Forecasting non stationary nonlinear time series data using hybrid models with support vector regression

Nouri Omar Alhammali[1], Amany Moussa Mohamed[2] and Ahmed Amin Elsheikh[3]

[1]*Faculty of Graduate Studies for Statistical Research, Cairo, Egypt*

[2]*Faculty of Graduate Studies for Statistical Research, Cairo, Egypt*

[3]*Faculty of Graduate Studies for Statistical Research, Cairo, Egypt*

Abstract: *Forecasting non stationary time series data has been a difficult and complicated task using the classical statistical predictive models such as linear regression (LR), autoregressive moving average (ARMA), Kalman filter techniques (KFT), exponential smoothing (ES) and other econometric models. A single model may not be sufficient to identify all the characteristics of non-stationary data. The purpose of this paper is to develop a hybrid model that combines three different decomposition methods (EMD, EEMD, SEMD) with support vector regression (SVR) to overcome the difficulty facing the single predictive models .The decomposition methods have the ability to analyze non-linear and non-stationary data by separating them into several components at different resolutions, while SVR is very robust with small training data and high-dimensional problem. The proposed hybrid models are evaluated using extensive simulation experiments under different conditions (sample size, time series model, prediction steps). Results show that the three hybrid models (EMD-SVR, EEMD-SVR, and SEMD-SVR) hybrid model is able to produce accurate forecasting results. The best accuracy is achieved by SEMD-SVR and EEMD-SVR. Results from real data application showed that SEMD-SVR was more accurate for ten steps ahead, whereas EEMD-SVR was more accurate for one step ahead. Furthermore, the outcomes demonstrated that the three suggested models perform better than the hybrid Ensemble Empirical Mode Decomposition with neural network (EEMD-NN), the hybrid Empirical Mode Decomposition with neural network (EMD-NN) and the hybrid complete ensemble empirical mode decomposition with support vector regression (CEEMDAN-SVR).*

Keywords: Empirical Mode Decomposition, Support Vector Regression, Simulation, Temperature Data.

1. INTRODUCTION

In statistical literature, many researchers have proposed various methodologies to improve the short and long term forecasting accuracy during the past decades, see [1],[2] [3]. These methods can be classified into two categories, namely, the classical statistical methods and the artificial intelligence (AI) based algorithms. The classical statistical methods mainly include multiple linear regression(MLR), autoregressive moving average (ARMA), Kalman filter techniques (KFT) and exponential smoothing (ES), see [4], and [5]. Artificial intelligence methods include several methods such as neural network (NN), Support vector regression (SVR), and echo state network (ESN) …etc see [6], [7] and [8]. Although these methods can provide some valuable improvements in terms of forecasting accuracy, most of these models are linear predictors, which have difficulties in forecasting the hard nonlinear and non-stationary behavior of time series data. Normalization techniques are sometimes used to improve forecasting accuracy [9]; however, the traditional normalization methods make assumptions that do not hold for most time series [10]. Hence, there is a real need to find a suitable methodology deals with nonlinear and non-stationary time series data. In literature, several researchers have utilized different hybrid methodologies to address the problem of nonlinearity and non-stationary. By [11], it was suggested a time series forecasting model combining neural networks (NN) and ARIMA models. They concluded that a hybrid technique benefits from the special advantages of NN models and Autoregressive Integrated Moving Average (ARIMA) in both linear and nonlinear modeling. Another hybrid model introduced by [12] which involved Autoregressive Integrated Moving Average (ARIMA) and Support vector regression (SVR) in order to improve forecasting accuracy. The proposed technique performed better than the logit/probit models. One the other hand, lots of data decomposition methods such as the wavelet transform (WT) and the empirical mode decomposition (EMD), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and variational mode decomposition (VMD), have been utilized to forecast non-stationary time series data, see [13] and [14]. Among these techniques, EMD and its extensions have been popular due to the great ability on solving non stationarity and nonlinearity problems; see [15] and [16]. Further, [17] proposed a hybrid EMD/EEMD-ARIMA model for long-term stream flow forecasting. When compared to the EEMD-ARIMA and ARIMA models, the EMD-ARIMA hybrid model performs best in projecting high and moderate stream flow values and matches best with the observations. One the same way, [18] compared (EMD/EEMD/SEMD) with ARIMA model for forecasting temperature recorded data. Results showed that SEMD/ARIMA model is more accurate than EMD/ARIMA and EEMD/ARIMA. Later on, as in [19] it was applied a hybrid (EMD-SVR) based model for to forecast the directional movements of electricity load demands and evaluates the performance on three load datasets. Results revealed that the hybrid EMD-SVR outperforms the single SVR model. Similarly, [20] forecasted price series using (EMD-SVR) model. The methodology's efficiency and predictability were tested using the Chilli wholesale pricing index (WPI) dataset as an example. The findings showed that the performance of the suggested model was much better than that of the standard SVR. In this context, [21] proposed a similar hybrid forecasting method (EMD-SVR). According to the findings, when compared to SVR, the new suggested hybrid prediction model, EMD-SVR, may significantly enhance prediction accuracy. By the appearance of EEMD and SEMD, new hybrid models have been introduced. In [22], it was utilized a hybrid (SEMD-NN) for forecasting Egypt stock market. By the criteria of some statistic loss functions, SEMD-NN outperformed Holt-winters family model, empirical mode decomposition based on neural network (EMD-NN) and ensample empirical mode decomposition and neural network (EEMD-NN) in improving forecast accuracy.

Considering the previously mentioned, it is almost universally agreed in the forecasting literature that no single method is best in every situation. This is mainly because real-world data are frequently complicated, in nature and any single model may not be able to capture different patterns equally well. This has motivated to develop an ensemble model i.e. combination of time series model and machine learning technique which deals with both linear and nonlinear pattern and improve forecasting accuracy.

