# Application of Empirical Mode Decomposition to the detection of Sudden Ionospheric Disturbances by monitoring the signal of a distant Very Low Frequency transmitter

by Lionel Loudet[1]

First release: 30 October 2009 — Version 2: 5 December 2009

## Contents



## Abstract

Sudden Ionospheric Disturbances (SIDs) result from an abnormal increase of the ionization density of the D region of the ionosphere originating from x-ray solar flares. They can affect radio communications and possibly lead to a radio blackout.

A classic means to detect the most powerful solar flares and resulting SIDs is to monitor the amplitude and phase of distant Very Low Frequency (VLF) transmitters.

The Hilbert-Huang transform (HHT) and the Empirical Mode Decomposition (EMD) have offered a breakthrough in the analysis of nonlinear and non-stationary data. This signal processing technique has been used with great success in many applications where traditional methods do not apply.

This document describes how the Empirical Mode Decomposition (EMD) technique helps in the detection of SIDs from recordings of the signal levels of VLF stations. The author has written software to apply EMD to actual data.

By comparing with GOES spacecraft measurements of solar x-ray flux, the study confirms that a careful choice of an Intrinsic Mode Function (IMF) allows extracting the uncommon changes of the signal strength resulting from x-ray solar flares.

With respect to the analysis of raw measurements, this study describes how a more precise timing of the disturbance is obtained. In addition, the extraction of faint changes from the noise is improved. The detection threshold appears reduced from the usual C1 to approximately B5.

## Introduction

The propagation of Very Low Frequency (VLF) radio communications is affected by the ionospheric disturbances. They are triggered by the sudden release of high-energy radiation from x-ray solar flares. Measurement of the amplitude and/or phase of the signal of a distant transmitter is a means to detect from the ground the most powerful solar flares and to characterize their effects in the Earth vicinity.

Firstly, this paper presents background information on x-ray solar flares and the specificities of VLF signal propagation. It shortly describes how the ionosphere can be affected by solar flares and create Sudden Ionospheric Disturbances (SIDs). The author's monitoring station, used to record the signal strength data, is briefly presented.

Then, the basic principles of the Hilbert-Huang Transform (HHT) and its key part, the Empirical Mode Decomposition (EMD), are explained. The algorithm is detailed and the software written by the author to analyze the data is described.

Finally, the EMD has been applied to actual data representative of typical recordings. Several examples, from a quiet day with no event to a very active day with many disturbances, are presented. The objective is to illustrate how the various IMFs resulting from the EMD are affected by SIDs of various importance. This paper describes how the components that correctly represent the propagation anomalies are chosen. By comparing with GOES spacecraft measurements of x-ray flux, the validity and consistency of the analysis is demonstrated.

When compared with the analysis of raw measurements, this method lessens the ambiguities in the detection of abnormal conditions. Also, the timing of the disturbance is more precisely obtained, especially for the end time that is usually difficult to determine with confidence, the exponential decay of the event leading to an unclear transition between the disturbed and normal propagation conditions. In addition, the results suggest that, with respect to the traditional data reduction method, the detection of fainter events is achievable. Very small disturbances, barely visible on original recordings, have been successfully extracted. Flares as small as B5—instead of the usual limit around C1—seem to produce, under some conditions, small VLF propagation disturbances.




1 Contact: http://sidstation.lionelloudet.homedns.org/email-en.xhtml

## Background

### X-ray Solar Flares

A solar flare results from a sudden release of energy stored in coronal magnetic fields. It produces a burst of radiation across the electromagnetic spectrum, from radio waves to gamma rays. The x-ray flux received at the Earth increases significantly within a few minutes and then returns to normal values during the recovery phase that lasts from a few tens of minutes to several hours. These x-rays are absorbed before reaching the ground. X-ray detectors on the GOES spacecraft are used to measure the x-ray flux in the Earth vicinity. The solar flares are classified—A, B, C, M or X—according to their x-ray brightness in the wavelength range 1 to 8 Å.

| Class | Peak Intensity $I$ in W/m² between 1 and 8 Å |
|---|---|
| **A** | $I \leqslant 10^{-7}$ |
| **B** | $10^{-7} \leqslant I < 10^{-6}$ |
| **C** | $10^{-6} \leqslant I < 10^{-5}$ |
| **M** | $10^{-5} \leqslant I < 10^{-4}$ |
| **X** | $I \geqslant 10^{-4}$ |

*Table 1: X-ray solar flares classification. Each class has a peak flux ten times greater than the preceding one.*

X-class flares are the most powerful. They can trigger radio blackouts across the whole planet. M-class flares can cause brief radio blackouts affecting mainly polar regions. Other flares have few noticeable consequences on Earth.

The peak intensity of a flare is indicated by a number after its class. For instance, a M3.1 flare has a peak intensity of $3.1 \cdot 10^{-5}$ W/m².

Depending on the Sun activity, solar flares occurrence varies from several per day to a few per month. It follows the 11-year solar cycle. Figure 1 shows the evolution of x-ray flares during the solar cycle 23.

### VLF Signal Propagation

Very Low Frequency (VLF) radio waves are used for military communications with submarines near the surface, for radio-navigation beacons and for time signals. The propagation characteristics in this part of the electromagnetic spectrum are somewhat different from those observed at higher frequencies.

In the daytime, the lowest part of the ionosphere—the D region—is created through a ionization process resulting from the solar radiation: the Lyman-α emission line (1215.67 Å) ionizes mainly the nitric oxide (NO). The VLF wavelengths are so long that they are conducted in the Earth-ionosphere waveguide (EIWG) between the Earth's surface and the D region. The propagation is very stable. Uncommon variations are used to observe the way the ionosphere is affected by x-rays flares from the sun.

At night, the D region disappears and the waves are refracted by the higher E and F regions. The "reflection" coefficient is higher and leads to increased signal strengths. The monitoring of SIDs is not possible during this period of time.

A typical signal level plot for a quiet day is presented in Figure 2. The sunrise and sunset patterns of the signal amplitude correspond to the transition between the nighttime refraction of the signal and the daytime waveguide propagation mode.

### Sudden Ionospheric Disturbance

A Sudden Ionospheric Disturbance results from an increased ionization density in the D region caused by a solar flare, or possibly by a GRB. When a flare occurs, the

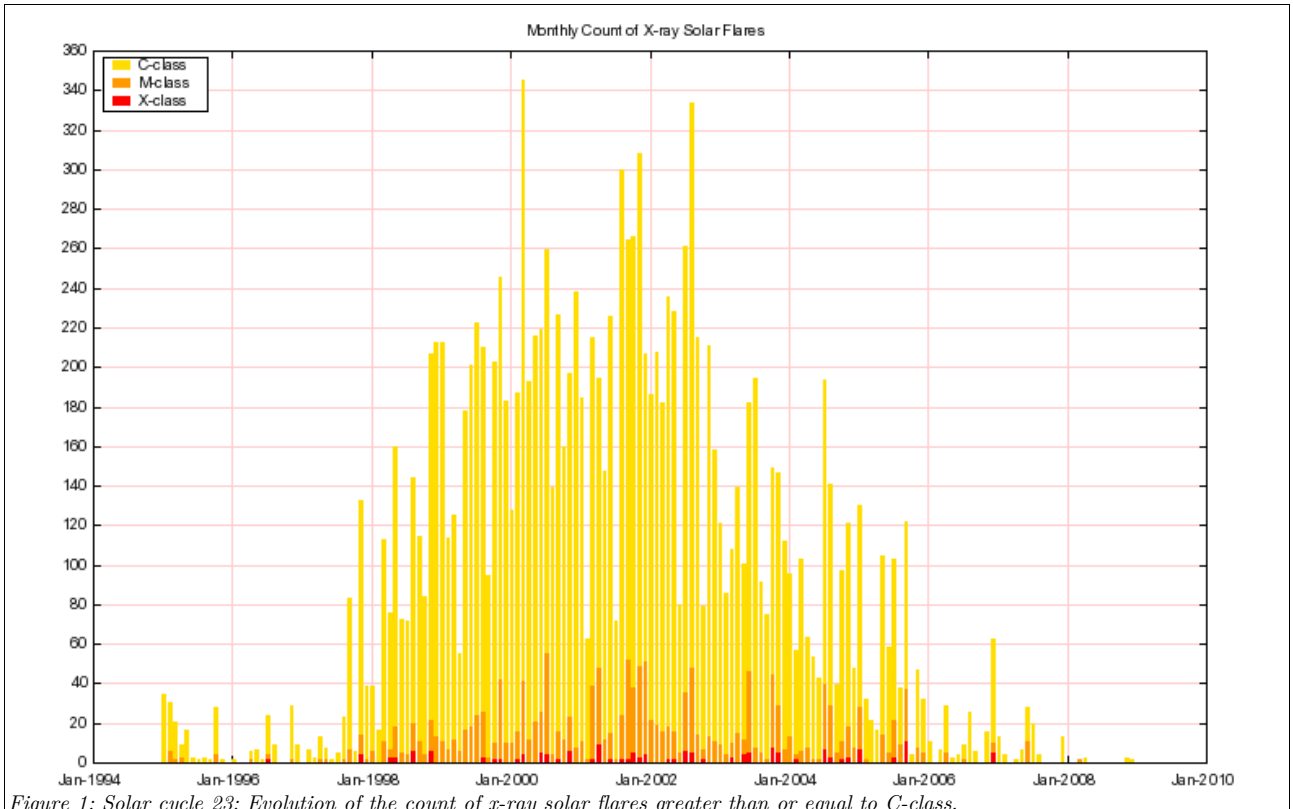


*Figure 1: Solar cycle 23: Evolution of the count of x-ray solar flares greater than or equal to C-class. Several flares per day occur during the solar maximum.*

