

A Comparison of Two Hybrid Approaches for Improving Neural Network “Simulation Study”

Mohammed Abou Elfettouh Ghazal¹, Asaad Ahmed Gad Elrab², Wafa Hamed Abd Allah^{1,*}

¹Mathematical Statistics, Faculty of Science, Damietta University, Damietta, Egypt

²Mathematics Department, Faculty of Science, Al-Azhar University, Cairo, Egypt

Abstract Traditional statistical models work on the assumption of linearity and stationary of time series. Machine learning models such as Artificial Neural Network (ANN) and Support Vector Regression (SVR) suffer the problem of over-fitting and are sensitive to parameter selection, respectively. This paper propose two models that integrates Statistical Empirical Mode Decomposition (SEMD) and ANN and Ensemble Empirical Mode Decomposition and ANN for improve the weakness of ANN. SEMD and EEMD are an adaptive technique that shifts the non-stationary and non-linear time series data till it becomes stationary. In first stage, the data is decomposed into a smaller set of Intrinsic Mode Functions (IMFs) and residuals using EEMD and SEMD. In the next stage, IMFs and residue are taken as the inputs for the ANN model. The methodology was compared with EEMD-ANN and SEMD-ANN models. The results suggest that the SEMD-ANN is better than EEMD-ANN.

Keywords Empirical mode decomposition, Statistical empirical mode decomposition, Ensemble empirical mode decomposition, Artificial neural network, Support vector regression, Discrete wavelet transform, Intrinsic mode functions

1. Introduction

Machine learning models such as Artificial Neural Network (ANN) and Support Vector Regression (SVR) suffer the problem of over-fitting and are sensitive to parameter selection, respectively [1], [2]. Hence, a slightest improvement in the prediction accuracy attracts researchers. Recently, several signal processing techniques like Discrete Wavelet Transform (DWT) and Empirical Mode Decomposition (EMD) have been used for decomposing the series in time-frequency domain and time domain, respectively [5], [6]. The process of decomposition can be considered as a data pre-processing technique for machine learning models. EMD, proposed by [4], and its extensions SEMD and EEMD decomposes a signal into a set of adaptive basis functions called Intrinsic Mode Functions (IMFs). It uses Huang-Hilbert Transform (HHT) to decompose the non-stationary and non-linear time series. Unlike DWT, it does not require the a priori information about the series, i.e., the scale of decomposition. Though DWT can handle non stationary data, it still requires linear generating process and suffers from leakage between the scales [3]. In any model building process, especially machine learning process, selection of appropriate features is considered to be an

important step. There are two types of features: (1) feature vector approach, (2) scale-based approach. Time series analysis traditionally seek out for appropriate model to fit the data; this is difficult by the fact that the data is typically non-stationary, with non-linear relationships between past and future values and conduct occurring instantaneously at diverse time scales. Adaptive data analysis methods or signal analysis methods such as Empirical Mode decomposition (EMD).

Review: Statistical Empirical Mode Decomposition

Statistical Empirical Mode Decomposition are a techniques designed to decompose a signal into its intrinsic modes [9], with each mode constrained to a limited frequency band and has seen extensive practice in the part of financial analysis. What makes EMD, SEMD, and EEMD attractive in financial analysis is that it is an empirically based method that is a posteriori and adaptive, permitting the data to speak for itself. No a priori assumptions are requisite, as is the case with traditional time-frequency methods such as Fourier or wavelet analyses. The time-frequency components obtained from above mentioned tools can simplify this task by allowing one to examine the series for one intrinsic mode function (IMF) at a time and over time horizons that are ideal for the particular IMFs. While adaptive data analysis tools are traditionally used to analyses the individual modes of a time series, usage of the method as a filter has also been recognized [4], [8]. An advantage of EMD, SEMD and EEMD-filtering are that the data still retains its nonlinearity and non-stationary, which is not the

* Corresponding author:

ayah.elkadeki@yahoo.com (Wafa Hamed Abd Allah)

Received: Jul. 15, 2020; Accepted: Aug. 23, 2020; Published: Sep. 26, 2020

Published online at <http://journal.sapub.org/statistics>

case when using conventional filtering techniques.

