



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

PNA

Probability, Networks and Algorithms



Probability, Networks and Algorithms

Empirical mode decomposition: a new tool for S-wave detection

P.J. Oonincx

REPORT PNA-R0203 JANUARY 31, 2002

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).

CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

Copyright © 2001, Stichting Centrum voor Wiskunde en Informatica

P.O. Box 94079, 1090 GB Amsterdam (NL)

Kruislaan 413, 1098 SJ Amsterdam (NL)

Telephone +31 20 592 9333

Telefax +31 20 592 4199

ISSN 1386-3711

Empirical Mode Decomposition: A New Tool for S-Wave Detection

P.J. Oonincx

CWI

P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

ABSTRACT

Seismic signals consist of several typically short energy bursts, waves, exhibiting several patterns in terms of dominant frequency, amplitude and polarisation. Amongst others, a significant wave is the S-wave. To detect such S-waves one can use conventional techniques that are based on physical differences between the several waves appearing in the seismic signal. In this paper we combine a conventional technique with the recently introduced Empirical Mode Decomposition, a non-linear signal analysis tool. By means of two examples a comparison of this new approach is made with both conventional techniques and the wavelet-based approach, which was proposed by the same author in an earlier paper.

2000 Mathematics Subject Classification: 42C99, 86A15, 94A12, 94A13.

Keywords and Phrases: Empirical Mode Decomposition, data-driven, adaptive filtering, seismology, feature detection, phase pickers.

Note: Work carried out under project PNA 4.2 and supported financially by the Technology Foundation (STW), project no. CWI.4616.

1. INTRODUCTION

A feature of a seismic earthquake signal is that it is built up by several different seismic waves, characterizing the type of signal. In this paper we are only interested in two of these waves, namely the primary wave (P-wave) and the secondary or shear wave (S-wave). The P-wave is the first wave (in time) that reaches the earth's surface when traveling from the epicenter of an earthquake towards a seismic station where it is recorded. In particular, we revisit the classical problem of detecting automatically the S-wave (its onset time, i.e., its first appearance in the seismogram), once the moment in time at which the P-wave arrives at the earth's surface (P-wave onset time) is known [2, 3, 9]. An accurate estimate of these arrival times is important for determining the type and location of the seismic event.

The S-wave arrival time is determined in a three-component seismogram, representing motion on a ground detector in three mutually orthogonal directions, two in the horizontal plane (x-y plane) and one vertical direction (z-axis). An example of a three-component seismogram is depicted in Figure 1. The seismogram in this figure is the result of a local event, the type of event we will concentrate on in this paper.

The detection of the S-wave arrival time is mostly based on some physical differences between the P-wave and the S-wave, as described thoroughly in [1, 4]. For our problem the most obvious property is the difference in onset times and its difference in frequency behaviour. The secondary wave is always delayed as compared to the primary wave onset time. Comparing the frequency spectra of both waves, we see that the P-wave appears at higher frequencies than the S-wave. Especially the latter property will be the base of our approach.

A more fundamental property is the fact that P-waves compress volumes and S-waves deform volumes. Furthermore, the S particle motion, i.e. the direction of the S-wave when it arrives at the earth's surface after travelling from the epicenter of an earthquake, is contained in a plane perpendicular to the direction of the P particle motion, called the S-plane. This property only holds if reflections at the earth's surface may be ignored or when the waves arrive in a direction almost perpendicular to the earth's surface. In our problem the latter assumption is justified. In our case, it is known that the P-wave travels along the travel direction of the seismic event, unless the medium is anisotropic. The phenomenon of the existence of an

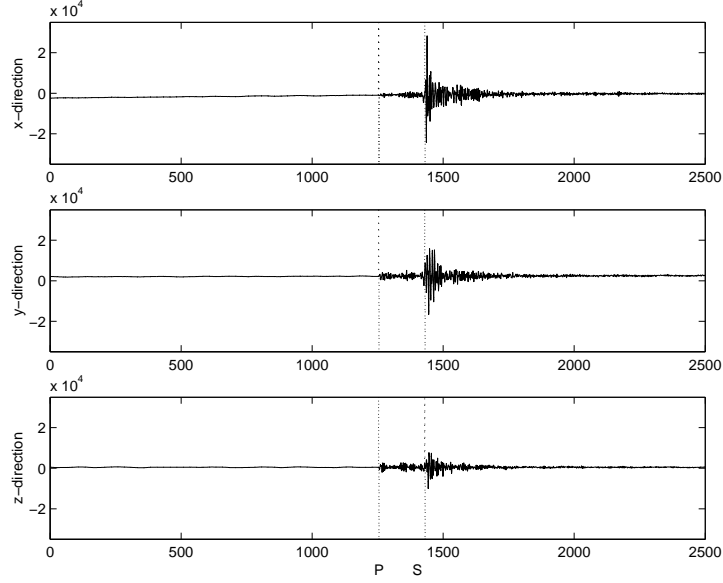


Figure 1: A three-component seismogram, with P- and S-wave onset times picked by an analyst at $i_P = 1255$ and $i_S = 1430$ respectively.

S-plane is one of the basis criteria for the approaches in [1, 4].

In [7] we already presented an automatic S-wave detection algorithm that combines traditional geophysical methods to detect S-waves as described before and the discrete wavelet transform. The idea behind this combination is to analyse the three components of a seismogram at several scales. By doing this, in many cases we can isolate S-waves in certain frequency bands and analyse only the signals in these bands with the traditional geophysical tools. However, the problems of choosing the appropriate frequency bands in which the S-wave appears. A sophisticated method was introduced in [7]. Here a combination of energy distribution functions and de-noising techniques were used to solve this problem.

