DOR: 20.1001.1.27170314.2021.10.2.2.3

Research Paper

Improving the Performance of Forecasting Models with Classical Statistical and Intelligent Models in Industrial Productions

Maryam Bahrami[1], Mehdi Khashei[1, 2*], Atefeh Amindoust[1]

[1]Department of Industrial Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran

[2]Department of Industrial and Systems Engineering, Isfahan University of Technology, Isfahan, Iran

[*]Email of Corresponding Author: Khashei@cc.iut.ac.ir

Received: May 19, 2021; Accepted: August 1, 2021

Abstract

The capability to receive and deliver customer demand on time in today's competitive world is a significant concern for all industries. In particular, demand management has entered a new era with many companies competing in the last decade. Customer demand management is one of the contemporary issues. The main goal of demand management is to improve supply chain effectiveness, and it is important to note that it is complementary to distribution management and product demand management. Therefore, demand forecasting is essential. For this purpose, in this study, the modeling of the combination structure using autoregressive integrated moving average models and multilayer perceptron neural networks in the field of demand with benchmark data is investigated. The data sets used in this study are two well-known benchmarks of the total product revenue of the Taiwan machinery industry and the sales volume of soft drinks. Eviews and Matlab software have been used to determine the unknown parameters of the proposed model. The experimental results of the research show that the performance of the proposed hybrid model is more accurate than its single components. In addition, results indicate that intelligent models can perform better than classic statistical models.

Keywords

Seasonal Autoregressive Integrated Moving Average (SARIMA), Seasonal Artificial Neural Network (SANN), Demand Forecast

1. Introduction

Supply chain management is related to a set of solutions that integrates chain members (suppliers, manufacturers, retailers, and customers) and its main aim is to reduce system costs as well as to increase customer service levels. It is the main reason that why managing customer demand in this chain is very important [1]. One of the most challenging problems of organizations facing customer demand is to change the demand value, which causes uncertainty on the producer side (production value). That also will be made the uncertainty in the purchase of parts and the whole supply chain. Companies try to alleviate this uncertainty in a way that one of the most common ways is to predict customer demand based on their previous demand [2]. Predictive science nowadays has become an integral part of most organizations for making decisions appropriate and proper planning. The importance of prediction in today's world is not hidden because accurate and timely forecasting can help managers and analysts make better quality decisions and formulate more effective strategies. The accuracy and quality of prediction models are one of the most important factors influencing the choice of prediction methods.

It is the most important reason that today, despite numerous predictive methods, efforts to achieve more accurate methods have never stopped [3]. The accuracy of the predictions depends on the model and structures chosen for the prediction. Predictive models in the field of time series are divided into two categories. Forecasting using single models and using hybrid models. Single models include classic statistical and intelligent models. Combined models are created by combining these two models. Time series data have linear and non-linear patterns. When models both linear and nonlinear patterns, the accuracy of the model performance increases. Single classical statistical models can model the linear pattern of data. Intelligent models such as neural networks can model nonlinear patterns. The combination of these two models improves the accuracy of prediction performance by modeling linear and nonlinear patterns [4]. In the following, several studies in the field of forecasting using single and combined models are presented.

Duerr et al. [5] used two single AR and ARIMA models to predict water demand. The presented models are compared with each other and with the Machine learning methods model. The results show that the performance of smart (intelligence) models is better than classical statistical models. On the medium and long-term forecast of energy demand, Ahmad and Chen [6] propose AdaBoost, ANN, NAEMI, MLRM models. This study compares the models presented with each other and the MLRM AdaBoost BoostedT, BaggedT, and ARIMAX models. The results show that the predictions made with the proposed models are almost identical. Ahmad, and Chen [7] to predict short and medium-term forecasting of cooling and heating load demand in a building environment, used single models of Tree Bagger (TB), Gaussian process regression (GPR), Bagged Tree (BaggedT), Neural Network (NN), Multiple Linear Regression (MLR), Boosted Tree (BoostedT). In this study, seven days, 14 days, and one-month data were used for six models and compared with each other. The results show that precisions of TB, Boosted, GPR, NN, and Bagged are better than the MLR model.

