Pak.j.stat.oper.res. Vol.19 No. 3 2023 pp 551-568 DOI: http://dx.doi.org/10.18187/pjsor.v19i3.4247

Pakistan Journal of Statistics and Operation Research

A Modified Weighting System for Combined Forecasting Methods Based on The Correlation Coefficients of the Individual Forecasting Models

Chantha Wongoutong[1*]

* Corresponding Author

Department of Statistics, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand.
fscictw@ku.ac.th

Abstract

Herein, a modified weighting for combined forecasting methods is established. These weights are used to adjust the correlation coefficient between the actual and predicted values from five individual forecasting models based on their correlation coefficient values and ranking. Time-series datasets with three patterns (stationary, trend, or both trend and seasonal) were analyzed by using the five individual forecasting models and three combined forecasting methods: simple-average, Bates-Granger, and the proposed approach. The MAPE and RMSE results indicate that the proposed method outperformed the others, especially when the time-series pattern was stationary and improved the forecasting accuracy of the worst and best individual forecasting models by %37–35and %10–7, respectively. Moreover, the proposed method showed improvements in MAPE and RMSE of around %20–18and %11–9compared to the simple-average and Bates-Granger methods, respectively. In addition, the combined forecasting methods outperformed the individual forecasting models when analyzing non-stationary data. Remarkably, the performances of the proposed and Bates-Granger methods were almost the same, with improvements in MAPE and RMSE in the range of %2–1on average. Therefore, the proposed method for creating weights based on the correlation coefficients of the individual forecasting models greatly improves combined forecasting methods.

Key Words: Forecasting model, Combined forecasting method, Correlation coefficient, Model weighting

1. Introduction

In the past decades, many statisticians have risen to the challenge of improving forecasting accuracy, and yet all these research efforts have merely led to the conclusion that no single forecasting method outperforms all of the others in all situations (Li et al. 2005). One approach for improving forecast accuracy is combining forecasts from two or more different forecasting models starting with the benchmark work of Bates and Granger (1969), who introduced combinations of forecasts as a ubiquitous way of improving forecasting accuracy. Moreover, combining forecasts is a very useful approach when selecting the most accurate forecasting method is difficult. For instance, Bunn (1989) noted that this improves forecasting accuracy by utilizing multiple sources of information and computing resources and defined this approach as "data-intensive forecasting". Furthermore, many authors have indicated that combining linear forecasts is generally more accurate than individual forecasts (Makridakis and Hibon 2000; Stock and Watson 2004; Patton and Sheppard 2009; Costantini and Pappalardo 2010; Martins and Werner 2012; Thaithanan and Wongoutong 2020). According to Clemen (1989), many techniques have been developed to perform combining forecasts. Nevertheless, until now, the results have been unanimous: combining forecasts leads to increased accuracy. Likewise, Makridakis and Hibon (2000) conducted the M3-competition that involved forecasting 3003 time-series datasets and concluded that the accuracy of the combined forecasts by using various methods outperforms the individual forecasting models being combined. Therefore, the combination of forecasting methods has become an essential strategy that has been used in many applications (Makridakis et al. 2018).

Combining forecasts is associated with the performance consistency of their individual forecasting models, and combining at least three individual forecasting models results in more accurate forecasting Armstrong (2001). According to Yang (2004), the two main directions for combining forecasts are for adaptation and improvement: the first targets the best individual performance among the pool of forecasting candidates while the aim of the second is to significantly outperform each forecasting candidate. When combining the forecasts generated by two or more individual forecasting models, it is vital to decide the weight assigned to each of them. The most popular among the combined forecasting methods is the weighted linear combination where the weights assigned to the individual forecasting models are either equal or decided according to some rigorous mathematical rule. Typical linear combined forecasting methods are simple-average, trimmed-average, winsorized-average, median, error-based, outperformance, variance-based pooling, etc. (Jose and Winkler 2008; Lemke and Gabrys 2010). The R package ForecastComb provides 15 popular simple, regression-based, and eigenvector-based estimation methods for creating combined forecasts.

Even though linear combination techniques are easy to understand and implement, they ignore the relationship between the actual and forecasted values. Several researchers have suggested changing each individual forecasting model's weight in the combination while checking for non-stability in the process (Deutsch et al. 1994; Chan et al. 2004; Timmermann 2006). Assigning different weights using linear correlation in forecasting combinations is a possible alternative (Martins and Werner 2012). Many researchers have worked on combining forecasting models by using correlation. For example, Diebold (1988) considered serial correlation in a least-squares framework by restricting the sum of the coefficients to 1. Likewise, Coulson and Robins (1993) included a lagged dependent variable beyond the forecasting candidates by focusing on the specific case of combining two forecasts with the combination error following a first-order autoregressive (AR(1)) process. They concluded that a parsimonious method for incorporating the dynamics is achieved by using a lagged dependent variable but not lagged forecasts. Moreover, Deutsch et al. (1994) created regime switches by using coefficients to weight the models.

However, only limited studies have endeavored to clarify the perspective of using correlation in forecasting combinations, so there is a strong need for further developments in this area. To realize the current study, a literature review was systematically performed with the goal being to list the methods for combining predictions and identify existing methods using linear correlation coefficients in their structures. Hence, the focus of this study is on developing weights by using linear combination techniques on five individual forecasting models (average, decomposition, Box-Jenkins, artificial neural network, and support vector machine). In the proposed combined forecasting method, the weighted linear combination of the individual forecasts is achieved by using a function of the correlation coefficient and the rank order of the correlation coefficients between the forecasted and actual values as weights. Finally, the performance of the proposed method was compared with the five individual forecasting models and the traditional combined simple-average and Bates-Granger methods.

The rest of this paper is organized as follows. In Section 2, the data used in the study are presented, while Section 3 provides the individual forecasting models, combination forecasting methods, and the proposed method used in the study. The experimental study to compare the methods is described in Section 4. The results of the experimental study and a discussion are given in Section 5. Finally, conclusions on the study are included in Section 6.

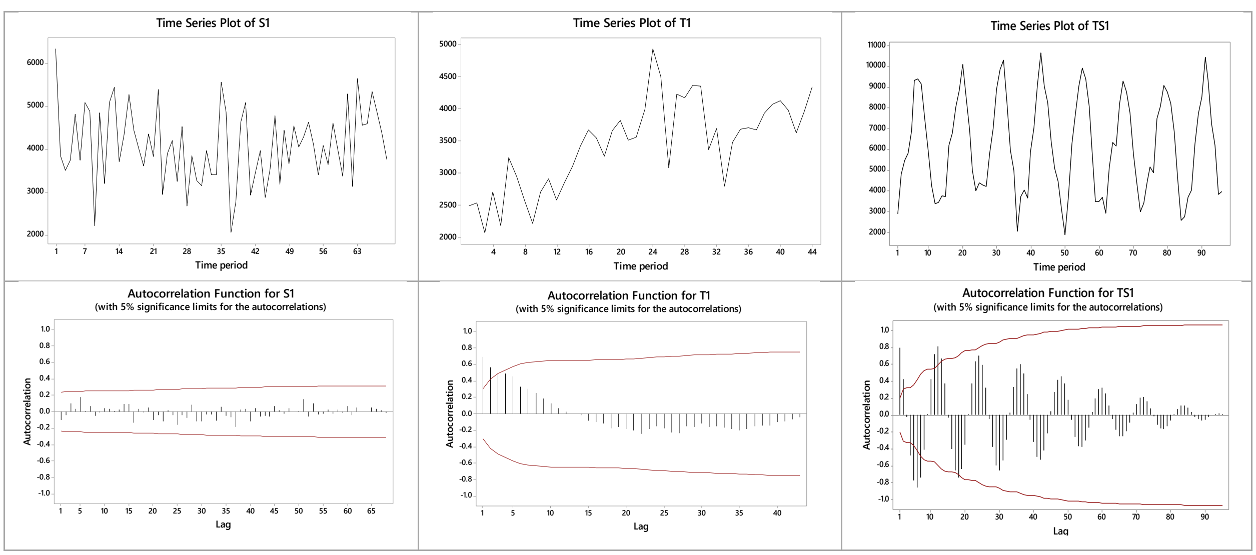
2. The data used in the study

Thirty real time-series datasets were used in this study to determine the performance of the proposed method. This consisted of 10 datasets for each of three types of time-series patterns: stationary, trend, and trend and seasonality. All of the datasets were obtained from the M3 competition conducted by Makridakis and Hibon (2000); they are freely available and any researcher can use the data without requiring permission. Brief details of these datasets are reported in Table 1, and time-series plots and autocorrelation plots for some of them are presented in Figure 1.

