Appl Intell (2012) 37:511–519
DOI 10.1007/s10489-012-0344-1

An enhanced hybrid method for time series prediction using linear and neural network models

Purwanto · C. Eswaran · R. Logeswaran

Published online: 4 April 2012

Abstract The need for improving the accuracy of time series prediction has motivated researchers to develop more efficient prediction models. The accuracy rates resulting from linear models such as linear regression (LR), exponential smoothing (ES) and autoregressive integrated moving average (ARIMA) are not high as they are poor in handling the nonlinear time series data. Neural network models are considered to be better in handling such nonlinear time series data. In the real-world problems, the time series data consist of complex linear and nonlinear patterns and it may be difficult to obtain high prediction accuracy rates using only linear or neural network models. Hybrid models which combine both linear and neural network models can be used to obtain high prediction accuracy rates. In this paper, we propose an enhanced hybrid model which indicates for a given input data which choice is better between the two options, namely, a linear-nonlinear combination or a nonlinear-linear combination. The appropriate combination is selected based on a linearity test of data. From the experimental results, it is found that the proposed hybrid model comprising linear-nonlinear combination performs better than other models for the data that have a linear relationship. On the contrary, the hybrid model comprising nonlinear-linear combination performs better than other models for the data that have a nonlinear relationship.

Keywords Exponential smoothing · Linear regression · ARIMA · Neural network · Enhanced hybrid method

Purwanto (✉) · C. Eswaran
Faculty of Information Technology, Multimedia University, 63100 Cyberjaya, Malaysia
e-mail: mypoenk@gmail.com

C. Eswaran
e-mail: eswaran@mmu.edu.my

Purwanto
Faculty of Computer Science, Dian Nuswantoro University, 50131 Semarang, Indonesia

R. Logeswaran
Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Malaysia
e-mail: loges@mmu.edu.my

1 Introduction

The accuracy of a prediction method depends on both the model and the complexity of the data and hence it is important to choose the best model based on the complexity of data.

Real-world time series data often contain both linear and nonlinear patterns. As such, using only linear or only nonlinear models has limitations in handling the relationships among the data. Some of the popular models used in time series prediction include autoregressive integrated moving average (ARIMA) [1, 2], linear regression [3, 4], neural networks [5–7] and even hybrid models combining ARIMA and neural network models in some applications [8].

An enhanced hybrid method for time series prediction is proposed in this paper. In this method, a linearity test is first applied to decide whether the pattern of the input data is linear or nonlinear. Depending on the result, we then apply a linear-nonlinear combination or a nonlinear-linear combination depending on whether the data has a linear or nonlinear pattern. To illustrate the capabilities of the proposed method, it is implemented for the prediction of morbidity of tuberculosis (MTB), birth and immigration. The proposed hybrid method evaluates the performances of four linear models, namely, linear regression (LR), exponential smoothing (ES, single and double), and autoregressive integrated moving average (ARIMA), and selects the best model for the input data. Nonlinear prediction is undertaken using a multi-layer perceptron (MLP) neural network. The optimal configuration of the MLP network is determined based on the input data.

The paper is organized as follows: Sect. 2 reviews the various models used in this work; Sect. 3 presents the proposed enhancement procedure for time series prediction; Sect. 4 describes the data used in this work; Sect. 5 describes experimental results; Sect. 6 provides the performance evaluation of the different models; Sect. 7 contains the concluding remarks.

2 Models used

The models used in this work can be classified into three groups, namely, linear models, neural network based models and hybrid models.

The following are some of the popular linear models that are used for prediction [9–11].

(i) Exponential smoothing (ES) models (Single ES and Double ES)
(ii) Linear regression (LR) model
(iii) Autoregressive integrated moving average (ARIMA) model

The neural network model used in this work is the multi-layer perceptron (MLP) network which is the most common neural network model used in prediction [8, 12]. The MLP networks are generally good at fitting non-linear time series data. The advantage of MLP is the model can capture the non-linear patterns of time series.

The time series prediction output (y_t) of the MLP is computed as [12]:

$$y_t = \beta_0 + \sum_{j=1}^{q} \beta_j \psi\left(\gamma_{j0} + \sum_{i=1}^{p} \gamma_{ji} y_{t-i}\right) + e_t \tag{1}$$

where $\beta_0, \beta_1, \ldots, \beta_q$ are a bias on the j-th unit, $\gamma_{10}, \ldots, \gamma_{qp}$ are the connections weights or parameters of the MLP, $y_{t-1}, y_{t-2}, y_{t-3}, \ldots, y_{t-p}$ are actual data, p is the number of input nodes, q is the number of hidden nodes and $\psi(.)$ is activation function. Several activation functions ψ are commonly used in MLP. These include sigmoid, bipolar sigmoid and hyperbolic tangent.

