ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 10(47), DOI: 10.17485/ijst/2017/v10i47/108645, December 2017* ISSN (Online) : 0974-5645

Artifact Removal from Sleep Hipnograms using Complete Ensemble Empirical Decomposition with Adaptive Noise

Vijayakumar Gurrala[1]*, Padmasai Yarlagadda[1] and Padmaraju Koppireddi[2]

[1]Departmentof Electronics and Communication Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad – 500090, Telangana, India; vijayakumar_g@vnrvjiet.in, padmasaiy@gmail.com
[2]Department of Electronics and Communication Engineering, JNTU Kakinada, Kakinada – 533003, Andhra Pradesh, India; padmaraju_k@yahoo.com

Abstract

Objectives: Many sleep related disorders can hamper the performance of an individual and there is an absolute need to research upon such disorders. **Methods /Statistical Analysis:** The neuro physiologists record an electrical activity of the brain through Electroencephalogram (EEG), since the emphasis of study is sleep. Artifacts are the recorded activities which are not the records of cerebral origin. As the EEG record itself runs for longer periods, it suffers from some artifacts. In this article Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMD-AN) method is proposed for removing the artifacts from the contaminated sleep EEG. **Findings:** In this article we consider Index of Orthogonality (IORT), Percentage Error in Energy (PEE) and Number of sifting iterations to assess the performance of CEEMD-AN method over the earlier methods are explored. **Application/Improvement:** The artifact free data gives best results for neuro physiologists to identify the disorder.

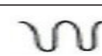
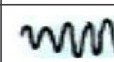
Keywords: Artifacts, CEEMD-AN, EEG Waves, Noise Adaptive Data Analysis, Sleep EEG

1. Introduction

The sleep EEG signal is recorded by placing electrodes on scalp at predefined points identified by physician. The signal is extracted by either multichannel or single channel data acquisition system. The recorded EEG has 5 waves of different frequencies and patterns of low amplitude in the order of 2 micro volts to 50 micro volts. The 5 bands of EEG are Delta, Theta, Alpha, Beta and Gamma. Each band of frequency of EEG corresponds to different activity of brain and which gives a better understanding of sleep stages. Table 1 shows the sub bands of EEG data, its frequencies, amplitude levels and shapes.

In clinical neurophysiology the brain waves such as EEG, Hipnogram, PolySomnogram (PSG) waveforms and its discharges are specified by parameters such as Amplitude, Duration, Frequency, Morphology i.e. the shapes and configuration of the discharges, Latency, Location of discharge, Reactivity that effects the dis-

Table 1. Sub bands of EEG data

EEG wave	Frequency (Hz)	Amplitude (pV)	Shape
DELTA	0.5 to 3	5-250	
THETA	4 to 7	20-100	
ALPHA	8 to 13	20-120	
BETA	14 to 30	5-50	
GAMMA	31 to 60	-10	—

*Author for correspondence

charge. Although EEG records brain activity, mostly it is contaminated with the activities not related to cerebral origin. The waves which are not originated from cerebral are termed as artifact[8]. These are mainly classified as Physiologic and Extraphysiologic artifacts. Physiologic artifacts are the activities of body other than brain whereas the Extraphysiologic artifacts are from equipment and environment.

1.1 Physiologic Artifacts

Muscle Activity-Myogenic potential discharges virtually in clinical EEG termed as Electromyograms (EMG). This is the most common muscle activity[17] of clenching of jaw muscles (chewing), which produces potentials of shorter duration than the brain of frequency 50–100 Hz. It is a rhythmic burst of muscle activity[15]; Glossokinetic Artifact- Apart from EMG[19], tongue functions like dipole, the mobility of tip acts as -ve with respect to base. The frequency of this artifact is usually observed in the delta range; Electrocardiogram (EKG or ECG) Artifacts-In some individuals the potential variations of heart over the surface of scalp creates an artifact; Eye Movements[11]-EEG recordings affected with eye movements useful in sleep stages analysis. Eyeball can also act as dipole with its cornea is positive pole, retina is negative pole. This dipole action produces an alternating current. A blink[16] can cause deflections between cornea and retina. The blinks are observed during vertical movement of eye. These eye blinking artifacts[12] are termed as ElectroOccularogram (EOG) artifacts[9]; Respiration Artifacts-Respiration produces artifacts of two types. One of it is in a slow, rhythmic activity like respiration body movements and the other one is sharp waves synchronous with inhalation and exhalation; Skin Artifacts – The metals of electrodes reacts with Sweat means sodium chloride and lactic acid and produce baseline sways of 0.5 Hz.