Table 1. Details of the datasets used in the study.

Type	Data set	M3 Competition Code	Time Period	Size	Datas et	M3 Competition Code	Time Period	Size
Stationary	S1	N1449 (Q)	1990Q1-2007Q1	69	S6	N234 (Y)	1950-1993	44
	S2	N2126 (M)	1978M1-1998M12	144	S7	N235 (Y)	1947-1993	47
	S3	N1453 (Q)	1990Q1-2007Q1	69	S8	N1426 (Q)	1990Q1-2007Q1	69
	S4	N1442 (Q)	1990Q1-2007Q1	69	S9	N1447 (Q)	1990Q1-2007Q1	69
	S5	N1472 (Q)	1990Q1-2007Q1	69	S10	N1987 (M)	1979M1-1990M12	144
Trend	T1	N239 (Y)	1950-1993	44	T6	N1944 (M)	1982M1-1993M12	144
	T2	N1882 (M)	1981M12-1998M11	144	T7	N1946 (M)	1982M1-1993M12	144
	T3	N972 (Q)	1980Q1-1992Q4	52	T8	N1232 (Q)	1980Q1-1993Q1	53
	T4	N198 (Y)	1947-1993	47	T9	N2014 (M)	1979M1-1990M12	144
	T5	N993 Q)	1980Q1-1992Q4	52	T10	N210 (Y)	1947-1987	41
Trend and Seasonality	TS1	N2789 (Q)	1890Q1-2003Q4	96	TS6	N756 (Q)	1984Q1-1994Q4	44
	TS2	N2784 (Q)	1857Q1-1880Q4	96	TS7	N1890 (M)	1982M1-1993M12	144
	TS3	N829 (Q)	1984Q1-1994Q4	44	TS8	N2012 (M)	1979M1-1990M12	144
	TS4	N863 (Q)	1977Q1-1992Q4	64	TS9	N2013 (M)	1979M1-1990M12	144
	TS5	N931 (Q)	1966Q1-1977Q1	45	TS10	N2015 (M)	1979M1-1990M12	144

Note. The letter in parentheses in the M3 competition code indicates quarterly (Q), monthly (M), or yearly (Y) data collection.



(a) (b) (c)

Figure 1. Time-series plots (above) and the autocorrelation plots (below) of datasets (a) S1 (stationary), (b) T1 (trend), and (c) TS1 (trend and seasonality).

3. Forecasting models

Finding a model that represents reality and predicts efficaciously is the main objective of forecasters, and several forecasting methods have been developed using various approaches to achieve this. Assembling several individual forecasting models to improve prediction is known as combined forecasting, which has been shown to outperform forecasting by using the individual approaches in the majority of cases (De Menezes et al. 2000). Armstrong (2001) claimed that the number of individual forecasting models to optimize combinatorial efficacy is five. Therefore, in this study, three combined forecasting methods (including the proposed method) were created by using five individual time-series forecasting models. The individual forecasting models and combined forecasting methods briefly described in this section are available in the R statistical package.

3.1 Individual forecasting models

3.1.1 Averaging models

1) Simple moving average

The simplest model, the simple moving average, is effective when the time-series data are assumed to be stable over time with no trend or seasonality (Svetunkov and Petropoulos 2018). Thus, the simple moving average model can be used to forecast the next value(s) in a time series depending on the average over specified *k* periods of the previous values for which each point is assigned an equal weight (*1/k*). The formula for this is

$$\hat{y}_t = \frac{1}{k}\sum_{i=1}^{k} y_{t-i}, \tag{1}$$

where y_t is an actual value, $\hat{y}_t$ is the forecast for period *t*, and *k* is the length of the simple moving average.

2) Double moving average

This approach is better for when there is a trend in a time-series dataset. A trend in the data means that the observation values tend to either increase or decrease over time. The double moving average model requires calculating the moving average and then calculating the second moving average using the first moving average values as observations (Khairina et al. 2021). The formula used for forecasting in the period *t*+*m* of the double moving average method is

$$\hat{y}_{t+m} = a_t + b_t m.. \tag{2}$$

The term $a_t = 2M_t - M'_t$ is the interception in period *t*, where $M_t = \frac{1}{k}\sum_{i=1}^{k} y_{t-i}$ is the first average value in period *t*, *k* is the order, and $M'_t = \frac{1}{k}\sum_{i=1}^{k} M_{t-i}$ is the second average value in period *t*. The term $b_t = \frac{2}{k-1}(M_t - M'_t)$ denotes the trend value in period *t*.

3.1.2 Exponential smoothing models

Exponential smoothing is a class of time-series forecasting methods for univariate data in which current values are given relatively more weight in forecasting than older observations. Exponentially decreasing weights are explicitly used in exponential smoothing models (Brown 1956; and Holt 2004).

1) Single exponential smoothing

The simplest approach is single exponential smoothing for data that is stable over time without seasonality or trends. The forecasting method in period *t* is formulated as

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t, \tag{3}$$

where, $\hat{y}_{t+1}$ is the forecast in period *t+1*, $\hat{y}_t$ is the forecast in period *t*, y_t is the actual value in period *t*, and α is a smoothing parameter between 0 and 1.

2) Double exponential smoothing

This is similar to the single exponential smoothing method except that additional weighting is used to detect trends in the data (Shastri et al. 2018). The forecasting method in period *t+p* is derived as follows:

$$\hat{y}_{t+p} = L_t + pT_t, \tag{4}$$

where smoothing parameters α and β are between 0 and 1. Meanwhile, level smoothing factor trend smoothing factor and *t*represents the estimated forecast value at time $L_t = \alpha Y_t + \left(1-\alpha\right)\left(L_{t-1} + T_{t-1}\right)$ *t*.represents the value of the slope at time $T_t = \beta\left(L_t - L_{t-1}\right) + \left(1-\beta\right)T_{t-1}$

3) Triple exponential smoothing (Holt-Winters)

This method can handle a univariate time series with seasonality by simply finding the central value and adding the effects of slope and seasonality (Brown 1956; Holt 2004; Winters 1960). There are two variations of the Holt-Winters method depending on the nature of the seasonal component: additive or multiplicative (Montgomery et al. 2008; Wongoutong 2021). Applying the additive Holt-Winters model when forecasting for period *t+p* can be obtained as follows:

$$\hat{y}_{t+p} = (L_t + T_t p) + S_{t-S+p}, \tag{5}$$

where level smoothing factor $L_t = \alpha(y_t - S_{t-S}) + (1-\alpha)(L_{t-1} + T_{t-1}),$ trend smoothing factor $T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1},$ and seasonality smoothing factor $S_t = \gamma(y_t - L_t) + (1-\gamma)S_{t-S}$.

The multiplicative Holt-Winters model for forecasting in period *t+p* can be obtained as follows:

$$\hat{y}_{t+p} = (L_t + T_t p)S_{t-S+p}, \tag{6}$$

where level smoothing factor $L_t = \alpha \dfrac{y_t}{S_{t-S}} + (1-\alpha)(L_{t-1} + T_{t-1}),$ trend smoothing factor $T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1},$ and seasonality smoothing factor $S_t = \gamma \dfrac{y_t}{L_t} + (1-\gamma)S_{t-L}$. In both models, *S* is the seasonality length , y_t refers to the actual value in time period *t*, and the values of smoothing parameters $\alpha, \beta,$ and γ are set to between 0 and 1.