The linear models such as exponential smoothing (single ES and double ES), linear regression (LR) and ARIMA and neural network models have achieved success in their respective linear and nonlinear domains. Linear models may not yield accurate results for complex nonlinear problems, and neural network models yielded mixed results in linear problems [13, 14]. When the data is known to contain mixed linear and nonlinear characteristics, or when the characteristics of the relationships is unknown, a hybrid methodology [14, 15] that employs both linear and nonlinear modeling capabilities would be more likely to achieve accurate prediction results.

In general, time series data is composed of a linear autocorrelation structure and a non-linear component as shown below [14]:

$$y_t = L_t + N_t \tag{2}$$

where y_t is the original time series data at time t, L_t is the linear component and N_t represents the non-linear component. The errors or residuals can be calculated using ARIMA model [14]:

$$e_t = y_t - \hat{L}_t \tag{3}$$

where e_t is errors or residuals, $\hat{L}_t$ is the predicted value and is computed using linear model. The error sequence e_t is then applied neural network to compute the predicted value of the non-linear component $\hat{N}_t$.

The hybrid model for time series prediction is computed as:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \tag{4}$$

where $\hat{N}_t$ is the predicted value of N_t and is computed using the neural network with input the residuals.

3 The proposed enhancement method

The proposed enhancement procedure applies a linearity test on the time series data and if the result is positive linear, a linear-nonlinear hybrid (LNH) model is selected. On the other hand, if the linearity test result is negative linear, a nonlinear-linear hybrid (NLH) model is selected. Several linear models are evaluated for identifying the best linear model (BLM) and this model is used along with a multilayer perceptron neural network (MLP) in the LNH and NLH models. To identify BLM, the performances of four linear models, namely Exponential Smoothing (single and double), Linear Regression and ARIMA are evaluated. In other words, the LNH model uses BLM first followed by MLP and the NLH model uses MLP first followed by BLM. The proposed hybrid method is illustrated in the flowchart shown in Fig. 1 which involves three steps as explained below:

Step I: *Preprocessing*. Preprocessing is required to clean the data (such as removing missing values and outliers in the data).

Step II: *Linearity test*. We use Ramsey RESET test (regression equation specification error test) method for testing the linearity of the data [16]. If the result of this test is positive linear, a LNH model will be selected. On the other hand, if the result is negative linear, a NLH model will be used as shown in Fig. 1.

The Procedure of the RESET test can be illustrated as follows [16]:

(a) Make a regression y_t on x_t, suppose the model first estimated by:

$$y_t = f_t + \hat{e}_t, \quad \text{where } f_t = \beta x_t,\ t = 1, 2, \ldots, n \tag{5}$$

(b) Add the linear model in (a), thereby the Eq. (5) becomes:

$$y_t = f_t + \alpha_2 f_t^2 + \cdots + \alpha_2 f_t^k + v_t, \quad \text{where } k \geq 2 \tag{6}$$

(c) Test for functional form with an F test

- The null hypothesis (H_0: $\alpha_2 = \cdots = \alpha_k = 0$) is that the correct specification is linear.
- The alternative hypothesis (H_1: $\alpha_2 \neq \cdots \neq \alpha_k \neq 0$) is the correct specification is non-linear.

The F-statistics is computed as:

$$\text{F-statistics} = F_{(k-1,n-k-1)} = \frac{(SSR_a - SSR_b)/(k-1)}{SSR_b/(n-k)} \tag{7}$$

where *SSR* is the sum of squared residuals.

The Ramsey RESET test result obtained as follows:

- If the F-statistics is greater than the F-critical value, then we **reject** the null hypothesis. It means that the true specification is non-linear (negative linear).
- If the F-statistics is less than the F-critical value, then we are **unable to reject** the null hypothesis. The result suggests that the true specification is linear (positive linear).

If the linearity test yields positive linear result, the following steps are used.

Step III(A): Apply the *Hybrid model* with *linear-nonlinear combination*. In this step, there are four phases as explained below:

- Phase I: *Select the best linear model*. Based on the time series data, linear models are first used to predict the linear component $\hat{L}_t$. Comparing the performance results of the four linear models, namely, ES (single and double), linear regression (LR) and autoregressive integrated moving average (ARIMA) models, the best model is selected. We compute time series prediction ($\hat{L}_t$) using different weights for ES (single and double) models. Whereas for the ARIMA model, we calculate time series prediction using different parameters. The performance measures used in this phase are root mean square error, mean absolute error and mean absolute percentage error.
- Phase II: *Compute the residual.* The residual or error series e_t is obtained using Eq. (3).
- Phase III: *Neural network model.* The residuals obtained in Phase II are applied to a neural network model (MLP) to determine $\hat{N}_t$ as shown in Eq. (1). To obtain the optimum architecture configuration for MLP, different values of input and hidden layer neurons as well as different activation functions such as hyperbolic tangent, bipolar sigmoid and sigmoid functions are tested. The neural network model for the residuals (errors) will be computed as:

$$f(e) = \beta_0 + \sum_{j=1}^{H} \beta_j \psi \left(\gamma_{j0} + \sum_{i=1}^{m} \gamma_{ji} e_i \right) \tag{8}$$

  where m: number of input nodes, H: number of hidden layers. The forecast in Eq. (8) denote as $\hat{N}_t$.
- Phase IV: *Combining the best linear and neural network models.* In this phase, we combine the best linear model result (phase I) and the optimum neural network result (phase III) to get the overall prediction of the hybrid model as shown in Eq. (4).

If the linearity test yields a negative result, the following steps are used.

Step III(B): *Hybrid model (nonlinear-linear combination)*. In this step, the input data is first applied to the neural network model (MLP) to calculate $\hat{N}_t$, and the residual e_t which is applied to the best linear model to obtain $\hat{L}_t$. The final prediction result is obtained as $\hat{y}_t = \hat{N}_t + \hat{L}_t$.

4 Data used

Four data sets are used in this study for evaluating the performance of the proposed enhanced hybrid model. The first and second data is related to morbidity. The morbidity of tuberculosis in Indonesia (MTB-I) data and morbidity of tuberculosis in Zambia (MTB-Z) data were collected from World Health Organization (WHO) [17]. Morbidity refers to the total number of people who suffer a certain disease, such as Malaria, Tuberculosis. The MTB-I during the period 1990–2007 and MTB-Z during the period 1990–2007 are used in this study. The data set for the years (1990–2005) are used for training and the data set for the years (2006–2007) are used for testing.

The third data used is the monthly New York City birth (NYB) data for the period from January-1946 to December-1959 [18]. NYB data comprises 168 data points. For data analysis, the first 158 data points are used for training and the remaining 10 data points are used for testing.

The fourth data pertains to the annual immigration into the United States (AIUS) from 1820 to 1962, with 143 data points [18]. The first 135 data points are used for training and the remaining 8 data points are used for testing.

The descriptive statistics of the four data sets such as the minimum, maximum, mean and standard deviation are shown in Table 1.