1.2 Extraphysiologic Artifacts

Electrode Artifacts - It is an electrode “pop” i.e. due to an abrupt change in impedance; it produces a single or multiple sharp waveforms. It is easily identified and having the appearance of vertical transient without much change in background activity. Sometimes this change is less and artifact may be a delta wave of low-voltage; Alternating Current (power line or 60 Hz) Artifact - This happens when active electrode impedance becomes large between electrode and the ground of amplifier. The artifact from power lines is eliminated by proper grounding; Movements in the Environment - The movements around patient by others generate artifacts; it is capacitive or electrostatic origin. In addition, some artifacts are due to interference. Radiation from electronic devices like radio, T.V, hospital paging systems etc. may overload EEG amplifiers. Interference with the recording system done by the usage of cutting or coagulating electrodes in operating rooms. The touch or hit of electrodes produce odd waveforms, the repetitive movement of head also generates rhythmic artifacts.

Photic Stimulation - When flashing lights are seen occipital cortex gives response of visual evoked potential during EEG recording and produce physiologic and extraphysiologic artifacts. These are called as Photic Stimulation artifacts; Photocell (photoelectric) artifacts are also considered as photic stimulation and these artifacts will disappear when one can block light from electrode.

The so far mentioned artifacts are undesirable in the recorded data hence these artifacts should be removed[10] from the collected data and the artifact free data is used as a base for classification of physiological disorders associated with sleep[1,23].

2. Materials and Methods

2.1 Fourier Transformation (FT)

Fourier analysis is one of the best tools to get spectral analysis and it applies to linear and stationary data. It involves an infinite number of exp (jwt) waves in the expansion of a signal x (t), which is not completely localized in time domain.

2.2 Short-Time Fourier Transformation

To get time dependency in FT, a simple solution is that the signal is pre-windowing at a particular time instants and obtain its Fourier transform. Repeat this process for each step in time; assume that in all windows the signal is stationary. The resulted time dependent spectrum (spectrogram) is Short Time Fourier Transform (STFT).

2.3 Wavelet Transformation (WT)

To avoid these short comings the wavelet analysis in continuous or in discrete is applied[5]. In Wavelet Transform correlate the signal with a mean - zero function called mother wavelet. Wavelet transform can provide signal amplitude spectrum in both time and frequency domains. The limitations of wavelet method are:

- Due to the fixed basis functions, the result is significantly influenced by mother wavelet and at each instants of time the shape does not match with the considered data.
- A high frequency component can be resolved in time domain, but similar is not done in frequency domain, vice versa[13].

Though it has restrictions this analysis becomes popular for non-stationary data, although, it is basically a linear technique[14].

2.4 Empirical Mode Decomposition (EMD)

Hilbert-Huang transformation (HHT) for decomposing non-stationary time-dependent data is called as Empirical Mode Decomposition (EMD) method. This method is used to process non-stationary and non-periodic signals[2]. In this adaptive technique[4] each signal is represented by different Intrinsic Mode Functions (IMF)[5].

Consider a signal to be processed is $s(t)$, it can be decomposed into sum of IMFs by using EMD process. Each IMF should have equal number of zero crossings and is symmetric with respect to its local mean. The IMFs are determined by a step by step iterative process called sifting. After decomposition the signal $s(t)$ is expressed as:

$$s(t) = \sum_{j=1}^{N} IMF_j(t) + r_N(t), \tag{1}$$

$$s(t) = \sum_{j=1}^{N+1} \mathrm{IMF_j}(t), \tag{2}$$

Where N is the number of IMFs, which are nearly orthogonal to each other, and all have nearly zero mean , and $r_n(t)$ is the final residue.

$$r_n(t) = IMF_{j+1}(t), \tag{3}$$

The criterion to determine IMF is given by

- The number of extremas and zero crossings should be either equal or differ by at most one.
- At any instants of time, the mean value of envelope defined by the local maxima and minima is zero.

