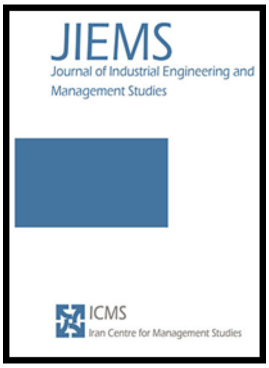
JIEMS
Journal of Industrial Engineering and Management Studies

Vol. 7, No. 2, 2020, pp. 187-201

DOI: 10.22116/JIEMS.2020.205148.1307

www.jiems.icms.ac.ir



Weighted MLP-ARIMA series hybrid model for time series forecasting

Zahra Hajirahimi[1,*], Mehdi Khashei[1]

Abstract

With the increasing importance of forecasting with the utmost degree of accuracy, utilizing hybrid frameworks become a must for obtaining more accurate and more reliable forecasting results. Series hybrid methodology is one of the most widely-used hybrid approaches that has encountered a great amount of popularity in the literature of time series forecasting and has been applied successfully in a wide variety of domains. In such hybrid methods is assumed that there is an additive relationship among different components of time series. Thus, based on this assumption, various individual models can apply separately on decomposed components, and the final forecast can be obtained. However, developed series hybrid models in the literature are constructed based on the decomposing time series into linear and nonlinear parts and generating linear-nonlinear modeling order for decomposed parts. Another assumption considered in the traditional series model is assigning equal weights to each model used for modeling linear and nonlinear components. Thus, contrary to traditional series hybrid models, to improve the performance of series hybrid models, these two basic assumptions have been violated in this paper. This study aims to propose a novel weighted MLP-ARIMA model filling the gap of series hybrid models by changing the order of sequence modeling and assigning weight for each component. Firstly, the modeling order is changed to nonlinear-linear, and then Multi-Layer Perceptron Neural Network (MLPNN) - Auto-Regressive Integrated Moving Average (ARIMA) models are employed to model and process nonlinear and linear components respectively. Secondly, each model's weights are computed by the Ordinary Least Square (OLS) weighting algorithm. Thus, in this paper, a novel improved weighted MLP-ARIMA series hybrid model is proposed for time series forecasting. The real-world benchmark data sets, including Wolf's sunspot data, the Canadian lynx data, and the British pound/US dollar exchange rate data, are elected to verify the effectiveness of the proposed weighted MLP-ARIMA series hybrid model. The simulation results revealed that the weighted MLP-ARIMA model could obtain superior performance compared to ARIMA-MLP, MLP-ARIMA, as well as the ARIMA and MLPNN individual models. The proposed hybrid model can be an effective alternative to improve forecasting accuracy obtained by traditional series hybrid methods.

Keywords: series hybrid model; weighted MLP-ARIMA model; auto-regressive integrated moving average (ARIMA); multi-layer perceptron neural network (MLPNN); time series forecasting.

Received: October 2019-14
Revised: October 2020-11
Accepted: October 2020-18

* Corresponding author; khashei@cc.iut.ac.ir
[1] Department of Industrial and Systems Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran.

1. Introduction

Accurate forecasting in the time series field plays a vital role in many research areas. Due to single models' insufficiencies in modeling and processing various behaviors in real-world data, combining different models grows into a well-accepted alternative for improving the forecasting accuracy. It is documented in the literature that hybridization can enhance individual models' forecasting power (Granger and Ramanathan 1984; Clemen, 1989; Timermann, 2006). This is because of the complicated nature of real-world time series. With their limitations in covering all existing data patterns, individual models are not sufficient for obtaining reasonable accurate forecasting results. The main idea behind combining different models is to covering the deficiency of individual models in modeling and processing several patterns in data by employing multiple individual models simultaneously.

The research on implementing the hybrid models and yielding improved forecast has rapidly developed in the recent literature of time series forecasting. A growing number of research studies used the concept of hybridization to overcome the shortcoming and disadvantages of individual models to improve forecasting accuracy. (Hajirahimi and Khashei, 2016) proposed three combination architectures of the ARIMA and ANN models for financial forecasting. Empirical results of forecasting the benchmark data sets indicated that hybrid models could generate superior results in comparison with both ARIMA and ANN models in forecasting stock prices. (Liu et al., 2019) proposed the novel hybrid model based on integrating empirical mode decomposition (EMD), moving average filter (MAF), least squares support vector regression (LSSVR), and quadratic, exponential smoothing (QES) for electrical energy consumption forecasting related to the cement grinding process. They demonstrate that the proposed model can extract each algorithm's advantages and outperform other forecasting models such as single LSSVR and single Bach Propagation Neural Network (BPNN). (Hu et al., 2019) developed a hybrid framework for short-term electric load forecasting models based on the GA-PSO algorithm and BPNN model's hybridization. The GA-PSO algorithm is used to optimize the parameters of BPNN. The numerical results demonstrated that the GA-PSO-BPNN hybrid model is superior to the other hybrid forecasting models and also BPNN individual model. (Aasim et al., 2019) developed a new Repeated WT based Auto-Regressive Integrated Moving Average (RWT-ARIMA) model, which has improved accuracy for very short-term wind speed forecasting. Comparing the proposed RWTARIMA model with the benchmark model for very short-term wind speed forecasting, the ARIMA model and WT-ARIMA model have been proved the proposed model's forecasting power for very short-term wind speed forecasting. Some recent studies utilized hybrid models for time series forecasting is briefly reviewed in Table 1. These large numbers of studies verified that today's specific attention focused on developing hybrid models since obtaining precious forecasting has become an essential and challenging task in a wide variety of time series forecasting areas. In a word, the literature review demonstrated the importance of combination theory in time series forecasting areas.