In this paper we use a new mathematical method, Empirical Mode Decomposition (EMD), which provides a new way of isolating S-waves before analysing them with conventional geophysical techniques. We will compare this technique with the proposed wavelet-based technique in [7].

2. GEOPHYSICAL METHODS FOR S-WAVE DETECTION

In this section we briefly consider a classical method to detect S-waves in seismic data, as proposed in [4]. We remark that this method was also used and discussed extensively in [7]. The classical strategy to detect onset times in seismic data is to construct one or more so-called characteristic functions. These are discrete-time functions, with some specific properties at the time sample, at which a certain wave appears in the seismogram.

2.1 Characteristic functions based on a cross-power matrix

Through this section we will look at a three-component seismic signal as $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$. The time-dependent N -point cross-power matrix for such signal u is then defined as follows.

Definition 2.1 Let $N \in \mathbb{N}$ and $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$. Then the N -point cross-power matrix of u at $i \in \mathbb{Z}$ is given by

$$M_{N,u}(i) = \begin{pmatrix} \langle u_1, u_1 \rangle_i & \langle u_1, u_2 \rangle_i & \langle u_1, u_3 \rangle_i \\ \langle u_2, u_1 \rangle_i & \langle u_2, u_2 \rangle_i & \langle u_2, u_3 \rangle_i \\ \langle u_3, u_1 \rangle_i & \langle u_3, u_2 \rangle_i & \langle u_3, u_3 \rangle_i \end{pmatrix}, \quad (2.1)$$

for $i \in \mathbf{Z}$, with

$$\langle u_n, u_m \rangle_i = 1/N \sum_{k=i}^{i+N-1} u_n(k) \cdot u_m(k),$$

for all $i \in \mathbf{Z}$.

Note, that before computing the cross-power matrices of a given signal, we first create a signal u , with zero mean at each component. This is done by subtracting the means of the components from the seismic signal. The reason for doing this is to neglect possible offsets without seismic cause. These can be generated by the measurement equipment.

From the definition Obviously, $M_{N,u}(i)$ is a Gram matrix and therefore positive semi-definite. So, the eigenvalues of $M_{N,u}(i)$ are real and positive,

$$\lambda_1(i) \geq \lambda_2(i) \geq \lambda_3(i) \geq 0,$$

and the eigenvectors $v_1(i)$, $v_2(i)$ and $v_3(i)$ can be chosen to form an orthonormal basis in \mathbb{R}^3 . The following two characteristic functions are based on these eigenvalues and eigenvectors. Hereby we assume $M_{N,u}(i) \neq O$, for any $i \in \mathbf{Z}$, which is quite a realistic assumption. We observe that $M_{N,u}(i) = O$ is equivalent with $u_1(k) = u_2(k) = u_3(k) = 0$, for $k = i, \dots, i + N - 1$.

2.1.1 Deflection angle: Let $v_1(i)$ denote the eigenvector of $M_{N,u}(i)$ corresponding to $\lambda_1(i)$, i.e., the eigenvector that represents the direction of the particle motion at time sample i with most seismic energy. Let i_P be the P-wave onset time. Then $v_1(i_P)$ is the direction of the P particle motion. The deflection angle is defined by

$$\kappa_1(i) = \frac{2}{\pi} \arccos \left(\frac{|(v_1(i), v_1(i_P))|}{\|v_1(i)\| \cdot \|v_1(i_P)\|} \right). \quad (2.2)$$

We observe that $\kappa_1(i_P) = 0$. Furthermore, since the direction of the S-wave particle motion $v_1(i_S)$ is perpendicular to $v_1(i_P)$, κ_1 attains its maximum 1 at the S-wave onset time i_S .

2.1.2 Degree of polarization: Following [8], the degree of polarization is defined by

$$\kappa_2(i) = \frac{(\lambda_1(i) - \lambda_2(i))^2 + (\lambda_1(i) - \lambda_3(i))^2 + (\lambda_2(i) - \lambda_3(i))^2}{2 \cdot (\lambda_1(i) + \lambda_2(i) + \lambda_3(i))^2}. \quad (2.3)$$

This characteristic function can be used both for detecting P- and S-wave onset times, since for all different onset times seismic energy is concentrated along one single direction. In practice we have for all these onset times $\lambda_1(i) \gg \lambda_2(i)$, which means that we may expect maxima for κ_2 at both i_P and i_S , the S-phase arrival time.

We observe that for all characteristic functions κ introduced in this section we have

$$0 \leq \kappa(i) \leq 1 \quad \forall i \in \mathbf{Z} \quad (2.4)$$

In the sequel we will use a combination of three characteristic functions $\kappa_1, \kappa_2, \kappa_3$ such that

$$\prod_{m=1}^3 \kappa_m^2(i) = \max_{n \in \mathbf{Z}} \prod_{m=1}^3 \kappa_m^2(n) \iff i = i_S. \quad (2.5)$$

The square product of the characteristic functions will be used to emphasize the maximum values attained in each function at i_S and to reduce other (local) maxima, related to features in the signal other than the S-phase arrival. We have already met two candidates to be used in this product of functions. In the following subsection we introduce the third function that can be used in (2.5).

2.2 Rotation and the Energy Ratio

For computing the third characteristic function we first transform the three-component seismic data into a three-component signal representing motion in the longitudinal direction and in two transversal directions. The longitudinal direction is the direction of the P particle motion ($v_1(i_P)$). The transversal directions are mutually orthogonal and are chosen in the plane perpendicular to the longitudinal direction ($\text{span}\{v_2(i_P), v_3(i_P)\}$). This transversal plane is the S-plane, since the direction of the S particle motion is in the S-plane.

