Wavelets statistical denoising (WaSDe): individual evoked potential extraction by multi-resolution wavelets decomposition and bootstrap

Besma Benchabane, Moncef Benkherrat, Boris Burle, Franck Vidal, Thierry Hasbroucq, Salah Djelel, Aissa Belmeguenai

HAL Id: hal-02401454
https://hal.science/hal-02401454

Submitted on 10 Dec 2019

Wavelets Statistical Denoising (WaSDe): Individual Evoked Potential Extraction by Multiresolution Wavelets Decomposition and Bootstrap

Besma BENCHABANE[1]* , Moncef BENKHERRAT [2][3], Boris BURLE [3],Franck VIDAL[3] , Thierry HASBROUCQ [3] ,Salah DJELEL[4] ,Aissa BELMEGUENAI [4]

[1] LASA , Badji Mokhtar-Annaba University , P.O. Box 12, 23000 Annaba , Algeria
[2] ECAM-EPMI, 13 bd de l'Hautil 95092, Cergy-Pontoise cedex, France
[3] Aix Marseille Univ, CNRS, LNSC, Marseille, France
[4] Universite 20 Aout de Skikda , Algerie
* E-mail: souma.31.07@hotmail.com

Abstract: The present study aims at developing a method to extract single sweep event-related potentials obtained with Eriksen's flanker task. Unlike previous methods, no a priori assumptions on the characteristics of signal and noise are necessary. The method is based on the wavelet decomposition, bootstrap and a statistical determination of the reliable frequency coefficients across the individual signals at each time point: significant coefficients will be conserved, whereas the other ones will be set to zero. After removing the unsystematic coefficients (i.e. the noise), the signal is reconstructed, allowing to keep only the components of the event-related potentials. The performances of the method are evaluated with both simulated data and real event-related potential recordings, and compared with other methods.

1 Introduction

The Event-Related Potentials (ERPs) was first used in the sixties and are still in use to evaluate the cognitive functions. They are also used to examine neurological disorders and psychiatry disorders such as dementia, Alzheimer's disease and autism [1]. The P300 potential has been widely used in brain computer interfacing (BCI) [2], due to the easiness of its observation in discriminative tasks. Recently, the ERP has been used to analyse the behavior of an operator executing complex tasks such as piloting an aircraft [3],[4] and [5].

ERPs are responses of the brain to visual or auditory stimuli and they are recorded on the scalp. The ERPs can be defined as the voltage changes of spontaneous electroencephalogram (EEG) activity. The ERPs are contaminated with large amplitude of EEG. The signal-to-noise Ratio (SNR) of recorded ERPs is poor and its value is low [6], [7]. To improve the SNR, the ERPs are commonly averaged [6]. The reduction of the noise component by using the average procedure is proportional to the square root of the number of trials. This procedure supposes the ERPs synchronous, time-locked responses and not correlated with EEG. These assumptions are not necessarily accurate. If the latency of the individual ERPs is not constant over the repetitions, the shape and the latency of the averaged signals will not represent the individual latencies and shapes of the ERPs signals. Another limitation comes from the fact that inter-repetition variability is entirely lost in the averaged procedure. It is known that the attention of the subject modulates the amplitude of the visual evoked potentials. Let us imagine that, during the course of an experiment, the subject focuses less and less attention to the task, the amplitude of the individual ERP signal should decrease. This cannot be captured using the average of the individual signals. Thus in order to improve our estimation of the ERP response recorded by non-invasive electrodes, one needs to develop methods that do not rely on averaging techniques and by allowing to estimate the parameters of the ERPs on a trial-to-trial basis. In this respect, some methods have been proposed. Most of them consider that ERPs are a stationary signal. The principle followed is a parametric approach. More conventional approaches use digital filtering [8],[9] and [10]. The

main inconvenience of this method is the overlap of the spectrum of EEG background activity and ERP signal. Some authors proposed to apply parametric identification techniques in an attempt to denoise a single trial [11]. When the SNR is poor, the efficiency of these methods decrease. Another drawback of these methods is the use of averaging all the trials as a model. The ERPs are non-stationary signals. To avoid stationary assumptions, some authors suggest using wavelet transform to improve the SNR of single trials. [12] proposed an algorithm using multiresolution and wavelet transform to extract single evoked potential from ongoing background. This algorithm presents large errors. Wang et al [13] proposed to use Donoho's method without using the average of all trials as a template. Quiroga et al [14] first suggested an algorithm using average as a template, this was improved and automatized by Ahmadi et al [15]. These methods are efficient when the SNR is greater than zero. Recently [16] , [17] and [18] have proposed similar techniques and algorithms to denoise individuals ERPs signals. The first authors use the Independent Component Analysis (ICA) technique and the second authors use the Principal Component Analysis (PCA). The two methods use the average as a template. During recording, the amplitude, the latency and the morphology of the signals varies across the trials [19]. In this case, the average is not representative of the single trials. The above authors present the results obtained by applying the methods on simulated signals with SNR greater than 0 dB whereas the SNR of the raw data is lower than 0 dB [6] and [7]. To be effective the user of the ICA technique needs to know a priori the number of sources, however, this information is not always available. Girol-dini et al [20] proposed a statistical method to detect the ERP signal based on Pearson's correlation and the average. Besides the use of the average, this method requires small time-delay variation or jitter. Other authors proposed to use adaptive algorithm such as Least Mean Square (LMS) algorithm [21]. This method gives better results when the SNR is high. The drawback of the adaptive filter method is the risk of instability of the LMS algorithm when the amplitude of the ERP signal varies over trials. To be effective the adaptive filter requires a judicious choice of these parameters such as the convergence parameter and the number of coefficients. In [22], the authors
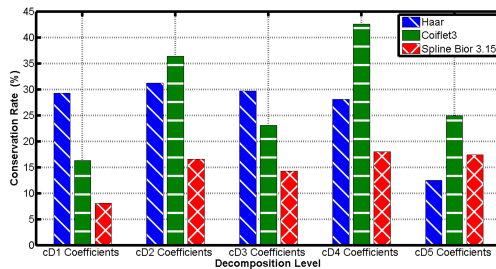
Fig (1)   Conservation rate of wavelets coefficients after threshold-
ing for different mother wavelets

proposed the adaptive Fourier model and the adaptive Walsh model.
The parameters of the Fourier model and Walsh model are adjusted
by LMS algorithm. However, the Fourier model and Walsh model
are not adapted for modeling transient signal like ERPs.