Table 1. Recent studies established hybrid models for time series forecasting

Authors	Field of study	Proposed hybrid model
(Zhang et al., 2019)	short-term wind speed forecasting	Variational Mode Decomposition (VMD) - Wavelet Transform (WT)-principal component analysis(PCA)-BPNN- Radial Basis Function Neural Network (RBFNN)
(Yuan et al. ,2019)	Prediction interval of wind power	Beta distribution function based particle swarm optimization (PSO) - long short-term memory (LSTM) neural network
(Lan et al., 2019)	Day-ahead spatiotemporal solar irradiation forecasting	Discrete Fourier Transform (DFT)-PCA-Elman Neural Network (ENN)
(Kumar Handar, 2019)	Stock market prediction	wavelet-adaptive network-based fuzzy inference system (WANFIS)

The series hybrid method, which is introduced for the first time by (Zhang, 2003), has been effectively applied for time series modeling and forecasting. (Zhang, 2003) supposed that time series is composed of linear and nonlinear mixture patterns. Thus, decomposing time series into linear and nonlinear components and choosing the ARIMA and Artificial Neural Network (ANN) models for processing separated linear and nonlinear patterns, respectively, the final hybrid ARIMA-ANN model is constructed. (Zhang, 2003) demonstrated that the proposed ARIMA-ANN model could improve the performance of individual models in real-world data forecasting. Since then, many researchers have applied various series hybrid models in several time series forecasting domains. (Singh et al., 2019) proposed a hybrid ARIMA-ANN series model for wind power forecasting. The experimental results showed that the series hybridization of ARIMA and ANN models provides better forecasting results as compared to the two models employing separately for wind power forecasting. (Zhang et al., 2018) designed a series hybrid framework using ARIMA and Support Vector Regression (SVR) models for forecasting the emergency patient flow. The obtained results showed that the proposed model's prediction capability is outperformed compared with individual models such as ARIMA and SVR models. Some recent studies used series hybrid models are summarized in Table 2.

Table 2. A literature review of recent series hybrid models developed for time series forecasting

Authors	Type of Series hybrid model	Field of study
(Chakraborty et al. ,2019)	ARIMA-Neural Network Autoregressive (NNAR)	Forecasting dengue epidemics
(Suhermi et al. ,2018)	ARIMA-Deep Neural Network (DNN)	Roll motion prediction
(Velasco et al. ,2018)	ARIMA-ANN	Load forecasting
(Naveena et al. ,2017)	ARIMA-ANN	Forecasting the Price of Robusta Coffee in India
(Mukaram and Yusof, 2017)	SARIMA-ANN	Solar radiation forecasting
(Moeeni et al., 2017)	SARIMA-Genetic Programming (GP)	Monthly reservoir inflow forecasting
(Babu and Sure, 2016)	ARIMA-ANN	Real-world time series

Literature shows that the series hybrid model is one of the most widely-used hybridization categories developed to improve forecasting accuracy. However, all these series hybrid models are proposed to improve forecasting accuracy; these hybridization procedures have two shortcomings in modeling and forecasting. Accordingly, despite a large literature on time series forecasting that uses the series hybrid models, two highlighted gaps could be extracted in developed series hybrid models. 1) all series hybrid models followed the linear-nonlinear modeling order, and 2) the equal weight is considered for linear and nonlinear parts. Besides, the literature review on the series hybrid models used in previous works shows that the most prevalent model used is ARIMA-ANN. Besides, all of the researchers follow the basic assumption considered in Zhang's model regarding ignoring each model forecast's appropriate weight. Thus, regarding these existing gaps in the literature and improving the performance of series hybrid models, in this paper, the novel weighted MLP-ARIMA is proposed for time series forecasting. The architecture of the proposed model consists of the following parts:

1) Decomposing a time series into nonlinear and linear components.
2) Applying the MLP and ARIMA models for processing separated patterns.
3) Assigning an exact optimum weight to each model by employing the OLS algorithm
4) Generating the final series hybrid forecast with the sum of weighted forecasts.

Motivated by existed gaps in the literature, the principal aim of the current study is to and enhance the forecasting accuracy of series hybrid models by proposing the novel weighted MLP-ARIMA model, which concenters on alleviating the two following shortcoming of previous series hybrid models:

1) Considering only the linear-nonlinear sequence modeling
2) Assigning equal weight to decomposed linear and nonlinear components

This study's major innovation lies in solving the above deficiencies of series hybrid models to improve the traditional series ARIMA-ANN model's forecasting performance. Thus, this paper attempts to develop a novel weighted MLP-ARIMA model for time series forecasting. In the proposed model, in contrast to conventional ARIMA-ANN models developed in the literature, the sequence of modeling is changed to MLP-ARIMA. Furthermore, despite the widely adopted meta-heuristic algorithms (Zaree et al., 2020; Iravani et al., 2020), in this study the OLS method is proposed for assigning appropriate weights of MLP and ARIMA models. The rest of this paper is organized as the following sections: in section 2, the ARIMA and MLP models' main concept is briefly explained. Section 3 describes the main idea and implementation steps of the proposed model. The experimental results of the case study are presented in Section 4. Section 5 provides the comparison obtained analysis. Section 6 gives the conclusions.

2. Methodologies

The individual models applied for time series forecasting problems can be represented to two main classes: statistical and intelligent methods. The most broadly-used approaches in these two classes of individual models are ARIMA and ANN models, which are described in the following two subsections, respectively.

2.1. Autoregressive Integrated Moving Average (ARIMA) models

ARIMA model is a repetitive procedure of forecasting future values of time series that similar to other statistical models by processing the historical data generate forecasting value of the variable. ARIMA model is composed of two main modeling parts, Auto-Regressive (AR) and Moving Average (MA), which are integrated together and construct ARIMA models. The ARIMA model is formulated in Eq. (1).

$$\phi(B)\nabla^d (y_t - \mu) = \theta(B) a_t \tag{1}$$

Where, y_t is the actual value in time t, a_t is the withe noise which is assumed to be independently and identically distributed with a mean of zero and constant variance of σ^2.