Candelieri et al. [8] presented a single SVM model for short-term forecasting of water demand. At first, the clustering is done on the data and predicted by the proposed SVM model. The proposed model with clustered data is compared with the model with non-clustered data. Comparison results show that clustering data improves the results. In order to forecast spares demand from equipment failures in a changing service logistics context, a single Bayesian network model is presented by Boutselis and Mc Naught [9]. The BN that learned from training data performed best, followed by a hybrid BN design combining expert elicitation and machine learning, and then a logistic regression model different LSO configurations were simulated to create a test dataset, and the simulation results were compared with the various forecasts. Ahmad and Chen [10] have used single models compact decision tree (CDT), Fit k-nearest classifier (FitcKnc) linear regression model (LRM), and stepwise linear regression model (stepwise-LRM) for short-term forecasts of energy demand. The results show that the predictions made with the proposed models are almost identical. Short-term electricity demand forecasts by Cam et al. [11] in three phases, the regression method was used. In the first phase, the data is clustered, and the results are calculated. The second phase data is clustered according to the data strategies, and the results are analyzed. In the third phase, the results were present based on the strategy and without clustering the data. The results show that the prediction of clustered data is more accurate. Bikcora et al. [12] presented a hybrid series-based model for predicting daily power consumption with ARMA-GARCH models. In this study, the proposed model is compared with its regression model. The results of comparing single and composite structures show that the proposed model with the combination structure offers better performance.

Han et al. [13] for predicting urban rail vehicle spare parts demand have used the combination structure of SVM and BP neural network models. Proposed composition structure by using metrics MAE, MSE, MAPE, and MSPE compared with single models. The results show that the proposed combination structure is more accurate and performs better than single-component models. Arunraj and Ahrens [14] used hybrid models for daily food sales forecasting. Proposed models in this research are seasonal autoregressive integrated moving average with external variables (SARIMAX), SARIMAX using multiple linear regression (SARIMA-MLR), SARIMA, and Quantile Regression (SARIMA-QR), multi-layered perceptron neural network (MLPNN). The results of the research show the suggested composition structures of performance better than single models. Mohammadi et al. [15] have been used for forecasting emergency supply-demand from single and hybrid models they are ARIMA–ANN, ARIMA–ANN, and GA–APSO. The mean square error (MSE) was used to compare the models. The research findings show that hybrid models perform better than single models. Aburto and Weber [16] have used from combining SARIMAX and ANN models for demand forecasts. In addition, the results of the research show that the proposed compound structure is more accurate and performs better than its single-component models. Some recent papers with single and hybrid structures are presented in Table 1.

Table 1. Some recently published papers with single and hybrid structures

Author(s)	Year	Type of Model	Used Component(s)	Domain	Description
Al-Fattah & Aramco [17]	2021	Hybrid	Artificial intelligence GANNATS	crude oil demand forecasting	Proposing a hybrid model consisting of the genetic algorithm, neural network, and data-mining approach for time-series models.
Sharma & George [18]	2021	Single	Artificial neural network	for the order-driven vendor recommendation system	Proposing a neural network model for the order-driven vendor recommendation system that ranks the suppliers as per the requirements of the order has been proposed
Hess et al. [19]	2020	Hybrid	Statistical forecasting & machine learning	Meal Demand Forecasting	Proposing an approach incorporating both statistical and machine learning methods for intermittent meal demand with a double-seasonal pattern forecasting.
Perez et al. [20]	2020	Single	ARIMA, Artificial neural network	Electricity Demand Forecasting	Proposing a novel algorithm to forecast big data time series
Van & Mes [21]	2020	Hybrid	ARIMA & Artificial neural network	Demand forecasting	Proposing a hybrid model with ARIMA and MLP models
Oey et al. [22]	2020	Single	demand & supply planning (D&SP)	Demand Forecasting	proposing a model for demand forecasting with supply planning
Xu et al. [23]	2019	Hybrid	SARIMA, Support Vector Regression	Forecasting Aviation industry demand	proposing a hybrid model with SARIMA-SVR

According to the studies presented in the field of forecasting, it is observed that there are two general approaches to forecasting 1) Forecasting with single models 2) Forecasting with hybrid models.