In this paper, we will present a new method to improve the SNR of
individual ERP trial. The method associate multiresolution wavelet
and bootstrap techniques. In this study, we show its performance
with simulated and real data as well as performances of other similar
methods.

2   Method

Each measurement signal representing brain electrical activity is
expressed as the sum of the event-related potentials and the EEG
background activity as:

$$x_i(n) = S_i(n) + b_i(n) \tag{1}$$

$i = 1, 2, \ldots, M \, ; 0 \le n \le N - 1$

where $S(n)$ is the noise-free signal at the discrete-time n and the
$b(n)$ is the noise component at the discrete-time n. M and N are
respectively the number of trials and the number of samples per trial.
The index i is the trial index number.

The proposed method is based on the association of two tech-
niques, wavelet transform and bootstrap method, which we call
Wavelets Statistical Denoising (WaSDe). The wavelet transform,
bootstrap method and the WaSDe algorithm are described below.

2.1   Wavelets transform

The principle underlying the wavelet transform is the description of
the temporal evolution of a signal through different time scales, by
providing information about the local regularity. The wavelet trans-
form is based on simple principles. Let's consider first a function
not infinite in time and which can be non-zero over a short period
of time. This function is called the analyzing wavelet. The wavelet
transform consists of, at any given time position, dilating or com-
pressing the analyzing wavelet by a scale factor, and to compute the
weighted product of the analyzing wavelet with the signal, for each
value of the scale factor. The weighted product is called the wavelet
coefficient. The wavelet coefficient will be high if there is a match
between the frequency of the analyzing wavelet and the signal to
analyze. Thus, a short-lasting activity will be detected at a low scale
factor, and inversely, a long-lasting activity will be detected at a large
scale factor.

The continuous wavelet transform (CWT) of a signal x is defined
for each times $t \in \mathbb{R}$ as:

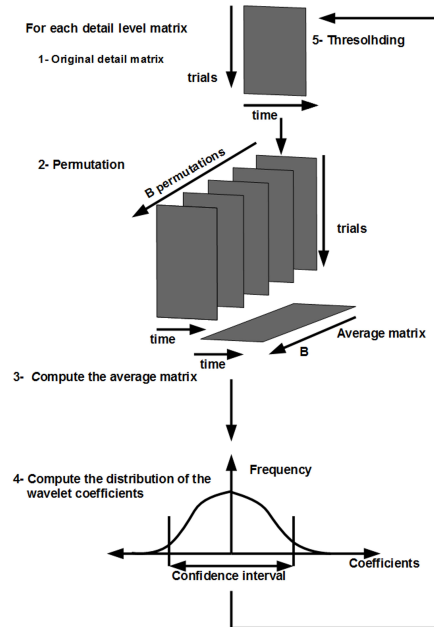$$W_x(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \psi\left(\frac{t - b}{a}\right) x(t) \, dt \tag{2}$$

Fig (2)   Flowchart of the WaSDe method.

where ψ is the wavelet function, a is the scale factor and b is
the time shift. The factor $\frac{1}{\sqrt{a}}$ was introduced to guarantee energy
preservation.

2.2   Multiresolution Wavelet Transform

The CWT is redundant and not efficient for algorithm implemen-
tations. To avoid redundancy and to increase the efficiency of
algorithm implementations, the multiresolution wavelet transform
(MWT) was introduced. It is based on the theory developed by Mal-
lat [18], it is defined at discrete scales a and discrete times b by
choosing the dyadic (basis 2) set of parameters $a = 2j$ and $d = k2j$
, where $j \in \mathbb{Z}$ and $k \in \mathbb{Z}$ (Z being the integer set). Mathematically,
the MWT has the desirable feature that the signal is decomposed
on an orthonormal basis. Interpretation of signals is facilitated by
their decomposition into coarse components a_{2j} (low frequency)
and detail components d_{2j} (high frequency), where j is the scale,
or resolution, at which the signal is analyzed. These components are
obtained at successively finer scales; thus, at each successive scale
j, some of the details are reduced and only the coarser features are
retained. In frequency domain, this process is analogous to lowpass
filtering to obtain coarse components or an approximation of a sig-
nal, while the residual details could by readily obtained by passing
the signal through nonoverlapping bandpass filters. As this process
is repeated at successive scales, the discrete signal $x(n)$,$n \in \mathbb{Z}$, is
filtered by various lowpass and bandpass filters to obtain coarse and
detail components, respectively [23] and [24].

Decomposition of coefficients in a wavelet orthogonal basis are
computed with a fast algorithm developed by [23] and [24]. This
algorithm cascades a discrete convolutions with h and g, and sub-
samples by a factor of two the output. h and g are respectively
coefficients of a lowpass filter and a bandpass filter. There are two
basis functions in wavelet domain. The first is the primary wavelet

2

$\psi(t)$ and the second is the scaling function $\phi(t)$. Decomposition of coefficients can be calculated as follows [23] and [24]:

$$a_{2j+1}(n) = \sum_{k \in \mathbb{Z}} h(k - 2n)\, a_{2j}(k) \qquad (3)$$

$$h(n) = 2^{\frac{-1}{2}} \int_t \psi(\tfrac{1}{2}t)\, \phi(t - n)\, dt \qquad (4)$$

$$d_{2j+1}(n) = \sum_{k \in \mathbb{Z}} g(k - 2n)\, d_{2j}(k) \qquad (5)$$

$$g(n) = 2^{\frac{-1}{2}} \int_t \phi(\tfrac{1}{2})\, \phi(t - n)\, dt \qquad (6)$$

a_{2j} and d_{2j} represent respectively the approximation and details coefficients at different scales j.