$\varphi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$, $\theta(B) = 1 - \sum_{j=1}^{q} \theta_j B^j$ are polynomials in B of degree p and q, $\varphi_i \quad (i = 1,2,...,p)$ and $\theta_j \quad (j = 1,2,...,q)$ are model parameters, $\nabla = (1 - B)$, B is the backward shift operator, p and q are integers and denote as orders of the model, and d is an integer and named as an order of differencing. The modeling formulation of ARIMA models regarding the (Box-Jenkins, 1976) methodology always contains three repetitive stages, including model identification, parameter estimation, and diagnostic checking. These steps are described as follows in detail.

1) In the first step, we are exploring for the actual values of p (the number of autoregressive), d (the number of differencing), and q (the number of the moving average). For this purpose, Box-Jenkins (1976) proposed the autocorrelation function (ACF) and the partial autocorrelation function (PACF) to recognize the order of the ARIMA.
2) After choosing a specific ARIMA (p, d, q), the parameters identified in the previous phase should be estimated using the ordinary least squared (OLS) method.
3) At last, the diagnostic checking of model adequacy is necessary to check whether the selected model is satisfactory for modeling and predicting time series or not. This step is usually done because of this fact that another model may be existed to provide better modeling of historical data. Therefore, for identifying the final structure, checking several diagnostic statistics and plots of the residuals is done to discover the most appropriate structure.

If the model is not adequate, a new structure of the ARIMA model will be constructed, and the three previous steps should be repeated until the best structure is found.

2.2. Multi-Layer Perceptron Neural Networks (MLPs) for forecasting time series

One of the most commonly used intelligent models that has was successfully applied in many domains, especially time series modeling and forecasting, is the ANN model. The main reason for this popularity is that ANN models extrapolate the underlying data generation without any assumption of the model form. Besides, another highlighted feature of neural networks is that they are universal approximators that can approximate a large class of functions accurately. There are various ANN model architectures available in the literature. Although neural networks have a similar structure but based on how to design, the distinction between different types of neural networks has been created. Single hidden layer feed-forward network (also known as multilayer perceptrons) is the most widely used neural network model architecture for time series modeling and forecasting. In this paper, this type of neural network is used for all nonlinear modeling. The MLP network consists of three layers, including input, hidden, and output layers (Khashei and Bijari, 2010).

Input layer phase: For the time series problems, an MLP is fitted with the past lagged value of actual data ($y_{t-1},...,y_{t-p}$) as an input vector. Therefore, the input layer is composed of p nodes that are connected to the hidden layer.

Hidden layer phase: The hidden layer makes an appropriate connection between the input and output layers. MLP models that are designed in this paper have a single hidden layer with q nodes. In this step, one of the crucial tasks is to determine the form of the activation function (g) to identify the relationship between the input and output layers. Neural networks support a wide range of activation functions such as linear, quadratic, tanh, and logistic. The logistic is a commonly used function as the hidden layer transfer function that is shown in Eq. (2).

$$g(x) = \frac{1}{1 + exp(-x)} \tag{2}$$

Output layer phase: In this step, by choosing an activation transfer function and the appropriate number of nodes, the output of the neural network is used to forecast the future values of time series. In this paper, the output layer of designed neural networks contains only one node because the one-step-ahead forecasting is considered. Also, the linear function as the nonlinear activation function is introduced for the output layer. The formula of the relationship between the input and output layer is presented in Eq. (3).

$$y_t = w_0 + \sum_{j=1}^{q} w_j . g\ [w_{0,j} + \sum_{i=1}^{p} w_{i,j} . y_{t-i}] + \varepsilon_t \tag{3}$$

Where, $w_{i,j} (i = 0,1....,p, j = 1,2,....,q)$ and $w_j (j = 0,1,2.....,q)$ are referred to as connection weights.

3. Formulating the proposed weighted MLP-ARIMA model

The broadly-established idea in hybridization is to decompose time series into some main components. Based on this idea, the literature on forecasting time series with hybrid models is growing with concentrating on developing series hybrid models following Zhang's model modeling procedures. Thus, this paper's main objective is to propose a weighted MLP-ARIMA hybrid model in contrast to traditional series hybrid models, which are considered linear-nonlinear modeling order with assuming equal weight for each model forecast. In the proposed model, the modeling order of series models is changed and assign exact optimum weights to each model forecast, which are not addressed in previous works. The modeling steps of the proposed model is explained in detail as follows:

1) ***Decomposing time series into nonlinear and linear components***: based on the basic idea of series hybrid models, at the first step, the time series is considered as an additive function of nonlinear and linear parts as follows:

$$y_t = N_t + L_t \quad (t = 1,2,...,m) \tag{4}$$

Where, y_t, N_t and L_t denote as actual data, nonlinear and linear components respectively and *m* denotes as a number of data existed in time series.

2) ***Modeling each component by MLP and ARIMA models:*** firstly, the MLP model is applied to the original time series to process all nonlinear patterns as formulated in Eq. (5).

$$\hat{N}_t = \sum_{j=1}^{q} w_j . g\ [w_{0,j} + \sum_{i=1}^{p} w_{i,j} . y_{t-i}] \tag{5}$$

After processing the nonlinear patterns by MLP model, the remained linear relationships that are latent in the residual of the MLP model are modeled by defining MLP's residuals as input data for the ARIMA model, as shown in Eq. (6).

$$\hat{L}_t = \sum_{i=1}^{p} \varphi_i y_{t-i} - \sum_{j=1}^{q} \theta_j \hat{e}_{MLP,t-j} \tag{6}$$

3) Assigning weights to obtained forecasts: in this step in contrast to other series hybrid models developed in the literature, the exact optimum weight of MLP and ARIMA forecasts are computed by the OLS algorithm given in Eq. (7) and (8), respectively.

$$\hat{W}_{\hat{N}_t}=\frac{\left(\sum_{t=1}^{m}y_t\hat{f}_{\hat{N}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\right)-\left(\sum_{t=1}^{m}y_t\hat{f}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}_{\hat{L}_t,t}\hat{f}_{\hat{N}_t,t}\right)}{\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)-\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)^{2}} \tag{7}$$

$$\hat{W}_{\hat{L}_t}=\frac{\left(\sum_{t=1}^{m}y_t\hat{f}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)-\left(\sum_{t=1}^{m}y_t\hat{f}_{\hat{N}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}_{\hat{L}_t,t}\hat{f}_{\hat{N}_t,t}\right)}{\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)-\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)^{2}} \tag{8}$$

Where, $\hat{f}_{\hat{N}_t,t}$ and $\hat{f}_{\hat{L}_t,t}$ are denoted as the forecasted value obtained by MLP and ARIMA models, respectively.