This paper suggests the use of hybrid methods in which the original data are decomposed into a set of intrinsic mode function (IMF) components and one residue, which can improve the accuracy of forecasting .The most common used methods are Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD) and Statistical Empirical Mode Decomposition (SEMD), See [23] and [24]. The principal idea is hybridizing each (EMD, EEMD, SEMD) with SVR, namely creating the (EMD-SVR, EEMD-SVR, SEMD-SVR) models, to receive better solutions. The proposed models have the capability of smoothing and reducing the noise (inherited from EMD, EEMD, and SEMD), the capability of filtering dataset and improving forecasting performance (inherited from SVR). See [25],[26],[27] and [28].To show the applicability and superiority of the proposed methods, a simulation study has been conducted under different scenarios, in addition to real data application. The contribution of this study will add an important scientific source to the statistical literature regarding modeling nonlinear and non-stationary time series data. The results obtained from simulation and real data applications may provide a clear picture on the most accurate estimation method for modeling non stationary time series data. In other words, it will be useful in determining the best methods that should be used to model non-stationary data. In this way, specialized and non-specialized researchers can easily analyze their data.

The rest of this paper is organized in the following manner: in section 2 we present the methods and material related to our work. Section 3 is devoted for the proposed methods. Results and discussion are in Section 4. Finally, conclusions are drawn in section 5.

2. Methodology

This section introduces our statistical methodology for forecasting non stationary and nonlinear time series data. It consists of two main stages, starting by decomposition process, followed by predictive process. At the decomposition process, we employee three different decomposition techniques namely EMD, EEMD, and SEMD. At the second stage we utilize a advance powerful predictive model namely, support vector regression. Next, we shall present the details for the two stages and show how to combine them to get our hybrid forecasting models.

2.1 Empirical Mode Decomposition(EMD)

Empirical Mode Decomposition (EMD) is a method of time-frequency domain signal decomposition that has become widely used for output-only modal identification of structures, see [25]. The EMD method decomposes a multi-component signal into a sequence of oscillatory waveforms known as IMFs that are meant to be single-frequency components. An IMF is a function that meets two criteria: (i) In the entire data set, the number of extrema and the number of zero-crossings must either be identical or differ by no more than one, and (ii) at any point, the local maxima and local minima on the envelope represent its mean value, which is zero. Sifting is the process of extracting an IMF. Assuming $\mathrm{y(t)}$ is the signal that needs to be decomposed, the key EMD processes are as follows:

1. Connect every local minimum and maximum by utilizing a cubic spline to extract the lower and upper envelopes.
2. Determine the value of $m_1(t)$, which is the mean of the upper and lower envelopes.
3. Find the difference between the mean $m_1(t)$and signal $\mathrm{y(t)}$ and the,$i_1(t) = \mathrm{y(t)} - m_1(t)$, which could be the first IMF.
4. Determine whether $i_1(t)$ fits the two IMF requirements given above. If $i_1(t)$meets both conditions to be an IMF, then $i_1(t)$is the first IMF of the original signal $\mathrm{y(t)}$.
5. If$i_1(t)$does not match the IMF criteria, the sifting process will be repeated, but this time the $i_1(t)$ will be treated as the original signal until it meets the two IMF conditions.
6. After subtracting the original signal from the IMF, the sifting procedure is repeated to deconstruct the data into n IMFs.

Finally, the signal $\mathrm{y(t)}$may be written as follows:

$$\mathrm{y(t)} = \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{i_j(t)} + \mathrm{r_n(t)} \qquad (1)$$

Where $\mathrm{i_j}$ (t) ($\mathrm{j} = 1, 2, 3, \dots \dots, \mathrm{n}$) represents the original signal$y(t)$'s IMFs, and$r_n(t)$is a $y(t)$ residue. Each IMF should, ideally, just have one frequency component. Occasionally, one IMF will have many frequency components, which are known as mode-mixing. For more details, see [21].

2.2 Ensemble Empirical Mode Decomposition (EEMD)

Ensemble Empirical Mode Decomposition (EEMD) uses a noise-assisted data analysis method suggested by [25], which has a uniform time frequency spaceat various scales. When the signal is added to the uniform white background, the signals with different scales are automatically projected onto proper scales of reference established by the white noise in the background. The artificially created white noise has been removed, and the recorded signal with numerous frequency components is projected onto appropriate reference scales. The resultant decomposition retains its physical uniqueness while the EEMD also solves the mode mixing problem. It comprises mostly of the following steps:

- Set the ensemble number and amplitude of white noise added sequence.
- The white noise is added to the signal $y(t)$ that was measured.

$$y_k(t) = y(t) + w_k(t) \qquad (2)$$

Where $w_k(t)$ is $k - th$ white noise,$y_k(t)$ is $k - th$the signal's measurement sequence.

• Apply EMD to $y_k(t)$, then decompose in to n IMFs.

$$\mathrm{y_k(t)} = \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{i_{k,j}}\,\mathrm{(t)} + \mathrm{r_k(t)} \qquad (3)$$

•Repeat step 2 and 3 , say m times, ($i.e.$ k = 1,2,3, . . , m) using different white noise sequences, maintaining the standard deviation of the simulated white noise ismaintained constant at 7, see [25].It should be mentioned that the number of ensembles ($i.e.$ m) must be set a priori in EEMD. However, through increasing the ensemble's sample count, the extra white noise's influence can be decreased to a negligible level. Generally, an ensemble size of a few hundred leads to a perfectresult.