The algorithm of EMD summarized as:

Algorithm 1: Algorithm of EMD,

For i = 1, 2…n

- Set IMFi (t) = S (t).
- Determine Local maxima and local Minima of the signal IMFi (t) by using cubic spline method. And assign local maxima to eu (t) and minima to el (t).
- Calculate envelopes mean e (t) by averaging eu (t) and el (t).
- Check the condition of IMFs and if not satisfied repeat the steps from 2 to 4 until the obtained IMF satisfies the above mentioned criterion.
- Store IMF.
- Subtract IMF from signal S (t) and obtain residue r (t) of S (t).

But this EMD method[6] has the drawbacks of:

- Frequent appearance of mode mixing i.e. EMD cannot identify IMF's whose frequencies are very close to each other.
- It shows undershoot and overshoot problems in the formation of upper and lower envelope.

2.5 Ensemble Empirical Mode Decomposition (EEMD)

The difficulties in EMD were overcome by a method called Ensemble Empirical Mode Decomposition (EEMD). It is a Noise-Assisted Data Analysis (NADA) method[3]. This method uses statistical characteristics of white noise[20] and this noise will be added to the processed signal. The added white noise spread uniformly over the entire time-frequency space. Now the processed signal seems to be a projection over white noise background. For such signals if EEMD is applied, it will produce noisy results[7]. EEMD[24] produces IMF's of ensemble mean of noise added EEG signal. This ensemble approach can clearly separate the scale naturally and reduce signal intermittency without any prior subjective criterion selection. For getting meaningful IMF's, noise is cancelled in ensemble trials during decomposition. Scale separation property of EMD is used in EEMD. EEMD eliminates the above mentioned drawback of mode mixing by the addition of White noise. Here Artifacts are treated as noise in the contaminated EEG and by using the decomposition method of EEMD, the signal will be reconstructed.

The algorithm of EEMD summarized as:

Algorithm 2: Algorithm of EEMD,

- Consider a white noise and add it to the signal S(t).

$$S_i(t) = S(t) + n_i(t)$$

- By using EMD algorithm, Obtain IMFs of noise added signal and store them. Repeat step 1 and step 2 for I = 1, 2,…..N with different realization of noise each time.
- Obtain the ensemble average of corresponding IMFs and residue of the decompositions as the final result.

2.6 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMD-AN)

Though EEMD process can overcome the problems mentioned in EMD process like mode mixing with the ensemble mean of IMFs. But due to the addition of noise to the signal S(t), the effect of noise is still remain in IMFs . To get a noisy free IMFs and burden of high computational load[22], a method called Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMD-AN) is proposed in this article.

The algorithm of CEEMD-AN summarized as:

Algorithm 3: Algorithm of Complete Ensemble Empirical Mode CEEMD-AN

- Decompose the noisy added signal $S(t) + w_0\, \varepsilon^i(t)$ to get first mode by EMD process:

$$C_1(t) = \frac{1}{N}\sum_{i=1}^{N} C_1^i\ ,\ \ i\,\epsilon\{1\ ,N\}$$

Where w_0 is the amplitude of the added white noise, $\varepsilon(t)$ is the white noise with unity variance.

- Determine the difference signal; $r_1(t) = r(t) - C_1(t)$.
- Add noise to $r_1(t)$ and decompose $r_1(t) + w_1 E_1(\varepsilon^i(t))$ to obtain first mode and define second mode by $C_2(t) = \frac{1}{N}\sum_{i=1}^{N} \ E_1\left(r_1(t) + w_1 E_1\left(\varepsilon^i(t)\right)\right)$.
- For K=2,….k calculate kth residue and obtain first mode. Define (k +1)th mode as follows $C_{k+1}(t) = \frac{1}{N}\sum_{i=1}^{N} \ E_1\left(r_1(t) + w_k E_k\left(\varepsilon^i(t)\right)\right)$.

Where $E_j(.)$ is a function to extract jth IMF decomposed by EMD.

- Repeat step 4 until the residue contains no more than 2 extrema. The residue mode is then defined as $R(t) = S(t) - \sum_1^k C_k(t)$.

Thus the reconstructed signal y(t) can be expressed as:

$$y(t) = \sum_{i=1}^{k} \ C_k(t) + R(t), \qquad (4)$$

With this criterion the problem of noise and the computational load effects in EEMD process were overcome by CEEMD-AN.