4) ***Constructing the final hybrid forecast:*** In the last stage, the final forecasts of the proposed weighted MLP-ARIMA model are obtained as follows:

$$\hat{f}_{weighted\ MLP-ARIMA,t}=W_{\hat{L}_t}\hat{L}_t+W_{\hat{N}_t}\hat{N}_t \qquad (t=1,2,...,m)$$

$$=\frac{\left(\sum_{t=1}^{m}y_t\hat{f}^{2}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)-\left(\sum_{t=1}^{m}y_t\hat{f}_{\hat{N}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\hat{f}_{\hat{N}_t,t}\right)}{\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)-\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)^{2}} \tag{9}$$

$$+\frac{\left(\sum_{t=1}^{m}y_t\hat{f}^{2}_{\hat{N}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\right)-\left(\sum_{t=1}^{m}y_t\hat{f}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}_{\hat{L}_t,t}\hat{f}^{2}_{\hat{N}_t,t}\right)}{\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\right)\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)-\left(\sum_{t=1}^{m}\hat{f}^{2}_{\hat{L}_t,t}\sum_{t=1}^{m}\hat{f}^{2}_{\hat{N}_t,t}\right)^{2}}$$

where, $W_{\hat{L}_t}$ and $W_{\hat{N}_t}$ are denoted as weights of ARIMA and MLP models obtained by OLS algorithm, respectively, and $\hat{f}_{weighted\ MLP-ARIMA,t}$ is referred to as the hybrid forecasting result.

4. Experiment and result analysis

In this section, the weighted MLP-ARIMA model's performance is proved by applying it to forecast the three widely-used real-world data sets, including Wolf's sunspot, the Canadian Lynx, and the British pound/US dollar exchange rate time series. These data sets are well-studied choices employed for evaluating different hybrid models constructed with ARIMA and MLP models for time series forecasting (Zhang et al., 1998; Khashei and Bijari, 2012; Hajirahimi and Khashei, 2019). First, the data sets and evaluation criteria are presented. Subsequently, the obtained numerical results are reported.

4.1. Data sets

The sunspot data considered in this paper covers the annual number of sunspots from 1700 to 1987. Two remarkable characteristics of the sunspot data set are nonlinearity and non-gaussianity. This data set is frequently used for assessing nonlinear models (Hipel and McLeod, 1994). The data set from 1700-1920 (221 data points) are utilized as a training data set, and remained data samples (67 observations) used as testing samples. The Lynx data set was utilized in a large number of time series forecasting and modeling problems, which is the annual record of the number of the Canadian Lynx from 1821 to 1934 (Stone and He, 2007). The Lynx data set consists of 114 annual observations from 1821 to 1934. In this time series, the first 100 data samples are used as a training data set, and the last 14 data points are marked as test data points. The last data set is the exchange rate between the British pound and the US dollar. The exchange rate data set employed in this study covers weekly observations from 1980 to 1993 (Meese and Rogoff, 1983). The plot of sunspot, Lynx, and exchange rate data sets are depicted in Fig. 1 to 3, respectively. The data sets descriptions are given in Table 3 in detail. It should be noted that all MLP and ARIMA models are programmed with MATLAB software package R2013a and Eviews 9 software, respectively.

Table 3: The description of data sets

Data set	Sample size	Training set	Tet set
Sunspot	288	221(1700-1920)	67 (1921-1987)
Lynx	114	100 (1821-1920)	14 (1921-1934)
Exchange rate	731	676(1980-1992)	52(1993)