sunlit side of the Earth is hit by high-energy radiation after a propagation time of about 8 minutes. It will penetrate the D region and ionize all constituents of the air, including $N_2$ and $O_2$. This higher electron density will lead to a rapid increase of radio-wave absorption, especially in the upper MF and lower HF ranges (short wave fading), possibly causing a radio blackout. For VLF (3–30 kHz), the "reflection" coefficient increases, leading to a stronger sky wave. Due to interferences between ground wave (direct propagation) and sky wave (refracted by the D region), the signal strength can increase or decrease during an event. As soon as the event ends, the electrons in the D region recombine and signal strengths return to normal levels. The total duration of the disturbance is from a few minutes to a few hours.

The ionization and recombination processes are not instantaneous. Consequently, the disturbance will be detected with a slight delay with respect to the arrival time of the x-ray photons (as measured by GOES satellites) and will last after the end of the flare. Several parameters influence this effect, such as the intensity and the shape of the x-ray flux or the pre-flare ionization level. A delay of 1 to 5 minutes after the start of the flare is usually observed. The ionospheric disturbance can last for an hour or more after the end of the flare.

As an example, a day with two solar events is shown Figure 3.

### Description of the SID monitoring station

The author's station is located in the South of France and monitors the VLF transmitter DHO38 in the North of Germany at a distance of about 1200 km. The signals are received through a loop antenna. The receiver contains an order-4 active filter centered on the transmitter frequency (23.4 kHz). A linear detector (full-wave rectifier and peak detector) is used to get the signal amplitude value. This amplitude value is then filtered (the filter time constant is around 1 minute) and converted through a 12-bits analog-to-digital converter. The station is referenced under the AAVSO (see [4]) observer ID A-118.

Taken into account the distance between the transmitter and the monitoring station, the sky wave propagation path has only one hop. The sub-reflective point is located in the Northeast part of France. The duration of the SID detection period (daytime) varies from 8 hours during winter to 16 hours during summer.

This monitoring technique allows detecting solar flares of class C1 and higher (above 1 µW/m²).

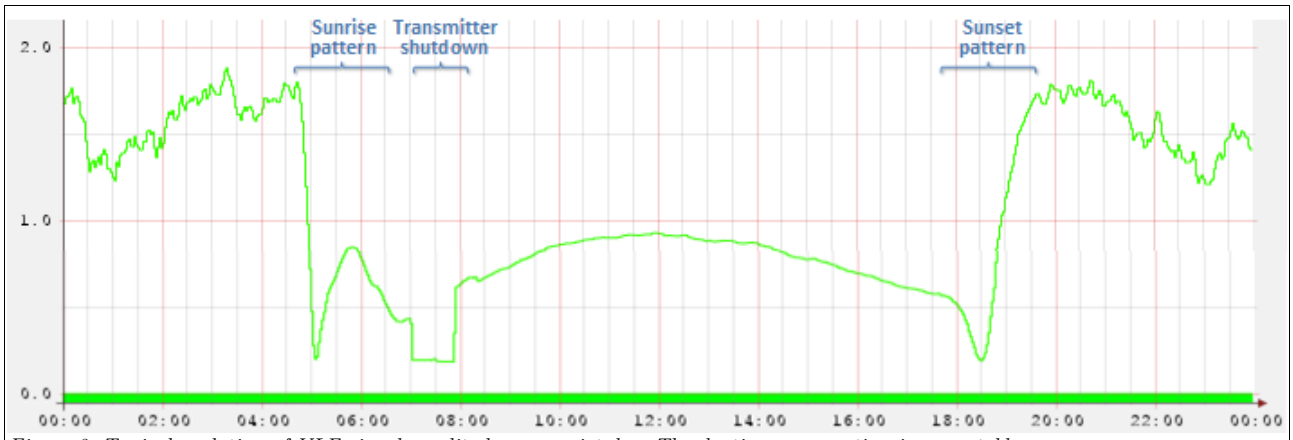


*Figure 2: Typical evolution of VLF signal amplitude on a quiet day. The daytime propagation is very stable.*


*Figure 3: Effect of two solar flares. The SIDs led to sudden increase of the received signal followed by a slow decay.*

## The Hilbert–Huang transform

### Introduction

The Hilbert-Huang transform (HHT) is a signal processing technique proposed by Huang et al. in 1998. The main advantage of this technique over Fast-Fourier transform (FFT) or Wavelet analysis is that it can analyze nonlinear and non-stationary data. Any frequency analysis of such data generates lots of harmonics and leads to an artificially wide frequency spectrum. Consequently, filtering in the frequency domain cannot be applied without artifacts and deformations, possibly leading to erroneous conclusions.

The HHT uses the empirical mode decomposition (EMD) to decompose the signal into a finite sum of orthogonal eigenmodes called intrinsic mode functions (IMF). The Hilbert spectral analysis is then applied to each IMF to obtain instantaneous frequency information. The EMD is the main novelty of the method. It allows decomposing the dataset into a small number of IMF components (usually, about 10). Each IMF represents only one oscillatory mode for which an instantaneous frequency value can be given.

By definition, an intrinsic mode function (IMF) is a function that satisfies two 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; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. This second condition is a means to ensure that the local mean of the data is zero, thus guaranteeing the meaningfulness of the instantaneous frequency analysis.

An energy-frequency-time distribution can then be obtained by applying the Hilbert spectral analysis to the set of IMFs. This part of the HHT is not needed for our purpose.

The HHT is used in biomedical applications, chemistry engineering, oceanography, meteorology, seismic studies, etc.

For those interested in more details regarding the HHT, the reference paper [2] is worth being read.

EMD Algorithm

The empirical mode decomposition is based on a process called "sifting":

1. Initialization: $r_0(t)=X(t)$ where $X(t)$ is the data to process, $j=1$
2. Sifting Process for the extraction of the $j$-th IMF:
   a. Initialize $h_0(t)=r_j(t)$, $k=1$ ;
   b. Determine local maxima and minima of $h_{k-1}(t)$ ;
   c. Calculate upper $E_{max}(t)$ and lower $E_{min}(t)$ envelopes of $h_{k-1}(t)$ through cubic spline interpolation;
   d. Calculate the mean of the two envelopes $m_{k-1}(t)=\frac{E_{max}(t)+E_{min}(t)}{2}$ ;
   e. Subtract the mean to obtain a "proto-IMF" $h_k(t)=h_{k-1}(t)-m_{k-1}(t)$ ;
   f. Check if IMF criteria completeness (see Note ①) are met for $h_k(t)$ .
   If yes, then $imf_j(t)=h_k(t)$ ;
   else repeat the sifting process (at step 2b) with $k=k+1$ .
3. Calculate the new signal under examination $r_j(t)=r_{j-1}(t)-imf_j(t)$ ;
4. Check if $r_k(t)$ meets the EMD completion criteria (see Note ②).
   If yes, calculate the next IMF (step 2) with $j=j+1$ ;
   else the decomposition is finished.

At the end of the decomposition, the following DATA are available:

- a set of $j$ IMF components: $imf_1(t)$ … $imf_j(t)$ ;
- the residue: $r_j(t)$ .

Note ①: The criteria for IMF validity are:

- a predefined number of consecutive siftings lead to proto-IMF with the same number of zero crossings, minima and maxima, and the number of zero-crossings and extrema are equal or at most differ by one;
- no more IMF can be extracted (No more than one maximum + one minimum);
- The standard deviation of the proto-IMF falls below a given threshold.