This article proposes an ANN model applied to data filtered with a novel adaptive data tools-filtering technique for multi-stage prediction. The SEMD and EEMD-filtered

ANN model will be tested and will be compared with an EMD-ANN applied to unfiltered data as well as to a random walk model in terms of accuracy of predictions and simulated returns on an investment.

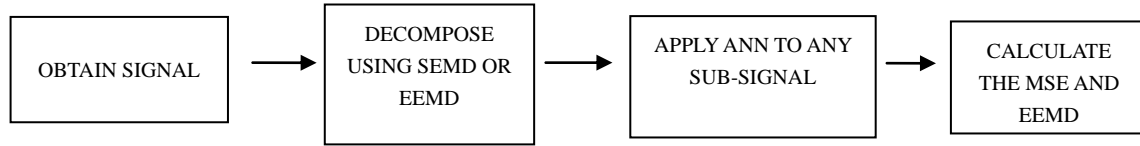


Figure 1. Flow Chart of Hybrid SEMD-ANN and EEMD-ANN Approach

In former, the IMFs at a particular time point are taken as a feature vector, while in the latter, the IMFs are predicted independently at each level and the results are combined [7]. The paper compared hybrid SEMD-ANN and EEMD-ANN techniques signal is first decomposed to different sub-series (IMFs) using SEMD and EEMD. Then, these sub-series are independently using ANN. In brief, scale-based approach has been used. The hybrid SEMD-ANN model integrates the benefits of both EEMD-ANN.

Section 2 will summarize hybrid SEMD-ANN and EEMD-ANN. Section 3 will introduce the EEMD method, illustrate more details of the drawbacks associated with mode mixing, present concepts of noise-assisted data analysis and of noise-assisted signal extraction, and introduce the EEMD in detail and statistical empirical mode decomposition SEMD. Section 4 will display the usefulness and capability of the EEMD-ANN and SEMD-ANN through simulation study. A summary and conclusion will be presented in the final section.

2. Hybrid SEMD-ANN, EEMD-ANN and EMD-ANN Framework

The flow chart of the SEMD-ANN and EEMD model is shown in figure 1. The detailed steps are enumerated below:

1. The original series is decomposed into a set of different sub-series using Statistical Empirical Mode Decomposition and Ensemble Empirical Mode Decomposition.
2. Each sub-series is combined separately using ANN.
3. Calculate the Mean Square Error (MSE) to compare between the results.

2.1. Steps of ANN

Network structure, input data format and training algorithm play an important role in the performance of the neural network. Network structure: Here, based on previous studies [5], [6], a three-layer resilient feed forward neural network consisting of input layer, hidden layer and output layer, is considered to predict the non-linear time series. Input data format: Each decomposed series should be predicted individually using ANN model. Training Algorithm: In this study, Resilient Back Propagation (RBP) is used for training the neural network. compared to

commonly used Back Propagation algorithm, RBP helps to achieve superior performance [6].

3. Methodology

3.1. Ensemble Empirical Mode Decomposition

To mitigate the scale part-ion difficulty without presenting a subjective intermittence test, a new noise-assisted data analysis techniques is invented, the Ensemble EMD (EEMD), which outlines the true IMF components as the mean of an ensemble of trials, each containing of the signal plus a white noise of finite amplitude. With this ensemble methodology, we can undoubtedly separate the scale certainly without any a priori subjective criterion selection. This new method is based on the insight gleaned from recent studies of the statistical properties of white noise (Flandrin, et al., 2004, and Wu and Huang, 2004), which indicated that the EMD is excellently an adaptive dyadic filter bank 1 when applied to white noise. More critically, the new approach is encouraged by the noise added analyses initiated by Flandrin et al. (2005) and Gledhill (2003). Their results established that noise could support data analysis in the EMD. The main idea of the EEMD is simple: the added white noise would populate the whole time-frequency space uniformly with the constituting components of different scales separated by the filter bank. When signal is added to this uniformly distributed white background, the bits of signal of different scales are mechanically projected onto proper scales of reference recognized by the white noise in the background. Of course, each single may create very noisy outcomes, for each of the noise-added decompositions consists of the signal and the added white noise. Since the noise in each trial is changed in separate trials, it is canceled out in the ensemble mean of enough trails. The ensemble mean is treated as the true answer, for, in the end, the only persistent part is the signal as more and more trials are added in the ensemble.