The signal is reconstructed as follows:

$$a_{2j}(n) = \sum_{k \in \mathbb{Z}} h(n - 2k)\, a_{2j+1}(k) + \sum_{k \in \mathbb{Z}} g(n - 2k)\, d_{2j+1}(k) \qquad (7)$$

At scale $j = 0$, $a_{2^0} = x(k)$

2.3 The choice of the wavelet

Most applications of wavelet bases exploit their capacity to efficiently approximate particular classes of functions with a few non-zero wavelet coefficients [25]. This is true not only for data compression but also for denoising. The design ψ must therefore be optimized to produce to a maximum number of wavelet coefficients that are close to zero [25] and [26]. This depends mostly on the regularity of the signal, the number of vanishing moments of ψ and the size of its support. We chose quadratic B-Splines (bior 3.15), for their similarity to the general shapes of evoked potentials. B-spline functions are probably the simplest functions with small supports that are the most efficient for both software and hardware implementation and permit the best reconstruction of the signal [14],[15] and [17].

To justify our choice of the wavelet, we compared the capacity of three wavelets: Haar, coiflet 3 and spline bior 3.15 to compress the energy of the signal in a limited but large wavelet coefficients. In order to compare the three wavelets, we applied the WaSDe algorithm to denoise the real signals. The evaluation is based on the coefficient ratio kept after thresholding. The figure 1 shows that the wavelet spline bior 3.15 is the one that has kept the least number of wavelet coefficients. This result confirms our choice.

2.4 Bootstrap

The bootstrap is a computer-intensive technique proposed and introduced by Efron in 1979 [27, 28]. This technique is applied when the size of data set is small and when the repetition of the experiment is difficult to achieve. The purpose of the bootstrap is to replace the unknown distribution of original data set with empirical distribution. The bootstrap is a statistic inference technique based on the sampling distributions. The method consists in reusing the original data set through resampling several times, to generate the bootstrap sample. The new bootstrap sample has the same size as the original data set. The bootstrap procedure consists in generating hundreds of new data sets by sampling with a replacement from our original data set. Each data sample can be drawn once, twice, or more or not at all. The second step consists in calculating the bootstrap distribution and the statistics parameters such as mean, standard deviation, variance and the confidence interval. In this study, we propose a new bootstrap procedure to determine the empirical distribution without modifying the original data set. We describe below the new bootstrap procedure and the WaSDe algorithm that we used to denoise ERPs signals.

2.5 WasDe Algorithm

Denoising is one of the major applications of MWT algorithm. It is performed both in frequency and time domain. Its resolution capabilities are beyond those of classical methods. The basic algorithm for denoising by using MWT is simple and proceeds in three steps:

1. wavelet decomposition of the signal,
2. thresholding the wavelet coefficients according a criterion,
3. reconstruction of the signal by using the remaining detail coefficients and the coarse coefficients.

The second step is the most important. The main problem is to determine the threshold in order to dissociate the coefficients related to the background activity (noise) from the coefficients related to the event-related potentials. The proposed method is based on the property of the orthogonal wavelet which compresses the energy of the signal in a relatively low number of large coefficients. On the contrary, the energy of the noise is spread across the whole transform and provides small coefficients. Thus, in the wavelets domain, signal and noise can be dissociated. In typical ERPs measurements and experiments, several trials are recorded under the same conditions. Each trial is wavelet decomposed into L detail levels $(D_1,\ldots,D_L)$, and the approximation A_L. For each decomposition level, the wavelet coefficients of all the sweeps are stored in a matrix whose horizontal rows represent the trials and vertical rows represent successive time points. With $L = 5$, we obtain six matrices $(D_1,D_2,D_3,D_4,D_5,A_5)$. In order to separate the signal-related coefficients and noise-related coefficients in each matrix of each level, we estimate the empirical distribution of the coefficients, through random permutation of the coefficients of each detail level matrix several times. For each random permutation, we obtain a new matrix. We compute the marginal mean values of each column, thus we obtain a mean vector. The mean vector will be stored in a new matrix called average matrix. The procedure is repeated with each new random permutation. The number of rows of the average matrix is equal to the number of random permutations. One thus obtains the empirical distribution of the mean of the coefficients from the mean matrix, and from there computes the inferior threshold $q1$ and the superior threshold $q2$ corresponding to a confidence set at 0.05. The coefficients whose significance is outside this confidence interval $[q1\ q2]$ are kept while the other coefficients are set to zero. This threshold is applied to the wavelet coefficients for each detail levels. After thresholding the wavelet coefficients, one can then calculate the inverse wavelet transform to obtain the denoised signal. We chose to use soft threshold because in addition to removing the coefficients with the absolute value below the threshold, it shrinks the kept wavelet coefficients by the value of the threshold. Indeed, the SNR of the recordings is low; the noise can also contaminate the large coefficients related to the ERPs signals. A recent study confirms our choice[29]. We apply the soft thresholding only on the coefficients of the details levels ($D_j, j = 1 : 5$).

The soft thresholding is defined as follows:

$$\begin{cases} 0 & q1 \le d_{j,n} \le q2 \\ d_{j,n} - q2 & d_{j,n} > q2 \\ d_{j,n} + q1 & d_{j,n} < q1 \end{cases}$$

We describe the WaSDe algorithm as follows:

M signals are decomposed in five scale levels by using the wavelet multiresolution decomposition. The coefficients are stored in the corresponding matrix whose number of rows is M. The Details matrix are noted D_j, where $j = 1 : 5$ and the coarse approximation is noted A_5.

For each detail level matrix, repeat the below steps 1, 2 and 3, B times:

1. Permute the matrix coefficients by using the bootstrap method.
2. Calculate the average of the coefficients matrix.
3. Store the average in the average matrix, whose number of rows is B.
4. Construct the empirical distribution of the average matrix.

3

5. Find the 95 % bootstrap confidence interval $[q1, q2]$.
6. Set to zero the coefficients inside the confidence interval of the original matrix D_n.
7. Compute the inverse transform, thus obtaining the denoised single ERP.