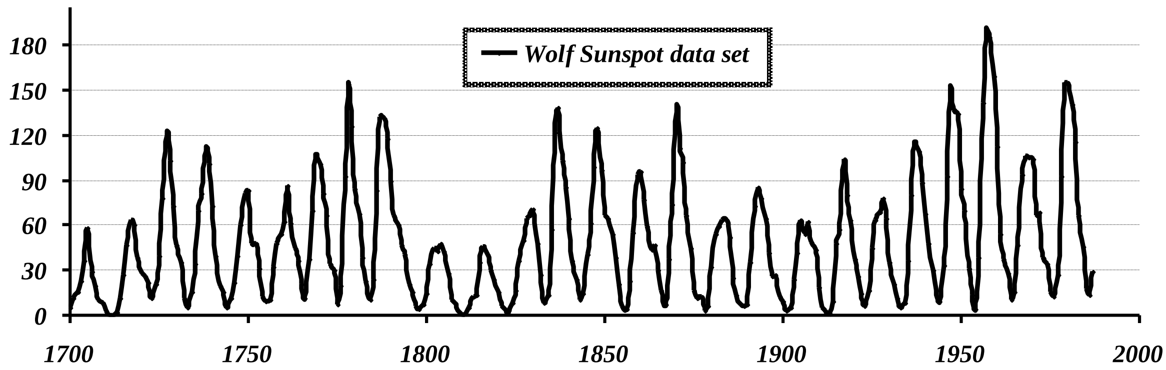


Figure 1. The annual sunspot data from 1700 to 1987

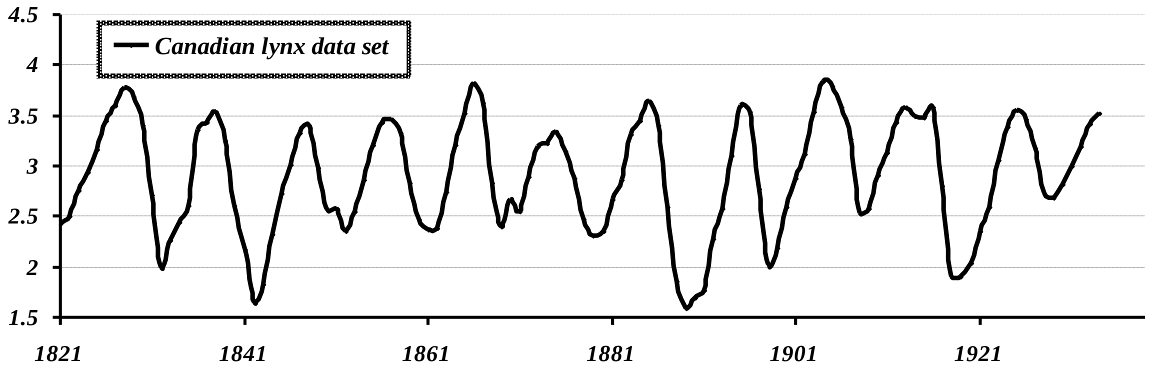


Figure 2. The annual Canadian lynx data from 1821 to 1934

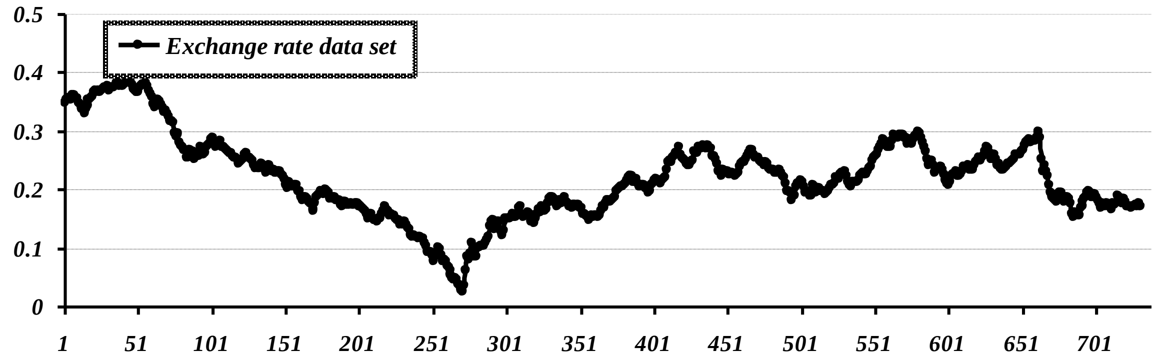


Figure 3. The weekly British pound/US dollar exchange rate data from 1980 to 1993

4.2. Evaluation criteria

In this paper, to evaluate the forecasting capability of the proposed model, four main criteria, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE), and Sum of Squared Error (SSE) are proposed which are formulated in Eq. (7) to (10), respectively.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 \tag{7}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|e_i| \tag{8}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|e_i|}{y_i} * 100\% \tag{9}$$

$$SSE = \sum_{i=1}^{N}(e_i)^2 \tag{10}$$

4.3. Sunspot data set forecasting result

According to the proposed algorithm, at the first stage, the time series is considered as nonlinear and linear parts. Then the MLP model is applied to the original time series. The best fitted MLP model, which is selected based on the best performance in the test data set, is the network with four input, four hidden neurons, and one neuron in the output layer. Then, the MLP model residual is computed and given to ARIMA to model the remained linear patterns in MLP's residuals. At the next stage, the weight of each model forecast is calculated by the OLS algorithm. At the last stage, the final hybrid forecast of the proposed model is obtained by Eq. (11).

$$\hat{y}_{weighted\ MLP-ARIMA,t} = W_{\hat{N}_t}\hat{N}_t + W_{\hat{L}_t}\hat{L}_t \quad (t = 1,2,...,m) \tag{11}$$

The plots of actual value against the weighted MLP-ARIMA models forecasted value in test samples is given in Fig. 4.

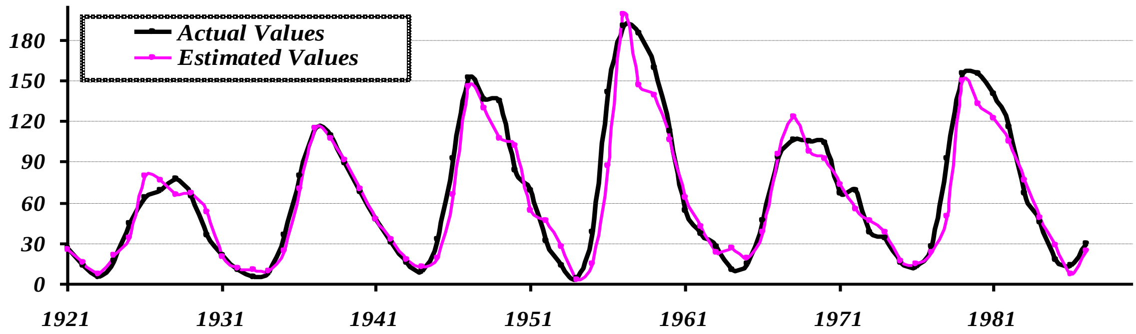


Figure 4. The actual value against the estimated value of the weighted MLP-ARIMA model for sunspot data set (test sample)

4.4. Lynx data set forecasting result

After decomposing time series into nonlinear and linear components in a similar fashion, the MLP model with the structure $N^{(7,5,1)}$ is designed to model nonlinear patterns. Then in the next stage for linear modeling, the ARIMA is applied to MLP's residuals. Finally, the final forecasting results employ the OLS weighting algorithm constructed by Eq. (11). The forecasted value obtained by the weighted MLP-ARIMA model is plotted in Fig. 5 for testing samples.