The critical theory advanced here is based on the following observations:

1. A collection of white noise cancels each other out in a time space ensemble mean; therefore, only the signal can survive and persist in the final noise-added signal ensemble mean.
2. Finite, not infinitesimal, amplitude white noise is necessary to force the ensemble to exhaust all possible

solutions; the finite magnitude noise makes the different scale signals reside in the corresponding IMF, dictated by the dyadic filter banks, and render the resulting ensemble mean more meaningful.

3. The true and physically meaningful answer of the EMD is not the one without noise; it is designated to be the ensemble mean of a large number of trials consisting of the noise-added signal.

This EEMD anticipated here has employed all these important statistical characteristics of noise. We will prove that the EEMD utilizes the scale separation principle of the EMD, and enables the EMD method to be a truly dyadic filter bank for any data. By adding finite noise, the EEMD removes mode mixing in all cases mechanically. Therefore, the EEMD represents a major improvement of the EMD method.

3.2. Statistical Empirical Mode Decomposition

New approach called (Statistical Empirical Mode Decomposition) discovered in 2012 by (Donghoh Kim, Kyungmee O Kim and Hee-Seok Oh), [10] as an extension of EMD. The suggested algorithm is intended for considering noisy signals that are used in the area of statistics. Thus, we call the proposed algorithm statistical EMD (SEMD). Officially, the SEMD procedure can be specified as follows:

1. ((Modified sifting) Take a signal x to be decomposed, and extract the first mode $h_{\lambda,1}$ by using a smoothing technique.

(A-1) Classify the local maxima (minima) Z of the signal $h_{\lambda,1}^0$, where $\mu_{\lambda,1}^0$, is the original signals x .

(A-2) Build an upper envelop μ_0^{\wedge} ((lower envelope l_0^{\wedge}) by applying a smoothing technique with a smoothing parameter

λ to the maxima (minima) z .

(A-3) Calculate the local mean $m_{\lambda} = \frac{1}{2}(\mu_{\lambda} + l_{\lambda})$ by the average of both the envelopes, and then get a candidate intrinsic mode $h_{\lambda,1}^1 = h_{\lambda,1}^0 + m_{\lambda}$.

(A-4) Replicate steps (A-1)–(A-3) for the signal $h_{\lambda,1}^0$ until the signal $h_{\lambda,1}^j$ at the j th iteration satisfies the IMF conditions.

(A-5) Decompose the signal $x = h_{\lambda,1} + r_{\lambda}$ where $h_{\lambda,1}, r_{\lambda}$ is defined as the limite of $h_{\lambda,1}^j, r_{\lambda}$ is the remaining signal.

2. If the remaining signal $r = x - h_{\lambda,1}$ has an intrinsic oscillation mode, then or can be further decomposed by conventional sifting.

4. Simulation Study

In this simulation, the software package R was employed to evaluate classical ANN, and the proposed combined method, SEMD-ANN. The following conditions were set at **Table 1**.

- (1) Four different test functions.
- (2) Two different samples size (50, and 100).