Single models include classic statistical and intelligent models, and hybrid models incorporate single statistical and intelligent classic models. The results of the literature review on single and hybrid models show that the predictions of hybrid models are more accurate and quality than single models. The accuracy and quality of the forecasts will increase by improving the models and taking into account all the data properties. Therefore, using more accurate models can increase the performance of prediction. Studies show the autoregressive moving average and neural network is that most of the models are used for forecasting. These models model linear and nonlinear patterns in the data well. According to the literature of forecasting models, using hybrid structures of demand forecasting for industrial products with international benchmark data less done. Therefore, in this research, modeling, and design of a hybrid structure using two single models to increase the accuracy of predictions has been done. Since the data used in this study have a seasonal pattern, the seasonal autoregressive moving average and the seasonal neural network are used for the first time. The following of this research includes three sections, the first section describes the features of the models used (SARIMA and SMLP) the second part presents the proposed combination structure, and the designed and presents the numerical results for two series of data on the production value of the Taiwan machinery industry and the sales volume of soft drinks and the final section will present the conclusions of the study.

2. Single models

Single prediction models include two categories: classic statistical models and intelligent models. Based on linear or nonlinear patterns in systems, their classical statistical models are divided into linear and nonlinear categories. If the structures in the time series data are perfectly linear, the classical linear statistical models can model and predict these time series correctly. Classical nonlinear statistical models are designed only for modeling specific nonlinear patterns. Therefore, these models alone cannot provide a good prediction. Intelligent models are one of the most dynamic areas of research in the present period that have the potential to detect unknown patterns. Such models are capable of storing large volumes of information and data and analyze and model the patterns in the data with high accuracy and without the limitations used in classical statistical models. Despite the single methods for forecasting, researchers are always trying to use more comprehensive approaches, provided more accurate forecasts. The combination of different models or the application of hybrid models has been considered to increase the accuracy of predictions and to improve the quality level of results in a different scenario. The models used in this study are SARIMA and SANN, which are the most accurate single models in the field of forecasting. In this section, the main concepts of Seasonal Moving Average Regression and Seasonal Multilayer Perceptron Neural Network are presented.

2.1 Seasonal Autoregressive integrated moving average (SARIMA) models

Time series models have been widely used in various sciences, such as engineering and economics. A time series is a set of observations or recorded values of a variable that is arranged in time. In addition to predictability, time series models show important information about the time-dependent change. Such models have a regular search algorithm (pattern recognition, parameter estimation, and pattern verification) to select the appropriate model. The Zt is the time series that is generated by an autoregressive moving average by average μ of the Box-Jenkins model.

$$\emptyset(\mathrm{B})(1-\mathrm{B})^{\mathrm{d}}(\mathrm{Z_t}-\mu)=\theta(\mathrm{B})\mathrm{a_t} \quad (1)$$

$$\theta(\mathrm{B})=1-\theta_1\mathrm{B}-\theta_2^2-\cdots-\theta_q\mathrm{B}^q \quad (2)$$