The seismic signal is transformed into the basis $\{v_1(i_P), v_2(i_P), v_3(i_P)\}$ by

$$\begin{pmatrix} u_L(i) \\ u_Q(i) \\ u_T(i) \end{pmatrix} = V(i_P) \begin{pmatrix} u_1(i) \\ u_2(i) \\ u_3(i) \end{pmatrix}, \quad (2.6)$$

with

$$V(i_P) = (v_1(i_P) \mid v_2(i_P) \mid v_3(i_P))^T. \quad (2.7)$$

The third characteristic function we use in (2.5) is the fraction of energy in the S-plane to the total amount of energy in the signal, given by

$$\kappa_3(i) = \frac{\sum_{n=i}^{i+N-1} (u_Q(n)^2 + u_T(n)^2)}{\sum_{n=i}^{i+N-1} (u_L(n)^2 + u_Q(n)^2 + u_T(n)^2)}, \quad (2.8)$$

for some $N \in \mathbb{N}$. This definition can also be rewritten as

$$\kappa_3(i) = 1 - \frac{\sum_{n=i}^{i+N-1} (v_1(i_P), u(n))^2}{\sum_{n=i}^{i+N-1} (u(n), u(n))}, \quad (2.9)$$

which shows how κ_3 depends on i_P . Note that we may expect a minimum $\kappa_3(i_P) = 0$ and a maximum $\kappa_3(i_S) = 1$. Furthermore, κ_3 satisfies (2.4).

In [7] we discussed stability of these characteristic functions in a mathematical sense. We proved stability for all presented characteristic functions with respect to computational and truncation errors in the cross-power matrices $M_{N,u}$ and measurement errors in u . We also considered errors in detecting S-wave onset times due to incorrectly determined P-wave arrival times i'_P . It has been shown that an incorrect P-wave onset time i'_P does not affect κ_2 . Concerning κ_1 we showed that $i'_P - i_P$ has to be small in order not to obtain large deviations in κ_1 , but also $\lambda_1(i_P) \neq \lambda_2(i_P)$ should be satisfied and u should be a signal of bounded variation. For κ_3 we computed an expression for the error due to incorrect P-wave onset times. This expression also is a good estimate for the error in κ_3 .

3. EMPIRICAL MODE DECOMPOSITION

The empirical mode decomposition (EMD) was first introduced by Huang et al. [6]. The principle of this technique is to decompose adaptively a given signal $f(t)$ into oscillating components. These components are called intrinsic mode functions (IMF) and are obtained from the signal f by means of an algorithm, called sifting. The essence of this algorithm is to identify the IMF by characteristic time scales, which can be defined locally by the time lapse between two extrema of an oscillatory mode or by the time lapse between two zero crossings of such mode. The idea is then to extract for each mode locally the highest frequency oscillations out of f . To compute also the frequency behaviour of each IMF in time, Huang proposed to use the instantaneous frequency of each IMF. However to calculate instantaneous frequencies we have to ensure that each IMF is symmetric with respect to its local mean, otherwise unwanted fluctuations

in the instantaneous frequency will be induced by asymmetric waveforms in the IMF.

In some sense the EMD can be seen as a type of adaptive wavelet decompositions, which was used [7] for this problem. Each IMF replaces then the detail signals of f at a certain scale or frequency band. However, the EMD is adaptive since the frequency subbands in which the IMF live are built up as needed to separate the different oscillating components of f . Furthermore, the EMD does not use any pre-determined filter or wavelet functions. It is a fully data driven method.

3.1 The Sifting Procedure

A procedure called sifting was proposed by Huang et al. [6] to decompose a sampled signal $v \in l^2(\mathbb{Z})$ by means of the EMD. The sifting procedure is based on two constraints, namely (i) each IMF has the same number of zero crossings and extrema; and (ii) each IMF is symmetric with respect to the local mean. Furthermore, it assumes that v has at least two extrema.

The sifting algorithm for $v \in l^2(\mathbb{Z})$ reads as follows:

1. Initialise: $r_0 = v$ (the residual) and $j = 1$ (index number of IMF),
2. Extract the j -th IMF:
3. (a) Initialise $h_0 = r_{j-1}$, $i = 1$,
 (b) Extract local minima/maxima of h_{i-1} ,
 (c) Compute upper envelope and lower envelope functions x_{i-1} and y_{i-1} by interpolating respectively local minima and local maxima of h_{i-1} ,
 (d) Compute $m_{i-1} = (x_{i-1} + y_{i-1})/2$ (mean of envelopes),
 (e) Update $h_i := h_{i-1} - m_{i-1}$ and $i := i + 1$,
 (f) Calculate stopping criterion

$$SD_i = \sum_{k=-\infty}^{\infty} \frac{|h_{i-1}(k) - h_i(k)|^2}{h_{i-1}(k)^2},$$

- (g) Decision: Repeat step (b)-(f) until $SD_i < \varepsilon$ and put then $\tilde{v}_j = h_i$ (j -th IMF)
4. Update residual $r_j = r_{j-1} - \tilde{v}_j$,
5. Repeat step 1-3 with $j := j + 1$ until the number of extrema in r_j is less than two.

Looking at the sifting algorithm from a mathematical point of view many questions may arise. First of all the type of interpolation scheme (3c) is of interest. In [6] Huang proposed to use cubic spline interpolation on non-equidistant sampled data. Although in practice this might be an efficient way for interpolation, the most appropriate way to interpolate would start with an interpolation scheme that matches the regularity of the signal v . Furthermore, in the algorithm a decision criterion has been built in. Also in [6] it was suggested to take $0.2 < \varepsilon < 0.3$, which gives a good performance of the algorithm in practice. However from a mathematical point of view the convergence of SD_i should be studied before picking a value for ε . A final remark concerns the determination of the local extrema (3b). For time-continuous signals it is obvious how to pick local extrema, but if only samples of such signals exists one should be careful. In this paper we have the following approach for determining extrema. First we compute a difference signal d given by $d(i) = v(i+1) - v(i)$. Then an extreme value is assigned if one of the following expressions hold:

- $d(i-1) * d(i) < 0$,
- $d(i) = 0$ and $d(i-1) * d(i+1) < 0$.