- Calculate the final IMF as the ensemble average of each deconstructed IMF, $\bar{\mathrm{i}}_{\mathrm{j}}(\mathrm{t}) = \frac{1}{\mathrm{m}} \sum_{\mathrm{k}=1}^{\mathrm{m}} \mathrm{i}_{\mathrm{k,j}}(\mathrm{t})$ (4)

$$\mathrm{x(t)} = \sum_{j=1}^{m} \bar{\iota}_j(t) + \bar{\mathrm{r}}_m(t) \quad (5)$$

Where $\bar{\iota}_j(t)$ is the $i-th$ IMF that is the ensemble mean of the corresponding IMFs, which were calculated from m white noise sequences, and $\bar{\mathrm{r}}_m(t)$is the residues' average.

2.3 Statistical Empirical Mode Decomposition (SEMD)

The Statistical Empirical Mode Decomposition (SEMD) is a function that conducts empirical mode decomposition for the sifting process using spline smoothing rather than interpolation. The smoothing parameter is automatically determined by cross-validation .As in [23], the SEMD can be explained as follows:

A. (Modified sifting): Consider the signal X to be decomposed, then employ a smoothing approach to extract the first mode, $h_{1,\lambda}$.

(A-1) Find the local maximum (minimum) z of the signal $h_{1,\lambda}^{0}$, where $h_{1,\lambda}^{0}$represents the original signal x.
(A-2) Create an upper envelope $\hat{u}_\lambda$ (lower envelope$\hat{\rho}_\lambda$) by smoothing the maxima (minima) z using a smoothing technique and a smoothing parameter $\hat{\rho}_\lambda$.
(A-3) Calculate the local average $m_\lambda = \frac{1}{2} (\hat{u}_\lambda + \hat{\rho}_\lambda)$ by averaging the contents of both envelopes, then find a candidate intrinsic mode$h_{1,h}^{1} = h_{1,h}^{0} - m_\lambda$.
(A-4) Repeat the steps (A-1)–(A-3) for the signal $h_{1,h}^{1}$ until the signal $h_{1,h}^{j}$at the jth iteration satisfies the IMF conditions.
(A-5) Decompose the signal x = $h_{1,\lambda} + r_\lambda$, where $r_\lambda$is the remaining signal and $h_{1,\lambda}$ is defined as the limit of $h_{1,h}^{j}$.

B. (Conventional sifting) If the remaining signal r_λ = x – $h_{1,h}^{1}$ has an intrinsic oscillation mode, then $r_\lambda$may befurther decomposed by conventional sifting.

2.4 Support Vector Regression:

The primary idea of SVR is to execute linear regression and reduce structural risks in the high-dimensional feature space that is produced by mapping the original input via a predetermined function $\emptyset(x_i)$.When a collection of samples $[x_i, y_i]$ is given, with i=1,2,…..N ,y_i is the output and x_i, is the input. The goal is the output

$$f(x) = w^T \emptyset(x) + b \quad (6)$$
$$R[f] = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} l(x_i, y_i \ , f(x_i)) \quad (7)$$

Where b is bias, W is regression coefficient and the penalty coefficient is C. $R[f]$is the structure risk, while $l(x_i, y_i \ , f(x_i))$ represents the loss function. The corresponding constrained optimization problem can be formulated as:

$$min \frac{1}{2} \|w\|^2 + \mathrm{C} \sum_{i=1}^{N} (\xi_i + \xi_i^*) \quad (8)$$
$$s.t. \, y_i - w^T \emptyset(x) - b \leq \varepsilon + \xi_i^* \quad (9)$$
$$w^T \emptyset(x) + b \leq \varepsilon + \xi_i^* \quad (10)$$
$$\xi_i, \xi_i^* \leq 0, i = 1,2, \dots, n$$

Where$(\xi_i + \xi_i^*)$refer to the slack variables. The Lagrange multiplier was added, you may write the regression function as:

$$\mathrm{f(x)} = \sum_{\mathrm{i}=0}^{\mathrm{N}} (\alpha_{\mathrm{i}} - \alpha_{\mathrm{i}}^*) \mathrm{K}(\mathrm{x_i}, \mathrm{x}) + \mathrm{b} \quad (11)$$

Where (α_i, α_i^*) are the Lagrange multipliers that meet the criteria $\alpha_i \leq 0, \alpha_i^* \geq 0$ and $\sum_{i=0}^{N}(\alpha_i - \alpha_i^*) = 0$. $K(x_i, x)$ is the kernel function conforming to Mercer's theorem.
The Support Vector Regression (SVR) has drawn close attention due to its high generalization in solving practical problems such as nonlinearity, small samples and over-fitting situations. The SVR is a learning machine relies on the structural risk minimization inductive principle to achieve the generalized performance. Unlike other Regression models that try to minimize the error between the real and predicted value, the SVR tries to fit the best line within a threshold value. It uses a subset of training points in the decision function (called support vectors), so it is also memory efficient. Due to the above mentioned properties, the SVR has been successfully applied to various fields see [28], [29], [30]. These were the reasons behind the use of SVR in our combined methods.

2.5 The Proposed hybrid Methods:

In this section, we describe the combined (SEMD, EEMD, EMD/SVR) approaches, for forecasting problem.
Given a time series data, the training phase of the method consists of the following steps:

SEMD-SVR

1) SEMD is applied to the original time series in order to identify the IMFs, denoted as $\tilde{y}^r$, in addition to the residue.
2) Having obtained the IMFs, the second step is to apply SVR for each of the extracted IMF, and for the residue as well, getting the predictions (F1, F2, …,Fn, Fr)
3) The third step is to find the final forecast which is summation of the predictions obtained from the second step:

$$\hat{y} = \sum_{j=1}^{n} F_j + F_r \qquad (12)$$

EEMD-SVR

For this combined method, the same previous procedure is applied, except, EEMD is applies instead of SEMD.