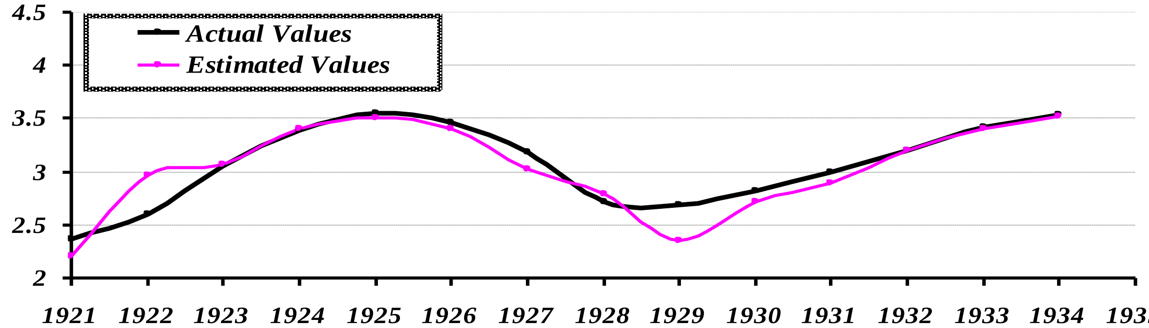


Figure 5. The actual value against the estimated value of the weighted MLP-ARIMA model for Lynx data set (test sample)

4.5. Exchange rate data set forecasting result

In line with the weighted MLP-ARIMA modeling procedure, the best fitted MLP model for the nonlinear decomposed part is a network composed of seven input, six hidden neurons, and the output layer consists of one neuron. In the next stage, the ARIMA model is established to process linear relationships hidden in MLP's residuals. Finally, the hybrid forecasting result is yielded by Eq. (11). The graphical representation of the weighted MLP-ARIMA model's forecasting results for the exchange rate testing data set is illustrated in Fig. 6.

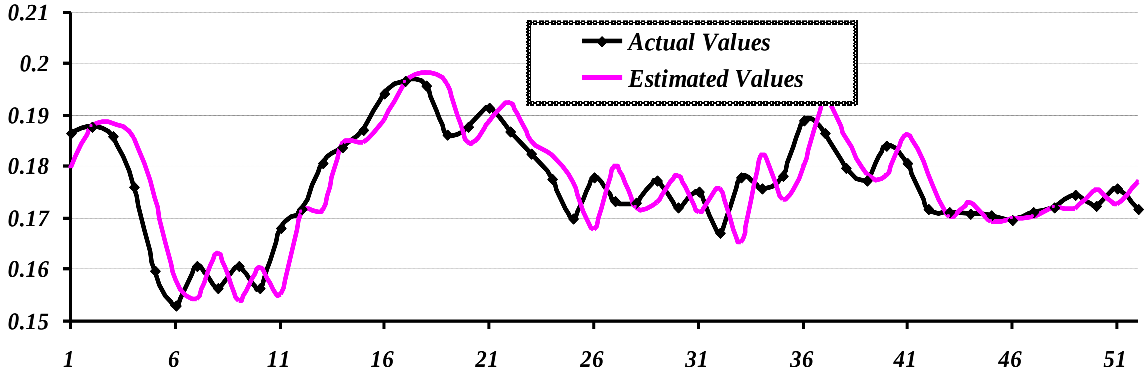


Figure 6. The actual value against the estimated value of the weighted MLP-ARIMA model for exchange rate data set (test sample)

Table 4 shows the weighted MLP-ARIMA model's performance result for all three cases forecasting in terms of the MSE, MAE, MAPE, and SSE error indexes.

Table 4. The forecasting results of weighted MLP-ARIMA hybrid model

Data set	Training sample			
	MSE	MAE	MAPE	SSE
sunspot	206.151	9.703	*0.508	44021
Lynx	0.024*	0.132	0.050	2.517
exchange rate	2.919	0.003	0.019	127.01

*Note: It should be multiplied by 10^{-5}.

It should be noted that normalization is a necessary prior process for data sets entry into the MLP models. The input data are normalized based on the following formula:

$$y_{normalized,t} = \frac{y_t - y_{min,t}}{y_{max,t} - y_{min,t}} \qquad (t = 1, 2, \ldots, m) \tag{12}$$

Then for reaching the optimum number of hidden nodes in designing MLP models, several structures of MLP models are implemented in which the number of hidden nodes has fluctuated from 2 to 20 nodes. Consequently, the optimum structure is selected regarding minimum MSE in the test data set. Hence, the logistic function is applied as a transfer function in both hidden and output layers in all cases. The tainlm (Levenberg-Marquardt) is selected as a training algorithm in designing all MLP models used in weighted MLP-ARIMA models. Additionally, the maximum 1000 iteration is adjusted as convergence criteria for training MLP models. Table (5) summarized the procedure of implementing different topologies of MLP models for finding the optimum structure in sunspot data set forecasting.

Table 5. The performance of the different topologies of the MLP models for forecasting sunspot data set