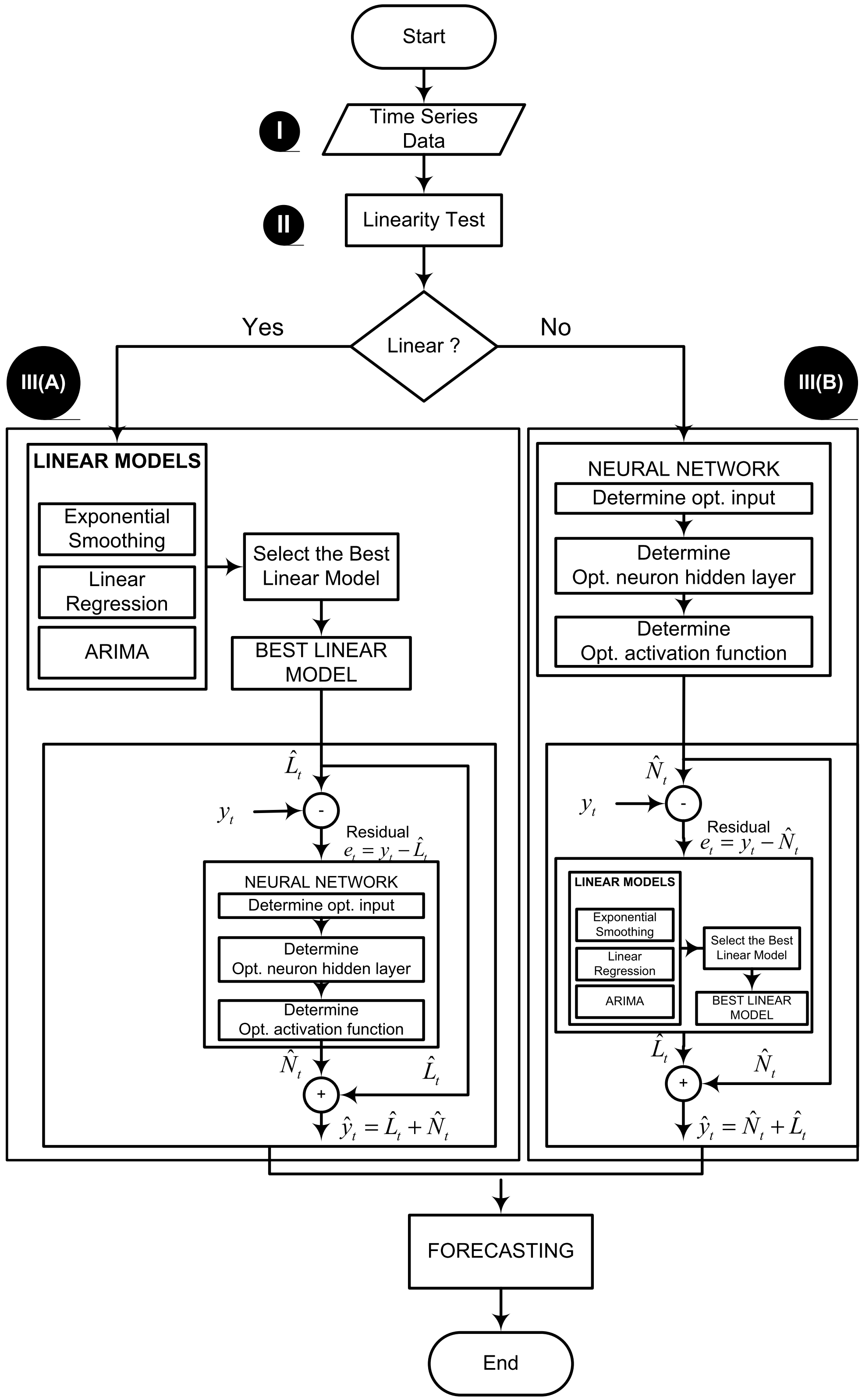


Fig. 1 Flowchart of an Enhanced Hybrid Method

Table 1 Descriptive statistics of the four data sets

Name	Min	Max	Mean	Std. Dev.
MTB-I	244.2	442.8	336.1	64.5
MTB-Z	297.0	652.0	535.9	100.0
NYB	20.0	30.0	25.06	2.32
AIUS	6.354	878.587	253.669	218.192

5 Experimental results

In this work, all of the model such as exponential smoothing (single and double), linear regression, ARIMA, neural network and proposed hybrid models are employed to predict the time series data. Three performance measures are used in determining prediction efficiency, namely root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). These measures have been used by many researchers to compare the accuracy of their models with other known models [8, 19, 20].

The first performance measure is root mean square error (RMSE), which is used to compare to predict value with actual value. The RMSE is computed as:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{n}} \tag{9}$$

The second performance measure is mean absolute error (MAE). The MAE is defined as:

$$MAE = \frac{\sum_{t=1}^{n}|Y_t - \hat{Y}_t|}{n} \tag{10}$$

And then, the third performance measure is mean absolute percentage error (MAPE), a measure of relative overall fitness. This performance measure is defined as:

$$MAPE = \frac{\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t} \cdot 100}{n} \tag{11}$$

where $\hat{Y}_t$ is the predict value, Y_t is the actual value and n is the number of observations.

5.1 Results obtained using data with linear pattern

In this section, we test the performance of the proposed model based on the four set's data. For linearity test, we apply Ramsey RESET test and it yields a positive linear result for MTB-I and NYB data. Hence the linear-nonlinear combination (left hand side of the flowchart shown in Fig. 1) is selected. The steps involved in computing the prediction result ($\hat{y}_t$) for MTB-I and NYB data are explained below. To determine the best linear model, the performance measures (MAE, RMSE and MAPE) are obtained for different linear models.

Table 2 Performance measures using Exponential Smoothing (ES) for MTB-I data

Models		Performance measures MTB- I data		
		MAE	RMSE	MAPE
Single ES (α)	0.20	39.27	42.49	12.84
	0.40	24.31	25.70	7.62
	0.60	17.65	18.86	5.42
	0.80	14.56	15.42	4.40
	1.00	12.59	13.54	3.75
	1.30	11.67	13.39	3.22
	1.50	**10.00**	**12.56**	**2.90**
	1.60	10.15	13.06	2.90
Double ES		**3.27**	**6.09**	**1.14**

Table 3 Performance measures using Linear Regression

Performance measures	Values
MAE	3.16
MAPE	1.11
RMSE	5.38

5.1.1 Results for MTB-I data

The results of performance measures obtained for MTB-I data using ES models are shown in Table 2. We compute prediction using single ES models with different weights of smoothing (α) and Double ES.

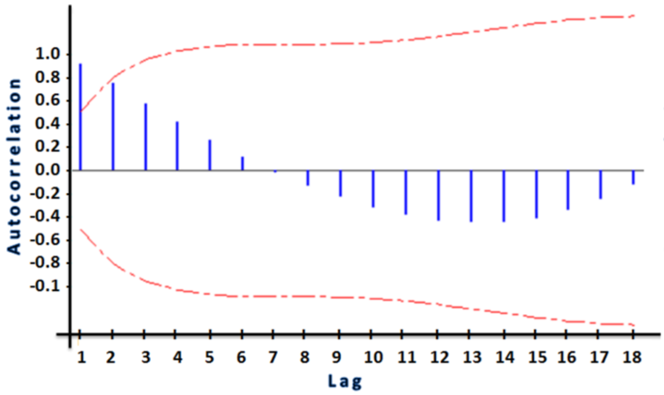
From Table 2, it is seen that the performance measures (MAE, RMSE and MAPE) of the Double ES model are lower than those of other models.