So that, θ_1, θ_2, ..., θ_q and $\emptyset_1$, $\emptyset_2$, ..., $\emptyset$p are polynomials from B From Degree p and q. B is a regression operator, p, d, q integers, and Zt are the observed values of time series. Some time series also have seasonal phenomena which are repeated after a regular time. The Seasonal Autoregressive Moving Average model can be used to model these time series. These seasonal models are consistent with the general structure of non-seasonally autoregressive moving average models. They were called by the Box-Jenkins [24] as multiplicative seasonal models and briefly as SARIMA(p, d, q)(P, D, Q)s are displayed. In this structure, p, d, q represents the non-seasonal component of the model, and P, D, Q represents the seasonal component of the model. Using the backward transfer operator B, the general form is shown as follows:

$$\varphi_p(\mathrm{B})\emptyset_P(\mathrm{B^s})\nabla^s\nabla_s^d\mathrm{Z_t}=\theta_q(\mathrm{B})\vartheta_Q(\mathrm{B^S})\mathrm{a_t} \quad (3)$$

In this model, φ_1, φ_2, ..., φ_p and $\emptyset_1$, $\emptyset_2$, ..., $\emptyset_P$ and θ_1, θ_2, ..., θ_q and ϑ_1, ϑ_2 , ..., ϑ_Q are polynomials with orders of p and P, q and Q respectively. In this model, the coefficients p, q denote non-seasonal order, and P, Q denotes seasonal order of self-regression and moving average processes. In this model, ∇_s^d is Seasonal Differential Operator and ∇^s is a non-seasonal differential Operator. Generally, time series models consist of four main stages. These are the steps 1) Correct determination of model order 2) Estimation of model parameters 3) Pattern recognition correctly and 4) Forecasting with the selected model.

2.2 Seasonal Multilayer Perceptron neural networks (SMLP)

A multilayer perceptron (MLP) is a class of feedforward neural networks (FNN). The term MLP is used ambiguously, sometimes loosely to refer to any feed-forward ANN, sometimes strictly to refer to networks composed of multiple layers of perceptrons (with threshold activation). Multilayer perceptrons are sometimes colloquially referred to as "vanilla" neural networks, especially when they have a single hidden layer. An MLP consists of at least three layers of nodes: an input layer, a hidden layer, and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function.MLP utilizes a supervised learning technique called backpropagation for training. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron.

It can distinguish data that is not linearly separable. Cybenko [25], Hornik et al. [26], and Funahashi [27] have shown that ANN is a universal function approximator. In a seasonal time series, the recurrence of some recognizable patterns after some regular intervals reminds us that a seasonal time series forecasting problem can be thought of as a function approximation problem. So, it could be said that ANN can learn seasonality in the data structure without removing the seasonal effect from the series, and with a proper ANN structure, it can make successful forecasting. Using the ‘‘s” parameter for determining the input and output neurons number may help to make better predictions.

The ‘‘s” parameter presents the series structure, such as; monthly, quarterly, etc. For monthly time series s=12 and for quarterly time series s = 4. Representing the number of input and output neurons with parameters ‘‘s” can increase the prediction performance of ANN in seasonal time series forecasting. In this kind of network structure, seasonal period observations are values of input neurons, and (i+1)th seasonal period observations output neurons' values. Each seasonal period is composed of several observations. Equation 4 gives the mathematical expression of the output of the SANN.

$$Y_{t+l} = \alpha_1 + \sum_{j=1}^{m} w_{jl}\, f(\sum_{i=0}^{s-1} v_{ij} Y_{t-i} + \theta_j).,\, l = 1,2, \dots, s. \quad (4)$$

where, Y_{t+l} (l = 1,2,. . . , s) represents the predictions for the future s periods; Y_{t-i} (i = 0,1,2,. . . , s _ 1) are the observations of the previous s periods, V_{ij} (i = 0,1,2,. . . , s _ 1 ; j = 1,2,. . . ,m) are weights of connections from input layer neurons to hidden layer neurons, W_{jl} (j = 1,2,. . . , m ; l = 1,2, . . . , s) are weights of connections from hidden layer neurons to output layer neurons, a_l (l = 1,2,. . . , s) and h_j (j = 1,2, . . . ,m) are weights of bias connections and f is the activation function. According to the proposed SANN, the number of input and output neurons should be 12 for monthly time series and 4 for quarterly time series for better forecasting. The number of network parameters is equal to the number of connections between the neurons and the bias terms. The number of network parameters should be adjusted according to the training set size to avoid memorization instead of learning.