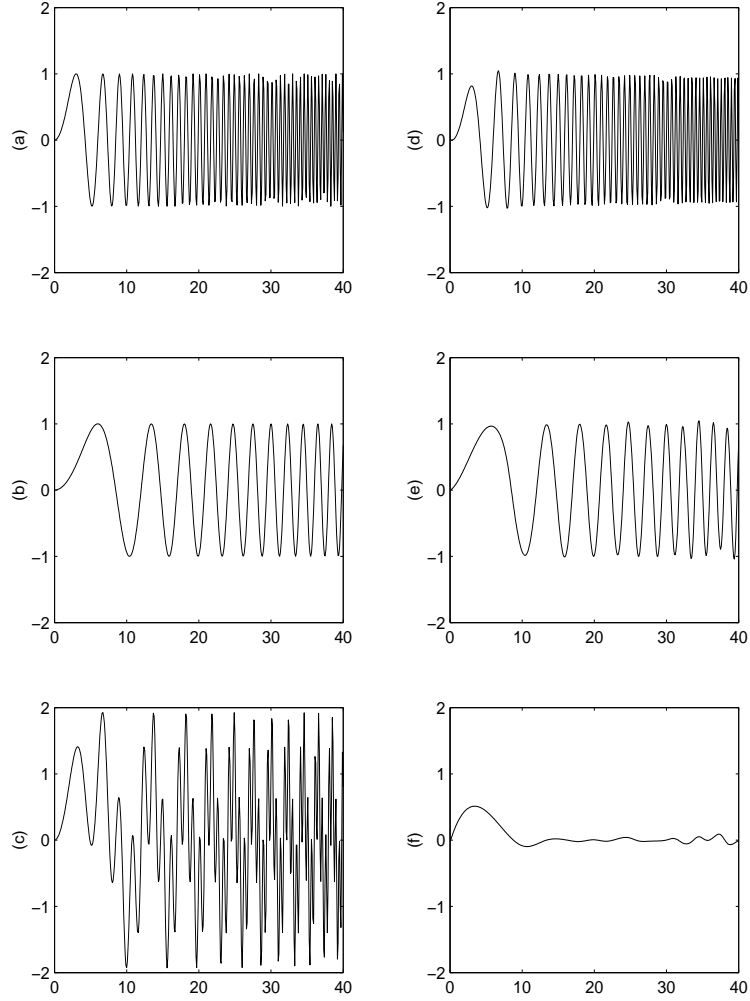


Figure 2: Decomposition of a signal consisting of chirps: a) $\sin(\pi t^2/20)$, b) $\sin(\pi t^2/80)$, c) $\sin(\pi t^2/20) + \sin(\pi t^2/80)$, d) first IMF, e) second IMF f) residual r_2 .

We observe that we will study these mathematical issues in more detail in a forthcoming paper.

The result of the sifting procedure is that v will be decomposed into intrinsic mode functions \tilde{v}_j , $j = 1, \dots, N$ and a residual r_N , i.e.,

$$v = \sum_{j=1}^N \tilde{v}_j + r_N, \quad (3.1)$$

with r_N the residual sequence that arises as the residual in the sifting process with less than 2 extreme values. We observe that in general $(\tilde{v}_k, \tilde{v}_l)_2 \neq 0$, which would be the case if \tilde{v}_j were the detailed signals of an orthogonal wavelet decomposition. As a consequence the energy in v generally differs from the total amount of energy in all \tilde{v}_j and r_N . Furthermore, the sequences \tilde{v}_j are constructed in such way, that locally frequencies decrease for subsequent IMF \tilde{v}_j . In the next subsection we present a way of analysing the frequency behaviour of all IMF over time. By doing this we create a time-frequency representation of our decomposition.

As an example to illustrate the EMD, we take two chirp-like signals $\sin(\pi t^2/20)$ and $\sin(\pi t^2/80)$ and sum them into one signal. These three signals have been depicted in Figure 2a, b and c. Since we have a linear

relation between time and frequency for these signals it will be hard to separate the signals again with a frequency dependent method. A wavelet decomposition will spread each component over several detailed signals (frequency bands). Recovering each of the two signals has then become an almost impossible task. In Figure 2e, f we see that the EMD recovered both signals. The highest frequent component was covered in the first intrinsic mode function, while the second IMF seems to be a good representation of $\sin(\pi t^2/20)$. After obtaining the second IMF the sifting process was stopped. The residual r_2 has been depicted in Figure 2f. This picture confirms that both chirp-like signals were covered well by the first two IMF. However, as we can see in this picture as well, close to the boundaries some information is left. This is due to the so-called not-knot cubic spline interpolation at the boundaries.

3.2 The Hilbert Transform

Since the index number j of each \tilde{v}_j does not give any explicit information on the frequency contents of such mode, we would like to complete our approach by mapping \tilde{v}_j into phase space. Especially for non-stationary signals it is important to have knowledge about the frequency behaviour of all \tilde{v}_j over time. Note that for detail signals in a wavelet decomposition the index number already gives us some information about the frequency behaviour of this particular detail signal, since the index number are linked to certain frequency bands.