The values of the performance measures obtained with LR model for MTB-I data are shown in Table 3.

Autoregressive integrated moving average(ARIMA) models assume that the data is stationary. If data is not stationary, it is made stationary by performing differencing. We compute autocorrelation of the MTB-I data to check whether the data is stationary.

From Fig. 2(a), it is seen that there is a blue bar that goes beyond the red line, indicating that differencing process needs to be done. Autocorrelation function for MTB-I data after performing differencing process one time is shown in Fig. 2(b) and there is no blue bar that goes beyond the red line (stationary data).

Subsequently, we compute prediction using ARIMA models with different parameter (p, d, q) values with d equal to 1. Table 4 shows that the performance measures (MAE, RMSE and MAPE) of the ARIMA (2, 1, 2) are the lowest for MTB-I data.



(a) Preliminary data

(b) After differencing one time

Fig. 2 Autocorrelation function for MTB-I data

Table 4 Performance measures using ARIMA models

Models	Performance measures MTB-I data		
	MAE	RMSE	MAPE
ARIMA(1,1,0)	3.03	4.86	1.04
ARIMA(0,1,1)	2.77	4.48	0.94
ARIMA(0,1,2)	54.58	64.51	17.03
ARIMA(1,1,2)	2.89	4.67	0.97
ARIMA(2,1,1)	2.57	4.51	0.90
ARIMA(2,1,0)	2.89	4.77	0.99
ARIMA(2,1,2)	**2.55**	**4.43**	**0.90**
ARIMA(2,1,3)	2.59	4.68	0.92
ARIMA(3,1,1)	2.91	4.67	1.00

Table 5 Performance measures using Neural Network models for residual of MTB-I data

Models	Activation function	Performance measures		
		MAE	RMSE	MAPE
NN(3,10,1)	Hyperbolic tangent	0.8548	1.1025	55.1098
NN(4,10,1)	Sigmoid	3.1471	4.2886	101.5324
NN(5,10,1)	Bipolar sigmoid	3.1327	3.6383	122.1917
NN(5,10,1)	Sigmoid	3.2108	4.2139	98.0415
NN(5,10,1)	**Hyperbolic tangent**	**0.3360**	**0.4233**	**52.2965**
NN(6,10,1)	Bipolar sigmoid	1.1555	1.6783	109.0282
NN(6,10,1)	Sigmoid	3.4570	4.4844	116.2814
NN(6,10,1)	Hyperbolic tangent	0.3946	0.5341	72.7637
NN(5,11,1)	Hyperbolic tangent	0.6320	0.9229	68.4932

From Tables 2, 3 and 4, it is found that the best linear model for MTB-I data is ARIMA(2,1,2). Using this model, the prediction estimate $\hat{L}_t$ and the residuals $e_t = y_t - \hat{L}_t$ are obtained. To obtain the predicted value of e_t (i.e. $\hat{N}_t$), the residual series e_t is applied to a neural network model.

To determine the optimum configuration for the MLP network, different values of input and hidden layer neurons are tested. Similarly, different activation functions such as hyperbolic tangent, bipolar sigmoid and sigmoid functions are also tested.

The performance results obtained for the residual series e_t of MTB-I data is shown in Table 5. From Table 5, it is found that the optimum neural network model has 5 input neurons, 10 hidden layer neurons and one output neuron (in abbreviated form, NN(5,10,1)) and uses hyperbolic tangent activation functions for the hidden and output layers yielding the minimum values of MAE, MAPE and RMSE. The optimum neural network configuration is shown in Fig. 3. The proposed hybrid model combines the best linear model, namely ARIMA(2,1,2) and neural network, namely NN(5,10,1). The performance measures using proposed hybrid model for MTB-I data are shown in Table 6.

Table 6 Performance measures obtained using the LNH model for MTB-I data

Performance measures	Values
MAE	0.34
MAPE	0.11
RMSE	0.18

To determine the consistency of the proposed model, different training and testing windows for MTB-I data are used. The MTB-I data set for the years (1990–2004) are employed for training and the remaining data (2005–2007) are used for testing. In this case also, the linearity test, namely, Ramsey RESET test yields a positive linear result as obtained for the window employed earlier, thus confirming the consistency of the model.