3. Results and Discussions

3.1 Data Acquisition

To Preprocess (Denoising) the EEG signal, first we should have either online data or the real data taken from the subject and the added noise is considered as an additional artifact information. To apply the proposed CEEMD-AN process, the data set is acquired from PhysioNet[18] website which is a publicly available data. This EEG Sleep-EDF database is a collection of 61 Polysomnograms (PSGs) from Fpz-Cz channel instead of C4-A1/C3-A2 and the corresponding hypnograms contain sleep pattern information of each subject. Each EEG signal was sampled at 100 Hz with different base and gain values. To get the signal in practical units of micro volts (uV) there should need a conversion to be done by subtracting the base from the signal and divide with gain[18]. This results the signal with an ordinate of units in micro volts and abscissa in time samples. In this work we processed a signal named 'SC4001E0-PSG_edfm' epoch of duration 60 seconds and 6000 samples is considered. Throughout the implementation we used the toolbox available in[21] and all algorithms were implemented in MATLAB R2015a. Figure 1 shows the Processed EEG PSG signal.

3.2 Intrisnsic Mode Functions(IMFs)

By applying the algorithms of EEMD and CEEMD-AN to the PSG signal, the obtained IMFs were presented in

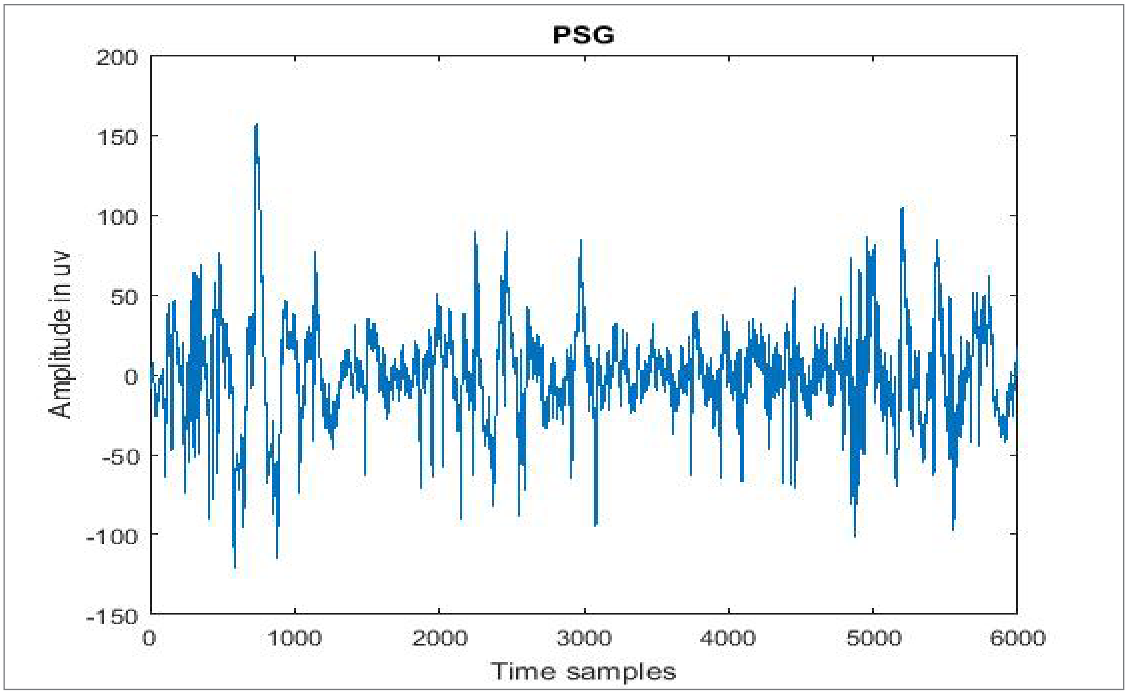


Figure 1. The EEG PSG Signal.

this section. In each plot of IMFs the first row of subplot is the processed EEG PSG signal and the remaining rows are the corresponding IMFs. By satisfying the conditions of IMFs, EMD decomposition suffers from the problems of whenever the frequencies of IMFs are close, it cannot identify them and overshoot, under shoot problems were existed while extracting maxima and minima of envelopes. By considering these issues for better processing we considered EEMD and CEEMD-AN, the advantage of CEEMD-AN over EEMD in terms of some parameters is listed in the following sections. The spectral characteristics were not considered as the point of interest in the article is on the performance evaluation of EEMD, CEEMD-AN methods in terms of time series.