To analyse non-stationary signals in both time and frequency many transforms can be used, e.g. the spectrogram (windowed Fourier) and the Wigner distribution [5]. In Huang's paper [6] the Hilbert transform was introduced as a tool to obtain frequency content information for all separate modes. For $f \in L^2(\mathbb{R})$ the Hilbert transform reads

$$\mathcal{H}[f](\omega) = \frac{1}{\pi} p.v. \int_{-\infty}^{\infty} \frac{f(t)}{\omega - t} dt, \quad (3.2)$$

with $p.v. \int$ meaning the Cauchy principal value of this integral. We observe that the integral in the right hand side is well defined, which follows straightforwardly from the fact that

$$\frac{1}{\pi} p.v. \int_{-\infty}^{\infty} \frac{f(t)}{\omega - t} dt = \text{Im} \left(\int_0^{\infty} \hat{f}(t) e^{-i\omega t} dt \right), \quad (3.3)$$

for any real-valued $f \in L^2(\mathbb{R})$. Here \hat{f} denotes the Fourier transform of $f \in L^2(\mathbb{R})$, given by

$$\hat{f}(\omega) = l.i.m._{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-N}^N f(t) e^{-i\omega t} dt,$$

with $l.i.m.$ meaning the limit in mean. This relation is also the starting point for building an analytic signal F corresponding to f , namely

$$F(t) = f(t) + i \mathcal{H}[f](t).$$

By doing this we identify the real-valued function f by a complex trace that is uniquely given by

$$F(t) = A(t) e^{-i\phi(t)},$$

with $A(t)$ the amplitude and $\phi(t)$ the phase of $F(t)$ as functions over time. These functions are also called instantaneous amplitude and instantaneous phase. To obtain frequency information of f we simply take the first derivative of $\phi(t)$, i.e.,

$$\omega(t) = \frac{d\phi(t)}{dt}, \quad (3.4)$$

the instantaneous frequency.

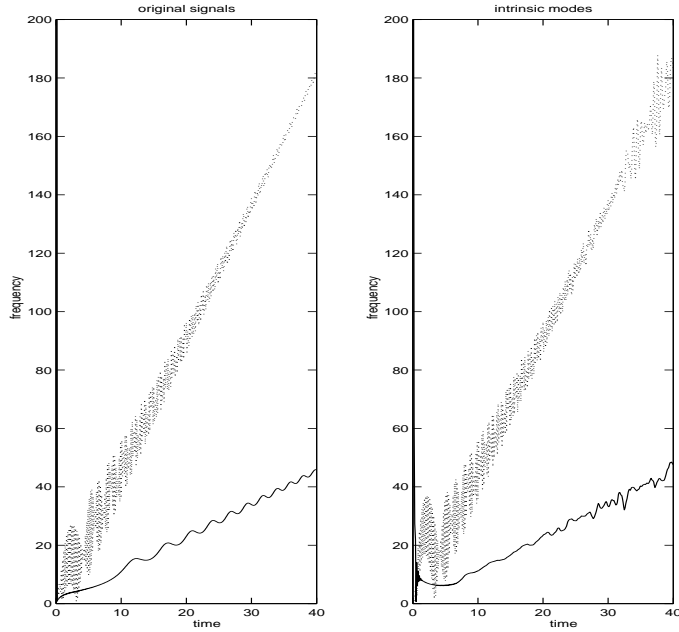


Figure 3: Instantaneous frequencies of signals/modes in Figure 2: chirp-like signals (left) and first two modes (right).

In our approach we deal with sampled data $v \in l^2(\mathbf{Z})$ instead of functions $f \in L^2(\mathbb{R})$. Therefore, if we want to study the frequency behaviour of all \tilde{v}_j we first have to apply an interpolation scheme to obtain time-continuous intrinsic modes before computing their Hilbert transforms. Of course for piecewise constant interpolation the Hilbert transform of v falls down to the imaginary part of the discrete Fourier transform of $\hat{v}(k)$, $k = 1, 2, \dots$, with \hat{v} the discrete Fourier transform of v . In this paper we will not discuss the discrete Hilbert transform any further, since for our seismic problem we are only interested in the decomposition of seismic data into intrinsic modes.

To illustrate how a time-frequency analysis of all IMF is constructed by means of the Hilbert transform, we computed the Hilbert transform of the sampled chirp-like signals in Figure 2a,b. Next we constructed analytic signals and computed the instantaneous frequencies of the sampled signals. The instantaneous frequencies $\omega(t)$ have been depicted in the left hand side of Figure 3 for both $\sin(\pi t^2/20)$ (solid line) and $\sin(\pi t^2/80)$ (dotted line). Clearly we see a linear behaviour of $\omega(t)$ for both signals, as was also expected. In the right hand side of Figure 3 the instantaneous frequency $\omega(t)$ has been depicted for the modes \tilde{v}_1 (dotted) and \tilde{v}_2 (solid). Here we see that the intrinsic modes are not restricted in frequency to some predetermined frequency band. Furthermore, locally in time we see decreasing frequencies for increasing index numbers j of the modes. Finally we observe that both pictures in Figure 3 are very similar. Differences appear mostly near the borders, where the not-knot cubic spline interpolation introduced some unwanted fluctuations.

3.3 EMD-Hilbert Transform vs. Fourier Analysis

The example of the combined chirp signal in the previous two subsections showed how the EMD differs from classical Fourier based techniques. Although the frequency behaviour of the signals changed linearly in time, the two chirp-like signals were identified as two separate signals in just two different modes. These modes showed the same time-frequency behaviour as the original functions. We already observed that this is an essential difference with Fourier and wavelet analysis, where the two modes would be separated into sub modes contained in decreasing/increasing frequency bands. In the following example we show what happens if the EMD is applied on a pure sine-wave, a function with a constant behaviour in frequency.