3. The proposed structure

It has been proven in the literature that compound structures extract and model more accurately the complex patterns in the data. In addition, applying such techniques to reducing prediction error and improving prediction accuracy which can eliminate the limitations and disadvantages of individual models by simultaneously applying their unique features. In the proposed composition structure, the time series is considered as a combination of two linear and nonlinear components according to the following relation

$$y_t = \mathrm{sum}\,(L_t\,, N_t) \quad (5)$$

Where, L_t and N_t represent the components in linear and non-linear time series, respectively, which should be estimated by single models. The implementation steps of the proposed compound structure are shown in Figure 1.

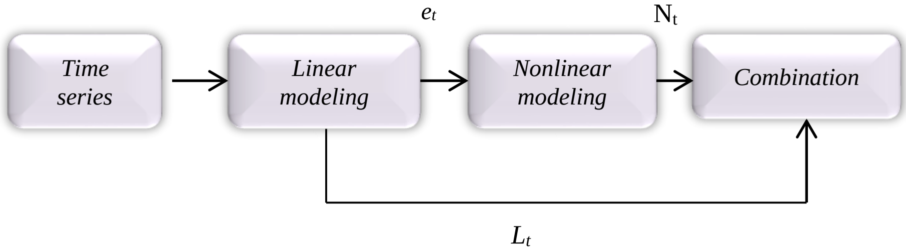

Figure 1. Implementation steps of the hybrid model

The proposed structure, with the unique advantages of two models of Seasonal Autoregressive Moving Average and Seasonal Artificial Neural Network, provides more accurate results in determining linear and nonlinear patterns than single models. This section presents the basic concepts of the proposed structures using single-use SARIMA and SMLP models.

3.1 The combination of SARIMA and SMLP models

In this combination structure, the time series data are linear model input and linear modeling is performed in this phase. Then the residuals generated from the linear modeling phase are modeled by the nonlinear model and the results of the nonlinear modeling phase are combined with the results of the linear modeling phase. Zhang [28] proposed this method. Linear data modeling is first performed by SARIMA, and the residuals of the linear model having nonlinear relationships are considered as inputs of the SMLP model. By the following formula e_t are the residuals of linear mode

$$e_t = y_t - L_t \tag{6}$$

Residues will be as inputs to the neural network model if there are n inputs to the neural network.

$$e_t = f(e_{t-1}, e_{t-2}, e_{t-3}, \dots . e_{t-n}) + \varepsilon_t \tag{7}$$

Finally, the final forecast is as follows

$$Y_t = N_t + L_t \tag{8}$$

In the above formula, N_t is a nonlinear component, and L_t is a linear component.

4. Appling linear/nonlinear hybrid models for demand forecasting

In this section, two abovementioned linear/nonlinear hybrid models are applied for the production value of the Taiwan machinery industry and the sales volume of soft drinks. The description of data sets, the procedure of hybrid models, and the designed hybrid models for each case are briefly presented in the next subsections.

4.1 Total production revenues of the Taiwan machinery industry data set

The first seasonal time-series data, which is considered in this investigation, contains the total production revenues of the Taiwan machinery industry [29,30] which exhibits strong seasonality and growth trends, as shown in Figure 2. This data set has 72 observations, corresponding to the period of January 1991 to December 1996, which is divided into two samples of training and testing. The training data set are 60 observations and the testing data set is 12 observations. It must be noted that in all intelligent models, the best-fitted structure, i.e., number of input and hidden neurons as well as their related transfer functions, is obtained using trial and error. In addition, all other parameters of the designed networks, i.e., learning rate, momentum, number of epochs, early stopping rate, goal accuracy, time of training, learning algorithms, are considered as their default value in Matlab.
The different stages of modeling based on the proposed structure using SARIMA and SMLP models for Taiwan data are as follows:

Stage I. (Linear modeling) In the first stage of the linear/nonlinear hybrid models using the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) and using the Eviews package software, the best-fitted model is SARIMA (0, 1, 1) $(1, 1, 1)_{12}$.