EMD-SVR

For this combined method, the same previous procedure is applied, except EMD is applies instead of SEMD.
The methodology for SEMD-SVR, EEMD-SVR and EMD-SVR are depicted in Figure 1.

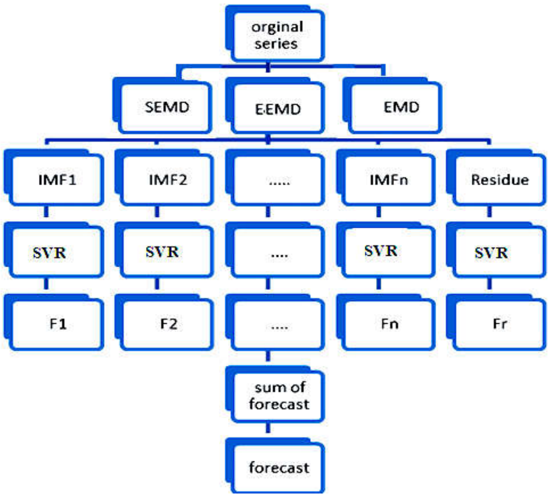

Figure 1.The chart of the methodology for SEMD-SVR, EEMD-SVR and EMD-SVR

4. RESULT and DISSCUSION

4.1 Simulations Results

The In order to evaluate the practical performance of the proposed hybrid models, a simulation study is conducted using the software package R with 1000 replications. The following conditions were set.

(1) Three different test functions, including different non stationary nonlinear time series models. See Table 1 and Figure 2.

(2) Four different sample sizes (30, 50, 100, 200). This choice is arbitrary, to represent small, medium and large sizes and to see the behavior of the model's performance as the sample size increases.

(3) For SVR, the kernel used in training and predicting was "Radial". The degree needed for kernel of type polynomial is the default order (order=3). However, one might consider changing these parameters. It is left for further investigations.

(4) Three different prediction steps (one step ahead, five steps ahead, ten steps ahead). Since the sample sizes are different, the percentages of training and test data will be different as well, see Table 2.

Table 1.Time series models used in simulation

Name	Formula	Source
heav	**Heav = 4 * sin(4 * pi * x) - sign(x - 0.3) - sign(0.72-x)**	**Donoho and Johnstone (1994)**
Fg1	**Fg1= 0.25 * ((4 * x - 2) + 2 * exp(-16 * (4 * x - 2)^2))**	Fan and Gijbels (1995)
Doop	**Dopp = (x * (1 - x))^0.5 * sin(2 * pi * 1.05/(x + 0.05))**	**Donoho and Johnstone (1994)**

Source: These functions are obtained from R Package (CVThresh)

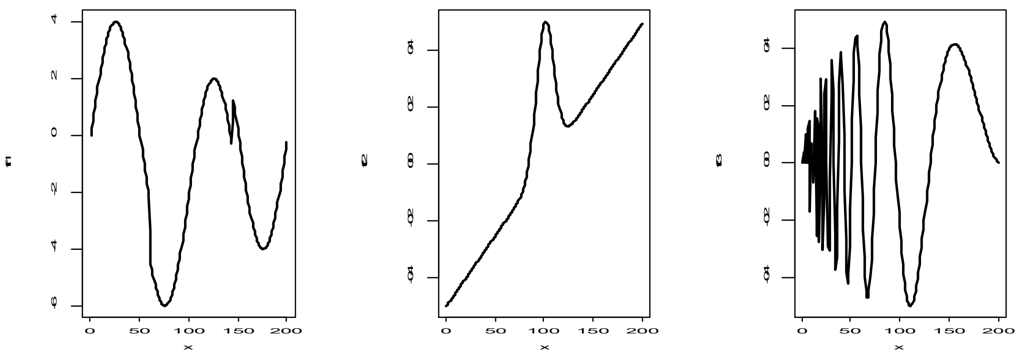

Figure 2.The Test functions used in simulation

Table 2.The percentages of training and test data used in simulation

N	1 step		5 steps		10 steps	
	Train %	Test %	Train %	Test %	Train %	Test %
30	97%	3%	83%	17%	67%	33%
50	98%	2%	90%	10%	80%	20%
100	99%	1%	95%	5%	90%	10%
200	99%	1%	97%	3%	95%	5%

(5) The Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE) are used to compare the performance of the hybrid models, as defined in these Equations

$$MSE = \frac{1}{N'}\sum_{i=1}^{N'}(x_i - \hat{x}_i)^2 \quad (13)$$

$$MAPE = \frac{1}{N'}\sum_{i=1}^{N'}\left|\frac{x_i - \hat{x}_i}{x_i}\right| \times 100 \quad (14)$$

$$MAE = \frac{1}{N'}\sum_{i=1}^{N'}|x_i - \hat{x}_i| \quad (15)$$

Where x_i refers to actual time series value, $\hat{x}_i$ is the estimated one, while N' is the size of test data.
The mean absolute percentage error (MAPE) is the most widely used forecasting accuracy measurement, see [31], [32] and [33]. MAPE has important, desirable features including reliability, unit-free measure, ease of interpretation, clarity of presentation, support of statistical evaluation, and the use of all the information concerning the error. Additionally, the MAPE is used not only for comparison purpose, but to determine the quality of the model forecasting as well. A MAPE value of < 10% indicates high accurate forecasting, 10% ≤ MAPE <20% indicates good forecasting, 20% ≤ MAPE <50% indicates reasonable forecasting, and MAPE ≥ 50% indicates inaccurate forecasting. Fortunately, our results for MAPE belong to the range (<10%) indicate high accurate forecasting.
In addition to MAPE, we used two extra criteria namely, Mean Absolute Error (MAE) and Mean Square Error (MSE).
Mean Absolute Error (MAE): The MAE is one of the most popular, easy to understand and compute metrics. Lower the value of the better is our forecast. The models which try to minimize MAE lead to forecast median.
Mean Square Error (MSE): The MSE is also among the popular methods used by statisticians to understand how well is forecast. The interpretation of the numbers is much more difficult in comparison to MAE. The models trying to minimize MSE lead to a forecast of the mean, See [34], for more different error's metrics.
According to the three comparison criteria, several interesting remarks have been drawn as following:
The first model