Table 1. Proposed combined method, SEMD-ANN

Test functions	Formula
1	$f(x) = \sin(\pi x) - \sin(2\pi x) + 0.5x$
2	$f(x) = \sin(\pi x) + \sin(2\pi x) - 0.5x$
3	$f(x) = \begin{cases} 10e^{-10x} + 2 & \text{if } x \leq 0.5 \\ 3 \cos(10\pi x) & 0.5 \leq x \leq 1 \end{cases}$
4	$f(x) = 5e^{(-10(x-0.5)^2)} + e^{(-10x)}$

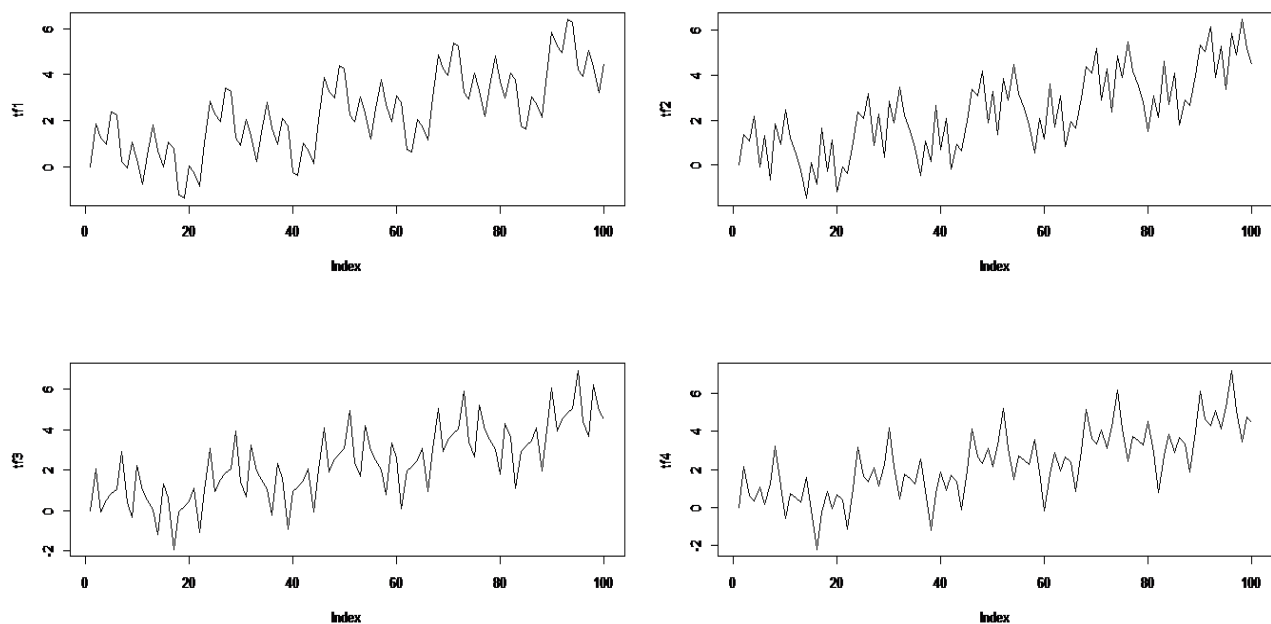


Figure 2. Presents the observations of each test function from Table 1. such as discontinuities or sharp bumps

From figure 2. we observed each function has some abrupt changing features such as discontinuities or sharp bumps, note that it is reasonable to assume a periodic boundary condition for test functions 1 and 2, 3 and 4 while 4, Datasets were simulated from each of the test functions with a two samples size of $n = 50$ and 100. And 4k-fold cross validation level of signal-to-noise were chosen. For each simulated dataset, the above two methods were applied to estimate the test function. In each case, 1000 replications of the two samples size $n = 50$ and 100 were made. Throughout the whole simulation the mean squared error (MSE) was used as

the numerical measure to assess the quality of the estimate.

5. Results and Discussion

Table 2 below shows that the suggested method performs better in terms of noise structures and test functions, under different samples size and 4k-fold cross validation. It is therefore understood that SEMD-ANN is highly recommended in handling non-linear and non-stationary signals.