3.1.3 Decomposition model

The decomposition model is a popular forecasting method for time-series data containing trend and seasonal patterns (Zhang et al. 2014). The main idea of this method is to analyze the four possible individual components in a time series: trend, cycle, seasonality, and irregular (Montgomery et al. 2015). In the decomposition model, each component's strength is estimated separately and then substituted into the model to explain the behavior of the time series in a straightforward manner as $y_t = f(T_t, S_t, C_t, I_t),$ where Y_t, T_t, S_t, C_t, and I_t are the time-series values (actual data) and the trend, seasonal, cycle, and irregular components for time period *t*, respectively. In general, the decomposition model depends on whether the nature of the seasonal component is additive or multiplicative. The respective mathematical expressions for the additive and multiplicative decomposition approaches are

$$y_t = T_t + S_t + C_t + I_t \text{ and } y_t = T_t \times S_t \times C_t \times I_t. \tag{7}$$

3.1.4 The Box-Jenkins model

The strategy used in the Box-Jenkins model is to predict a time series by using past values. This method begins with the assumption that the process that generated the time series can be approximated by using an autoregressive moving average (ARMA) model if it is stationary or an autoregressive integrated moving average (ARIMA) model if it is non-stationary (Box et al. 2015). The general ARIMA model is expressed as follows:

$$\phi_p(\mathrm{B})\Phi_P(\mathrm{B}^s)(1-\mathrm{B})^d(1-\mathrm{B}^s)^D y_t = \theta_q(\mathrm{B})\Theta_Q(\mathrm{B}^s)\varepsilon_t, \tag{8}$$

where d is the order of differencing, D is the order of seasonal differencing, and s is the number of seasons per year. The operator polynomials are

$$\phi_p(\mathrm{B}) = (1-\phi_1\mathrm{B}-\ldots-\phi_p\mathrm{B}^p),\ \theta_q(\mathrm{B}) = (1-\theta_1\mathrm{B}-\ldots-\theta_q\mathrm{B}^q),\ \Phi_P(\mathrm{B}^s) = (1-\Phi_1\mathrm{B}^s-\ldots-\Phi_P\mathrm{B}^{Ps}),$$

$$\Theta_Q(\mathrm{B}^s) = (1-\Theta_1\mathrm{B}^s-\ldots-\Theta_Q\mathrm{B}^{Qs}),\ (1-\mathrm{B})y_t = y_t - y_{t-1},\ (1-\mathrm{B}^s)y_t = y_t - y_{t-s}, \text{ and } \varepsilon_t \sim N(0,\sigma^2).$$

The maximum values of d, D, p, q, P, and Q should be 2 (Box et al. 2015), and so these operator polynomials are usually simple expressions.

The Box-Jenkins methodology consists of a four-step iterative procedure as follows:

1. **Model identification:** tentatively identify the appropriate Box-Jenkins model by using historical data by analyzing plots of the autocorrelation function and partial autocorrelation function, and then determine the appropriate model type for a specific situation by matching the observed correlations to the theoretical correlations.
2. **Parameter estimation:** estimate their values based on the maximum likelihood or minimum least-squares methods.
3. **Diagnostic checking:** use plots and statistical tests of the residual errors to determine the adequacy of the model fitting, and if need be, consider alternative models.
4. **Forecasting:** use the appropriate model for forecasting.

3.1.5 The artificial neural network (ANN) model

In recent years, time-series models using machine learning based on ANN have become vital alternatives due to their nonlinear modeling capability for data time-series forecasting (Lin and Lee 2013). The most popular ANN architecture in the forecasting domain are the multilayer perceptron (MLP), a class of feedforward ANN, which is a nonlinear autoregressive model. The *nnetar* function from the R package (2015) was used for fitting the ANN model for the time-series data. This function creates feedforward neural networks with a single hidden layer using lagged inputs for forecasting a univariate time series. A single layer of hidden units is enough to provide the desired accuracy in most forecasting situations. The nodes are connected to those in the immediate next layer in each layer by acyclic links (Zhang et al. 1998). The structure of a typical ANN with MLP architecture is shown in Figure 2, including a feedforward structure of an input layer, one or more hidden layers, and an output layer.

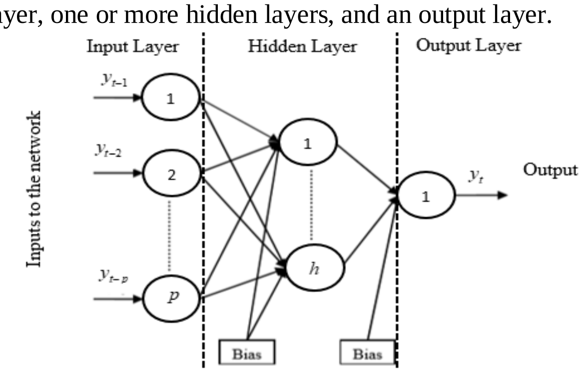


Figure 2. A typical ANN structure.

In a fully connected ANN model with *p* input, *h* hidden, and a single output node, the relationship between the inputs y_{t-i} (i=1, 2, …, p) and the output y_t is given by

$$y_t = G\left(\alpha_0 + \sum_{j=1}^{h} \alpha_j F\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-1}\right)\right), \tag{9}$$

where α_j and β_{ij} ($i = 1, 2, ..., p;\ j = 1, 2, ..., h$) are the connection weights, α_0 and β_{0j} are the bias terms, F and G are the network activation functions.

3.1.6 The support vector machine (SVM) model

Cortes and Vapnik (1995) first suggested the SVM method, which has been used in many fields such as data mining, classification, regression, and time-series forecasting. Okasha (2014) and Guo et al. (2011) also proposed using SVM to forecast time-series data. The main objective of the SVM model is to deduce specific decision rules with satisfactory generalization ability (support vectors) by choosing some specific subset of training data. In the SVM model, nonlinear mapping of the input space into a higher dimensional feature space is deployed, after which an optimally separating hyperplane is extracted. A set of mathematical functions defined as the kernel are used in SVM algorithms. The kernel function involves taking data as input and transforming them into the required form. Different SVM algorithms use different types of kernel functions: linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid. RBF is the most appropriate type of kernel function for time series (Karatzoglou et al. 2004), and so it was used in the present study.

The SVM regression algorithm can be applied to time-series forecasting by adopting a sliding time window defined by the set of time lags $\{k_1, k_2, ..., k_I\}$ used to build the forecast. For given time period *t*, the model inputs are $\mathbf{y} = \left(y_{t-k_I}, ..., y_{t-k_2}, y_{t-k_1}\right)$ and the desired output is y_t. In SVM regression, the input $\mathbf{y} = \left(y_{t-k_I}, ..., y_{t-k_2}, y_{t-k_1}\right)$ is transformed into a high *m*-dimensional feature space by using nonlinear mapping (ϕ) that depends on the kernel function. Subsequently, the SVM algorithm finds the best linear separating hyperplane in the feature space while tolerating a small error when fitting the data as follows:

$$y_t = w_0 + \sum_{i=1}^{m} w_i \phi_i(\mathbf{y}). \tag{10}$$

Figure 3 shows SVM regression with the ε -insensitive loss function adapted from Karatzoglou et al. (2004).

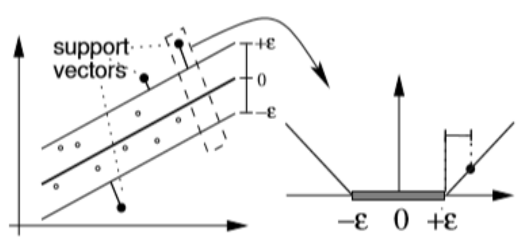


Figure 3. Linear SVM regression and the ε-insensitive loss function.