The value of the repetition number B is fixed to 1000 [28].

We also present a flowchart (Fig. 2) that shows the sequence of the different steps of the WaSDe algorithm.

3 Data processing

3.1 Simulation

For the simulation, the synthetic ERPs signals were constructed by superimposing 3 functions (Fig. 3-G) to simulated ERP signal. Sixty different sequences have been generated, each sequence containing a total of 512 samples. We further introduce, using a Gaussian distribution (variance = 5), a random fluctuation in the position of the peaks in order to mimic latency variability observed in the real data . The background EEG activity is simulated and added to the signal. The EEG signal was generated by a Matlab program proposed by Rafal Bogacz and Nick Yeung (Princeton University, December 2002). The EEG noise was scaled to alter the SNR. The mean value of the SNR is -10 dB.

3.2 Simulation parameters

In order to evaluate the performances of our method in simulation, we used three parameters. The first, the mean square error (MSE).The MSE was computed across the sixty sweeps after applying the WaSDe method. The mean square error (MSE) is defined as:

$$MSE = \frac{1}{N} \sum_{n=1}^{N} (S(n) - x(n))^2 \qquad (8)$$

where $N = 512$ is the total number of samples in each simulated waveform, $x(n)$ is the denoised signal and $S(n)$ is the original signal (without noise). The MSE is computed for each sweep. The second parameter is Signal Noise Ratio (SNR). The SNR is defined as:

$$SNR = 10 \log \left(\frac{\sum_{n=1}^{N} S^2(n)}{\sum_{n=1}^{N} (S(n) - y(n))^2} \right) \qquad (9)$$

The third parameter is the percentage root-mean-square difference (PRD).The PRD is defined as:

$$PRD = \sqrt{\left(\frac{\sum_{n=1}^{N} (S(n) - x(n))^2}{\sum_{n=1}^{N} S^2(n)} \right)} \times 100\% \qquad (10)$$

where $S(n)$ is the noise-free signal and $y(n)$ is the denoised signal. The WaSDe was compared with the progressive ensemble average (PEA). The PEA is defined as:

$$y_{PEA}(n) = \frac{1}{i} \sum_{j=1}^{M} x_j(n) \qquad (11)$$

with $j = 1, \ldots, M$.

Where i is the stimuli index number and M is the total number of stimuli or trials.

The PEA method consists, at each new recording denoted i, in performing the average of the previous recordings with the new recording i.

3.3 Results and analysis

Fig. 3 shows the contour plots of noisy signals, noise-free signals and the outputs of the WaSDe, Hermite [37], NZT [15] and Wang method [13]. We can observe that the signals denoised by using WaSDe method are similar to the noise-free signals and clearly recognizable in the contour plots unlike the other methods. More specifically, the two sharp peaks between 100 and 200 ms, the peak at 400 ms and the peaks between 500 ms and 650 ms are better reconstructed with the WaSDe method. These peaks are not easily identified in the noisy trials due to their similarity with background EEG activity. We can also show that the variability in amplitude and latency of the single signals across trials is preserved. The variability across trials is an important factor to investigate the brain disease such as depression [30]. The noise free signal is flat before 0 ms and after 900 ms.The contour plot of the WaSDe method shows that the noise is removed before 0 ms and after 900 ms. This result proves the efficiency of the WaSDe algorithm to separate the wavelet coefficients related to the ERP signal from those related to the noise unlike the other methods. Fig. 3 also shows the average of all noisy signals and the average of the denoised signals by using WaSDe method. We can observe that the average of the signals denoised by using the WaSDe method has a high SNR. Fig. 4 (a) compares the SNR of the simulated individual signals before and after denoising by using WaSDe, Hermite, NZT and Wang methods. The first point on the curve is the average of the first five SNR values and the second point represented is the average of the next five SNR values and so on for the other points. In Fig. 4 (a), the WaSDe method showed a significant improvement in the estimation of the single trials unlike the other methods. The average value of the SNR of the simulated data is -4.2 dB and the average
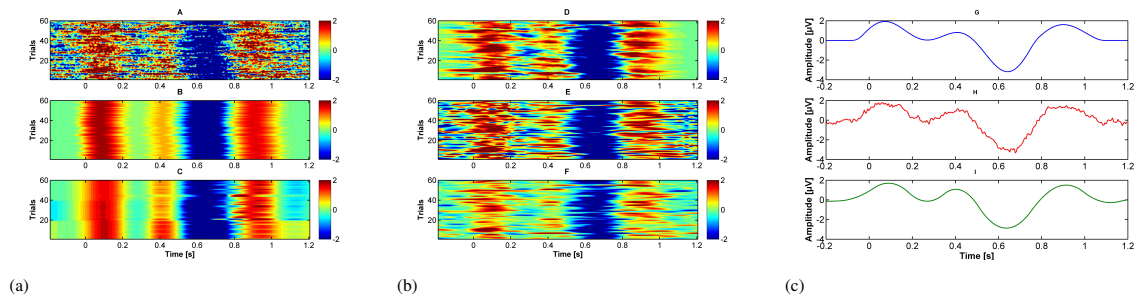

(a)   (b)   (c)

Fig (3)   Contour plot of the simulated noisy signals (A), noisy-free signals (B), denoised signals using the WaSDe method (C), Hemite method (D), NZT method (E) and Wang's approach (F) and the average of the noisy-free signals (G) , the average of the noisy signals (H) and the average of the denoised signals by using the WaSDe method (I)
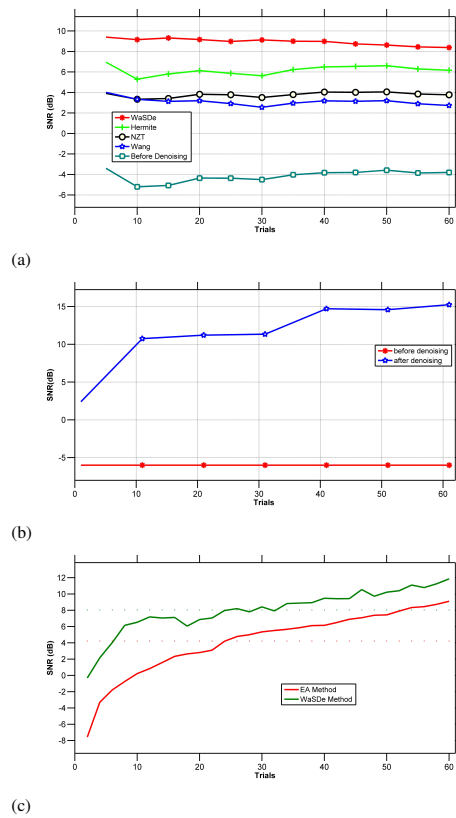
4