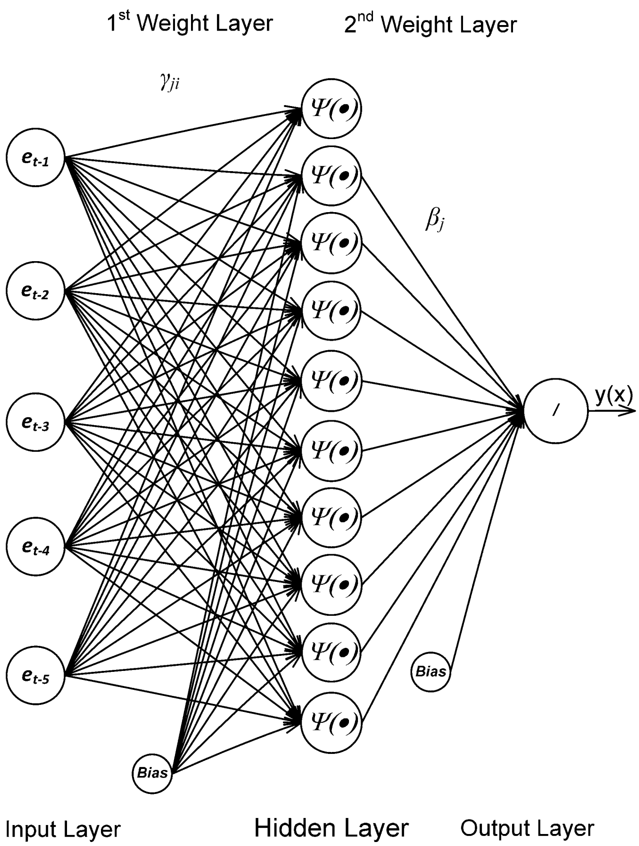


Fig. 3 The architecture configuration of neural network for residual of MTB-I data

Table 7 Performance measures using linear models for NYB data

Models	Performance measures		
	MAE	RMSE	MAPE
Linear regression	1.16	1.45	4.73
Single ES	1.01	1.27	4.08
Doubble ES	1.07	1.39	4.33
ARIMA (4,1,3)	**0.86**	**1.11**	**3.49**

5.1.2 *Results for NYB data*

The proposed hybrid model is also applied for NYB data. The comparison of performance measures using linear models for NYB data are shown in Table 7. From Table 7, the minimal values of performance measures is ARIMA(4,1,3) model. The model is used for time series prediction ($\hat{L}_t$) and the residuals $e_t = y_t - \hat{L}_t$ are obtained. To obtain the predicted value of e_t (i.e. $\hat{N}_t$), the residual series e_t are used as input and applied to a neural network model. The optimum configuration using neural network model has 30 input neurons, 15 hidden layer neurons and one output neuron (NN(30,15,1)) and uses hyperbolic tangent activation functions for the hidden. The performance measures using proposed hybrid model combining ARIMA(4,1,3) model and NN(30,15,1) model for the NYB data are given in Table 8.

Table 8 Performance measures obtained using the LNH model (ARIMA(4,1,3) and NN(30,15,1)) for NYB data

Performance measures	Values
MAE	0.259
MAPE	1.037
RMSE	0.314

Table 9 Performance measures using neural network models for MTB-Z data

Models	Activation function	Performance measures		
		MAE	RMSE	MAPE
NN(5,10,1)	Hyperbolic tangent	8.58	10.55	1.65
NN(6,10,1)	Hyperbolic tangent	8.32	10.34	1.50
NN(7,10,1)	**Hyperbolic tangent**	**6.23**	**8.16**	**1.03**
NN(8,10,1)	Hyperbolic tangent	9.30	11.23	1.54
NN(7,10,1)	Bipolar sigmoid	11.04	12.60	1.83
NN(7,10,1)	Sigmoid	29.77	39.37	5.18
NN(7,9,1)	Hyperbolic tangent	7.11	8.87	1.17
NN(7,11,1)	Hyperbolic tangent	6.76	8.56	1.12

5.2 Results obtained using data with nonlinear pattern

For MTB-Z and AIUS data, since the linearity test yields a negative result, we use nonlinear-linear combination in the hybrid model (right hand side of the flowchart shown in Fig. 1). Based on Fig. 1, the data is first applied to the neural network model to predict $\hat{N}_t$.

5.2.1 *Results for MTB-Z data*

To obtain the optimum configuration of the neural network model, different values of input and hidden layer neurons as well as different activation functions are tested. The results are shown in Table 9. The optimum neural network configuration for MTB-Z is obtained as NN(7,10,1). Using this model, we get the predicted values $\hat{N}_t$ and then we compute the residual values $e_t = y_t - \hat{N}_t$.

The residual values e_t are then applied to the best linear model to get time series prediction ($\hat{L}_t$). From the experimental results, it is found that the best linear model for residual of MTB-Z data is ARIMA(3,1,1).

The performance measures achieved for the MTB-Z data based on proposed hybrid model combining NN(7,10,1) model and ARIMA(3,1,1) model are given in Table 10.