 DOI: http://dx.doi.org/10.18187/pjsor.v19i3.4247

3.2 Combined forecasting methods

Improving forecasting accuracy by combining forecasts is a well-established procedure (Winkler and Makridakis 1983; Armstrong 2001; Thaithanan and Wongoutong 2020). Armstrong (2001) claimed that combining five individual forecasting models provides optimal efficacy. The combined forecasting method is associated with the performance consistency of each individual forecasting model and assigning combinatory weights. We consider $\mathbf{Y}=[\mathrm{y}_1,\mathrm{y}_2,\ldots,\mathrm{y}_N]^{\mathrm{T}}$ as the actual time series to be forecasted using *n* different individual forecasting models, $\hat{\mathbf{Y}}^{(i)}=[\hat{\mathrm{y}}_1^{(i)},\hat{\mathrm{y}}_2^{(i)},\ldots,\hat{\mathrm{y}}_N^{(i)}]^{\mathrm{T}}$ as its forecast obtained from the ith model $(i=1,2,\ldots,n)$, and $\hat{\mathbf{Y}}^{(c)}=[\hat{\mathrm{y}}_1^{(c)},\hat{\mathrm{y}}_2^{(c)},\ldots,\hat{\mathrm{y}}_N^{(c)}]^{\mathrm{T}}$ as the combined forecasted series of the original time series. In the present study, three combined forecasting methods were used to improve forecasting accuracy: simple-average, Bates-Granger, and the proposed method.

3.2.1 The simple-average method

This method assigns equal weights to all of the individual forecasting models. Although the simple-average method may appear to be a naïve approach for combining forecasts, more complex methods for combining forecasts do not often improve upon its accuracy (Clemen, .(1989It is well-documented that the simple-average is a robust combination method that is difficult to beat (Stock and Watson 2004; Timmermann, .(2006Assigning equal weights ($w_i=1/n$) to each of the individual forecasting models can be written as

$$\hat{y}_k^{(SA)}=w_1\hat{y}_k^{(1)}+w_2\hat{y}_k^{(2)}+\ldots+w_n\hat{y}_k^{(n)}=\sum_{i=1}^{n}w_i\hat{y}_k^{(i)}:w_i=1/n, \tag{11}$$

where $\hat{y}_k^{(i)}(i=1,2,\ldots,n;\ k=1,2,\ldots,N)$ and $\hat{y}_k^{(c)}(i=1,2,\ldots,n;\ k=1,2,\ldots,N)$ denote the individual forecasts and the simple-average combination forecast, respectively.

3.2.2 The Bates-Granger method

Bates and Granger (1969) introduce the idea of combining forecasts in their seminal paper. They use the diagonal elements of the estimated mean-squared prediction error matrix to compute combination weights $w_i=\dfrac{\hat{\sigma}_{(i)}^{-2}}{\sum_{j=1}^{n}\hat{\sigma}_{(\mathrm{j})}^{-2}}$, where $\hat{\sigma}_{(i)}^{-2}$ is the estimated mean-squared prediction error of the *i*$^{\text{th}}$ model. The combined forecast is then obtained as

$$\hat{y}_k^{(BG)}=w_1\hat{y}_k^{(1)}+w_2\hat{y}_k^{(2)}+\ldots+w_n\hat{y}_k^{(n)}=\sum_{i=1}^{n}w_i\hat{y}_k^{(i)}:w_i=\frac{\hat{\sigma}_{(i)}^{-2}}{\sum_{j=1}^{n}\hat{\sigma}_{(\mathrm{j})}^{-2}}, \tag{12}$$

wher $\hat{y}_k^{(i)}(i=1,2,\ldots,n;\ k=1,2,\ldots,N)$ and $\hat{y}_k^{(c)}(i=1,2,\ldots,n;\ k=1,2,\ldots,N)$ denote the individual forecasting models and Bates-Granger forecasting combination method, respectively.

3.2.3 The proposed method

When combining the forecasts produced by two or more models, it is vital to decide the weights assigned to each of them. The weights can be assigned by using linear correlation (Martins and Werner 2012). Therefore, in the present study, the proposed method for combining forecasts comprises five individual forecasting models whereby weights are assigned to each by using the correlation coefficient between the actual and forecasted values followed by ranking them. The steps to perform the weights assigned to each participating model in the proposed method are as follows:

1. Compute the correlation coefficient between the actual and the forecasted values for each model:

$$r_i = \frac{\sum_{k=1}^{N}(y_{ik}-\bar{y})(\hat{y}_{ik}-\bar{\hat{y}})}{\sqrt{\sum_{k=1}^{N}(y_{ik}-\bar{y})^2\sum_{k=1}^{N}(\hat{y}_{ik}-\bar{\hat{y}})^2}};\ i=1,2,3,4,5 \text{ and } k=1,2,3,\dots,N,$$

   where r_i is the correlation coefficient value between the actual and forecasted values for the i^{th} model and y_{ik} and $\hat{y}_{ik}$ refer to the actual and forecasted values for the i^{th} model.

2. Rank the r_i values from smallest to largest and order them as $r_1 < r_2 < r_3 < r_4 < r_5$.

3. Assign a weight to each forecasting model calculated as $w_i = \frac{(r_i)(i)}{\sum_{i=1}^{5}(r_i)(i)}$, where i = 1,2,3,4,5 and $\sum_{i=1}^{5} w_i = 1.$

4. Compute the forecasted values by applying

$$\hat{y}_k^{(proposed)} = w_1\hat{y}_k^{(1)} + w_2\hat{y}_k^{(2)} + w_3\hat{y}_k^{(3)} + w_4\hat{y}_k^{(4)} + w_5\hat{y}_k^{(5)} = \sum_{i=1}^{5} w_i\hat{y}_k^{(i)}. \tag{13}$$

4. The experimental study

The steps in the experimental study

Thirty real-world time-series datasets were used in this study, 10 each with stationary, trend, or both trend and seasonal characteristics. These datasets were used to assess the effectiveness of the individual forecasting models and the combined forecasting methods. Here are the overall steps used in this study, which are also illustrated as a flow chart in Figure 4.

Step 1. Plot each time-series dataset to detect the type of autocorrelation pattern: stationary, trend, or both trend and seasonal.

Step 2. Create the forecasting values for each time-series dataset with the five individual forecasting models according to the autocorrelation pattern (10 each for the stationary, trend, and both trend and seasonal patterns).

Step 3. Compute the accuracy as mean absolute percentage error (MAPE) and root-mean-squared error (RMSE) values for each time-series dataset by using the following models according to the autocorrelation pattern.

1. The simple moving average, single exponential smoothing, Box-Jenkins, ANN, and SVM models for the stationary pattern datasets.

2. The double moving average, double exponential smoothing, Box-Jenkins, ANN, and SVM models for the trend pattern datasets.

3. The Holt-Winters, decomposition, Box-Jenkins, ANN, and SVM models for the datasets with both trend and seasonal patterns.

Step 4. Apply the three combined forecasting methods: simple-average**,** Bates-Granger, and the proposed method to the forecasted values from five individual forecasting models according to the autocorrelation pattern.
Step 5. Compute the accuracy as the MAPE and RMSE values for each method for each time-series dataset.
Step 6. Compare the performances of the individual forecasting models and the combined forecasting methods for each autocorrelation pattern in terms of the MAPE and RMSE values.