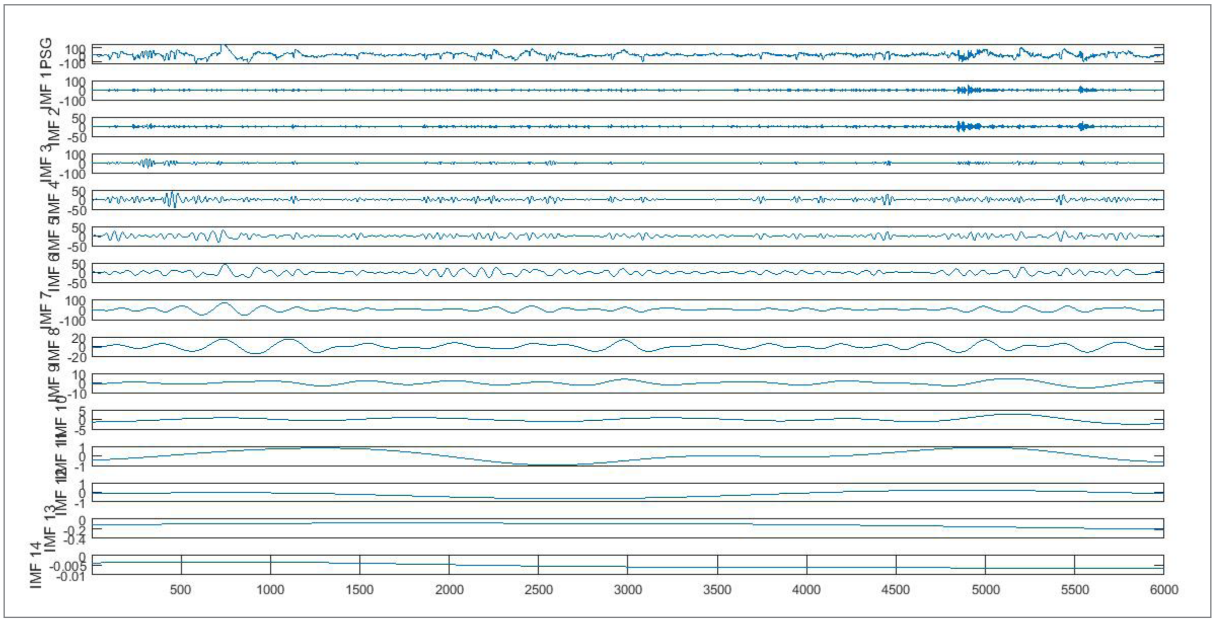


Figure 2. IMFs obtained from EEMD process

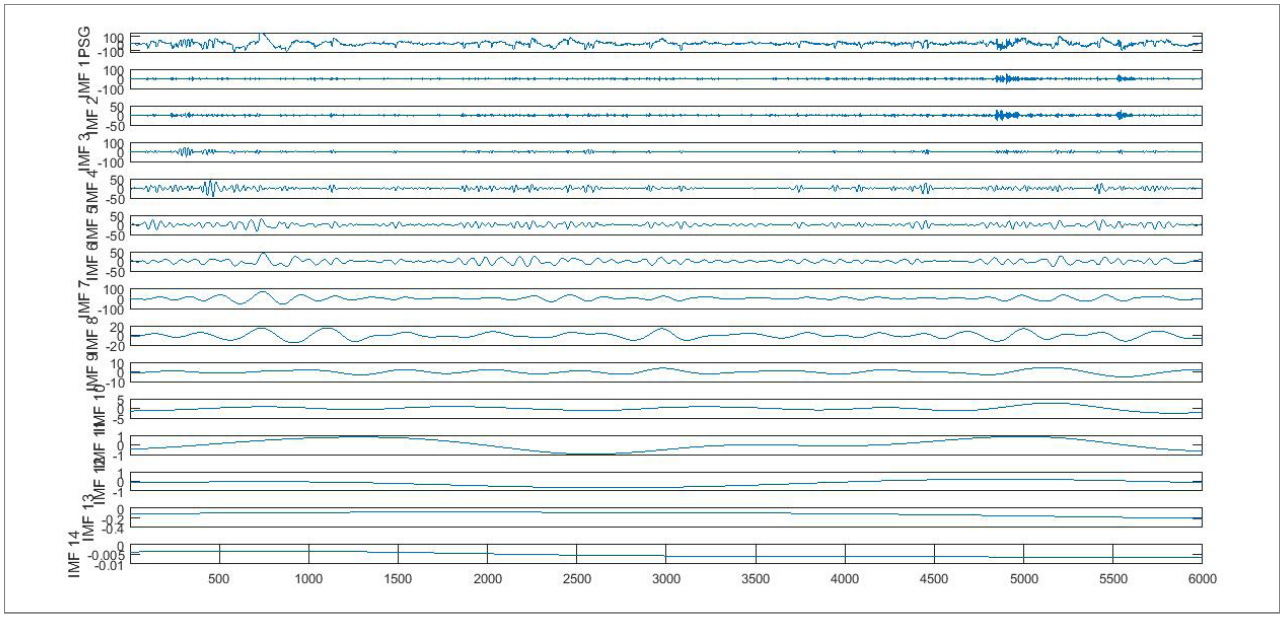


Figure 3. IMFs obtained from CEEEMD-AN process

Figure 2 and Figure 3 represents the IMFs obtained from EEMD and CEEMD-AN methods by assuming the standard deviation of noise is 0.2 and the ensemble size of 500. It is observed that almost both methods were producing some similar IMFs. The added noise is treated as artifact and that is to be removed from the contaminated PSG by using decomposition methods. The amplitude of the added noise (white noise), Wu and Huang suggested[2] to use small amplitude values for data dominated by high-frequency signals, and vice versa. These values varied to application of different non stationary signals.

3.3 Parameters

In this article we consider the following parameters to assess the performance evaluation of CEEMD-AN method.

- Index of Orthogonality (IORT).
- Percentage Error in Energy (PEE).
- Number of sifting iterations.

3.3.1 IORT

As EEG signal is decomposed into different IMFs, the reconstructed signal should almost resemble the original signal. Index of Orthogonality gives the information about amount of approximation of signal with its decomposed IMFs. The expression for IOT is given by:

$$\mathrm{IOT} \triangleq \frac{\sum_{j=1}^{n+1}\sum_{\substack{k=1\\ k\neq j}}^{n+1}\int_0^T y_j(t)y_k(t)\,dt}{\int_0^T s^2(t)\,dt}, \tag{5}$$

Where y(t) is the reconstructed signal from IMFs

3.3.2 PEE