(a)

(b)

(c)

Fig (4)   Comparison of the SNR of simulated data .obtained after denoising by using WaSDe, Hermite, NZT and Wang method (a), the evolution of the SNR as a function of trials (b) and the comparison of the SNR of the progressive average of the noisy trails and the progressive average of the denoised trials by using WaSDe method (c)
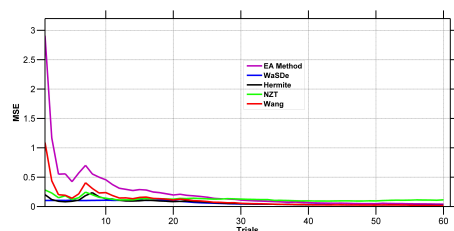

Fig (5)   Comparison of the MSE obtained after denoising by using WaSDe method , Hermite , NZT and Wang's method

value of the SNR produced by WaSDe method is +8.2 dB. Fig. 4 (b) shows the evolution of the SNR as a function of the number of trials . The level of the SNR remains stable despite the reduction in the number of signals. This result highlights an important advantage of the WaSDe method which is the reduction of the necessary number of signals to calculate the average with a high SNR. Reducing the number prevents the subject from mental fatigue [31]. This advantage is due to the use of the bootstrap technique. Fig. 4 (c), compares the SNR for each ERP trial denoised using WaSDe method (red curve) and the PEA method (green curve). We can observe that
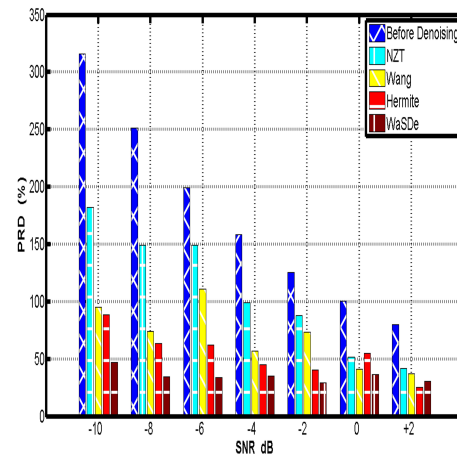

Fig (6)   Comparison of the PRD of the WaSDe, NZT, Hermite and Wang method

for the WaSDe method, the SNR reaches the positive value much faster, whereas the PEA reaches the positive SNR after ten trials. For the PEA method, the SNR continues to improve with the increase of the number of the trials. This is a predictable result as stated in the introduction section of this paper. The average value of the improvement of the SNR obtained by using the WasDe method compared to the PEA method is +12.5 dB. Given, the SNR obtained on those denoised data, it becomes very easy to estimate the single sweep relevant parameters (peak amplitude and latency).We evaluate the signal quality by computing the MSE based on the noise-free signal. In Fig. 5, plots show the comparison of the MSE produced by the WaSDe method and the other methods. The WaSDe method consistently produces a lower MSE unlike the other methods. We can observe that the MSE produced by the WaSDe method converges much faster to a minimum unlike the other methods and there is no improvement of the MSE with additional trials. Fig. 6 compares the PRD produced by the WaSDe method and the other's methods for different SNR levels. We can observe that the WaSDe method produces a lower PRD unlike the other methods when the SNR is inferior to zero and produces a comparable result when the SNR is greater than zero. It is known that the SNR of the ERPs is low; one can conclude that the WaSDe method is more adapted to extract the single-trial ERPs from noise.

4   Real data processing

We applied the WasDe method to real data, experimental EEG data. We analyzed the data set relative to the study of the Error Negativity waveform (Ne).

4.1   Error Negativity

The aptitude of a subject to detect its errors is necessary to regulate his behavior. When subjects commit an error in reaction time choice tasks, a negative wave is observed in EEG signal [32]. This wave is called error negativity(Ne). The signal is composed by a negative wave (Ne) followed by a positive wave (Pe).

4.2   Subjects

Ten subjects aged 20 to 31 years (mean: 25 years) volunteered for the experiment. All of them were right-handed and had normal or corrected-to-normal vision. According to the declaration of Helsinki,

each subject signed a written informed consent before its beging of the experiment [32]. All subjects were volunteers and gave written informed consent to the studies.

4.3   Task

The subjects performed an Eriksen's flanker task (Eriksen and Eriksen, 1974) [32]. On each trial, three letters were presented to subjects who had to respond to the central one (target) while ignoring the others (distractors). They ran 20 experimental blocks of 128 trials each.

4.4   Recording and preprocessing

Electroencephalographic activity (EEG) was recorded with 64 Ag/AgCl scalp electrodes (10-20 system positions, BIOSEMI Active-two electrodes, Amsterdam) and electromyographic activity (EMG) from the flexor pollicis brevis of each hand was recorded by paired surface Ag/AgCl electrodes. The sampling rate was 1024 Hz (filters: DC to 268 Hz, 3 dB/octave). The data were off line referenced to the left mastoid. The vertical and horizontal EOG was recorded in order to correct eye movement artifacts by the statistical method of Gratton et al [33].

The difficulties of removing the ocular artefacts from the EEG signal are the spectral overlaps between the EOG signal and the EEG signal and their shape similarity. Gratton et al. proposed an off-line subtraction method to remove EOG signals from EEG. They used the time linear regression technique to estimate the propagation factor between the EOG electrode and each EEG electrode. The propagation factor decreases from frontal to parietal electrodes. For each EEG electrode, the sum of the vertical and horizontal EOG signals weighted by their corresponding propagation factor is subtracted from the EEG signal.