Table 10 Performance measures using NLH model for MTB-Z data

Performance measures	Values
MAE	3.7197
MAPE	0.6179
RMSE	4.8561

Table 11 Performance measures using NLH (NN(20,15,1) + ARIMA(6,1,14)) model for AIUS data

Performance measures	Values
MAE	29.609
MAPE	20.074
RMSE	40.107

5.2.2 Results for AIUS data

For AIUS data, it is found that the optimum configuration of the neural network model is NN(20,15,1). This model determines the predicted values $\hat{N}_t$ first and then computes the residual values $e_t = y_t - \hat{N}_t$. The error or residual values e_t are then applied to the best linear model, namely ARIMA(6,1,14) to get time series prediction $(\hat{L}_t)$. The proposed hybrid model combining NN(20,15,1) model and ARIMA(6,1,4) model is applied to get time series prediction for the AIUS data. The performance measures using the proposed hybrid model for the AIUS data are shown in Table 11.

6 Comparison of models

The proposed hybrid model is compared with known linear and neural network models based on the prediction performances for the linear (MTB-I and NYB) and nonlinear (MTB-Z and AIUS) data. As shown before, for MTB-I data, the best linear model is ARIMA(2,1,2) and the optimum neural network configuration is NN(5,10,1). Table 12 shows a comparison of MAE, MAPE and RMSE values obtained using ARIMA(2,1,2), NN(5,10,1) and the proposed hybrid LNH and NLH models. We note from Table 12 that the hybrid model LNH combining ARIMA(2,1,2) and NN(5,10,1) gives the best results compared to all other models for MTB-I data which has a linear pattern.

For NYB data (linear pattern), the best linear model is ARIMA(4,1,3) and the optimum neural network configuration is NN(30,15,1). Hence, the hybrid model LNH combines ARIMA(4,1,3) model with NN(30,15,1) model. The comparison of MAE, MAPE and RMSE values obtained using individual models and the proposed hybrid model is shown in Table 13.

Table 12 Comparison of performance measures using individual and hybrid model for MTB-I data

Models	Performance measures		
	MAE	RMSE	MAPE
ARIMA(2,1,2) (BLM)	3.383	4.633	1.384
NN(5,10,1)	10.246	10.748	4.342
LNH(ARIMA(2,1,2)+NN(5,10,1))	**2.234**	**2.765**	**0.912**
NLH(NN(5,10,1)+ARIMA(2,1,2))	8.630	9.105	3.504

Table 13 Comparison of performance measures using individual and hybrid model for NYB data

Models	Performance measures		
	MAE	RMSE	MAPE
ARIMA(4,1,3) (BLM)	1.398	1.581	4.945
NN(30,15,1)	0.967	1.251	3.589
LNH (ARIMA(4,1,3)+NN(30,15,1))	**0.965**	**1.227**	**3.427**
NLH (NN(30,15,1)+ARIMA(4,1,3))	1.020	1.502	3.836

Table 14 Improvement achieved by proposed LNH model over the other models for MTB-I data

Models	MAE (%)	RMSE (%)	MAPE (%)
ARIMA(2,1,2) (BLM)	33.95	40.31	34.13
NN(5,10,1)	78.19	74.27	79.00
NLH	74.11	69.63	73.98

From Table 13, the hybrid model combining ARIMA(4, 1,3) and NN(30,15,1) gives the best results for NYB data which also has a linear pattern.

The improvements achieved with respect to MAE, RMSE, and MAPE values by the LNH model over other models for MTB-I data are shown in Table 14. Thus, it is shown that the proposed hybrid model LNH is able to achieve significant performance improvement over the other models for data with linear pattern.

Next, we make a comparison of the different models based on MTB-Z data which has a nonlinear pattern. For this data, the optimum neural network configuration is NN(7,10,1) and the best linear model is found to be ARIMA(3,1,1). The hybrid model NLH combines NN(7,10,1) model with ARIMA(3,1,1) model. The performance measures of the different models are shown in Table 15. From this table, it is seen that the best results are achieved by the NLH model for nonlinear pattern.

For AIUS data which has a nonlinear pattern, the hybrid model NLH combines NN(20,15,1) model with ARIMA(6, 1,4) model. The comparison of performance measures using individual and hybrid models for AIUS data is given Table 16.