Note ②: The criteria for EMD completion are:

- max number of iterations reached;
- the residue is monotonic (the residue has no more than one minimum + maximum);
- the last calculated IMF is monotonic (the IMF has no more than one minimum + maximum);
- no more IMF can be extracted (the residue or the last calculated IMF has no more than one minimum + one maximum);
- the residue gets too small with respect to a predetermined value;
- the last calculated IMF gets too small with respect to a predetermined value.

### EMD Software

The author has implemented the EMD algorithm in a C program file. The source code is small with less than 1 000 lines of code. The code is ANSI compliant except for the use of the `vsnprintf()` function. It compiles under Linux (`gcc` compiler) and under Windows (Visual Studio C++). The source code is freely available under the GNU General Public License. Source code and Windows binary can be downloaded at [7].

The decomposition algorithm is detailed above. A correct setting of the termination criteria parameters is important to ensure that the components resulting from the decomposition accurately represent the physical properties of the signal.

Here is a summary of the commands and options of the tool:

```
emd v1.0.0 (c) 2009, Lionel Loudet
DESCRIPTION:
  Empirical Mode Decomposition (EMD) of a signal.

USAGE:
  emd [-v(v...)] [-h] [-i <input_file>] [-o <output_file>]
      [-s <s_number>] [-d <std dev>] [-n <max imf count>] [-e <epsilon>]
      [-a <aberrant_limit]

  Options:
    v       Increase verbosity level.
    h       Shows this message.
    i       Input filename. Defaults to stdin.
            File must contain two columns (x and y) of real data.
    o       Output filename. Defaults to stdout.
            Output file contains the following columns:
            <data x> <data y> <imf 1> ... <imf N> <residue>.
    s       'S number'. Used as termination criteria.
            Strictly positive integer. Defaults to 2.
    d       Maximum Standard Deviation. Used as termination criteria.
            Strictly positive real number. Defaults to 0.3.
    n       Maximum number of intrinsic mode function (IMF) to search for.
            Strictly positive integer. Defaults to 16.
    e       Tolerance on RMS value of IMF. Used as termination criteria.
            Strictly positive real number. Defaults to 0.000 000 001.
    a       Threshold for aberrant values suppression. Difference between two
            consecutive data must not exceed the specified value.
            Aberrant values suppression is not enabled by default.
```

### Example of EMD

This section describes how the `emd` software processes actual VLF signal level data. By increasing the verbosity level (`-vvv` option), the software shows the steps of the decomposition:

```
$ emd -i dho38-20070710.txt -o dho38-20070710.out.txt -vvv
Reading file dho38-20070710.txt ...
Read 42250 entries.
Settings: S_number 2. Max Std Dev 0.30. Max IMF 16.
          Epsilon 1.00e-09. Aberrant Limit 1.80e+308.
Starting Analysis ...
IMF 1 - Step 1 [12628 zeros, 8189 min, 8189 max, Std Dev 1.000]
IMF 1 - Step 2 [16382 zeros, 8718 min, 8718 max, Std Dev 0.786]
IMF 1 - Step 3 [17385 zeros, 8913 min, 8913 max, Std Dev 0.921]
IMF 1 - Step 4 [17783 zeros, 9045 min, 9045 max, Std Dev 0.046]
Found IMF 1.
IMF 2 - Step 1 [5459 zeros, 3146 min, 3146 max, Std Dev 1.002]
IMF 2 - Step 2 [6255 zeros, 3262 min, 3262 max, Std Dev 0.928]
IMF 2 - Step 3 [6505 zeros, 3323 min, 3323 max, Std Dev 0.108]
Found IMF 2.
IMF 3 - Step 1 [1906 zeros, 1137 min, 1137 max, Std Dev 0.998]
IMF 3 - Step 2 [2263 zeros, 1189 min, 1189 max, Std Dev 0.605]
IMF 3 - Step 3 [2363 zeros, 1210 min, 1210 max, Std Dev 0.120]
Found IMF 3.
IMF 4 - Step 1 [651 zeros, 402 min, 402 max, Std Dev 0.997]
IMF 4 - Step 2 [795 zeros, 427 min, 427 max, Std Dev 0.571]
IMF 4 - Step 3 [845 zeros, 435 min, 435 max, Std Dev 0.185]
Found IMF 4.
IMF 5 - Step 1 [173 zeros, 124 min, 123 max, Std Dev 1.008]
IMF 5 - Step 2 [236 zeros, 143 min, 142 max, Std Dev 0.226]
Found IMF 5.
IMF 6 - Step 1 [43 zeros, 26 min, 27 max, Std Dev 1.001]
IMF 6 - Step 2 [52 zeros, 28 min, 29 max, Std Dev 0.138]
Found IMF 6.
IMF 7 - Step 1 [18 zeros, 8 min, 9 max, Std Dev 1.056]
IMF 7 - Step 2 [18 zeros, 8 min, 9 max, Std Dev 0.078]
Found IMF 7.
IMF 8 - Step 1 [6 zeros, 3 min, 2 max, Std Dev 0.510]
IMF 8 - Step 2 [6 zeros, 3 min, 2 max, Std Dev 0.021]
Found IMF 8.
Monotonic residue found.
Analysis done. Total of 8 IMF found.
Saving data to dho38-20070710.out.txt
Done
```

In this example, the sifting process leads to a decomposition into 8 IMFs. This value appeared consistent for the whole recordings of signal levels that were analyzed. We can see that few components are necessary to completely represent the original data.

An IMF is deemed valid its when standard deviation gets below a predefined threshold (0.3 in this case). When searching for a new component, from two to four siftings steps are necessary to obtain an IMF meeting the validity criteria.

This example shows how the number of extrema and zero-crossings reduces from IMF 1 to IMF 8, thus ensuring the convergence of the algorithm. Each new IMF contains a lower-frequency oscillatory mode. The decomposition is stopped when a monotonic residue is found.

## EMD of VLF amplitude measurements

### Example 1: Quiet day

In this section, we will analyze an example representative of a quiet day. Figure 4 shows the signal levels of DHO38 recorded on September 09, 2006.

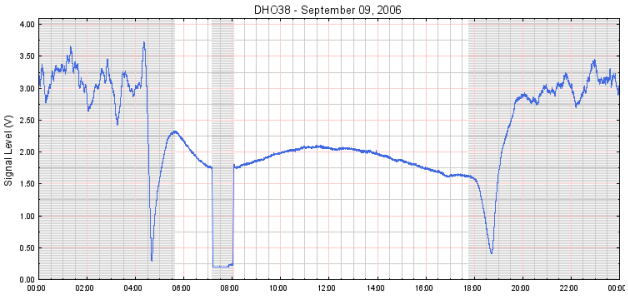


*Figure 4: Typical example of the evolution of the DHO38 signal amplitude on a quiet day (September 09, 2006) with stable propagation during the day.*

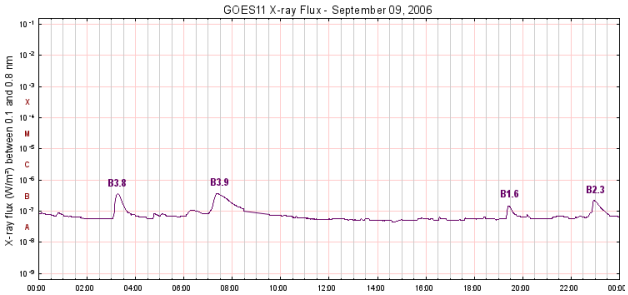


*Figure 5: GOES11 x-ray flux measurements on September 09, 2006.*

On that day, the x-ray flux has not exceeded B3.9 (see GOES data, Figure 5). Table 2 is the official list of the flares from the NGDC database ([3]). None of the flares triggered SIDs.

| Flare Time (UTC) | | | Class |
|---|---|---|---|
| Start | Max | End | |
| 03:04 | 03:16 | 03:26 | B3.8 |
| 07:10 | 07:22 | 07:46 | B3.9 |
| 19:19 | 19:25 | 19:34 | B1.6 |
| 22:50 | 22:57 | 23:09 | B2.3 |

*Table 2: List of x-ray flares on September 09, 2006. None of these flares triggered SIDs.*

Figure 7 shows how VLF signal amplitude data recorded during a quiet day is decomposed into 8 IMFs plus the residue.

One important point when applying the empirical mode decomposition algorithm is to validate its representativeness. The IMF components must represent the physical mechanisms embedded in the original signal. The criteria chosen to validate each IMF and to terminate the decomposition must be carefully determined. The sifting process has to be applied with care to be sure that the intrinsic oscillatory modes are extracted but that no processing artifacts are introduced. The empirical mode decomposition is somehow intuitive on that aspect, but the use of default settings usually gives excellent results.