1- The hybrid method (SEMD- SVR) is the best when the sample size is 30 and the prediction is for five steps ahead, as well as when the sample size is 50 and the prediction is for each of (1 and 10) steps ahead and also when the sample size is 100 and the prediction is for each of (1, 5 and 10) steps ahead.
2- The hybrid method (EEMD-SVR) is the best when the sample size is 30 and the prediction is for ten steps and also when the sample size is 50 and the prediction is for five steps and also when the sample size is 200 and the prediction is for each of (1, 5 and 10) steps ahead.
3- Hybrid method (EMD-SVR) is the best when the sample size is 30 and the prediction is for only one step.
4- In general, when evaluating the three hybrid methods, it appears that 50% of the simulation experiments in which the (SEMD-SVR) hybrid method outperforms the other methods. The hybrid method (EEMD-SVR) came in second place with a percentage of 41.67%, and finally the method (EMD-SVR) with a percentage of 8.33%.

Table 3.The simulation results for the first model

Size (n)	Criterion	hybrid models	Jump (h)		
			1	5	10
30	MSE	SEMD-SVR	7.244345	1.707047	14.3505
		EEMD-SVR	6.866971	3.100828	14.20504
		EMD-SVR	6.036794	3.524499	14.28761
	MAPE	SEMD-SVR	4.906697	0.397286	4.460718
		EEMD-SVR	4.880217	0.454506	4.469119
		EMD-SVR	4.54474	0.482814	4.318393
	MAE	SEMD-SVR	2.41765	1.153203	3.154286
		EEMD-SVR	2.357701	1.557204	3.131699
		EMD-SVR	2.159245	1.65985	3.140921
50	MSE	SEMD-SVR	7.006086	7.070684	3.283772
		EEMD-SVR	7.197738	6.209192	3.902278
		EMD-SVR	6.29743	6.29534	4.029156
	MAPE	SEMD-SVR	10.54437	44.13331	1.451814
		EEMD-SVR	11.08269	40.46244	1.454752
		EMD-SVR	10.9601	42.28653	1.563049
	MAE	SEMD-SVR	2.26579	2.218669	1.487795
		EEMD-SVR	2.452038	2.062356	1.648200
		EMD-SVR	2.419688	2.078763	1.667356
100	MSE	SEMD-SVR	4.058441	5.937194	6.929037
		EEMD-SVR	4.557971	6.24299	7.552217
		EMD-SVR	5.074741	6.856831	7.531743
	MAPE	SEMD-SVR	28.37897	10.85229	9.910298
		EEMD-SVR	29.44944	11.06491	9.131135
		EMD-SVR	29.63245	12.52468	9.953065
	MAE	SEMD-SVR	1.780268	2.143123	2.202621
		EEMD-SVR	1.912192	2.209373	2.319987
		EMD-SVR	2.016215	2.328571	2.313869
200	MSE	SEMD-SVR	2.342658	3.297585	4.713886
		EEMD-SVR	2.288167	3.230656	4.615499
		EMD-SVR	3.231589	4.392609	5.75379
	MAPE	SEMD-SVR	8.672174	14.03074	11.42612
		EEMD-SVR	8.409924	13.63662	11.25382
		EMD-SVR	10.47038	16.01157	12.53788
	MAE	SEMD-SVR	1.269726	1.536662	1.875392
		EEMD-SVR	1.256398	1.523083	1.856695
		EMD-SVR	1.533094	1.817879	2.100463

The second model

1- The hybrid method (SEMD-SVR) is the best when the sample size is (50, 100, or 200) and the prediction is for each of (1, 5, and 10) steps.
2- The hybrid method (EEMD-SVR) is the best when the sample size is 30 and the prediction is for each of (1, 5 and 10) steps.
3- The hybrid method (EMD-SVR) is not the best for all sample sizes and for all steps of the second model.
4- In general, it appears that (SEMD-SVR) outperformed the other two hybrid methods in 75% trails of our simulation experiments. The hybrid method (EEMD-SVR) came in second place.