Input layer	Hidden layer		Output layer			Structure	Performance (MSE)
Number of neurons	Number of neurons	Transfer function	Number of neurons	Transfer function	Training function		
3	2	logsic	1	purelin	tainlm	$N^{(3,2,1)}$	711.718
3	3	logsic	1	purelin	tainlm	$N^{(3,3,1)}$	688.163
3	4	logsic	1	purelin	tainlm	$N^{(3,4,1)}$	619.712
3	5	logsic	1	purelin	tainlm	$N^{(3,5,1)}$	594.993
3	6	logsic	1	purelin	tainlm	$N^{(3,6,1)}$	511.022
3	7	logsic	1	purelin	tainlm	$N^{(3,7,1)}$	489.409
3	8	logsic	1	purelin	tainlm	$N^{(3,8,1)}$	479.719
3	2	tansig	1	purelin	tainlm	$N^{(3,2,1)}$	657.717
3	3	tansig	1	purelin	tainlm	$N^{(3,3,1)}$	630.461
3	4	tansig	1	purelin	tainlm	$N^{(3,4,1)}$	573.058
3	5	tansig	1	purelin	tainlm	$N^{(3,5,1)}$	590.994
3	6	tansig	1	purelin	tainlm	$N^{(3,6,1)}$	521.485
3	7	tansig	1	purelin	tainlm	$N^{(3,7,1)}$	509.136
3	8	tansig	1	purelin	tainlm	$N^{(3,8,1)}$	498.713
4	2	logsic	1	purelin	tainlm	$N^{(4,2,1)}$	465.165
4	3	logsic	1	purelin	tainlm	$N^{(4,3,1)}$	409.026
4	4	logsic	1	purelin	tainlm	$N^{(4,4,1)}$	351.187
4	5	logsic	1	purelin	tainlm	$N^{(4,5,1)}$	356.730
4	6	logsic	1	purelin	tainlm	$N^{(4,6,1)}$	370.551
4	2	tansig	1	purelin	tainlm	$N^{(4,2,1)}$	473.237
4	3	tansig	1	purelin	tainlm	$N^{(4,3,1)}$	395.778
4	4	tansig	1	purelin	tainlm	$N^{(4,4,1)}$	367.043
4	5	tansig	1	purelin	tainlm	$N^{(4,5,1)}$	367.017
4	6	tansig	1	purelin	tainlm	$N^{(4,6,1)}$	383.326

5. Comparison analysis

In this section, the forecasting power and effectiveness of the proposed hybrid model are compared with its constituents, including ARIMA and MLP, and other series hybrid models such as ARIMA-MLP and MLP-ARIMA models in MAE and MSE evaluation criteria. Forecasting results in test data set for comparing the proposed model with other sunspot models, Canadian Lynx, and exchange rate data sets are summarized in Table 5. The comparison results given in Table 5 indicate that by evaluating obtained forecasting results, the following crucial points can be extracted:

1) Empirical results addressed in Table 5 represented that the proposed weighted MLP-ARIMA model outperformed the traditional MLP-ARIMA model in both indicators in all data sets. As an example, in the sunspot data set, the improvements obtained by the proposed weighted MLP-ARIMA model compared with the traditional MLP-ARIMA model are 4.44% and 1.54% in MAE and MSE error indexes, respectively.

2) Another highlighted result is that the weighted MLP-ARIMA model can improve the performance of the ARIMA-MLP model. For instance, in the lynx data set, the weighted MLP-ARIMA model's improvement compared with the ARIMA-MLP model is 3.56% in MAE and 0.64% in MSE terms.
3) Based on the numerical results, it is induced that the proposed weighted MLP-ARIMA model can improve the forecasting accuracy of both based components (ARIMA and MLP models) simultaneously. For example, in the exchange rate data set, the percentage improvements of the proposed model compared with the ARIMA model in MAE and MSE terms are 31.53% and 26.76%, respectively, and compared with the MLP model in MAE and MSE terms are 30.76% and 26.79%, respectively.

Table 5. Performance of the weighted MLP-ARIMA model in comparison with other models (test data)

Model	**Data set**	**MAE**	**MSE**
ARIMA	sunspot	13.03	306.08
MLP		13.54	351.19
ARIMA-MLP		12.78	280.15
MLP-ARIMA		10.15	213.98
Weighted MLP-ARIMA		**9.70**	**210.67**
ARIMA	Lynx	0.112255	0.020486
MLP		0.112109	0.020466
ARIMA-MLP		0.103972	0.017233
MLP-ARIMA		0.101394	0.020862
Weighted MLP-ARIMA		**0.100250**	**0.017121**
ARIMA	Exchange rate	0.0053597	4.52977
MLP		0.0052513	4.52657
ARIMA-MLP		0.0051212	4.35907
MLP-ARIMA		0.004956	3.641265
Weighted MLP-ARIMA		**0.003670**	**3.31391**

6. Conclusion

Hybrid models, due to their highlighted features in modeling complex behaviors and enhancing forecasting abilities of individual models, have been applied widely for time series forecasting problems. Series hybrid methodology is one of the most well-known hybridization methods which is used in numerous practical applications. However, these hybrid models encountered some limitations, including the fixed modeling order (linear-nonlinear) and assigning equal weight to each model forecasts. Thus, in this study, a novel weighted MLP-ARIMA series hybrid model is proposed for time series forecasting that integrates MLP and ARIMA models using a series of hybrid methodologies. In the proposed model, in contrast to other traditional series models, the time series is decomposed to nonlinear and linear components, and then the MLP model is used in the first stage for nonlinear pattern recognition. The other novelty addressed in the proposed model is to violate the equal weight assumption considered in other series hybrid models by employing the OLS algorithm to determine each model's weight. The real-world data sets, including Wolf's sunspot data, the Canadian lynx data, and the British pound/US dollar exchange rate, are utilized as the study cases for training and testing the proposed model. Experimental results proved that Compared with other employed hybrid forecasting models (ARIMA-MLP and MLP-ARIMA) and the ARIMA and MLP models, the proposed weighted MLP-ARIMA model could yield the highest accuracy. The proposed weighted MLP-ARIMA model can be suggested as an accurate series hybrid approach for time series forecasting.

Reference

Babu, CN., and Sure, P., (2016). "Partitioning and interpolation based hybrid ARIMA–ANN model for time series forecasting", *Sādhanā*, Vol. 41, pp. 695–706.

Box ,P., and Jenkins, G.M., (1976). Time Series Analysis: Forecasting and Control, Holden-day Inc., San Francisco, CA.

Chakraborty, T., Chattopadhyay, and S., Ghosh, I., (2019). "Forecasting dengue epidemics using a hybrid methodology", *Physica A*, Vol. 527.

Clemen, R. T., (1989). “Combining forecasts: A review and annotated bibliography”, *International Journal of Forecasting*, Vol. 5, pp. 559-583.

Granger, CWJ., and Ramanathan, R., (1984). "Improved Methods of Combining Forecasts", *Journal of Forecasting*, Vol. 3, pp. 197–204.