When decomposition is applied for time series there exists an energy leakage. The IMFs are not theoretically orthogonal and hence the value of IOT is small but not zero and sometimes very severe in simulation results. Hence the Percentage Error in Energy (PEE) was considered for the amount of energy leakage is defined as:

$$\mathrm{Pee} = \frac{(E_s - E_{IMF})}{E_s} \text{ X } 100 \tag{6}$$

Where Es is the energy of the signal given by:

$$E_s = \int_0^T s^2(t)\,dt \tag{7}$$

And E_{IMF} is the sum of the energies of all IMF modes.

$$E_{IMF} = \sum_{i=1}^{n+1}\int_0^T y_i^2(t)\,dt \tag{8}$$

3.3.3 Sifting Iterations

The process of sifting is to generate IMFs by satisfying the two conditions. These involve generation of upper and lower envelopes for getting local extrema, take the average of these two envelopes and subtract the envelope mean from the signal if the conditions of IMF are not satisfied after one iteration of aforementioned procedure; the same procedure is applied to the residue signal until properties of IMF are satisfied. The process of EEMD suffers from the high computation loading which the number of iterations required are high. Figure 4 shows box plots of EEMD and CEEMD-AN methods which gives information about the number of sifting iterations for each IMF mode. It is clearly evident that the y-axis of box plot scaling 0 to 800 in EEMD and 0 to 250 in CEEMD-AN methods. By visual inspection we can say that the iterations required were less in CEEMD-AN method.