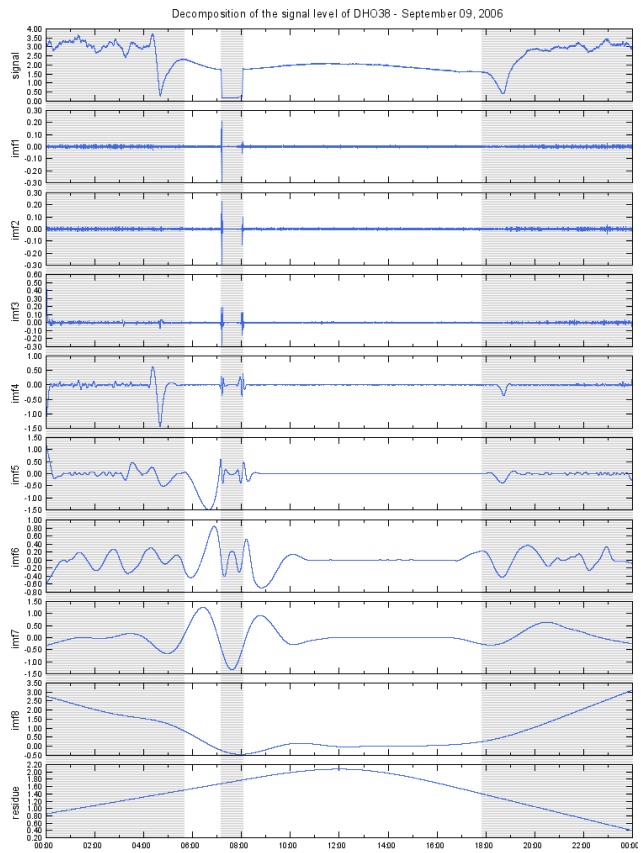


*Figure 7: EMD Analysis of VLF signal strength. Each IMF represents lower frequencies than the previous one.*

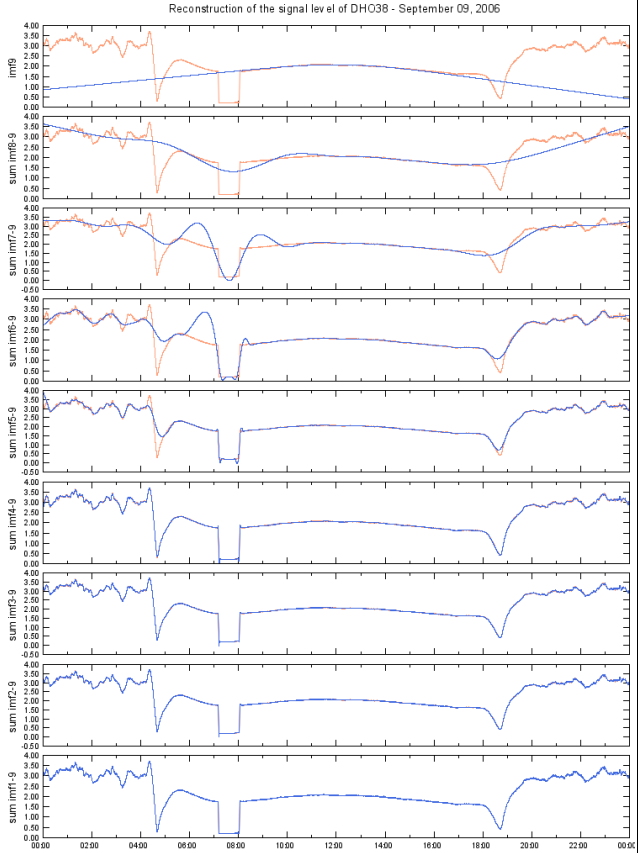


*Figure 8: Reconstruction of the original signal by adding successively high-order to low-order IMF components.*

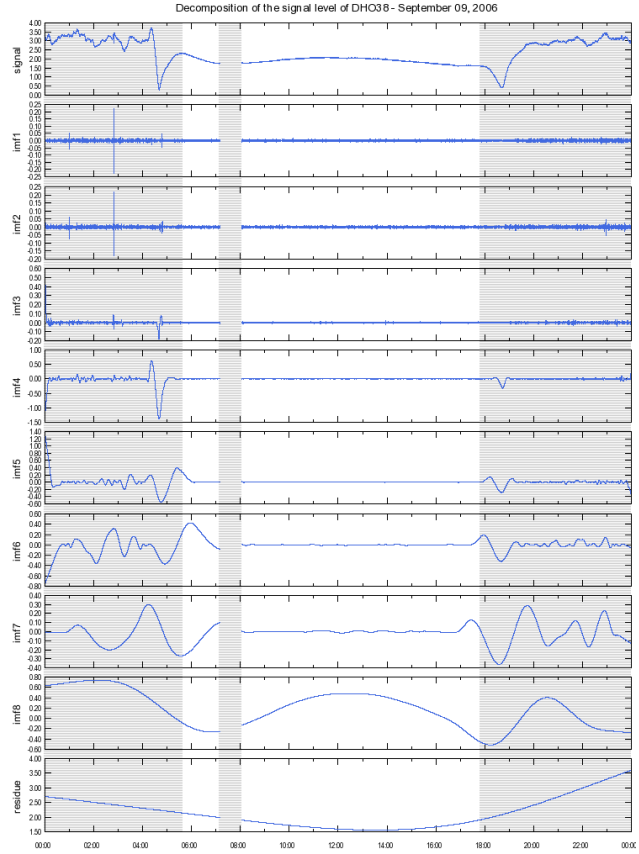


*Figure 6: EMD of VLF signal strength when ignoring data recorded during the transmitter shutdown.*

If we analyze more in depth the decomposition shown in Figure 7, we can see that:

- the low-order IMF 1, 2 and 3 contain only noise and artifacts from the sharp transitions of the signal at the beginning and end of the daily transmitter shutdown (from 07:00 to 08:00 UTC). We can consider that these three IMFs are the representation of the noise of the signal amplitude measurements. They will not contain information related to the presence of a disturbance and can be ignored.
- the IMF 4 to 6 contain oscillatory modes corresponding to the sunrise/sunset patterns and night-time fluctuations. The sharp transitions at 07:00 and 08:00 UTC still create artifacts. This subset of components represent mid-term fluctuations of the signal measurements that are clearly correlated with the underlying physical properties of the ionospheric propagation.
- the IMF 7 and 8 as well as the residue are the representation of the evolution of the propagation characteristics between the day and the night. These high-order components represent long-term fluctuations of the VLF signal propagation.

The original signal can be reconstructed by adding the IMFs and the residue (refer to Figure 8). The original data is shown in orange and the blue plots represent the successive sums of IMF components. At each step, another IMF representing a finer structure is added. At the end of the process, the original data is exactly recovered.

The EMD is a means to filter or de-noise data. If the components that contain only noise information are not included in the reconstruction process, the signal obtained is a filtered version of the original data. The more low-order

IMF components are excluded from the reconstruction process, the more aggressive is the filter.

The artifacts observed in the components around 07:00 and 08:00 UTC are created by the shutdown of the transmitter and can be suppressed by ignoring data recorded during that period of time. The cubic spline interpolation used by the decomposition process handles very well an irregular sampling in the dataset.
Figure 6 shows the components resulting from the empirical mode decomposition when suppressing all noise information between 07:00 and 08:00 UTC. This plot can be directly compared with Figure 7. One can clearly see that the artifacts are no longer visible.

Figures 9 and 10 detail this comparison for the IMF component 5.

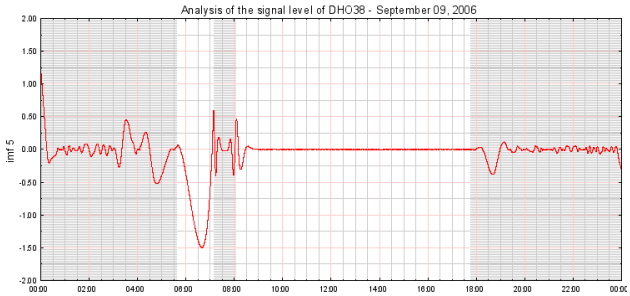

*Figure 9: Evolution of IMF 5 component when taking into account the values recorded during the transmitter shutdown. The component remains close to zero during the day, except for the artifacts before 07:00 and after 08:00.*

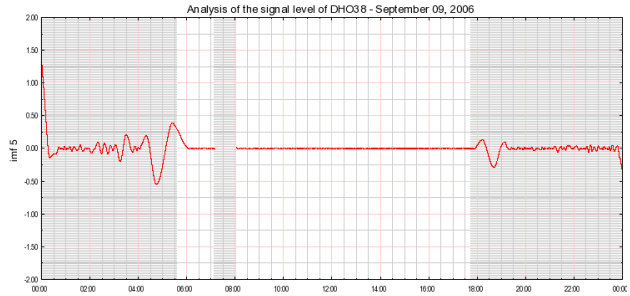

*Figure 10: Evolution of IMF 5 component when ignoring the values between 07:00 and 08:00. All artifacts disappeared.*

It is actually correct to ignore the measurements when there is no signal to detect and just the background noise. We then do not force the decomposition algorithm to model a non-natural phenomenon that has no link with signal amplitude evolutions resulting from changes in propagation conditions.
Consequently, we will use for this study datasets with data removed when the transmitter is shut down.

