cyclone and its associated front; these maxima have been labeled A–F on Fig. 9. For example, maximum A corresponds to the passage of the remnants of Hurricane Opal on 6 October 1995. The passage of a particular strong cold front on 11 November 1995 is highlighted with the letter G.

Because this data has a high temporal resolution, one would hope to see mesoscale signals in it as well as synoptic events. To this end, Figure 10 shows the instantaneous amplitude (in mb) as a function of time (19–29 October 1995) corresponding to disturbances with a period of 1, 2, 3, and 4 h. Again, using the daily weather maps, many of the significant events could be classified as a passage of either a front or trough. There are probably other mesoscale events present in this analysis but the lack of other independent data precludes their classification.

Taken together, Figs. 9 and 10 show how well the Hilbert-Huang transform can discern between various mesoscale and synoptic events.

4. Conclusion

The purpose of data analysis is to discover the physical processes underlying the observations. For generations scientists have only had Fourier analysis. Although this technique is quite good for periodic signals, it is quite inadequate for most meteorological
signals because they are nonlinear and aperiodic. Hilbert-Huang transforms provide a novel technique to analyze and interpret this data.

Recently several new techniques (short-time Fourier transforms, wavelets, and EOFs) have been developed to handle these aperiodic and nonlinear signals. To this list we now add the promising technique, commonly known as Hilbert-Huang transforms, because it is adaptive, local, complete, and nearly orthogonal. So far this technique has only been applied to climatological data sets. Here we tested the method on data sets which also contain a synoptic signal: sea level heights, incoming solar radiation, and barographic observations. These tests showed that the Hilbert-Huang transforms capture a wide variety of phenomena: the diurnal cycle, frontal passages, baroclinic instability, and the seasonal cycle. Therefore, we can use this technique to flag important weather events such as a flood or the passage of a squall line.

In this paper we have highlighted two methods for presenting Hilbert-Huang transforms. One method plots the instantaneous amplitude as a function of time and (instantaneous) period or frequency, the so-called spectrogram. These plots are useful in suggesting the location and nature of significant events in the original data. The second method is the marginal spectrum. It is similar to the popular concept of power spectrum in Fourier analysis. This gives a global picture of the data set and suggests which processes are linear and periodic and which are not.

Acknowledgements

The author would like to gratefully acknowledge Dr. N. E. Huang for allowing the use of his technique and providing considerable insight into the mechanics of how it works.
The barographic data was kindly provided by Dr. Stefano Grivet-Talocia.
References


Salisbury, J. I., and M. Wimbush, 2002: Using modern time series analysis techniques to


Figures

Figure 1: The computation of the first intrinsic mode function from a sea level data set. The solid line in frame (a) shows a portion of the original data with sea level height in m. The solid lines in frames (b) and (c) show the first intrinsic mode when the mode's extrema and zero crossings are equal or differ by one for the first and third time, respectively. Frame (d) illustrates the first IMF from the original data set (solid line) and when the original data is Shapiro filtered (dotted line). Frame (e) gives the signal after the first intrinsic mode has been removed.

Figure 2: The solid line gives the amplitude and period of the first IMF shown in Figure 1. The dashed line also gives the amplitude and period of the first IMF of the signal shown in Figure 1 but the data has first been filtered by a Shapiro filter.

Figure 3: Using sea level data, the curve fit when instantaneous amplitudes and periods over a 24 hour interval are grouped together to form a single Hilbert-Huang transform (solid line). The dotted line is the average of 24 curve fits to each hour's instantaneous amplitudes and periods. Some of the instantaneous amplitudes used in computing the averages are plotted as data points.

Figure 4: Average marginal spectrum derived from 48,000 hours of sea level observations at the mouth of the Chesapeake Bay.

Figure 5: Contour plot of the (natural) logarithm of the instantaneous amplitudes of sea level height in cm as a function of time and period. The amplitudes were time-averaged as described in the text over 480 h.
Figure 6: Marginal spectrum derived from 96,000 hours of direct and diffuse solar radiation measured on a horizontal surface at Norfolk, VA.

Figure 7: Contour plot of the (natural) logarithm of the instantaneous amplitudes of direct and diffuse solar radiation in W/m² as a function of time and period. The amplitudes were time-averaged as described in the text over 480 h.

Figure 8: Marginal spectrum derived from barographic observations near Champaign-Urbana, IL during the latter half of 1995.

Figure 9: Contour plot of the (natural) logarithm of the instantaneous amplitudes of surface pressure in mb as a function of time and period. The amplitudes were time-averaged as described in the text over 128 minutes.

Figure 10: Line plots of the instantaneous amplitudes of surface pressure in mb as a function for periods of (a) 1 h, (b) 2 h, (c) 3 h, and (d) 4 h. The letters T and F denote the passage of a trough or front as derived from daily weather maps. The amplitudes were time-averaged as described in the text over 16 minutes.
Figure 1: The computation of the first intrinsic mode function from a sea level data set.

The solid line in frame (a) shows a portion of the original data with sea level height in m. The solid lines in frames (b) and (c) show the first intrinsic mode when the mode's extrema and zero crossings are equal or differ by one for the first and third time, respectively. Frame (d) illustrates the first IMF from the original data set (solid line) and when the original data is Shapiro filtered (dotted line). Frame (e) gives the signal after the first intrinsic mode has been removed.
Figure 2: The solid line gives the amplitude and period of the first IMF shown in Figure 1. The dashed line also gives the amplitude and period of the first IMF of the signal shown in Figure 1 but the data has first been filtered by a Shapiro filter.
Figure 3: Using sea level data, the curve fit when instantaneous amplitudes and periods over a 24 hour interval are grouped together to form a single Hilbert-Huang transform (solid line). The dotted line is the average of 24 curve fits to each hour's instantaneous amplitudes and periods. Some of the instantaneous amplitudes used in computing the averages are plotted as data points.
Figure 4: Average marginal spectrum derived from 48,000 hours of sea level observations at the mouth of the Chesapeake Bay.
Figure 5: Contour plot of the (natural) logarithm of the instantaneous amplitudes of sea level height in cm as a function of time and period. The amplitudes were time-averaged as described in the text over 480 h.
Figure 6: Marginal spectrum derived from 96,000 hours of direct and diffuse solar radiation measured on a horizontal surface at Norfolk, VA.
Figure 7: Contour plot of the (natural) logarithm of the instantaneous amplitudes of direct and diffuse solar radiation in $\text{W/m}^2$ as a function of time and period. The amplitudes were time-averaged as described in the text over 480 h.
Figure 8: Marginal spectrum derived from barographic observations near Champaign-Urbana, IL during the latter half of 1995.
Figure 9: Contour plot of the (natural) logarithm of the instantaneous amplitudes of surface pressure in mb as a function of time and period. The amplitudes were time-averaged as described in the text over 128 minutes.
Figure 10: Line plots of the instantaneous amplitudes of surface pressure in mb as a function for periods of (a) 1 h, (b) 2 h, (c) 3 h, and (d) 4 h. The letters T and F denote the passage of a trough or front as derived from daily weather maps. The amplitudes were time-averaged as described in the text over 16 minutes.