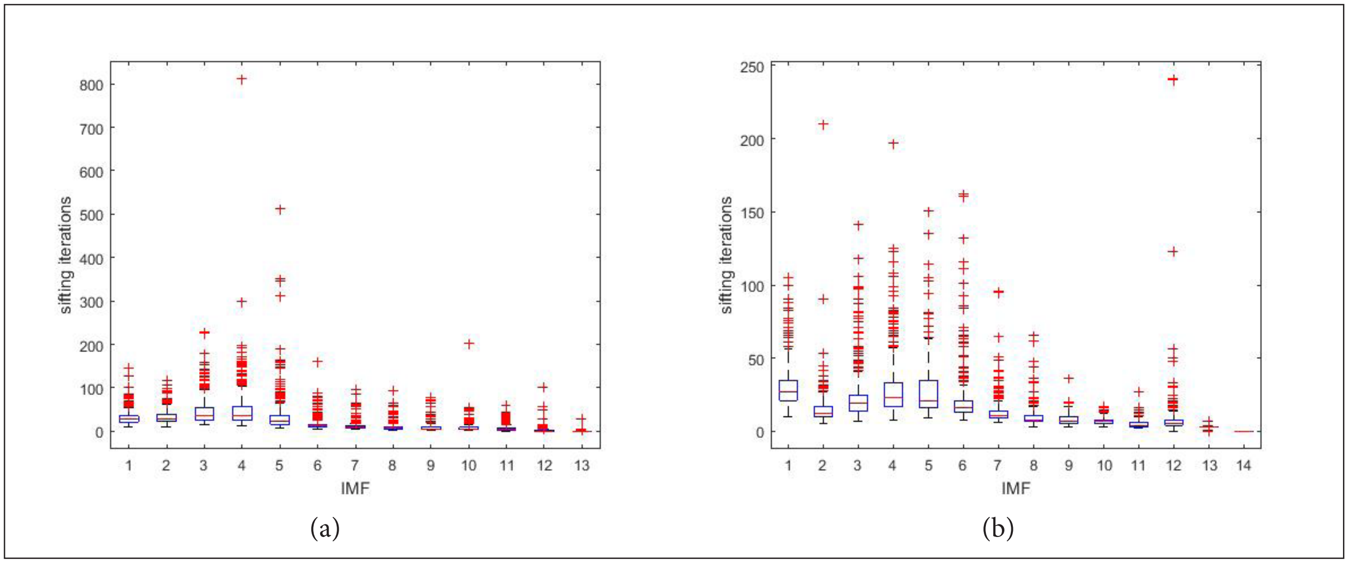


Figure 4. Boxplots of (a) EEMD and (b) CEEEMDAN processes.

The following Table 2 shows the performance assessment of CEEMD method over EEMD for processing EEG-PSG signal.

Table 2. Performance assessment of CEEMD method over EEMD

Processing Method	IORT	PEE	Iterations
EEMD	0.2155	3.28	124474
CEEMD-AN	0.1339	3.266	98582

There is a clear evident that:

- The IORT for EEMD is 0.2155 and for CEEMD-AN is 0.1339 which means by using CEEMD-AN method the index of orthogonality was increased.
- The PEE for EEMD is 3.28 and for CEEMD-AN is 3.266. There is almost an equal amount of energy leakage in both the methods but strictly said that there was an improvement in CEEMD-AN.
- The numbers of sifting iterations for EEMD are 124474 and for CEEMD-AN are 98582. Which shows that there were about 25892 less iterations were required in CEEMD-AN.i.e. CEEMD-AN required 20.8% of less iteration when compared to EEMD method.

4. Conclusion

Due to artifacts EEG PSG signal is non-linear and non- stationary and noisy. The proposed CEEMD-AN method in this article was successfully applied on Sleep EEG hypnogram data (Polysomnography) and observed that this method alleviated the problems of Index of Orthogonality, Percentage of Error in Energy and the Computation overload effects in EEMD method. Hence in this paper, CEEMD-AN method is proposed to remove artifacts from PSG signal. It shows improvement in results obtained by CEEMD-AN over EEMD. Future work shall include feature extraction from the artifact removed PSG signal which can gives best results for neuro physiologists to identify the disorders or abnormalities in sleep.

5. Acknowledgment

This work has been funded by University Grants Commissions-South Eastern Regional Office under Minor Research Project with reference number MRP-6620/16(SERO/UGC).