Table 4.The simulation results for the second model

Size (n)	Criterion	hybrid models	Jump (h)		
			1	5	10
30	MSE	SEMD-SVR	0.098524	0.162799	0.241736
		EEMD-SVR	0.097596	0.161649	0.219801
		EMD-SVR	0.104094	0.16256	0.256828
	MAPE	SEMD-SVR	1.334993	1.671993	3.96528
		EEMD-SVR	1.326191	1.539372	3.113121
		EMD-SVR	1.330007	1.655947	3.873418
	MAE	SEMD-SVR	0.250609	0.327434	0.413225
		EEMD-SVR	0.250046	0.327174	0.390669
		EMD-SVR	0.257265	0.328621	0.428886
50	MSE	SEMD-SVR	1.162163	1.237614	1.297024
		EEMD-SVR	1.173795	1.252895	1.341357
		EMD-SVR	1.197914	1.278627	1.347201
	MAPE	SEMD-SVR	3.137363	4.079726	3.625562
		EEMD-SVR	3.324551	4.374008	3.827155
		EMD-SVR	3.343869	6.41765	3.988018
	MAE	SEMD-SVR	0.868884	0.8833	0.908994
		EEMD-SVR	0.867135	0.887636	0.924143
		EMD-SVR	0.877519	0.90031	0.927552
100	MSE	SEMD-SVR	0.969649	1.169587	1.212121
		EEMD-SVR	0.992622	1.181465	1.247189
		EMD-SVR	1.011904	1.206323	1.249866
	MAPE	SEMD-SVR	4.78981	4.320906	4.637898
		EEMD-SVR	5.078806	4.984837	6.003655
		EMD-SVR	6.152788	4.821664	6.462972
	MAE	SEMD-SVR	0.790722	0.865054	0.879609
		EEMD-SVR	0.793224	0.869359	0.892284
		EMD-SVR	0.801262	0.877885	0.894936
200	MSE	SEMD-SVR	1.10848	1.105049	1.127035
		EEMD-SVR	1.128499	1.119555	1.148519
		EMD-SVR	1.133327	1.13362	1.155328
	MAPE	SEMD-SVR	2.818742	3.150766	3.19363
		EEMD-SVR	2.954661	3.250137	3.251609
		EMD-SVR	2.999221	3.835977	3.282703
	MAE	SEMD-SVR	0.83949	0.835964	0.846312
		EEMD-SVR	0.84615	0.842139	0.855041
		EMD-SVR	0.842742	0.845573	0.858012

The third model:

1- The hybrid method (SEMD-SVR) is the best when the sample size is (50, 100, or 200) and the prediction is for each of (1, 5, and 10) steps.
2- Hybrid method (EEMD-SVR) is the best when the sample size is 30 and the prediction is for both (1 and 5) steps.
3- The hybrid method (EMD-SVR) is the best when the sample size is 30 and the prediction is for 10 steps.
4- In general, when evaluating the three hybrid methods, it appears that 75% of the simulation experiments excelled in the (SEMD-SVR) hybrid method, and the (EEMD-SVR) hybrid method came in second place with a percentage of 16.67%, and finally the (EMD-SVR) method with a percentage of 8.33%.

Table 5.The simulation results for the second model

Size (n)	Criterion	hybrid models	Jump (h)		
			1	5	10
30	MSE	SEMD-SVR	0.117527	0.129328	0.268176
		EEMD-SVR	0.1107	0.12848	0.271894
		EMD-SVR	0.111277	0.138608	0.253571
	MAPE	SEMD-SVR	18.59379	7.217024	4.811387
		EEMD-SVR	12.39283	7.154187	4.426099
		EMD-SVR	19.55334	7.763637	4.420855
	MAE	SEMD-SVR	0.277138	0.289778	0.43828
		EEMD-SVR	0.267176	0.288181	0.441914
		EMD-SVR	0.26735	0.300593	0.422363
50	MSE	SEMD-SVR	0.967863	1.294558	1.299513
		EEMD-SVR	0.994671	1.31855	1.360462
		EMD-SVR	1.017952	1.300063	1.359382
	MAPE	SEMD-SVR	2.05108	3.23371	68.90272
		EEMD-SVR	2.353634	3.498248	88.51766
		EMD-SVR	2.093515	3.322041	62.26736
	MAE	SEMD-SVR	0.790027	0.905487	0.909456
		EEMD-SVR	0.792234	0.917417	0.930316
		EMD-SVR	0.804135	0.905800	0.931062
100	MSE	SEMD-SVR	0.967863	1.180922	1.251642
		EEMD-SVR	0.994671	1.18501	1.255211
		EMD-SVR	1.017952	1.198078	1.259066
	MAPE	SEMD-SVR	2.05108	3.56269	3.114822
		EEMD-SVR	2.353634	3.80556	3.461598
		EMD-SVR	2.093515	4.759192	3.698619
	MAE	SEMD-SVR	0.790027	0.86711	0.894951
		EEMD-SVR	0.792234	0.867525	0.897038
		EMD-SVR	0.804135	0.874631	0.897106
200	MSE	SEMD-SVR	1.106837	1.116265	1.140787
		EEMD-SVR	1.128389	1.128762	1.151889
		EMD-SVR	1.136593	1.135748	1.164765
	MAPE	SEMD-SVR	2.744174	1.84581	2.49725
		EEMD-SVR	2.848205	1.956768	2.681864
		EMD-SVR	3.020367	2.040381	2.981351
	MAE	SEMD-SVR	0.837834	0.840803	0.851683
		EEMD-SVR	0.845696	0.845768	0.855771
		EMD-SVR	0.847397	0.8485	0.861354

4.2 Application on real data

In addition to simulation experiments, we have evaluated the above three hybrid methods using Libyan temperature data. This data set is monthly recorded from Jan-1998 to Dec- 2022, see Figure 5. The data used in this study were collected from Libyan Center for Meteorology. Time series with trends, or with seasonality, are not stationary since the trend and seasonality will affect the value of the time series at different times. Therefore, it might be useful to check stationarity, trend and seasonality before applying the three hybrid methods.

1. **Stationarity Test:**

Results of Augmented Dickey-Fuller Test showed that this series is not stationary (Dickey-Fuller = -1.9145, p-value = 0.605).

2. **Trend Test:**

In addition to stationarity test we used Mann-Kendall Test for trend. Results conforms the presence of a significant positive trend in temperature over the 25 years, with the p-value (0.00084).