The next step now is to determine if the uncommon variations created by a sudden ionospheric disturbance are extracted by the analysis, and if so, which components contain this information.

**Example 2: Analysis of a day with one SID**

To determine how the components react to the uncommon variations of the signal amplitude, we will analyze a day with a single event.

Figure 11 shows the measurements made on June 05, 2007 showing a typical disturbance around 16:00 UTC.

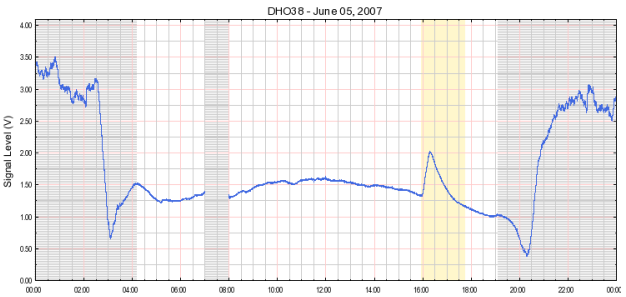


*Figure 11: June 05-2007 measurements of DHO38. The grayed areas indicate approximately the periods of time during which events cannot be detected. They correspond to nighttime and to transmitter shutdown. The SID is highlighted in yellow.*

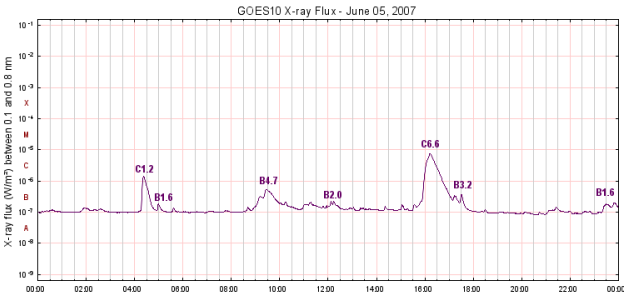


*Figure 12: GOES10 x-ray flux measurements on June 05, 2007 showing the correlation of the C6.6 event with the recorded SID.*

The x-ray flares of that day are listed in Table 3. The SID results from a C6.6 flare. The timing of the disturbance has been measured according to the AAVSO methodology (refer to [5]). The disturbance lasts about 1:40.

| Flare Time (UTC) | | | Class | SID time (UTC) | | |
|---|---|---|---|---|---|---|
| Start | Max | End | | Start | Max | End |
| 04:15 | 04:23 | 04:29 | C1.2 | | | |
| 04:56 | 04:59 | 05:03 | B1.6 | | | |
| 09:20 | 09:27 | 09:39 | B4.7 | | | |
| 12:04 | 12:07 | 12:09 | B2.0 | | | |
| 15:29 | 16:13 | 16:25 | C6.6 | 15:59 | 16:14 | 17:40U |
| 17:26 | 17:31 | 17:33 | B3.2 | | | |
| 23:20 | 23:30 | 23:40 | B1.6 | | | |

*Table 3: List of x-ray flares on June 05, 2007. A C6.6 flare has produced a SID.*

When those measurements are processed through the EMD algorithm, the components shown in Figure 13 are obtained. We clearly see that only the IMF 5 and 6 are affected by the event. The other components embed either noise or long-term oscillatory modes that have no link with the signal disturbance. This is consistent with what has been observed in the previous example.

Here, the IMF 5 is the best representation of the abnormal propagation conditions. This component remains close to zero for the whole daylight period, except during the sudden ionospheric disturbance.

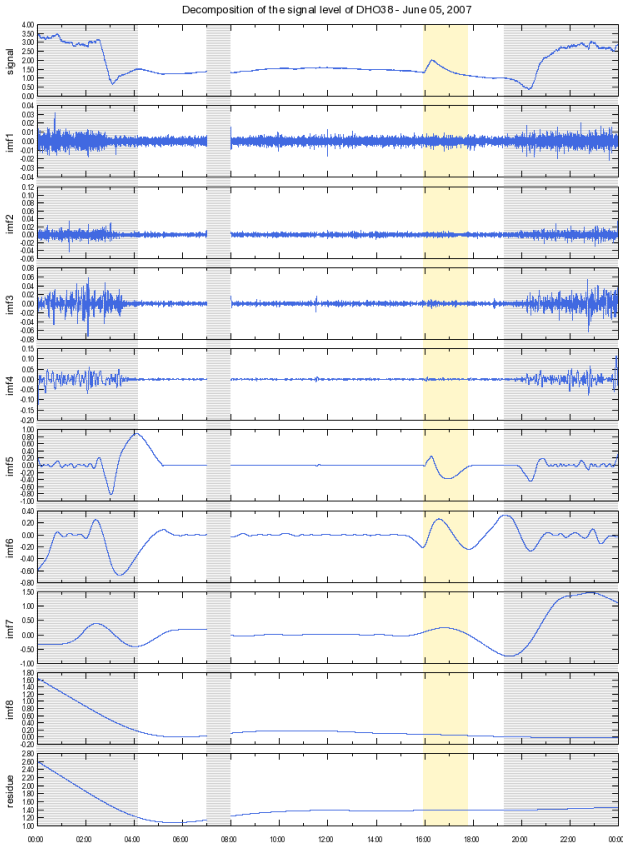


*Figure 13: Components resulting from the EMD of the measurements made on June 05, 2007.*

We then propose to use the IMF 5 as an indicator. We will now check if its combination with other IMF components (lower-order IMF 4 and higher-order IMF 6) could be a better indicator and give more information on the disturbance.

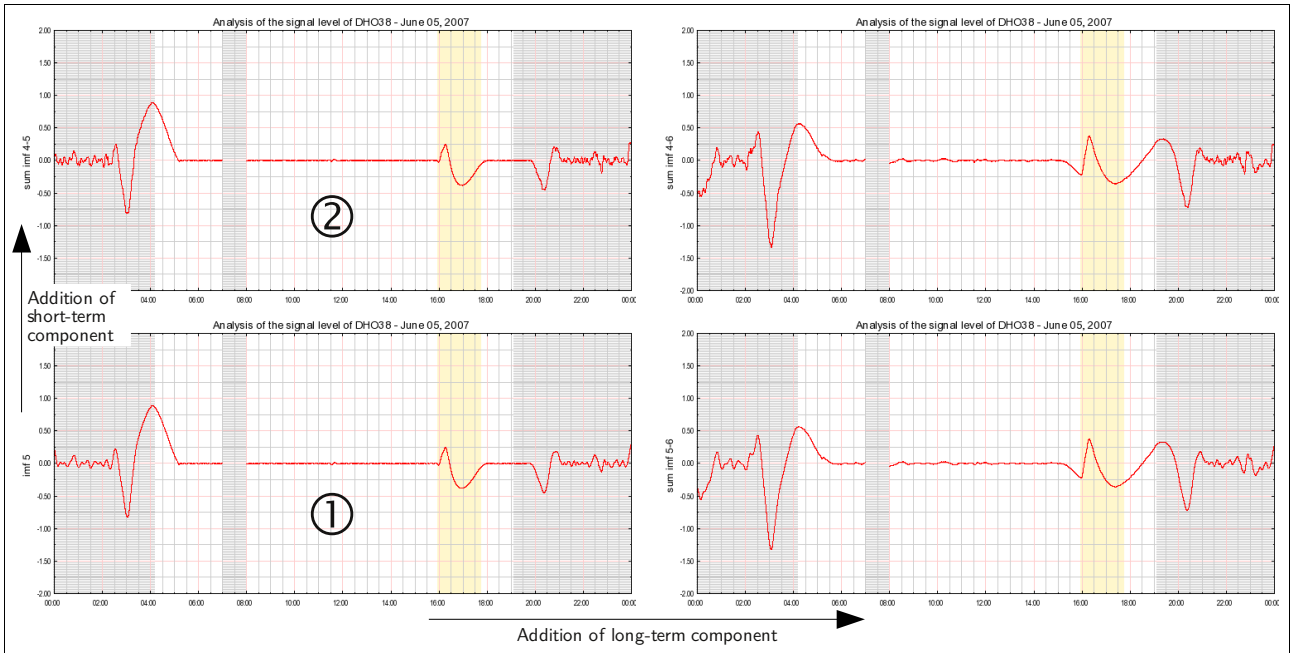


*Figure 14: Comparison of various indicators on June 05, 2007. The reference is the lower left plot (IMF 5). IMF 6 is added in the right-hand plots. IMF 4 is added in the top plots.*

Figure 14 compares several indicators:
① IMF 5 alone;
② IMF 5 + 4 (addition of next short-term component);
③ IMF 5 + 6 (addition of next long-term component);
④ IMF 5 + 4 + 6 (addition of next short and long-term components).