Example:

Take $f(t) = \sin(\omega_0 t)$, $\omega_0 \in \mathbb{R}^+$, $t \in \mathbb{R}$. We follow the sifting procedure for time-continuous functions. This yields

1. Initialise: $r_0 = f$ and $j = 1$,
2. Extract the first IMF:
3. (a) Initialise $h_0 = f$, $i = 1$,
(b) Extract local minima/maxima of f :

$$V_{\max} = \left\{ \left(\frac{4k+1}{2\omega_0} \pi, 1 \right) \mid k \in \mathbb{Z} \right\}, \quad V_{\min} = \left\{ \left(\frac{4k+3}{2\omega_0} \pi, -1 \right) \mid k \in \mathbb{Z} \right\}.$$

- (c) Upper envelope: $x_0(t) = 1$, lower envelope $y_0(t) = -1$,
- (d) Mean of the envelopes $m_0(t) = (x_0(t) + y_0(t))/2 = 0$,
- (e) Update $h_1 := h_0 - m_0 = h_0$ and $i := 2$,
- (f) Stopping criterion: $SD_1 = 0$
- (g) Decision: $SD_1 < \varepsilon$ for all $\varepsilon > 0$, so the procedure ends and $\tilde{f}_1(t) = h_1(t) = \sin(\omega_0 t)$ (first IMF)
4. Update residual $r_1 = r_0 - \tilde{f}_1 = f - f = 0$.

So the sifting procedure only generates one intrinsic mode \tilde{f}_1 , namely f itself. The next step is to compute the Hilbert transform of $\sin(\omega_0 t)$. Using some elementary calculus from Fourier analysis we obtain $\mathcal{H}[\sin(\omega_0 \cdot)](t) = \cos(\omega_0 t)$. The analytical signal corresponding to $f(t)$ is then given by

$$F(t) = \sin(\omega_0 t) + i \cos(\omega_0 t) = i(\cos(\omega_0 t) - i \sin(\omega_0 t)) = i(\cos(-\omega_0 t) + i \sin(-\omega_0 t)) = e^{-i(\omega_0 t - \pi/2)}.$$

So $\phi(t) = \omega_0 t - \pi/2$ and therefore the instantaneous frequency of $\sin(\omega_0 t)$ is given by

$$\omega(t) = \frac{d(\omega_0 t - \pi/2)}{dt} = \omega_0.$$

From this example we see that analysing a mono-frequency function, i.e., $\sin(\omega_0 t)$, the EMD only generates one intrinsic mode, namely the mono-frequency function itself. Furthermore, computing the instantaneous frequency $\omega(t)$ of such function via the Hilbert transform yields $\omega(t) = \omega_0$. We conclude that for these kind of functions the EMD-Hilbert approach falls down to the classical Fourier analysis.

4. S-WAVE DETECTION ALGORITHM

The idea to detect S-phase arrival times using the EMD is as follows. In a similar way as we did for the wavelet transform [7] we decompose the three-component seismic signal $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$ into ‘detail’ signals and use only one of these modes to compute the characteristic functions we presented in Section 2. For this algorithm we will restrict ourselves to local seismic data, i.e. obtained from waves that only had to travel during a short period of time and as a result where $i_S - i_P < 20$ seconds approximately.

In the case of band-pass filters (also wavelet filters) one should choose one or more bands which is likely to contain most of the S-wave frequency contents. Next characteristic functions can be applied on the signal parts in these bands. Mostly a sophisticated approach is needed to find these most optimal bands, e.g. in [7] we used an additional transform to certify that the chosen detail signals indeed contained mostly the S-wave. The idea of the EMD-approach is as follows. In the case of a local event we may assume that the neighbourhood of the S-wave onset will mostly contain both P-wave contents and S-wave contents. Other waves are assumed to have a later onset time and/or lower frequency contents as these two waves. These assumptions yield the following algorithm:

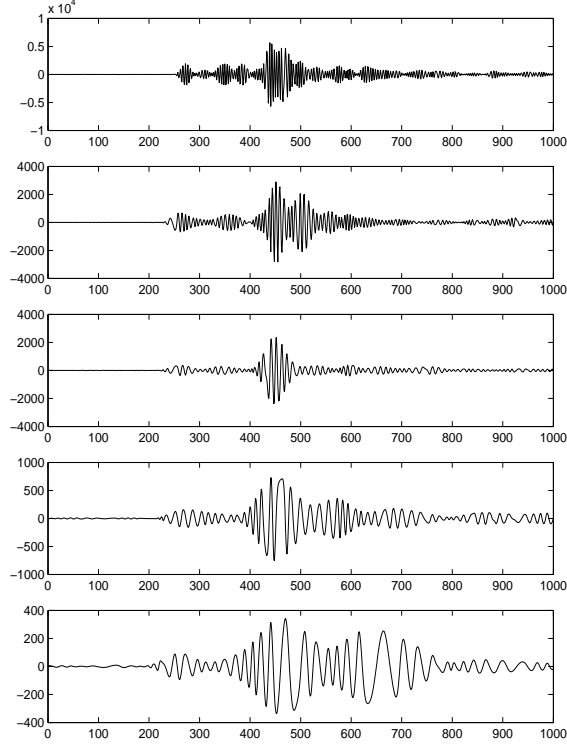


Figure 4: First 5 intrinsic modes of the local event in Figure 1 (vertical/z-direction).

- Compute the second IMF of each component $u^{(k)}$, $k = 1, 2, 3$ of the three-component seismogram. This yields $\tilde{u}_2^{(k)}$, $k = 1, 2, 3$,
- Compute the characteristic function κ , based on $\tilde{u}_2^{(k)}$, $k = 1, 2, 3$,
- Determine i_S from κ .