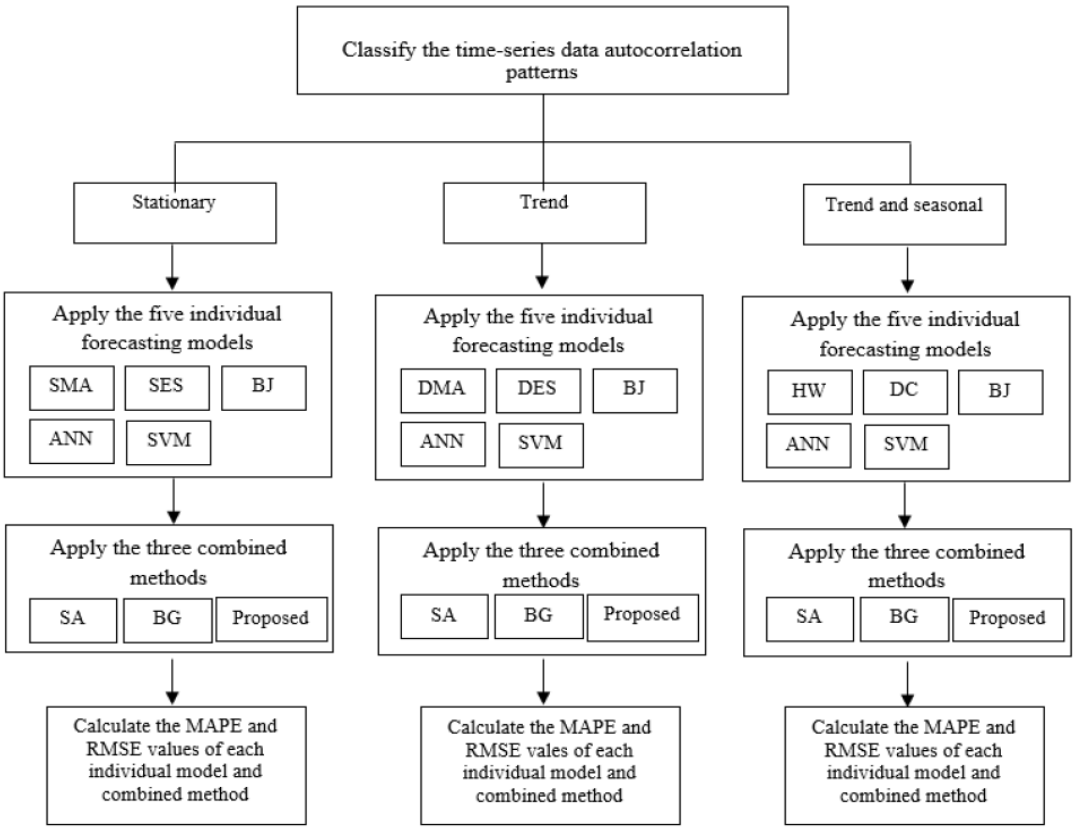


Figure 4 A flow chart of the experimental study of 30 time-series datasets with stationary, trend, or trend and seasonal patterns analyzed with five individual forecasting models and three combined forecasting methods. SMA, simple moving average; DMA, double moving average; SES, single exponential smoothing; DES, double exponential smoothing; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

The measures for forecasting accuracy

The most frequently used measures to identify the most accurate methods for time-series forecasting are MAPE and RMSE (Wongoutong 2020). These two error indices were used to verify the accuracy of time-series forecasting in this study. MAPE is a relative error measure using absolute values that can be used to compare the forecasting accuracy when using differently scaled time-series data. RMSE is an absolute error measure by using the square of the deviation that can prevent positive and negative deviation values from canceling each other out. MAPE and RMSE are respectively defined as

$$\text{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\left(\left|\frac{y_t-\hat{y}_t}{y_t}\right|\right), \quad \text{RMSE}=\sqrt{\frac{\sum_{t=1}^{n}(y_t-\hat{y}_t)^2}{n}},$$

where y_t and $\hat{y}_t$ are the true and predicted values at time t and n is the number of data points.

5. Results and discussion

The MAPE and RMSE values for the forecasting analysis of the time-series datasets are reported in Tables 2 and 3 for the stationary pattern, Tables 4 and 5 for the trend pattern, and Tables 6 and 7 for the trend and seasonal data pattern respectively. For all 10 real-world datasets of the stationary pattern (S1–S10), the simple moving average, single exponential smoothing, Box-Jenkins, ANN, and SVM forecasting models and the combined forecasting simple-average, Bates-Granger, and proposed methods produced average MAPE values of 15.336%, 15.100%, 14.216%, 11.325%, 12.782%, 12.141%, 11.130%, and 9.908% (Table 2) and average RMSE values of 922.61, 871.61, 862.84, 674.70, 814.71, 726.03, 669.07, and 609.35 (Table 3), respectively. These results indicated that the three combined forecasting methods outperformed the individual forecasting models. Especially, the proposed method achieved the lowest values for both accuracy measures, and it is evident that it quite considerably outperformed the other methods.

For the time-series datasets with the trend pattern (T1–T10), the double moving average, double exponential smoothing, Box-Jenkins, ANN, and SVM forecasting models and the combined forecasting simple-average, Bates-Granger, and proposed methods produced average MAPE values of 8.065%, 7.767%, 6.196%, 5.754%, 5.474%, 5.472%, 5.135%, and 5.041% (Table 4) and average RMSE values of 403.791, 336.345, 292.311, 275.590, 281.235, 274.570, 261.926, and 254.418 (Table 5), respectively. Once again, the three combined forecasting methods outperformed the individual forecasting models, and the proposed method provided lower values than the other combined forecasting methods.

For the time-series datasets with both the trend and seasonal patterns (TS1–TS10), the decomposition, Holt-Winters, Box-Jenkins, ANN, and SVM forecasting models and the combined forecasting simple-average, Bates-Granger, and proposed methods produced average MAPE values of 6.494%, 7.866%, 6.144%, 6.335%, 5.877%, 5.001%, 4.791%, and 4.696% (Table 6) and average RMSE values of 343.284, 437.498, 334.068, 331.952, 355.457, 277.649, 265.118, and 258.454 (Table 7), respectively. Once again, the three combined forecasting methods outperformed the individual forecasting models, and the proposed method provided lower values than the other combined forecasting methods. Thus, the proposed method demonstrated its superiority over the individual forecasting models as well as the other combined forecasting methods in all three autocorrelation pattern scenarios.

Table 2. MAPE values of the individual forecasting models and combined forecasting methods for the time-series datasets with the stationary autocorrelation pattern.

Data	Individual Forecasting Model					Combined Forecasting Method		
	SMA	SES	BJ	ANN	SVM	SA	BG	Proposed
S1	19.205	22.405	18.628	13.193	16.799	16.064	14.683	10.894
S2	17.235	15.787	14.162	16.419	14.666	14.562	14.419	14.020
S3	21.432	19.979	18.918	18.811	17.637	17.396	17.184	16.184
S4	23.464	21.930	24.255	18.763	19.792	18.771	18.160	16.002
S5	9.114	8.243	8.149	3.758	6.967	6.533	4.305	3.579
S6	11.073	10.553	10.378	7.782	8.545	8.024	7.138	6.762
S7	13.752	12.803	11.613	7.553	12.287	10.103	8.231	7.761
S8	16.196	17.028	14.958	11.128	12.472	12.080	10.341	8.637
S0	14.187	14.682	14.598	10.097	12.094	11.994	11.243	10.126
S10	7.704	7.585	6.500	5.742	6.557	5.883	5.591	5.116
Average	**15.336**	**15.100**	**14.216**	**11.325**	**12.782**	**12.141**	**11.130**	**9.908**

The underlined values infer the best performance. MAPE: mean absolute percentage error; SMA, simple moving average; SES, single exponential smoothing; BJ, Box-Jenkins; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

Table 3. RMSE values of the individual forecasting models and combined forecasting methods for the time-series datasets with the stationary autocorrelation pattern.