We can see that the addition of the next short-term component is of little effect and does not give more precise information on the event itself.
Also, the addition of the next long-term component has major drawback: the return to zero of the indicator after the event is no longer representative of the event duration.

The IMF 5 is the best representation of the disturbance of the signal amplitude.

One difficulty with the detection of SIDs from VLF amplitude measurements is the precise determination of the start and end times of the abnormal propagation conditions. A precise timing is necessary since the length of the disturbance directly determines the "importance" of the event (see [5]) on a scale from "1-" (disturbance shorter than 19 minutes) to "3+" (longer than 125 minutes).
It is often easy to obtain precisely the start time because the suddenness of the change gives little room for ambiguities. The max time corresponds to the inflection point of the amplitude measurement. Its determination can be more challenging, but a precision of one minute or so is usually achievable. The difficulty arise from obtaining the end time. Usually, the return to normal propagation conditions follows an exponential decay that stems from the recombination mechanism (refer to [1] for the mathematical model). It is consequently rather subjective to give a timing for the end of the SID event.

By observing Figure 14, we can notice the improvement offered by the EMD analysis with respect to the determination of the timing of the event. The return to zero of an IMF component is more evident than the return to a level corresponding to normal propagation conditions of the signal amplitude. In this example, the end of the event had initially been estimated at 17:40 UTC with uncertainties. In fact, the IMF 5 component returns to zero at 17:51 UTC, showing that the event importance had been underestimated by the analysis of the raw signal measurements.

**Example 3: Analysis of a day with two SIDs**

We will now apply the analysis to the data measured on April 05, 2006 (see Figure 15) with two SIDs resulting from C2.6 and C8.1 flares.

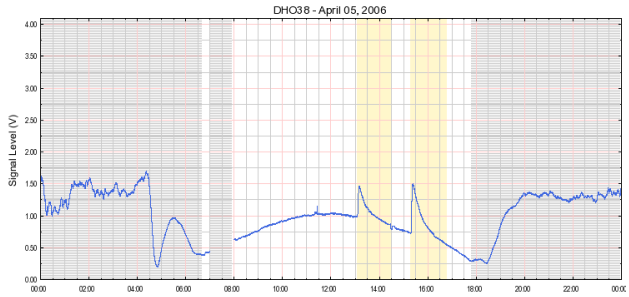


*Figure 15: Signal amplitude of DHO38 recorded on April 05, 2006. Two events resulting from C2.6 and C8.1 x-ray flares have been recorded.*

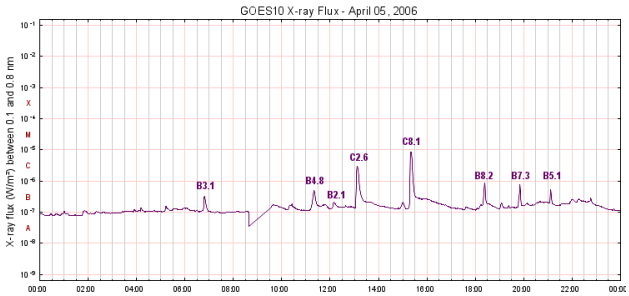


*Figure 16: GOES10 x-ray flux measurements on April 05, 2006*

| Flare Time (UTC) | | | Class | SID time (UTC) | | |
|---|---|---|---|---|---|---|
| Start | Max | End | | Start | Max | End |
| 06:44 | 06:49 | 06:52 | B3.1 | | | |
| 11:13 | 11:20 | 11:23 | B4.8 | | | |
| 12:06 | 12:09 | 12:14 | B2.1 | | | |
| 13:01 | 13:07 | 13:11 | C2.6 | 13:06 | 13:09 | 14:30U |
| 15:14 | 15:21 | 15:24 | C8.1 | 15:18 | 15:21 | 16:50U |
| 18:18 | 18:23 | 18:25 | B8.2 | | | |
| 19:45 | 19:51 | 19:53 | B7.3 | | | |
| 21:03 | 21:07 | 21:09 | B5.1 | | | |

*Table 4: List of x-ray flares on April 05, 2006. Two SIDs have been observed on that day.*

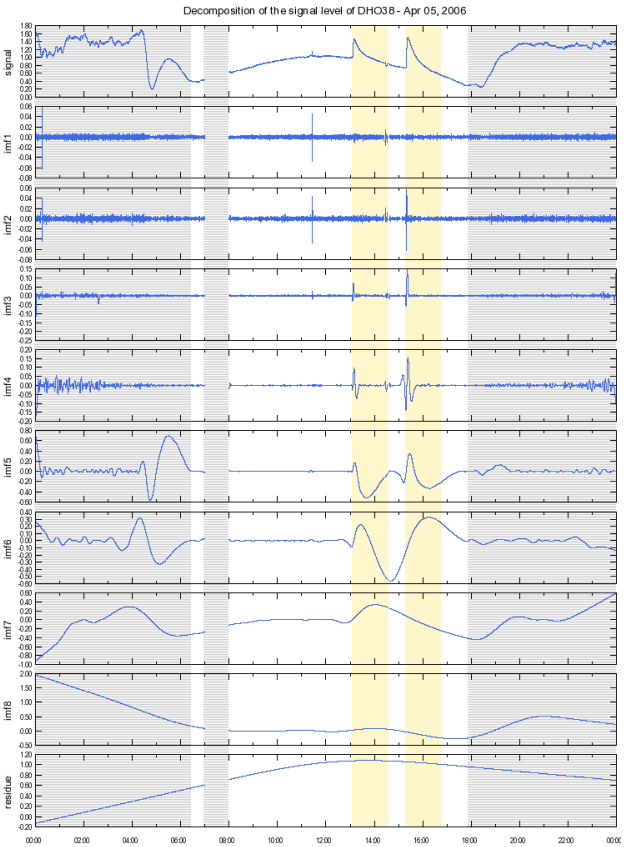


*Figure 17: EMD of the signal recorded on April 05, 2006.*

The components resulting from the empirical mode decomposition are presented in Figure 17. In this case, IMF 4, 5 and 6 show perturbations correlated with the disturbance.

The methodology for comparing the combination of several components is repeated here and presented in Figure 18. It confirms that the IMF 5 alone is still the best indicator and is sufficient to represent the SIDs. There are artifacts created by the proximity of the two flares but they do not prevent the analysis of the data.

By using the IMF 5 component, the end time of the second SID (initially determined approximately at 16:50 UTC) is rather close to 16:15 UTC. The EMD appears again of a great help in alleviating the uncertainties of the end time measurement.