3. **Seasonality Test:**

Results obtained from "seaste" package using "isSeasonal" function, returned "TRUE" indicated the presence of seasonality.
Accordingly, our three hybrid methods seems to be suitable, in which this series was decomposed into 3 IMF's by EMD, 3 IMF's by EEMD and 1 IMF's by SEMD.
Table 6 shows the numerical results of the mean squared error (MSE), the absolute mean squared error (MAPE) and the relative mean squared error (MAE). Obviously, for one step forecasting it can be seen that EEMD-SVR outperforms (EMD-SVR and SEMD-SVR), in addition to the two existing methods EMD-NN and EEMD-NN. For five steps forecasting results revealed that EMD-SVR outperforms the other four methods. For ten steps forecasting it has been observed that SEMD-SVR outperforms the other four methods. Although this result may seem somewhat different from what we obtained from simulation experiments regarding the superiority of the SEMD method and the EEMD method, the SEMD method is still superior in the case of 10 prediction steps. As for the EEMD method, it excelled in the case of one-step prediction. The reason behind this could be due to the nature and size of the data and the characteristics inherent in its seasonal changes.
A remarkable notice is that the two existing artificial hybrid methods (EMD-NN and EEMD-NN) performs worst compared to our three suggested methods for Libyan temperature data.

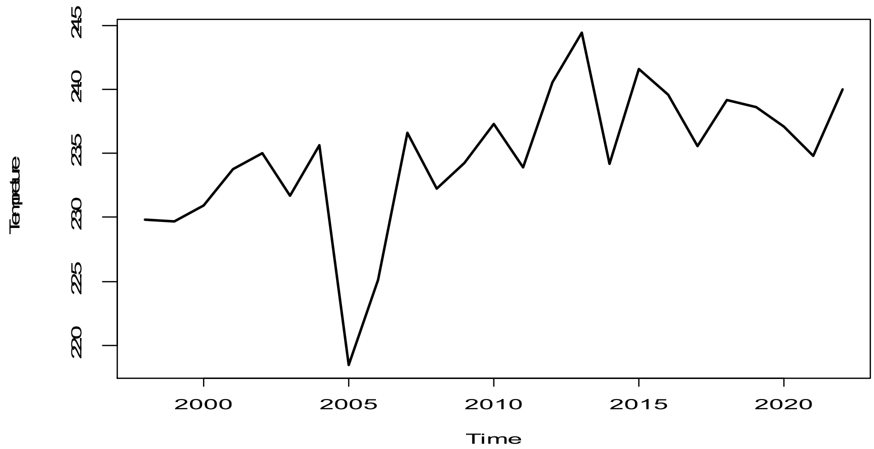

Figure 3.The Libyan temperature data from 1998 – 2022

Table 6.The comparing the accuracy of the proposed hybrid methods with some existing artificial hybrid methods

Jump (h)	1			5			10		
Criterion	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE
EMD-SVR	0.436985	0.018208	0.190956	0.186258	0.007797	0.05596	0.424103	0.017708	0.213546
EEMD-SVR	0.412914	0.017205	0.170498	0.394119	0.01652	0.188168	0.492991	0.020583	0.28261
SEMD-SVR	0.51593	0.021497	0.266184	0.23109	0.009664	0.084869	0.376424	0.015723	0.172277
CEMD- SVR	0.4551138	0.01896307	0.20712855	0.2177832	0.009106592	0.08435374	0.5114078	0.02134974	0.3040779
EEMD-NN	0.614319	0.025597	0.377388	0.824319	0.034619	0.866208	0.952693	0.040176	1.280283
EMD-NN	0.493566	0.020565	0.243607	0.258151	0.010864	0.117559	0.423182	0.017841	0.320656

5. CONCLUSION

In this paper we have presented three hybrid models which combines empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), and statistical empirical mode decomposition (SEMD) with support vector regression. The ultimate goal is getting accurate estimates for future time series values. Simulation experiments and real data application have proved that the proposed methodologies are expected to be easily implemented and can be used for different kinds of non-stationary and nonlinear time series data under a variety of sample sizes (n=30,50,100,200) and forecasting steps (h=1,5,10). The best accuracy was achieved by SEMD-SVR, then EEMD-SVR. Results from real-data application using Libyan temperature data revealed that SEMD-SVR was more accurate for ten steps ahead while EEMD-SVR was more accurate for one step ahead. Moreover, results showed that the three suggested models outperform EMD-NN, EEMD-NN and CEEMDAN-SVR.

REFERENCES

[1] *Mills, T. C. (1990). Time series techniques for economists. Cambridge university press.*

[2] *Percival, D. B. & Walden, A. T. (1993). Spectral analysis for physical applications. cambridge university press.*

[3] *Faragher, R. (2012). Understanding the basis of the kalman filter via a simple and intuitive derivation [lecture notes]. IEEE Signal processing magazine, 29(5), 128-132.*

[4] *Billah, B., King, M. L., Snyder, R. D., & Koehler, A. B. (2006). Exponential smoothing model selection for forecasting. International journal of forecasting, 22(2), 239-247.*

[5] *Gurney, K. (1997). An introduction to neural networks. CRC press.*

[6] *Awad, M., & Khanna, R. (2015). Efficient learning machines: theories, concepts, and applications for engineers and system designers (p. 268). Springer nature.*

[7] *Wang, H., Lei, Z., Liu, Y., Peng, J., & Liu, J. (2019). Echo state network based ensemble approach for wind power forecasting. Energy Conversion and Management, 201, 112188.*

[8] *Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. Advances in adaptive data analysis, 1(01), 1-41.*

[9] *Singla, Pardeep, Manoj Duhan, and Sumit Saroha. "Different normalization techniques as data preprocessing for one step ahead forecasting of solar global horizontal irradiance." In Artificial Intelligence for Renewable Energy Systems, pp. 209-230. Woodhead Publishing, 2022.*