4.5   Real data parameters

In order to evaluate the performance of the WaSDe method to extract the single ERP signal from noise, we estimate the SNR by using the method proposed by [34]. The method is based on the establishing of the correlation between successive noisy single trials.

Let two noisy trials,

$$x_1(n) = S_1(n) + b_1(n) \tag{12}$$

$$x_2(n) = S_2(n) + b_2(n) \tag{13}$$

and the correlation coefficient ,

$$r = \frac{\frac{1}{N}\sum_{n=1}^{N} x_1(n).x_2(n)}{\sqrt{\frac{1}{N}\sum_{n=1}^{N} x_1(n).\frac{1}{N}\sum_{n=1}^{N} x_2(n)}} \tag{14}$$

where N is the number of samples per trial.

The correlation between noises is low. The correlation between the signal and noise is also low. When the noise is reduced, the denominator becomes smaller and the numerator remains unchanged [36]. The SNR estimator called $SNRr$ is given by [34]:

$$SNRr = A\frac{r}{1-r} + B \tag{15}$$

where A and B are chosen to make $SNRr$ unbiased estimator [34] and [35], with

$$A = \exp\left(\frac{-2}{N-3}\right) \tag{16}$$

and :

$$B = \frac{1}{2}\left(1 - \exp\left(\frac{-2}{N-3}\right)\right) \tag{17}$$

4.6   Results and data Analysis

The single trials are visualized through the "ERP_ images". To construct ERP-images, the trials are then plotted as parallel colored lines, with the x-axis representing time, the y-axis representing the arranged trials, and a color code indicating the polarity and the intensity of the signal for each trial and each time point. The time zero indicates the onset of the stimulus (Fig. 7). Fig. 7 shows the real data before and after denoised the single trials by using WaSDe, Wang, PEA, NZT and Hermite method. In the WaSDe contour plot, we can observe clearly the Ne wave at about 300 ms and the Pe wave peaks at about 600ms post stimulus unlike the other methods. Fig. 8a shows the average of the raw data and the average of the denoised signals by using WaSDe method. The Ne and Pe are clearly recognizable at respectively 300 ms and 600 ms. The Fig. 8b shows the histogram of all trials before and after denoising by using WaSDe method. The shown distributions have standard deviations before and after denoising, with respective values 12.02 and 7.47. This shows the efficiency of WaSDe method in increasing SNR for single trials. Table 1 summarises the results of the SNRr estimator for the four compared methods. The SNRr is calculated by using the method proposed by Bezerianos et al. [34]. Each SNRr value in Table 1 is calculated as the average of fourteen SNRr values. The WaSDe gave the greatest SNRr improvement followed by the Hermite whereas the two wavelet approaches gave the worst SNRr enhancement. The results obtained after filtering the real data confirm the results obtained by simulations.

5   Discussion

Using the average of ERP responses to analyse the behavior of a subject during the execution of a cognitive task is not sufficient. An analysis of the single trials is essential. Besides having a poor SNR, the extraction of single trials presents other difficulties. Firstly, the number of recordings is relatively low; it varies between 30 and 60. This limitation is due to the complexity of certain cognitive tasks which can lead to the fatigue of the subjects especially those who are either sick or aged. The Second difficulty comes from the non-stationary nature of the ERP signals. In fact, changes in the degree of fatigue, habit, or the attention level nature of the subject can affect the magnitude and latency of ERP signals. To overcome these difficulties, we introduce a new method based on the dyadic wavelet decomposition associated to bootstrap method to improve the SNR of single trials. We used dyadic wavelet decomposition because wavelet basis is more adapted to transient signals like ERPs than Fourier basis. We used the bootstrap to empirically estimate the distribution of the coefficients, to avoid making assumptions about the form of the distribution of the wavelet coefficients. In [38], the authors assume the distribution of the wavelet coefficients Gaussian so as to simplify the analytical calculations. The second reason to use the bootstrap method is the improvement of the estimation of the distribution of the coefficients when the amount of samples is small. The third reason is the difficulty to record a large number of trials when the subject is sick or aged. Contrary to previous method, the WaSDe method can be applied with minimum assumptions about the characteristics of the signals. That is; no assumption regarding the noise distribution and no parameters to adjust. The WaSDe method is an automatic method that can be applied to any electrode. Several previous methods used the average as a template. In fact, many methods proposed require a priori information about the temporal and spectral characteristics of recordings. The performances of the ICA method depend on the choice of the number of sources. The results obtained in simulation show the efficiency of the method to filter the individual signals by improving the SNR and preserving the variability in amplitude and latency of the individual signals from trial-to-trial. The MSE shows the capacity of our method to reach a minimum level faster than the average method (PEA) and therefore the reduction in the number of records necessary to obtain the individual signals with high SNR. Reducing the number of trials avoids the subject mental fatigue and the decrease of his attention to the stimuli and consequently avoids the attenuation of the amplitude of

Table (1) Comparison of estimated SNR after denoising the real data between NZT, Wang, Hermite and WaSDe method

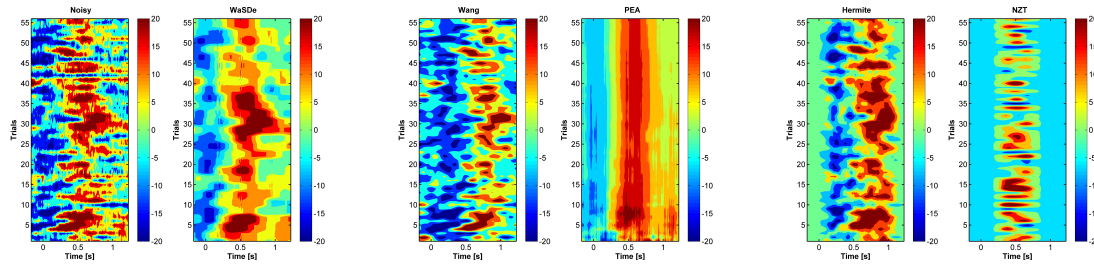|	SNRr	Improvement	SNRr	Improvement	SNRr	Improvement	SNRr	Improvement
Original	-6,739	-	-3,043	-	-2,314	-	0,854	-
NZT	-6,764	-0,025	-1,301	1,741	-1,626	0,687	1,514	0,660
Wang	0	6,739	-8,187	-5,144	3,321	5,635	0	-0,854
Hermite	11,175	17,914	0,950	3,993	10,014	12,328	12,687	11,832
WaSDe	25,238	31,977	7,487	10,531	22,691	25,006	26,809	25,954