Since the approach seems to be very simple, we should make a few remarks. First of all, for obtaining the second IMF we also have to compute the first IMF. Since these computations consist of the iterative process of deriving interpolating functions, this procedure is very time consuming. Furthermore, this approach only works under the above mentioned assumptions, where the wavelet-based approach in [7] also could be used for non-local seismic data. These are data for which the previous assumptions are not satisfied in general. Moreover, for very low-frequent data (like non-local S-waves) the window-size, that plays a role in κ_3 and in the covariance matrix used to determine κ_1 and κ_2 , should be larger than for high-frequent data. Within our wavelet approach this window-size was coupled to the index number of the scale in a canonical way. Since the IMF are not related to pre-determined frequency bands this coupling is not possible in the EMD-based algorithm.

In Figure 4 the first 5 IMF of the vertical component of the local event in Figure 1. Obviously, frequency contents decrease for higher indexed IMF. Furthermore the S-wave is recognisable rather well in the second image of this figure, representing the second IMF. This is exactly the IMF we use within our approach.

An alternative approach based on the EMD is to compute characteristic functions based on the first residual, i.e. the original seismic data with omission of the first IMF in each direction, the top image in Figure 4. By doing this one creates a time-dependent high pass filter. At each time sample all data are conserved with frequencies less than the frequency in the first IMF at that particular time sample. This is a kind of

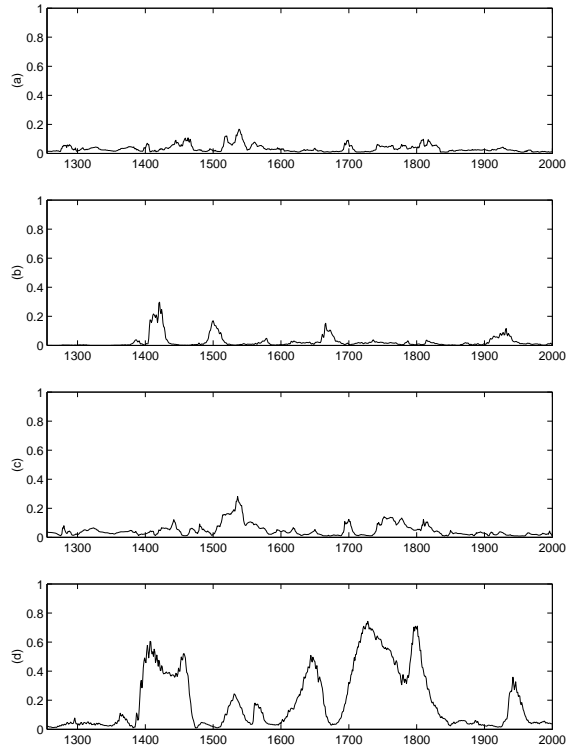


Figure 5: Characteristic functions of the seismic data in Figure 1: a) Cichowicz’s method, b) Cichowicz’s method and BPF, c) wavelet method, d) EMD method.

intermediate approach between our proposed algorithm and classical band-pass filtering before computing κ , which was also done in [4].

5. EXAMPLES AND RESULTS

In this section we demonstrate our algorithm by means of two examples that caused some problems when analysing with both the traditional method by Cichowicz [4] and the wavelet-based method we recently proposed [7]. Of course we also compare the results for these examples with the results for the same examples when using one of the above mentioned methods. Both signals were obtained from local events and sampled at a 40 Hz sampling rate. Furthermore, the window size N that plays a role in computing κ was set at 30 samples.

5.1 Example 1

In the first example we consider the seismic data as depicted in Figure 1. The onset times were picked by an analyst at sample $i_P = 1255$ (P-wave) and $i_S = 1420$ (S-wave). The P-wave onset time was used in Cichowicz’s approach, i.e., characteristic functions of Section 2 without preprocessing, to compute κ . This function κ has been depicted in Figure 5a. As we can see the S-wave onset time cannot be determined from this picture. In Figure 5b we see κ applied on the same data but now first filtered with a 0.6-6 Hz band pass Butterworth filter. By using this filter as a preprocessing step Cichowicz’s method became useful for picking. The deviation in the manual pick and the automatic pick with this method was less than 20 samples and so within 0.5 seconds. As for the first method also the wavelet approach is not able to pick the S-wave onset time correctly, which is clearly visible in Figure 5c. Finally we used the EMD-based algorithm proposed in this paper to detect the S-wave onset time. The characteristic function κ for this test has been depicted in Figure 5d. As one can see, the onset time i_S can be determined from this figure with just a small deviation of 20 samples (0.5 seconds) compared with the analyst’s pick.

Resuming, this first example shows how band pass filtering can enhance the performance of traditional

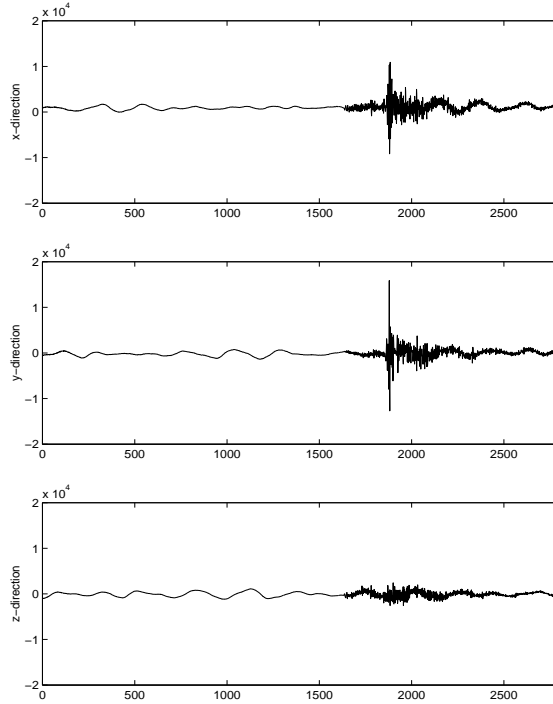


Figure 6: A three-component seismogram with microseismic noise. P- and S-wave onset times picked by an analyst at $i_P = 1570$ and $i_S = 1865$ respectively.

methods and that a very narrow banded filter can decrease the performance of traditional methods for some data sets. The latter is the case when using a wavelet-based approach for the seismic data, we used in this example.