Hajirahimi, Z., and Khashei, M., (2016). "Improving the performance of financial forecasting using different combination architectures of ARIMA and ANN models", *Journal of Industrial Engineering and Management Studies*, Vol. 3, pp. 17-32.

Hajirahimi, Z., and Khashei, M., (2019). "Weighted sequential hybrid approaches for time series forecasting", *Physica A*, Vol. 531, pp. 1-14.

Hipel, K.W., and McLeod, A.I., (1994). Time Series Modeling of Water Resources and Environmental Systems, Amsterdam, Elsevier.

Hu, Y., Li, J., Hong, M., Ren, Jingzheng., Lin, Ruojue., Liu, Yue., Liu, M., and Man,Yi., (2019). "Short term electric load forecasting model and its verification for process industrial enterprises based on hybrid GA-PSO-BPNN algorithm -A case study of papermaking process", *Energy*, Vol. 170, pp. 1215-1227.

Iravani, M., Bashirzadeh, R., and Tarokh, M.J., (2020). "Developing an urban congestion pricing model by considering sustainability improvement and using a multi-objective optimization approach", *Journal of Industrial Engineering and Management Studies*, Vol. 7, pp. 56-76.

Khashei, M., and Bijari, M., (2010). "An artificial neural network (p, d ,q) model for timeseries forecasting", *Expert Systems with Applications*, Vol. 37, pp. 479–489.

Khashei, M., and Bijari, M., (2012). "A new class of hybrid models for time series forecasting", *Expert System with Application*, Vol. 39, pp. 4344–4357.

Kumar Chandar, S., (2019). “Fusion model of wavelet transform and adaptive neuro fuzzy inference system for stock market prediction”, Journal of Ambient Intelligence and Humanized Computing, pp.1-9.

Lan, H., Zhang, Ch., Hong, Ying-Yi., He, Y., and Wen, Shuli., (2019). "Day-ahead spatiotemporal solar irradiation forecasting using frequency based hybrid principal component analysis and neural network", *Applied Energy*, Vol. 247, pp. 389–402.

Liu, Z., Wang, X., Zhang, Q., and Huang, Ch., (2019). "Empirical mode decomposition based hybrid ensemble model for electrical energy consumption forecasting of the cement grinding process", *Measurement*, Vol. 138, pp. 314–324.

Meese, R.A., and Rogoff, K., (1983). "Empirical exchange rate models of the seventies: do they/t out of samples?" , *Journal of International Economics*, Vol. 14, pp. 3–24.

Moeeni, H., Bonakdari, H., and Ebtehaj, Isa., (2017). "Monthly reservoir inflow forecasting using a new hybrid SARIMA genetic programming approach", *Journal of Earth System Science*, Vol. 126, pp. 1–13.

Mukaram, MZ., and Yusof, F., (2017). "Solar radiation forecast using hybrid SARIMA and ANN model: A case study at several locations in Peninsular Malaysia", *Malaysian Journal of Fundamental and Applied Sciences*, pp. 346-350.

Naveena, K., Singh, Subedar., Rathod, Santosha., and Singh, Abhishek., (2017). "Hybrid ARIMA-ANN Modeling for Forecasting the Price of Robusta Coffee in India", *International Journal of Current Microbiology and Applied Sciences*, Vol.7, pp. 1721-1726.

Singh, P. K., Singh, Nitin., and Negi, Richa., (2019). "Wind Power Forecasting Using Hybrid ARIMA-ANN Technique", *Ambient Communications and Computer Systems*, pp. 209-220.

Stone, L., He, D., (2007). "Chaotic oscillations and cycles in multi-trophic ecological systems", *Journal of Theoretical Biology*, Vol. 248, pp. 382–390.

Suhermi, N., Prastyo, DD., and Ali, B., (2018). "Roll motion prediction using a hybrid deep learning and ARIMA model", *Procedia Computer Science*, Vol. 144, pp. 251–258.

Timermann, A., (2006). Forecast Combinations, in Handbook of Economic Forecasting, Vol. 1, eds. G. Elliott, C. Granger, and A. Timmermann, Ams-terdam: Elsevier, pp. 135–196.

Velasco, LCP., Polestico, DLL., Macasieb, GPO., Reyes, MBV., and Vasquez, Jr. FB., (2018). "Load forecasting using autoregressive integrated moving average and artificial neural network", *International Journal of Advanced Computer Science and Applications*, Vol. 9, pp. 23–39.

Yuan, X., Chen, Ch., Jiang, M., and Yuan, Yanbin., (2019). "Prediction interval of wind power using parameter optimized Beta distribution based LSTM model", *Applied Soft Computing Journal*, Vol. 82.

Zaree, M., Kamranrad, R., Zaree, M., and Emami, I., (2020). "Project scheduling optimization for contractor's Net present value maximization using meta-heuristic algorithms: A case study", *Journal of Industrial Engineering and Management Studies*, Vol. 7, pp. 36-55.

Zhang, G. P., (2003). "Time series forecasting using a hybrid ARIMA and neural network model", Neurocomputing, Vol. 50, pp. 159–175.

Zhang, G., Patuwo, BE., and Hu, M., (1998). "Forecasting with artificial neural networks: the state of the art", *International Journal of Forecasting*, Vol. 14, pp. 35–62.

Zhang, Y., Chen, B., Pan, G., and Zhao, Y., (2019). "A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting", *Energy Conversion and Management*, Vol. 195, pp. 180–197.

Zhang, Y., Luo, L., Yang, J., Liu, D., Kong, R, and Feng, Y., (2018). "A hybrid ARIMA-SVR approach for forecasting emergency patient flow", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 10, pp. 3315–3323.

This article can be cited: Hajirahimi, Z., Khashei, M., (2020). "Weighted MLP-ARIMA series hybrid model for time series forecasting", *Journal of Industrial Engineering and Management Studies*, Vol. 7, No. 2, pp. 187-201.