Fig (7)   Real data from Fz electrode before denoising and the method's outputs

single ERPs signals. The results obtained on real data confirm the results obtained in simulation. Indeed, we can observe an improvement in the SNR of the individual signals as well as the preservation of dynamic changes in latency and amplitude of the signals after denoising useful to investigate the brain disease [30].

In summary, we propose the WaSDe method in order to denoise the single ERP signals. The WaSDe method has several advantages. First, it does not require any a priori knowledge about the features of the ERPs signal, secondly, thanks to the bootstrap technique, the WaSDe method does not require several trials to estimate the single-trial ERPs signals with a better signal-to-noise ratio. This advantage allows us to reduce the recording time and prevent the sujet from mental fatigue. Finally, the WaSDe method is an automatic method that require a minimum of adjustment. Unlike the ICA method, the WaSDe method does not require multichannel EEG recordings. The WaSDe denoise each signal independently of the signals from neighboring electrodes.

The WaSDe method requires computing power; this disadvantage is overcome by the every increasing power of current computers. Real time is not required in this application. The program was run on a PC equipped with an Intel i3 processor with 4GB RAM. The processing time for one channel is 0.8s. This processing time is reasonable in neurophysiology.

6   Conclusion

Given the low SNR obtained in ERPs data, averaging across trials has long been the only method to increase SNR. However, averaging induces important artifacts both in terms of shapes and latencies and prevents any trial-by-trial analysis. For this reason, it is important to develop methods that go beyond averaging and allow analysis of single sweeps. Here, we propose a method that associates multiresolution wavelet and bootstrap, in order to identify the activities that appear systematically, with a relatively low dispersion, around a specified event. Contrary to other available methods (including ICA [16] and [17]), the proposed method is not a "data_ mining" approach. Indeed, the method only allows one to study, at the single sweep level, what is observed on average. The method allows one to detect hidden information, not visible on average. However, by allowing a quantitative measure on every trial, it opens the possibility to correlate brain activities to performance, or to other brain activities, improving our understanding of the functional role of brain waves and / or brain areas.

(a)   (b)

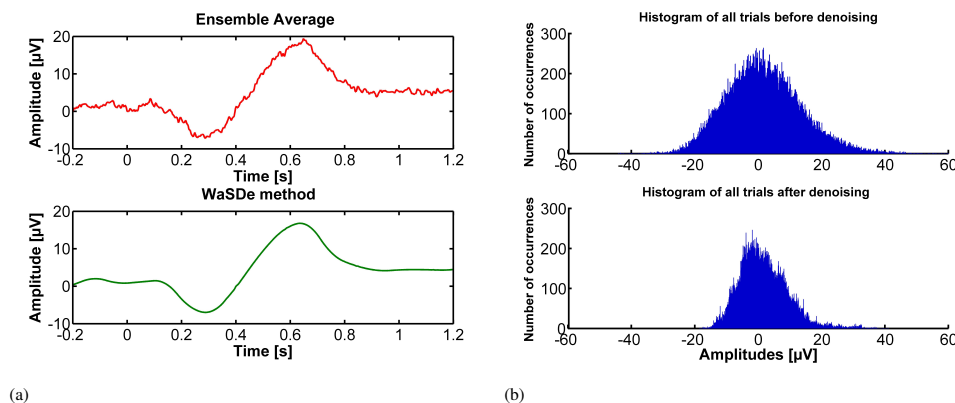Fig (8)   The average (a) and the histogram (b) of the all real noisy trials before and after denoising by using the WaSDe method

                                                                                                                    7

7 References

[1] Jeste, S.S., Nelson, C.A.: 'Event related potentials in the understanding of autism spectrum disorders: an analytical review', *Journal of autism and developmental disorders*, 2009, **39**, (3), pp. 495–510

[2] Mayaud, L., Congedo, M., Laghenhove, A.V. et al .: 'A comparison of recording modalities of p300 event-related potentials (erp) for brain-computer interface (bci) paradigm', *Neurophysiologie Clinique/Clinical Neurophysiology*, 2013, **43**, (4), pp. 217–227

[3] Dehais, F., Causse, M., Vachon, F.et al .: 'Failure to detect critical auditory alerts in the cockpit: evidence for inattentional deafness', *Human factors*, 2014, **56**, (4), pp. 631–644

[4] Scannella, S., Causse, M., Chauveau, N.et al . : 'Effects of the audiovisual conflict on auditory early processes', *International Journal of Psychophysiology*, 2013, **89**, (1), pp. 115–122

[5] Giraudet, L., St-Louis, M.E., Causse, M. 'Electrophysiological correlates of inattentional deafness: no hearing without listening'.In: HFES Europe Chapter Conference. 2012. pp. 89

[6] Sanei, S., Chambers, J.: 'EEG signal processing'. (John Wiley & Sons, 2013)

[7] Ifeachor, E.C., Jervis, B.W.: 'Digital signal processing: a practical approach'.(Pearson Education, 2002)

[8] Fridman, J., John, E.R., Bergelson, M. et al .: 'Application of digital filtering and automatic peak detection to brain stem auditory evoked potential', *Electroencephalography and clinical neurophysiology*, 1982, **53**, (4), pp. 405–416