Data	Individual Forecasting Model					Combined Forecasting Method		
	SMA	SES	BJ	ANN	SVM	SA	BG	Proposed
S1	873.72	865.76	1063.67	678.86	782.10	738.06	686.59	604.35
S2	937.00	860.62	793.92	810.95	813.56	756.20	748.42	709.83
S3	1622.55	1464.03	1418.21	1355.51	1445.52	1286.92	1257.94	1136.68
S4	1512.19	1469.14	1453.55	1235.57	1350.38	1229.69	1201.67	1111.42
S5	716.15	652.53	650.73	297.81	597.57	523.90	358.73	312.26
S6	921.51	860.18	830.57	605.33	805.07	677.39	608.18	569.94
S7	653.50	611.68	584.82	367.81	612.56	499.68	401.76	375.97
S8	627.67	615.07	607.87	460.27	571.58	514.55	482.67	437.36
S0	890.93	844.91	809.07	575.35	737.99	658.13	585.83	503.31
S10	470.87	472.13	415.95	359.55	430.76	375.81	358.86	332.34
Average	**922.61**	**871.61**	**862.84**	**674.70**	**814.71**	**726.03**	**669.07**	**609.35**

The underlined values infer the best performance. RMSE: root-mean-squared error; SMA, simple moving average; SES, single exponential smoothing; BJ, Box-Jenkins; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

Table 4. MAPE values of the individual forecasting models and combined forecasting methods for the time-series datasets with the trend autocorrelation pattern.

Data	Individual Forecasting Model					Combined Forecasting Method		
	DMA	DES	BJ	ANN	SVM	SA	BG	Proposed
T1	14.731	11.119	10.665	11.556	8.326	9.056	8.435	8.007
T2	1.148	0.730	0.745	0.827	0.682	0.726	0.698	0.690
T3	1.120	0.926	0.901	0.955	0.769	0.783	0.762	0.744
T4	14.098	22.240	11.447	6.901	8.123	7.910	6.896	6.812
T5	3.734	3.123	3.005	2.772	2.854	2.852	2.805	2.769
T6	17.094	14.900	12.194	9.728	11.930	12.170	11.338	11.136
T7	8.466	7.917	7.167	6.373	6.526	6.603	6.501	6.408
T8	2.952	2.695	2.634	3.314	2.145	2.228	2.206	2.207
T9	10.900	8.159	7.111	8.131	6.660	6.935	6.694	6.589
T10	6.403	5.863	6.088	6.987	6.721	5.003	5.013	5.050
Average	**8.065**	**7.767**	**6.196**	**5.754**	**5.474**	**5.427**	**5.135**	**5.041**

The underlined values infer the best performance. MAPE: mean absolute percentage error; DMA, double moving average; DES, double exponential smoothing; BJ, Box-Jenkins; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

 DOI: http://dx.doi.org/10.18187/pjsor.v19i3.4247

Table 5. RMSE values of the individual forecasting models and combined forecasting methods for the time-series datasets with the trend autocorrelation pattern.

Data	Individual Forecasting Model					Combined Forecasting Method		
	DMA	DES	BJ	ANN	SVM	SA	BG	Proposed
T1	652.476	475.337	456.317	474.166	424.244	411.119	393.039	<u>358.372</u>
T2	90.923	68.069	65.346	66.940	61.150	63.538	61.745	<u>59.807</u>
T3	76.021	61.505	58.434	59.154	51.690	52.502	51.195	<u>49.902</u>
T4	500.535	467.082	369.204	<u>249.900</u>	335.846	298.242	268.365	265.308
T5	250.034	203.020	199.696	181.095	190.688	186.472	182.810	<u>179.650</u>
T6	1016.762	889.997	688.674	<u>571.876</u>	679.656	691.334	639.244	627.792
T7	498.514	469.226	410.799	<u>372.069</u>	383.647	392.091	383.633	376.839
T8	176.723	144.522	142.349	183.565	135.536	126.907	126.812	<u>125.007</u>
T9	357.192	279.087	236.186	264.352	228.268	231.369	223.119	<u>214.443</u>
T10	418.734	305.605	296.107	332.786	321.622	292.130	289.298	<u>287.060</u>
Average	**403.791**	**336.345**	**292.311**	**275.590**	**281.235**	**274.570**	**261.926**	**254.418**

The underlined values infer the best performance. RMSE: root-mean-squared error; DMA, double moving average; DES, double exponential smoothing; BJ, Box-Jenkins; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

Table 6. MAPE values of the individual forecasting models and combined forecasting methods for the time-series datasets with both trend and seasonal autocorrelation patterns.

Data	Individual Forecasting Model					Combined Forecasting Method		
	DC	HW	BJ	ANN	SVM	SA	BG	Propose d
TS1	10.024	15.327	11.548	15.690	9.930	9.546	9.428	<u>9.097</u>
TS2	11.755	15.158	12.437	9.403	11.740	9.864	8.944	<u>8.574</u>
TS3	7.448	9.095	8.869	6.580	6.610	5.690	5.498	<u>5.326</u>
TS4	7.225	7.190	4.354	6.927	4.492	4.321	4.165	<u>4.165</u>
TS5	3.676	3.860	3.549	3.247	3.789	2.878	2.680	<u>2.601</u>
TS6	6.683	6.379	3.478	5.631	4.333	3.141	<u>2.985</u>	3.138
TS7	4.650	4.789	3.974	3.405	4.509	3.377	3.260	<u>3.163</u>
TS8	3.829	4.847	3.705	3.562	3.896	3.141	3.082	<u>3.078</u>
TS9	4.597	6.408	4.646	4.479	4.265	3.791	3.674	<u>3.668</u>
TS10	5.053	5.608	4.877	4.426	5.203	4.263	4.194	<u>4.148</u>
Average	**6.494**	**7.866**	**6.144**	**6.335**	**5.877**	**5.001**	**4.791**	**4.696**

The underlined values infer the best performance. MAPE: mean absolute percentage error; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

Table 7. RMSE values of the individual forecasting models and combined forecasting methods for the time-series datasets with both trend and seasonal autocorrelation patterns.

Data	Individual Forecasting Model					Combined Forecasting Method		
	DC	HW	BJ	ANN	SVM	SA	BG	Proposed
TS1	615.312	965.849	665.456	896.985	635.966	572.952	562.499	<u>552.666</u>
TS3	643.688	925.782	671.660	571.478	704.502	558.030	508.313	<u>486.110</u>
TS6	355.444	389.245	442.585	305.482	379.185	319.955	308.858	<u>299.939</u>
TS2	272.295	311.588	179.902	219.255	203.944	153.698	153.018	<u>152.666</u>
TS5	379.939	406.573	341.290	312.902	474.918	299.245	275.880	<u>264.713</u>
TS6	271.404	259.857	184.195	249.815	213.097	<u>149.592</u>	147.863	152.765
TS7	345.054	381.626	309.808	279.540	362.976	262.884	251.815	<u>242.338</u>
TS8	168.647	233.362	169.674	155.386	192.658	148.672	143.757	<u>141.169</u>
TS9	185.170	287.126	182.558	165.705	163.034	146.256	138.895	<u>135.791</u>
TS10	195.887	213.974	193.547	162.970	224.290	165.210	160.285	<u>156.387</u>
Average	**343.284**	**437.498**	**334.068**	**331.952**	**355.457**	**277.649**	**265.118**	**258.454**

The underlined values infer the best performance. RMSE: root-mean-squared error; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

As examples, bar charts of the MAPE values of stationary, trend, and trend and seasonal datasets S1, T1, and TS1 are shown in Figure 5 (a)–(c), respectively. These offer a clear visual demonstration of the superiority of the proposed method over the five individual forecasting models and the other two combined forecasting methods.