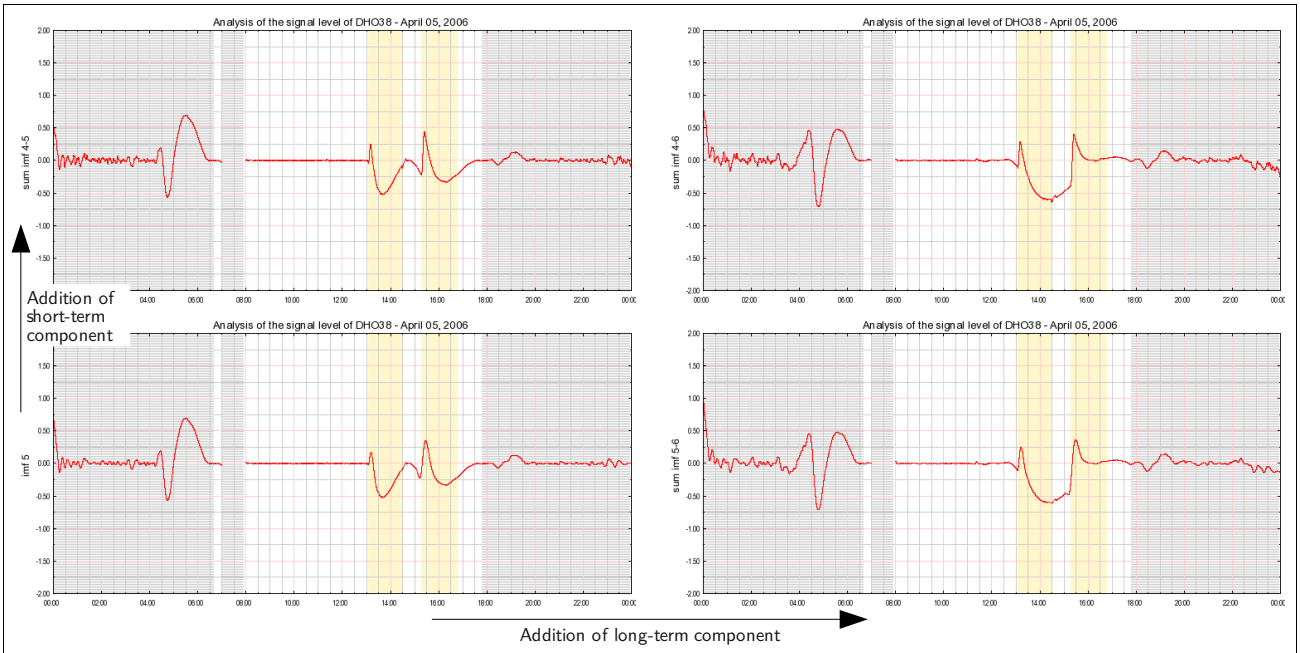


*Figure 18: Comparison of various indicators of the disturbances on April 05, 2006. The reference is the lower left plot (IMF 5). IMF 6 is added in the right-hand plots. IMF 4 is added in the top plots.*

It is interesting to note that the IMF 5 and 6 show a small change around 11:20 UTC. When referring to Table 4, a small B4.8 flare took place at that time. Upon careful examination of the original signal, we can see a very small signal increase lasting about 10 minutes that went unnoticed in the original analysis. A detailed plot is presented Figure 19 with the GOES measurements for comparison.

B-class flares usually affect so slightly the VLF propagation that their effects are hidden in the noise. One can see the great advantage of EMD. Since the components represent the physical properties of the signal, very small effects can be extracted from the noise. The signal-to-noise ratio of a carefully chosen IMF component is vastly improved over the that of the raw signal.

In the previous example (data of June 05, 2007), a flare of similar intensity (B4.7) went unnoticed in both the original signal and the IMF component. When looking more in depth to the GOES flux measurements, the B4.7 flare on June 05, 2007 has a broad shape (it lasts 19 minutes) whereas the B4.8 flare observed on April 05, 2006 has a peak that reaches quickly its maximum amplitude (the event lasts 10 minutes). Studies of the ionization/recombination mechanisms of the D region of the ionosphere and analysis of the influence of various flare shapes on the VLF propagation has been proposed in [1] § 6.5. The different shapes of those two flares of similar levels is a plausible explanation for the differences in the effects on the ionosphere.

This example suggests that, by applying EMD, the traditional limit of C1.0 for the ground detection of solar flares can be reduced to around B5.0. The exact limit depends on the shape of the flux evolution during the flare—sharp peaks affecting more the D region—and probably on the pre-flare ionization level—higher electron density increasing the sensitivity to a new flare.

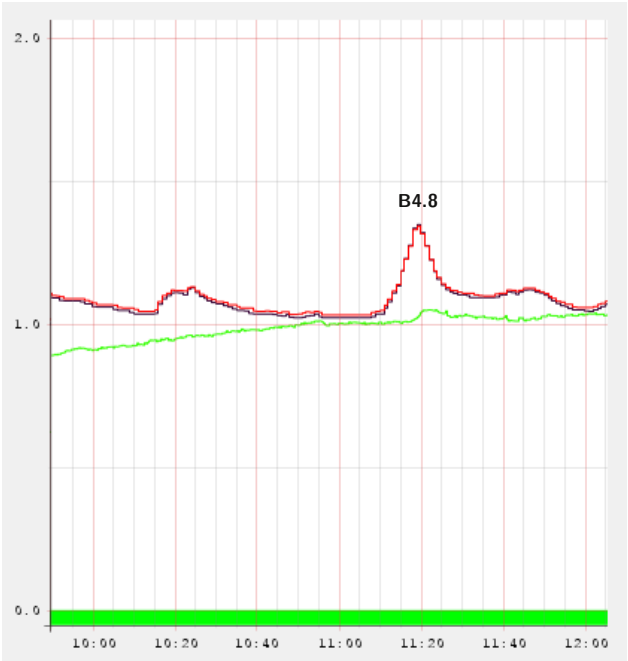


*Figure 19: Detail of the DHO38 signal level during the B4.8 flare on April 05, 2006 and correlation with GOES measurements. x-axis is the UTC time, y-axis represent the measured signal level in volts. The signal measurement is the green plot. Purple and red plots are GOES10 and GOES12 measurements shown on a logarithmic scale. B-class starts at y=1.0 and C-class starts at y=1.5.*

**Example 3: Active day**

The purpose of this section is now to check the behavior of the EMD algorithm in a more complex situation. A day with significant solar activity has been chosen. On July 10, 2007, five events have been detected by the station (see Figure 20). They were triggered by flares ranging from C1.2 to C8.2. To add to the difficulty, one flare took place just before the transmitter shutdown period and another one just after. Unusually, the transmitter has been shut down for only 10 minutes from 07:00 to 07:10 UTC.

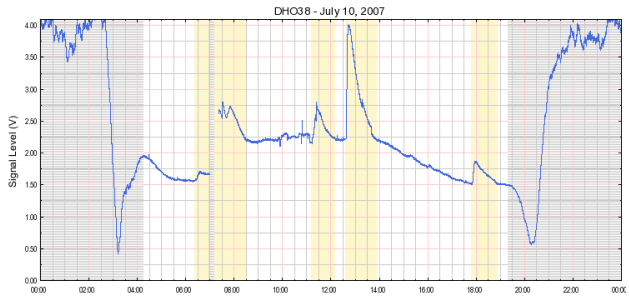


*Figure 20: Five SIDs have been detected on July 10, 2007.*

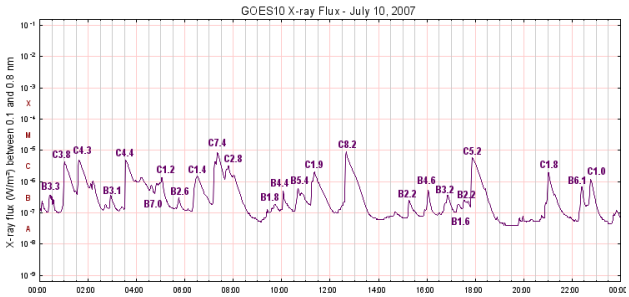


*Figure 21: GOES10 measurement of x-ray flux on July 10, 2007.*

Table 5 lists the numerous flares detected by GOES and the five SIDs that have been recorded.

Figures 22 and 23 show how behaves the algorithm under those difficult circumstances. The IMF 5 gives by itself many information. All of the five events are clearly visible on the IMF 5 component. Nevertheless, by using the IMF component, a more precise timing of the events can be obtained. The third event lasted 30 minutes less than initially reported and the fourth event lasted 35 minutes more.

When observing the details of the component, one could see small perturbations around 10:00 UTC, 10:40 UTC and 16:00 UTC. They are correlated to small flares B4.4, B5.4 and B4.6 that were not noticed on the raw data. This observation confirms the increased sensitivity offered by the analysis. Interestingly, the addition of the IMF 4 (upper-left plot on Figure 23) enhances the visibility of those three small events.

The addition of the IMF 6 (longer term component) allows to differentiate the two events (C7.4 and C2.8) after the transmitter shutdown, and to give a precise timing of the resulting disturbance.

| Flare Time (UTC) | | | Class | SID time (UTC) | | |
|---|---|---|---|---|---|---|
| Start | Max | End | | Start | Max | End |
| 00:21 | 00:25 | 00:31 | B3.3 | | | |
| 00:53 | 01:01 | 01:11 | C3.8 | | | |
| 01:32 | 01:37 | 01:46 | C4.3 | | | |
| 02:52 | 02:56 | 03:00 | B3.1 | | | |
| 03:29 | 03:33 | 03:43 | C4.4 | | | |
| 04:32 | 04:36 | 04:39 | B7.0 | | | |
| 04:46 | 05:03 | 05:05 | C1.2 | | | |
| 05:39 | 05:44 | 05:48 | B2.6 | | | |
| 06:17 | 06:31 | 06:39 | C1.4 | 06:26 | 06:32 | 07:00D |
| 07:08 | 07:21 | 07:28 | C7.4 | 07:12E | 07:31U | 08:30 |
| 07:39 | 07:48 | 07:52 | C2.8 | | | |
| 09:19 | 09:42 | 09:52 | B1.8 | | | |
| 10:00 | 10:04 | 10:06 | B4.4 | | | |
| 10:32 | 10:40 | 10:55 | B5.4 | | | |
| 11:11 | 11:21 | 11:29 | C1.9 | 11:14 | 11:23U | 12:10 |
| 12:35 | 12:40 | 12:46 | C8.2 | 12:37E | 12:41 | 13:55 |
| 15:11 | 15:16 | 15:22 | B2.2 | | | |
| 15:51 | 16:03 | 16:06 | B4.6 | | | |
| 16:41 | 16:51 | 16:57 | B3.2 | | | |
| 17:13 | 17:19 | 17:24 | B1.6 | | | |
| 17:28 | 17:32 | 17:46 | B2.2 | | | |
| 17:48 | 17:53 | 18:04 | C5.2 | 17:50 | 17:54 | 18:52 |
| 20:50 | 21:01 | 21:06 | C1.8 | | | |
| 22:15 | 22:23 | 22:28 | B6.1 | | | |
| 22:39 | 22:46 | 22:51 | C1.0 | | | |