[9] Boston, J.R., Ainslie, P.J.: 'Effects of analog and digital filtering on brain stem auditory evoked potentials', *Clinical Neurophysiology*, 1980, **48**, (3), pp. 361–364

[10] Wastell, D.G.: 'The application of low-pass linear filters to evoked potential data: filtering without phase distortion', *Electroencephalography and Clinical Neurophysiology*, 1979, **46**, (3), pp. 355–356

[11] Rossi, L., Bianchi, A. M., Merzagora, A. et al.: 'Single trial somatosensory evoked potential extraction with arx filtering for a combined spinal cord intraoperative neuromonitoring technique', *Biomedical engineering online*, 2007, **6**, (1), pp. 2

[12] Bartnik, E.A., Blinowska, K.J., Durka, P.J.: 'Single evoked potential reconstruction by means of wavelet transform', *Biological cybernetics*, 1992, **67**, (2), pp. 175–181

[13] Wang, Z., Maier, A., Leopold, D. A. et al.: 'Single-trial evoked potential estimation using wavelets', *Computers in Biology and Medicine*, 2007, **37**, (4), pp. 463–473

[14] Quiroga, R.Q., Garcia, H.: 'Single-trial event-related potentials with wavelet denoising', *clinical neurophysiology*, 2003, **114**, (2), pp. 376–390

[15] Ahmadi, M., Quiroga, R.Q.: 'Automatic denoising of single-trial evoked potentials', *NeuroImage*, 2013, **66**, pp. 672–680

[16] Darshan, I., Diaz, J., Zouridakis, G.: 'Consistency of the auditory evoked response: the presence of aberrant responses and their effect on n100 localization', *Journal of neuroscience methods*, 2012, **208**, (2), pp. 173–180

[17] Darshan, I., Zouridakis, G.: 'Single-trial evoked potential estimation: comparison between independent component analysis and wavelet denoising', *Clinical Neurophysiology*, 2007, **118**, (3), pp. 495–504

[18] Mowla, M. R., Ng, S. C., Zilany, M. S. et al.: 'Single-trial evoked potential estimation using iterative principal component analysis', *IEEE Sensors Journal*, 2016, **16**, (18), pp. 6955–6960

[19] Jarchi, D., Sanei, S., Mohseni, H.R. et al: 'Coupled particle filtering: A new approach for p300-based analysis of mental fatigue', *Biomedical Signal Processing and Control*, 2011, **6**, (2), pp. 175–185

[20] Giroldini, W., Pederzolni, L., Bilucaglia, M. et al : 'A new method to detect event-related potentials based on pearson's correlation', *EURASIP Journal on Bioinformatics and Systems Biology*, 2016, **2016**, (1), pp. 11

[21] Thakor, N.V.: 'Adaptive filterng of evoked potentials', *IEEE Transactions on Biomedical Engineering*, 1987, (1), pp. 6–12

[22] Thakor, N.V., Guo, X.R., Vaz, C. et al.: 'Orthonormal (fourier and walsh) models of time-varying evoked potentials in neurological injury', *IEEE transactions on biomedical engineering*, 1993, **40**, (3), pp. 213–221

[23] Mallat, S.: 'A theory for multiresolution signal decomposition: the wavelet representation', *IEEE transactions on pattern analysis and machine intelligence*, 1989, **11**, (7), pp. 674–693

[24] Mallat, S.: 'A wavelet tour of signal processing'. (Elsevier, 2009)

[25] Misiti, M., Misiti, Y., Oppenheim, G. et al.: 'Wavelets and their Applications'. (John Wiley & Sons, 2013)

[26] Chui, C.: 'An introduction to wavelets'. (Elsevier, 1992)

[27] Efron, B. et al .: 'Bootstrap methods: Another look at the jackknife', *The Annals of Statistics*, 1979, **7**, (1), pp. 1–26

[28] Efron, B.: 'Computers and the theory of statistics: thinking the unthinkable', *SIAM review*, 1979, **21**, (4), pp. 460–480

[29] Bhowmick, A., Chandra, M.: 'Speech enhancement using voiced speech probability based wavelet decomposition', *Computers & Electrical Engineering*, 2017, **62**, pp. 706–718

[30] Röschke, J., Wagner, P.: 'A confirmatory study on the mechanisms behind reduced p300 waves in depression', *Neuropsychopharmacology*, 2003, **28**, (S1), pp. S9

[31] Kato, Y., Endo, H., and Kizuka, T.: 'Mental fatigue and impaired response processes: event-related brain potentials in a go/nogo task', *International Journal of Psychophysiology*, 2009, **72**, (2), pp. 204–211

[32] Roger, C., Bénar, C.G., Vidal, F. et al.: 'Rostral cingulate zone and correct response monitoring: Ica and source localization evidences for the unicity of correct-and error-negativities', *Neuroimage*, 2010, **51**, (1), pp. 391–403

[33] Gratton, C.M. G., Donchin, E.: 'A new method for off-line removal of ocular artifact', *Electroencephalography and clinical neurophysiology*, 1983, **55**, (4), pp. 468–484

[34] Bezerianos, A., Laskaris, N., Fotopoulos, S. et al. : 'Data dependent weighted averages for recording of evoked potential signals', *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 1995, **96**, (5), pp. 468–471

[35] Davila, C. E., and Mobin, M. S.: 'Weighted averaging of evoked potentials', *IEEE Transactions on Biomedical Engineering*, 1992, **39**, (4), pp. 338–345

[36] Peng, C., et al . 'Novel wavelet threshold denoising method in axle press-fit zone ultrasonic detection'. In: AIP Conference Proceedings. vol. 1806. , 2017. p. 040007

[37] Boudiaf, M., Benkherrat, M., Mansouri, K.: 'Denoising of single-trial event-related potentials using adaptive modelling', *IET Signal Processing*, 2017, **11**, (7), pp. 846–853

[38] Benkherrat, M., Burle, B., Allain, S.et al 'Individual evoked potential extraction by multiresolution wavelets decomposition'. In: Computer as a Tool, 2005. EUROCON 2005. The International Conference on. vol. 1. , 2005. pp. 417–420