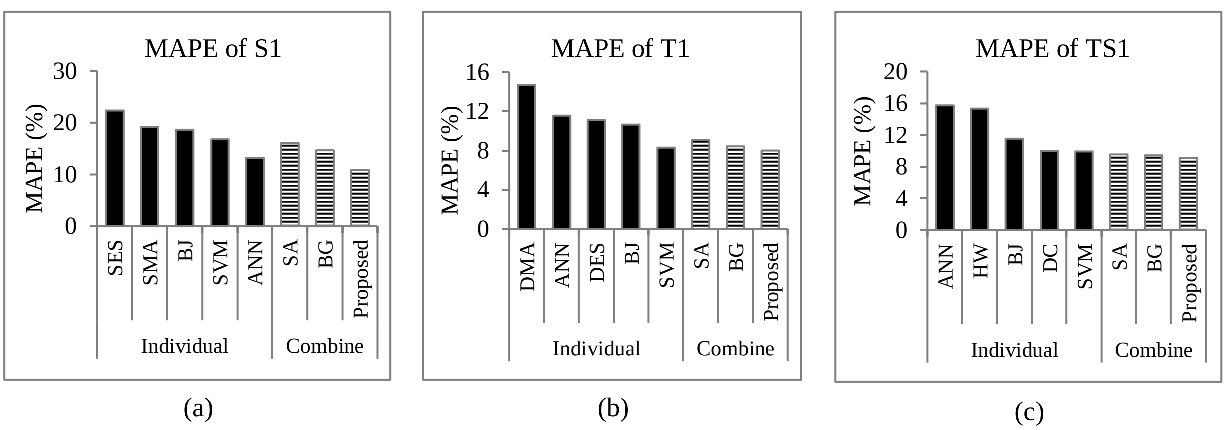


Figure 5. Bar charts of MAPE values for the forecasting analyses of the S1 stationary pattern dataset (a), the T1 trend pattern dataset (b) and the TS1 dataset with both trend and seasonal patterns. SMA, simple moving average; DMA, double moving average; SES, single exponential smoothing; DES, double exponential smoothing; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

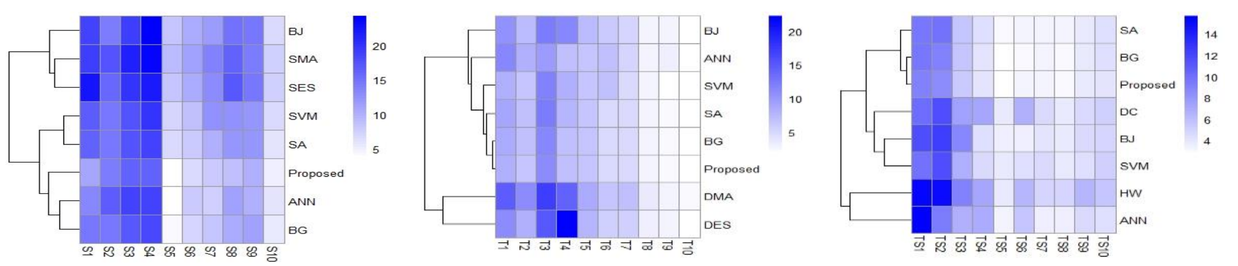
Classification clusters of the performances of the forecasting models and combined forecasting methods for time-series datasets with stationary, trend, or both trend and seasonal patterns are presented as a heatmap of their MAPE values in Figure 6 (a)–(c), respectively. In this study, clustering was achieved by using Euclidean distance and the complete linkage method from the hclust function in the R statistics package version 4.0.3 to find similar groups in each pattern. Next, the dendrograms clustering the algorithm branches were rotated so that the blocks of high and low expression values were near to in the expression matrix. Finally, visualization was realized by applying a color scheme to display the expression matrix. The tree branches were rotated to create blocks in which the individual values were the closest in both directions. These are color-coded as expression values.

For the stationary pattern datasets (S1–S10), the heatmaps in Figure 6 (a) clearly show the patterns picked out by the clustering algorithm as three clustering groups for the MAPE values (worst to best). The first group contains the Box-Jenkins, simple moving average, single exponential smoothing models, the second group contains the SVM model and simple-average method, and the third group contains the ANN model and the Bates-Granger and proposed methods.

For the trend pattern datasets (T1–T10), the heatmaps in Figure 6 (b) once again show three clustering groups; the first contains the double moving average and double exponential smoothing models, the second contains the Box-Jenkins and ANN models, and the third contains the SVM model and the simple-average, Bates-Granger, and proposed methods.

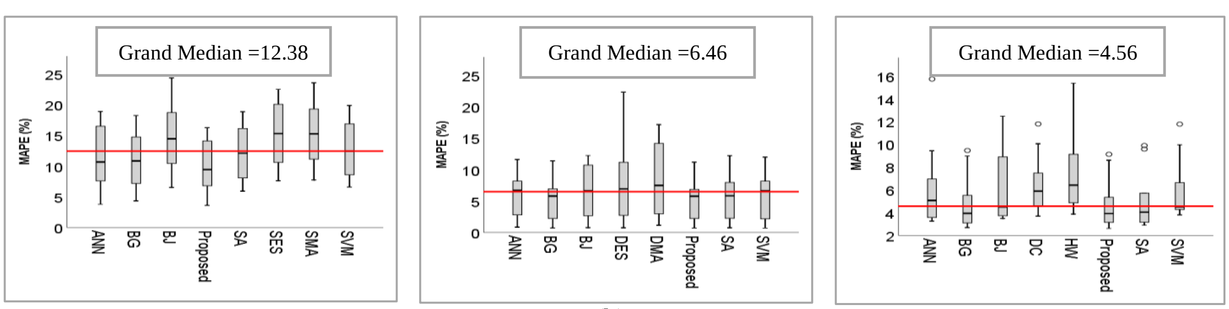
For the datasets with both trend and seasonal patterns (TS1–TS10), the heatmaps in Figure 6 (c) once again shows three clustering groups; the first contains the Holt-Winters and ANN models, the second contains the decomposition, Box-Jenkins, and SVM models, and the third contains the simple-average, Bates-Granger, and proposed methods. Thus, the proposed method was categorized into the same groups as Bates-Granger and the best individual forecasting models.

The performances of the forecasting models and combined forecasting methods in terms of their MAPE values when analyzing time-series data with stationary, trend, or both trend and seasonal patterns are presented as boxplots in Figure 7 (a)–(c), respectively. These show that the median of the MAPE of most of the individual forecasting models was above the grand median (except for the ANN model with the stationary pattern datasets), while the three combined forecasting methods provided MAPE medians below the grand median. Especially, the proposed method achieved the lowest MAPE median for all autocorrelation patterns.



(a) (b) (c)

Figure 6. Heatmaps of the MAPE values for (a) stationary pattern datasets S1–S10, (b) trend pattern datasets T1–T10, and (c) datasets with both trend and seasonal patterns TS1–TS10. SMA, simple moving average; DMA, double moving average; SES, single exponential smoothing; DES, double exponential smoothing; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.



(a) (b) (c)

Figure 7. Boxplots of the MAPE values for (a) stationary pattern datasets S1–S10, (b) trend pattern datasets T1–T10, and (c) datasets with both trend and seasonal patterns TS1–TS10. SMA, simple moving average; DMA, double moving average; SES, single exponential smoothing; DES, double exponential smoothing; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

The proposed method's overall improvement in MAPE and RMSE was calculated relative to the five individual forecasting models and the two other combined forecasting methods for the three autocorrelation patterns (stationary, trend, and both trend and seasonality) (Table 8). For the stationary data, the improvement in MAPE by the proposed method was 37.03%, 35.33%, 30.93%, 10.88%, and 24.23% over the simple moving average, single exponential smoothing, Box-Jenkins, ANN, and SVM models, respectively, and 19.89% and 10.90% over the simple-average and Bates-Granger methods, respectively. Similarly, the improvement in RMSE by the proposed method was 35.18%, 31.78%, 30.22%, and 7.35% over the simple moving average, single exponential smoothing, Box-Jenkins, ANN, and SVM models, respectively, and 17.67% and 9.08% over the simple-average and Bates-Granger, respectively. The results for the datasets with the trend pattern or with both the trend and seasonal patterns exhibited the same trend. It is once again evident that the proposed method outperformed the five individual forecasting models and the other two combined forecasting methods for all three autocorrelation patterns, particularly so for the stationary pattern. For the other two patterns, although the proposed method was better than the individual forecasting models, the Bates-Granger method was almost as effective as the proposed method.

Table 8. The percentage improvement in MAPE and RMSE by the proposed method over the five individual forecasting models and two combined forecasting methods.