Table 15 Comparison of performance measures using individual and hybrid model for MTB-Z

Models	Performance measures		
	MAE	RMSE	MAPE
ARIMA(3,1,1) (BLM)	39.748	41.082	7.638
NN(7,10,1)	28.474	33.977	5.554
NLH(NN(7,10,1)+ARIMA(3,1,1))	**17.533**	**19.840**	**3.404**
LNH (ARIMA(3,1,1)+NN(7,10,1))	70.967	71.571	13.568

Table 16 Comparison of performance measures using individual and hybrid model for AIUS data

Models	Performance measures		
	MAE	RMSE	MAPE
ARIMA(6,1,14) (BLM)	50.184	58.997	17.585
NN(20,15,1)	46.943	54.746	17.419
NLH (NN(20,15,1)+ARIMA(6,1,14))	**39.863**	**51.822**	**13.820**
LNH (ARIMA(6,1,14)+NN(20,15,1))	73.918	82.710	26.141

Table 17 Improvement achieved by proposed NLH model over the other models for MTB-Z data

Models	MAE (%)	RMSE (%)	MAPE (%)
ARIMA(3,1,1) (BLM)	55.89	51.71	55.43
NN(7,10,1)	38.42	41.61	38.71
LNH	75.29	72.28	74.91

From Table 16, it is seen that NLH model yields the minimum values of MAE, MAPE and RMSE for nonlinear pattern.

Table 17 shows the improvement achieved by the NLH over other models for MTB-Z data. From Table 17, it is clear that for data with nonlinear pattern, the proposed NLH model achieves significant performance improvement over other models.

7 Conclusion

This paper has discussed an enhancement method of hybrid models for the prediction of time series data. It is shown that applying appropriate hybrid combination (linear-nonlinear or nonlinear-linear) is important depending on the pattern of data. The proposed hybrid method will first test the linearity of the time series data to determine which hybrid combination, namely, linear-nonlinear or nonlinear-linear should be used for prediction. Four different types of data have been used to test the proposed hybrid method. From the experimental results, it has been found that the hybrid model (LNH) comprising linear-nonlinear combination out performs other models for data that have a linear pattern. On the other hand, for data with nonlinear pattern, the hybrid model (NLH) comprising nonlinear-linear combination performs better than the other models. In both the cases, a significant performance improvement has been achieved by the proposed models compared to other known linear and nonlinear models.

References

1. Javier C, Rosario E, Francisco JN, Conejo AJ (2003) ARIMA models to predict next-day electricity prices. IEEE Trans Power Syst 18:1014–1020
2. Katimon A, Demun AS (2004) Water use trend at universiti teknologi Malaysia: application of ARIMA model. J Pendidik Univ Teknol Malays 41:47–56
3. Moghaddas-Tafreshi SM, Farhadi M (2008) Linear regression-based study for temperature sensitivity analysis of Iran electrical load. In: IEEE international conference on industrial technology, pp 1–7
4. Li BJ, Hua HC (2007) The combined forecasting method of GM(1,1) with linear regression and its application. In: International conference on grey systems and intelligent services, pp 394–398
5. Sitte R (2002) Neural networks approach to the random walk dilemma of financial time series. Appl Intell 16:163–171
6. Neukukar V, Hamidi-Beheshti MT (2010) A local linear radial basis function neural network for financial time series forecasting. Appl Intell 33:352–356
7. Harri N, Teri H, Ari K, Juhani R, Kolehmaine M (2004) Evolving the neural network model for forecasting air pollution time series. Eng Appl Artif Intell 17:159–167
8. Faruk DO (2010) A hybrid neural network and ARIMA model for water quality time series prediction. Eng Appl Artif Intell 23:586–594
9. Santoso S (2009) Business forecasting: metode peramalan bisnis masa kini dengan minitab dan SPSS. Elex Media Komputindo, Jakarta
10. Le CT (2003) Introductory biostatistics. Wiley, New York
11. Peter JB, Davis RA (2002) Introduction to time series and forecasting. Springer, New York
12. Suhartono (2008) Feedforward neural networks untuk pemodelan runtun waktu. Dissertation, Gajah Mada University, Yogyakarta, Indonesia
13. Denton JW (1995) How good are neural networks for causal forecasting? J Bus Forecast 14:17–20
14. Zhang GP (2003) Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing 50:159–175
15. Roselina S, Siti MS, Hashim SZM (2008) Hybridization model of linear and nonlinear time series data for forecasting. In: Second Asia international conference on modelling & simulation, pp 597–602
16. Kim TH, Lee YS, Newbold P (2004) Spurious nonlinear regressions in econometrics. School of Economics, University of Nottingham
17. World Health Organization (2009) Global tuberculosis control: epidemiology, strategy, financing: WHO report 2009. WHO Press, Geneva
18. Hyndman RJ (n.d.) (2011) Time series data library. http://robjhyndman.com/TSDL. Accessed on October 10, 2011
19. Rojas I, Valenzuela O, Rojas F, Guillen A, Herrera LJ, Pomares H et al (2008) Soft-computing techniques and ARMA model for time series prediction. Neurocomputing 71:519–537
20. Lee YS, Tong LI (2011) Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. Knowl-Based Syst 24:66–72