5.2 Example 2

In the second example we consider the seismic data as depicted in Figure 6. Obviously these data consist of the several waves traveling from the earthquake's epicenter towards the earth's surface but also of some low-frequency noise, called microseismic noise. The onset times were picked by an analyst at sample $i_P = 15705$ (P-wave) and $i_S = 1865$ (S-wave). As in the first example the P-wave onset time was used in Cichowicz's approach to compute the characteristic function κ as depicted in Figure 7a. As in the first example the S-wave onset time cannot be determined from this picture. Using the same band pass filter as in the first example does not help as well to pick the S-wave onset time, which is clearly visible in Figure 7b. The wavelet-based and EMD-approaches both picked the S-wave onset time correctly, which can be seen in Figure 7c and d respectively. The wavelet approach picked $i_S = 1830$, which is 30 samples early, while the EMD approach picked $i_S = 1885$, which is 20 samples late. So both non-conventional approaches show a good performance for this data set, however one should observe that the wavelet-approach in [7] contains a sophisticated piece of algorithm that has been built in exclusively to deal with seismic data containing microseismic noise.

Resuming, the second example shows how necessary non-conventional techniques like wavelet decompositions and EMD can be for picking S-wave onset times, especially for data that contain microseismic noise. The newly proposed EMD approach shows its capacity to deal with non-trivial seismic data without necessarily using sophisticated de-noising techniques.

6. CONCLUSIONS

In this paper we presented the Empirical Mode Decomposition as an efficient tool for S-wave detection algorithms. We showed that the adaptive character of the EMD enables us to detect S-wave in cases, for

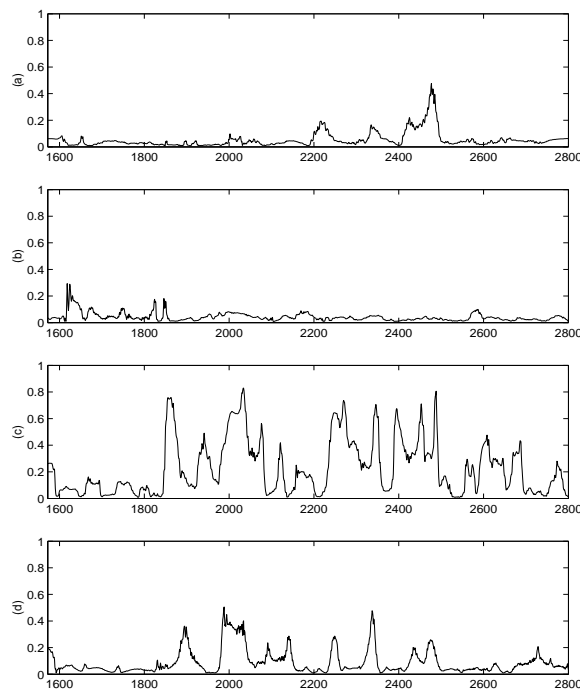


Figure 7: Characteristic functions of the seismic data in Figure 6: a) Cichowicz's method, b) Cichowicz's method and BPF, c) wavelet method, d) EMD method.

which it is not possible to detect them with classical Fourier-based methods. However, some remarks should be made on the comparison between this approach and the other approaches discussed in this paper. First of all we did not test our method on a huge database of seismic events, which was done for the other discussed methods. Furthermore, the EMD is a highly non-linear method that is very time consuming, especially compared to the Fourier- and wavelet-based methods. Finally, the wavelet-based approach has shown to be efficient as well for non-local event. Due to the character of the EMD-based approach it is very unlikely that this will also be the case for this new approach.

Future work will concentrate on finding sophisticated mathematics to describe the EMD procedure and to come to a less time consuming algorithm. Also testing the new approach on a huge sets of local events will be necessary to compare the proposed method with the existing algorithms.

ACKNOWLEDGEMENTS

The author likes to thank the KNMI for providing seismic data to test the algorithm.

REFERENCES

1. K. Aki and P.G. Richards, *Quantitative seismology*, Freeman and Company, San Francisco, 1980.
2. R.V. Allen, "Automatic earthquake recognition and timing from single traces", *Bull. Seism. Soc. Am.*, 68, 1521-1532, 1978.
3. M. Bear and U. Kradolfer, "An automatic phase picker for local and teleseismic events", *Bull. Seism. Soc. Am.*, 77(4), 1437-1445, 1987.
4. A. Cichowicz, "An automatic S-phase picker", *Bull. Seism. Soc. Am.*, 83(1), 180-189, 1993.
5. L. Cohen, *Time-frequency analysis*, Prentice Hall, New Jersey, 1995.
6. N. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proc. Roy. Soc. London A*, 454, 903-995, 1998.

7. P.J. Oonincx, "A Wavelet Method for Detecting S-Waves in Seismic Data", *Computational Geosciences*, 3, 111-134, 1999.
8. J. Samson and J. Olson, "Some comments on the descriptions of the polarization states of waves", *Geophys. J. R. astr. Soc.*, 61, 115-129, 1980.
9. R. Sleeman and T. van Eck, "Robust automatic P-phase picking: An on-line implementation in the analysis of broad-band seismogram recordings", *Phys. Earth Planetary Interior*, 113, 45-55, 1999.