Type	Stationary			Trend			Trend and Seasonal		
	Improvement by the Proposed Method (%)								
	Method	MAPE	RMSE	Method	MAPE	RMSE	Method	MAPE	RMSE
Individual	SMA	37.03	35.18	DMA	34.17	35.21	DC	39.03	39.59
	SES	35.33	31.78	DES	22.94	20.22	HW	27.92	26.11
	BJ	30.93	30.22	BJ	15.79	12.43	BJ	20.64	21.48
	ANN	10.88	7.35	ANN	13.58	9.88	ANN	21.9	18.43
	SVM	24.23	26.82	SVM	5.65	8.20	SVM	20.61	26.97
Combined	SA	19.89	17.67	SA	5.48	6.20	SA	5.18	5.81
	BG	10.9	9.08	BG	1.51	2.70	BG	1.3	2.02

SMA, simple moving average; DMA, double moving average; SES, single exponential smoothing; DES, double exponential smoothing; HW, Holt-Winters; BJ, Box-Jenkins; DC, decomposition; SVM, support vector machine; ANN, artificial neural network; SA, simple-average; BG, Bates-Granger.

6. Conclusions

In practice, it is quite common for one forecasting model to perform well in certain periods while other models perform better in other periods. Thus, it is a challenge to find a forecasting model that outperforms all other ones under all circumstances. One approach to improving the accuracy of forecasting to combine forecasts from two or more different forecasting models. Herein, we propose a new weighting system for combined forecasting methods by obtaining the correlation coefficients between the actual and predicted values from the individual forecasting models and ranking them.

In this study, time-series datasets with three autocorrelation patterns (stationary, trend, or both trend and seasonal) were used to evaluate the forecasting performance of the proposed method. As well as outperforming the individual forecasting models, it obviously outperformed the other combined forecasting methods, especially when the autocorrelation pattern was stationary. For this pattern, the improvement in MAPE and RMSE values was 35–37%

for the worst-performing individual forecasting model and 7–10% RMSE for the best-performing individual forecasting model, and for the combined forecasting methods, the improvement in MAPE and RMSE was 18–20% over the simple-average method and 9–11% over the Bates-Granger method. For the time-series datasets with either trend or both trend and seasonal patterns, the performances of the proposed and Bates-Granger methods were almost (an improvement in MAPE and RMSE of 1–2%). However, the Bates-Granger method is complex due to using the diagonal elements in the estimated mean-squared prediction error matrix to compute the combination weights, whereas the technique in the proposed method for computing these is much simpler. Therefore, the proposed method is a plausible alternative for creating weights for the individual forecasting models in combined forecasting methods.

Acknowledgments

The author is grateful for funding from the Thailand Institute of Scientific and Technological Research Institute (TISTR) and would like to thank the Basic Research Fund (BRF) of the Faculty of Science from Kasetsart University for supporting the fund in conducting this research.

References

1. Armstrong, J. S. (2001). Combining forecasts. *Principles of forecasting: A handbook for researchers and practitioners*, 417-439.
2. Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Journal of the operational research society*, *20*(4), 451-468.
3. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
4. Brown, R. G. (1956). Exponential smoothing for predicting demand. Arthur D. Little. *Inc., Cambridge, MA, USA*.
5. Bunn, D. (1989). Forecasting with more than one model. Journal of forecasting, 8(3), 161-166.
6. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20, 273-297.
7. Chan, C. K., Kingsman, B. G., & Wong, H. (2004). Determining when to update the weights in combined forecasts for product demand—an application of the CUSUM technique. European Journal of Operational Research, 153(3), 757-768.
8. Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. International journal of forecasting, 5(4), 559-583.
9. Costantini, M., & Pappalardo, C. (2010). A hierarchical procedure for the combination of forecasts. International journal of forecasting, 26(4), 725-743.
10. Diebold, F. X. (1988). Serial correlation and the combination of forecasts. Journal of Business & Economic Statistics, 6(1), 105-111.
11. De Menezes, L. M., Bunn, D. W., & Taylor, J. W. (2000). Review of guidelines for the use of combined forecasts. European Journal of Operational Research, 120(1), 190-204.
12. Deutsch, M., Granger, C. W., & Teräsvirta, T. (1994). The combination of forecasts using changing weights. International Journal of Forecasting, 10(1), 47-57.
13. Edward Coulson, N., & Robins, R. P. (1993). Forecast combination in a dynamic setting. Journal of Forecasting, 12(1), 63-67.
14. Guo, J., Zhou, J., Qin, H., Zou, Q., & Li, Q. (2011). Monthly streamflow forecasting based on improved support vector machine model. Expert Systems with Applications, 38(10), 13073-13081.
15. Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. International journal of forecasting, 20(1), 5-10.
16. Jose, V. R. R., & Winkler, R. L. (2008). Simple robust averages of forecasts: Some empirical results. International journal of forecasting, 24(1), 163-169.
17. Karatzoglou, A., Hornik, K., Smola, A., & Zeileis, A. (2004). kernlab-an S4 package for kernel methods in R. Journal of statistical software, 11(9).
18. Khairina, D. M., Khairunnisa, R., Hatta, H. R., & Maharani, S. (2021). Comparison of the trend moment and double moving average methods for forecasting the number of dengue hemorrhagic fever patients. Bulletin of Electrical Engineering and Informatics, 10(2), 978-987.

19. Li, G., Song, H., & Witt, S. F. (2005). Recent developments in econometric modeling and forecasting. Journal of travel Research, 44(1), 82-99.
20. Lemke, C., & Gabrys, B. (2010). Meta-learning for time series forecasting and forecast combination. Neurocomputing, 73(10-12), 2006-2016.
21. Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. International journal of forecasting, 16(4), 451-476.
22. Martins, V. L. M., & Werner, L. (2012). Forecast combination in industrial series: A comparison between individual forecasts and its combinations with and without correlated errors. Expert Systems with Applications, 39(13), 11479-11486.
23. Okasha, M. K. (2014). Using support vector machines in financial time series forecasting. International Journal of Statistics and Applications, 4(1), 28-39.
24. Patton, A. J., & Sheppard, K. (2009). Optimal combinations of realised volatility estimators. International Journal of Forecasting, 25(2), 218-238.
25. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). Introduction to time series analysis and forecasting. John Wiley & Sons.
26. Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. Journal of forecasting, 23(6), 405-430.
27. Svetunkov, I., & Petropoulos, F. (2018). Old dog, new tricks: a modelling view of simple moving averages. International Journal of Production Research, 56(18), 6034-6047.
28. Shastri, S., Sharma, A., Mansotra, V., Sharma, A., Bhadwal, A., & Kumari, M. (2018). A study on exponential smoothing method for forecasting. International Journal of Computer Sciences and Engineering, 6(4), 482-5.
29. Thaithanan, J., &Wongoutong, C. (2020). A combined forecasting model for predicting the number of road traffic accident deaths in Thailand. Advances and Applications in Statistics, 64(2), 143–163.
30. Timmermann, A. (2006). Forecast combinations. Handbook of economic forecasting, 1, 135-196.
31. Winkler, R. L., & Makridakis, S. (1983). The combination of forecasts. Journal of the Royal Statistical Society: Series A (General), 146(2), 150-157.
32. Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management science, 6(3), 324-342.
33. Wongoutong, C. (2021). The Effect on Forecasting Accuracy of the Holt-Winters Method When Using the Incorrect Model on a Non-Stationary Time Series. Thailand Statistician, 19(3), 565-582.
34. Wongoutong, C. (2021). Improvement of the Holt-Winters Multiplicative Method with a New Initial Value Settings Method. Thailand Statistician, 19(2), 280-293.
35. Yang, Y. (2004). Combining forecasting procedures: some theoretical results. Econometric Theory, 20(1), 176-222.
36. Zhang, X., Zhang, T., Young, A. A., & Li, X. (2014). Applications and comparisons of four time series models in epidemiological surveillance data. Plos one, 9(2), e88075.