6. References

1. Aboalayon AIK, Ocbagabir TH, Faezipour M. Efficient sleep stage classification based on EEG signals. IEEE Systems Applications and Technology Conference (LISAT); Long Island. 2014. p. 1–6. Crossref.
2. Huang NE. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society. 1998; 454(1971):903–95.
3. Wu Z, Huang NE. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Advances in Adaptive Data Analysis. 2009; 1(1):1–41. Crossref.
4. Flandrin P, Rilling G, Goncalves P. Empirical mode decomposition as a filter bank. IEEE Signal Process Letters. 2004. 11(2):112–4. Crossref.
5. Flandrin P, Goncalves P, Rilling G. Equivalent filter banks from interpretation to applications. Interdisciplinary Mathematical Sciences. 2005; 5:57–74.
6. Weng B, Blanco-Velasco M, Earner KE. ECG denoising based on the empirical mode decomposition. International Conference of the IEEE Engineering in Medicine and Biology; 2006. p. 1–4. Crossref.
7. Chang KM. Ensemble empirical mode decomposition for high frequency ECG noise reduction. BiomedizinischeTechnik/Biomedical Engineering. 2010; 55(4):193–201. PMid: 20569049. Crossref.
8. AndererPet. Artifact processing in computerized analysis of sleep EEG. Neuropsychobiology. 1999; 43(3):150–7.
9. Croft RJ, Barry RJ. Removal of ocular artifact from the EEG A review. Clinical Neurophysioly. 2000; 30(1):5–19. Crossref.
10. Schlogl A. Artefact detection in sleep EEG by the use of Kalman filtering. EMBEC'99 Proc Medical and Biological Engineering and Computing Supplement; 1999. p. 1648–9.
11. Noureddin B, Lawrence PD, Birch GE. Online removal of eye movement and blink EEG artifacts using a High Speed Eye Tracker. IEEE Transactions on Biomedical Engineering. 2012; 59(8):2103–10. PMid: 21278013. Crossref.
12. Shahabi H, Moghim S, Zamiri-Jafarian H. EEG eye blink artifact removal by EOG modeling and Kalmanfilter. 5th International Conference on Biomedical Engineering and Informatics (BMEI); 2012. p. 496–500.
13. Krishnaven V, Jayaraman S, Aravind S, Hariharasudhan V, Ramadoss K. Automatic identification and removal of ocular artifacts from EEG using wavelet transform. Measurement Science Review. 2006; 6(4):1–13.
14. IndiradeviKP,EliasE,SathideviPS,NayakDS,Radhakrishnan K. A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram. Computers Biology and Machine. 2008; 38(7):805–16. PMid: 18550047. Crossref.
15. Gao J, Zheng C, Wang P. Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis. Clinical EEG Neuroscience. 2010; 41(1):53–9. PMid: 20307017. Crossref.
16. Bizopoulos PA, Al-Ani T, Tsalikakis DH, Tzallas AT, Fotiadis DI, Koutsouris DD. An automatic electroencephalography blinking artefact detection and removal method based on template matching and ensemble empirical mode decomposition. Engineering in Medicine and Biology Society. 2013. p. 5853–6. PMid: 24111070.
17. Safieddine D, Kachenoura A, Albera L, Birot G, Karfoul A, Pasnicu A, Biraben A, Wendling F, Senhadji L, Merlet I. Removal of muscle artifact from EEG data comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches. EURASIP Journal on Advances in Signal Processing. 2012. Crossref.
18. PhysioNet PhysioBankATM. 2016. http://www.physionet.org/physiobank/database/sleep-edfx
19. Aydemir O, Pourzare S, Kayikcioglu T. Classifying various EMG and EOG artifacts in EEG signals. Przegląd Elektrotechniczny. 2012; 88(11a):218–22.
20. Wu Z, Huang NE. A study of the characteristics of white noise using the empirical mode decomposition method. Proceedings of the Royal Society. 2004; 460(2046):1597–611. Crossref.
21. Emd Tool box. 2016. http://perso.ens-lyon.fr/patrick.flandrin/emd.html
22. Torres ME, Colominas MA, Schlotthauer G, Flandrin P. A complete ensemble empirical mode decomposition with adaptive noise. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP); 2011. p. 4144–7. Crossref.
23. Tan EBD, Leong WY. Sleep disorder detection and identification. Journal of Biomedical Science and Engineering. 2012; 5:330–40. Crossref.
24. Zhu XJ, Lv SQ, Fan LJ, Yu XL. The EEG signal process based on EEMD. IEEE International Symposium on Intelligent Processing and Trusted Computing; 2011. p. 222–5.