Table 2. Indicated that SEMD-ANN is better than EEMD-ANN and EMD-ANN

Test Function 1	Test Function 2	Test Function 3	Test Function 4
MSE.SEMD-ANN	MSE.SEMD-ANN	MSE.SEMD-ANN	MSE.SEMD-ANN
1.897675	1.791206	2.046836	1.947746
Test Function 1	Test Function 2	Test Function 3	Test Function 4
MSE.EEMD-ANN	MSE.EEMD-ANN	MSE.EEMD-ANN	MSE.EEMD-ANN
1.954632	1.934515	2.744834	1.95810
Test Function 1	Test Function 2	Test Function 3	Test Function 4
MSE.EMD-ANN	MSE.EMD-ANN	MSE.EMD-ANN	MSE.EMD-ANN
2.954632	3.934515	3.744834	2.95810

As we mentioned previously in simulation sections experiment replicated 1000 times for each test functions and a two sample sizes were adopted for each test functions.

6. Conclusions

This paper submitted a two-stage method to minimize issues in conventional ANN. Coupling SEMD at early stages followed by conventional ANN combination is the crucial design of the proposed method. A simulation was carried out to test the empirical performance of the technique through simulation experiments. Findings from this study showed that SEMD can be improved and resolve issues of non-linear and non-stationary signals, and enhance MSE values.

Availability of Data and Materials

All data generated and analyzed during this study are included within this article.

ACKNOWLEDGEMENTS

This work of AbdAllah.W.H was supported by the Mathematical Statistics-Faculty of science- Damietta University of Egypt funded by the Arab Republic of Egypt Government (AREG). This research of Ghazal.M.A. was supported by Damietta University, Faculty of Science at the Mathematical Department funded by the by the Arab Republic of Egypt Government (AREG).

List of Abbreviations

ANN: Artificial Neural Network
DWT: Discrete Wavelet Transform
EMD: Empirical Mode Decomposition
EEMD: Ensemble Empirical Mode Decomposition
EEMD-ANN: Ensemble Empirical Mode Decomposition of Artificial Neural Network
HHT: Huang-Hilbert Transform
IMF: Intrinsic Mode Function
IMFs: Intrinsic Mode Functions
MSE: Mean Square Error
SVR: Support Vector Regression
SEMD: Statistical Empirical Mode Decomposition
SEMD-ANN: Statistical Empirical Mode Decomposition of Artificial Neural Network

REFERENCES

- [1] G.S. Atsalakis and K.P. Valavanis. Surveying stock market forecasting techniques part II: Soft computing methods. *Expert Systems with Applications*, 36 (Part 2): 5932 – 5941, 2009.
- [2] G.S. Atsalakis and K.P. Valavanis. Surveying stock market forecasting techniques part I: Conventional methods. In Zopounidis C., *Computation Optimization in Economics and Finance Research Compendium*, pages 49–104. Nova Science Publishers, Inc, New York, 2013.
- [3] P.M. Crowley. Long cycles in growth: Explorations using new frequency domain techniques with US data. *Bank of Finland Research Discussion Paper* 6, pages 1–52, 2010.

- [4] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, and H.H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A*: 454:903–995, 1998.
- [5] S. Lahmiri. Wavelet low- and high-frequency components as features for predicting stock prices with backpropagation neural networks. *Journal of King Saud University - Computer and Information Sciences*, 26:218–227, 2014.
- [6] H. Liu, C. Chen, H.Q. Tian, and Y.F. Li. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renewable Energy*, 48:545 – 556, 2012.
- [7] F. Murtagh, J.L. Starck, and O. Renaud. On neuro-wavelet modeling. *Decision Support Systems, Data mining for financial decision making* 37(4): 475 – 484, 2004.
- [8] M. Theodosiou. Forecasting monthly and quarterly time series using STL decomposition. *International Journal of Forecasting*, 27(4): 1178 – 1195, 2011.
- [9] G.D. Wu and S.L. Lo. Effects of data normalization and inherent-factor on decision of optimal coagulant dosage in water treatment by artificial neural network. *Expert Systems with Applications*, 37(7): 4974 – 4983, 2010.
- [10] Kim et al. Extending the scope of empirical mode decomposition by smoothing. *EURASIP Journal on Advances in Signal Processing*: 168, 1-17, 2012.