[10] *Ogasawara, Eduardo, Leonardo C. Martinez, Daniel De Oliveira, Geraldo Zimbrão, Gisele L. Pappa, and Marta Mattoso. "Adaptive normalization: A novel data normalization approach for non-stationary time series." In The 2010 International Joint Conference on Neural Networks (IJCNN), pp. 1-8. IEEE, 2010.*

[11] *Zhang, G. Peter. "Time series forecasting using a hybrid ARIMA and neural network model." Neurocomputing 50 (2003): 159-175.*

[12] *Ince, Huseyin, and Theodore B. Trafalis. "A hybrid model for exchange rate prediction." Decision Support Systems 42, no. 2 (2006): 1054-1062.*

[13] *Singla, Pardeep, Manoj Duhan, and Sumit Saroha. "A dual decomposition with error correction strategy based improved hybrid deep learning model to forecast solar*

irradiance." Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 44, no. 1 (2022): 1583-1607.

[14] *Singla, Pardeep, Manoj Duhan, and Sumit Saroha. "A point and interval forecasting of solar irradiance using different decomposition based hybrid models." Earth Science Informatics (2023): 1-18.*

[15] *Chuang, Chen-Chia, and Zne-Jung Lee. "Hybrid robust support vector machines for regression with outliers." Applied Soft Computing 11, no. 1 (2011): 64-72.*

[16] *Wu, Zhibin, Wen Zhang, and Xiaojun Zeng. "Exploring the short-term and long-term linkages between carbon price and influence factors considering COVID-19 impact." Environmental Science and Pollution Research 30, no. 22 (2023): 61479-61495.*

[17] *Wang, T. (2021). A combined model for short-term wind speed forecasting based on empirical mode decomposition, feature selection, support vector regression and cross-validated lasso. PeerJ Computer Science, 7, e732.*

[18] *Mohammed, Muhammad Baay, Abdslam K. Suliman, and Nuri Omer Ali. "Comparison of hybrid models in temperature prediction in the Libyan city of Sirt." Journal of Pure & Applied Sciences 18, no. 4 (2019).*

[19] *Bican, Bahadir, and Yusuf Yaslan. "A hybrid method for time series prediction using EMD and SVR." In 2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP), pp. 566-569. IEEE, 2014.*

[20] *H. V. Jansen, N. R. Tas and J. W. Berenschot, "Encyclopedia of Nanoscience and Nanotechnology", Edited H. S. Nalwa, American Scientific Publishers, Los Angeles, vol. 5, (2004), pp. 163-275.*

[21] *Das, Pankaj, Girish Kumar Jha, Achal Lama, Rajender Parsad, and Dwijesh Mishra. "Empirical mode decomposition based support vector regression for agricultural price forecasting." Indian Journal of Extension Education 56, no. 2 (2020): 7-12.*

[22] *Altıntaş, Atilla, and Lars Davidson. "EMD-SVR: a hybrid machine learning method to improve the forecasting accuracy of highway tollgates traveling time to improve the road safety." In International Conference on Intelligent Transport Systems, pp. 241-251. Cham: Springer International Publishing, 2020.*

[23] *Ghazal, Mohammed Abou Elfettouh, Asaad Ahmed Gad Elrab, and Wafa Hamed Abd Allah. Hybrid statistical empirical mode decomposition with neural network in time series forecasting."International Journal of Statistics and applied Mathematics (2019).*

[24] *Kim, D., Kim, K. O., & Oh, H. S. (2012). Extending the scope of empirical mode decomposition by smoothing. EURASIP Journal on Advances in Signal Processing, 2012, 1-17.*

[25] *Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. Advances in adaptive data analysis, 1(01), 1-41.*

[26] *Sun, H., Guo, W., Lan, Y., Wei, Z., Gao, S., Sun, Y., & Fu, Y. (2022). Black-Box Modelling and Prediction of Deep-Sea Landing Vehicles Based on Optimised Support Vector Regression. Journal of Marine Science and Engineering, 10(5), 575.*

[27] *Ren, Y., Suganthan, P. N., & Srikanth, N. (2014). A novel empirical mode decomposition with support vector regression for wind speed forecasting. IEEE transactions on neural networks and learning systems, 27(8), 1793-1798.*

[28] *Nava, N., Di Matteo, T., & Aste, T. (2018). Financial time series forecasting using empirical mode decomposition and support vector regression. Risks, 6(1), 7.*

[29] *Lee, Ming-Chi. "Using support vector machine with a hybrid feature selection method to the stock trend prediction." Expert Systems with Applications 36, no. 8 (2009): 10896-10904.*

[30] *Hu, Zhichao, Runfeng Zhang, Zhanna Zenkova, and Yue Wang. "Wind speed prediction performance based on modal decomposition method." In 2020 2nd International Conference on Information Technology and Computer Application (ITCA), pp. 736-741. IEEE, 2020.*

[31] *Armstrong, J. Scott, and Fred Collopy. "Error measures for generalizing about forecasting methods: Empirical comparisons." International journal of forecasting 8, no. 1 (1992): 69-80.*

[32] *Ren, Louie, and Yong Glasure. "Applicability of the revised mean absolute percentage errors (MAPE) approach to some popular normal and non-normal independent time series." International Advances in Economic Research 15 (2009): 409-420.*

[33] *De Myttenaere, Arnaud, Boris Golden, Bénédicte Le Grand, and Fabrice Rossi. "Mean absolute percentage error for regression models." Neurocomputing 192 (2016): 38-48.*

[34] *Singla, Pardeep, Manoj Duhan, and Sumit Saroha. "Review of different error metrics: a case of solar forecasting." AIUB Journal of Science and Engineering (AJSE) 20, no. 4 (2021): 158-165.*