*Table 5: Example of an active day. On July 10, 2007, 25 x-ray flares have been detected by GOES spacecraft. Five SID events have been recorded.*

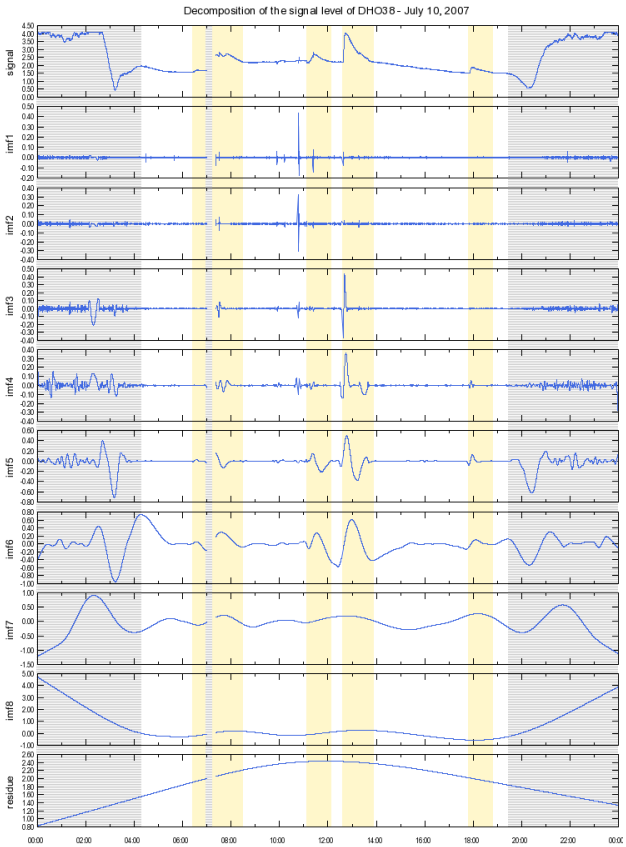


*Figure 22: Set of components obtained by the EMD algorithm on an active day.*

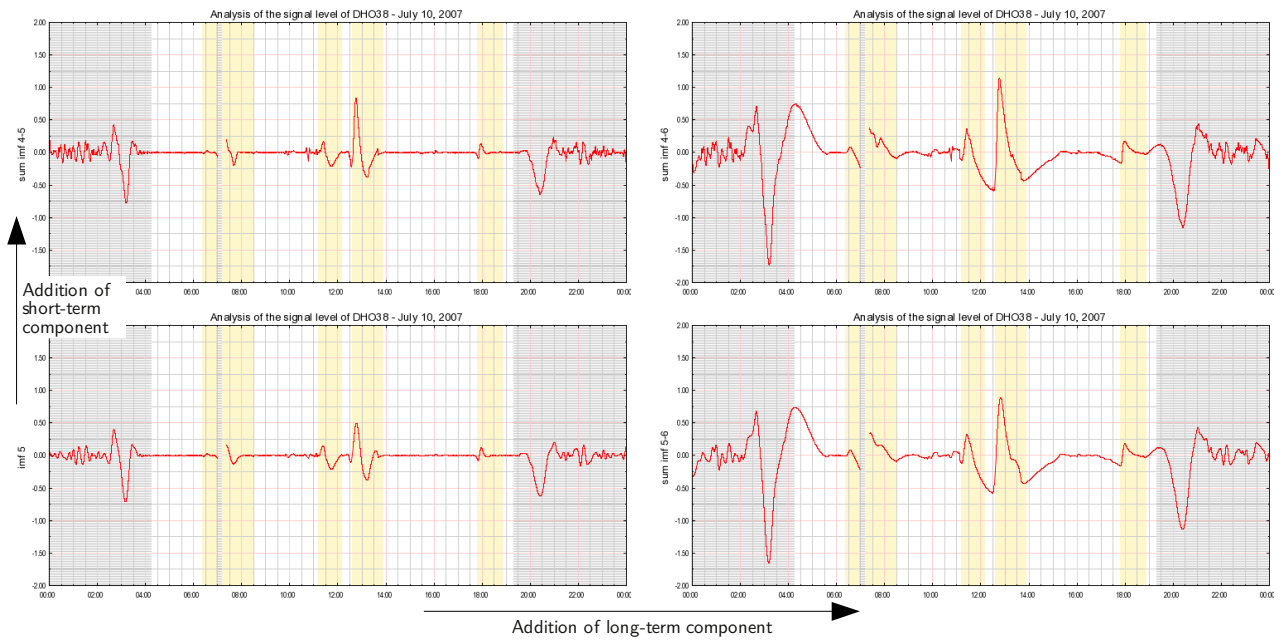


*Figure 23: Comparison of various indicators of the disturbances on July 10,2007. The reference is the lower left plot (IMF 5). IMF 6 is added in the right-hand plots. IMF 4 is added in the top plots.*

## Conclusion

The empirical mode decomposition represents raw measurements as a small number of components that embed the underlying physical properties of the signal. It is then a means to extract uncommon variations by filtering out noise and normal signal evolution. We have shown that this algorithm applies successfully to the recordings of the signal strength of VLF transmitters. A single IMF component appears sufficient to represent the amplitude variations resulting from ionospheric disturbances. The signals of DHO38 have been used as examples for this study. Since this transmitting station is shut down daily, there are artifacts created by the sharp signal transitions that could be mistaken by events. Consequently, the data recorded during the shutdown period—representing only noise—must be ignored from the analysis. The EMD handles perfectly an irregular sampling of a signal.

The method proved helpful in obtaining a more precise timing of the disturbances.

The confidence in the detected events is increased, and the method should lower the number of false positives and false negatives. Also, the sensitivity is improved and, on several examples, smaller perturbations—from flares as low as B5—have be detected.

The upcoming solar cycle 24 will give an opportunity to apply this method to real conditions.

Further work is necessary but an automatic processing of the data seems possible. This could be of a great help for stations that record simultaneously many stations.

Another study could be to determine how the method applies to the measurements of the signal phase.

## Acronyms

| | |
|---|---|
| AAVSO | American Association of Variable Star Observers |
| EIWG | Earth-Ionosphere Wave Guide |
| EMD | Empirical Mode Decomposition |
| FFT | Fast Fourier Transform |
| GOES | Geostationary Operational Environmental Satellite |
| GRB | Gamma-Ray Burst |
| HF | High Frequency |
| HHT | Hilbert–Huang transform |
| IMF | Intrinsic Mode Function |
| LF | Low Frequency |
| NGDC | National Geophysical Data Center |
| SID | Sudden Ionospheric Disturbance |
| VLF | Very Low Frequency |

## References

[1] (in French) Jean-Jacques Delcourt, Ionosphère : région D, sondage passif et propagation des signaux LF et VLF, Hermes Science Publications, 28 février 2003, 270 p. (ISBN 978-2746206564)

[2] Huang, et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." Proc. R. Soc. Lond. A (1998) 454, 903–995. Available at http://keck.ucsf.edu/~schenk/Huang_etal98.pdf

[3]Archive of solar flares (National Geophysical Data Center):
http://www.ngdc.noaa.gov/stp/SOLAR/ftpsolarflares.html#xray
2006 data:
ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_FLARES/XRAY_FLARES/xray2006/
2007 data:
ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SOLAR_FLARES/XRAY_FLARES/xray2007/

[4] AAVSO SID program:
http://www.aavso.org/observing/programs/solar/SIDMonOverview.shtml

[5] AAVSO rules for determining SID characteristics:
http://www.aavso.org/observing/programs/solar/reducing.shtml

[6] Description of the author's SID monitoring station with access to real-time measurements:
http://sidstation.lionelloudet.homedns.org/

[7] Binary and source files of the emd software:
http://anonsvn.lionelloudet.homedns.org/emd/tags/1.0.0/