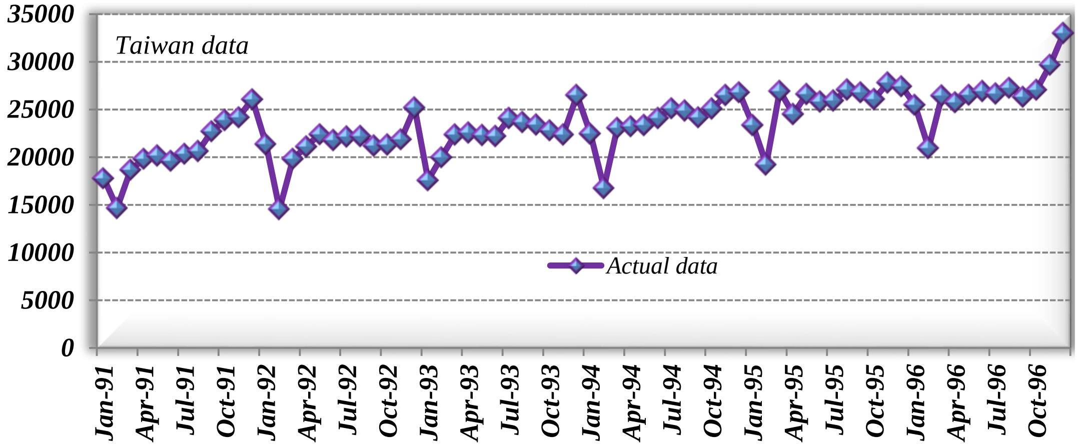

Figure 2. Total production value of Taiwan Machinery Industry (January 1991 to December 1996)

Stage II: (Nonlinear modeling) In order to obtain the optimum network architecture, based on the concepts of artificial neural networks design and using pruning algorithms in MATLAB 7 package software, different network architectures are evaluated to compare the SANNs performance. The best-fitted network which is selected, and therefore, the architecture which presented the best forecasting accuracy with the test data, is composed of seven inputs, three hidden and one output neurons (in abbreviated form, $N^{(7-3-1)}$). The structure of the network is shown in Figure 3.

Stage III: (Combination) In this stage, obtained results from stages I and II are combined. The estimated values of the SARIMA-SMLP model against actual values for all data Figure 4.

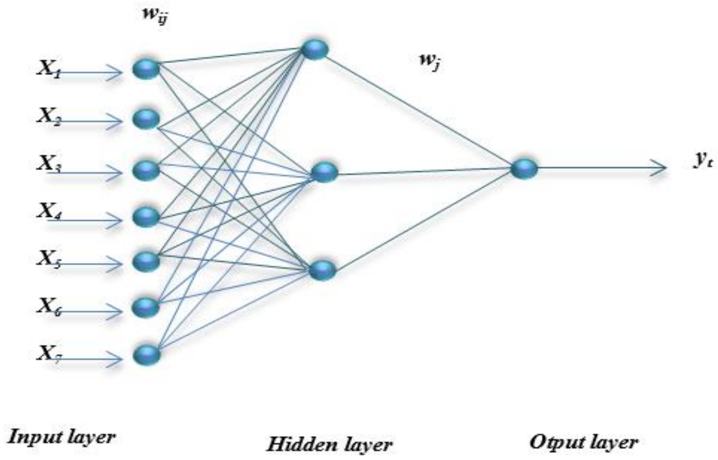

Figure 3. Structure of the best-fitted network (production value case), $N^{(7-3-1)}$

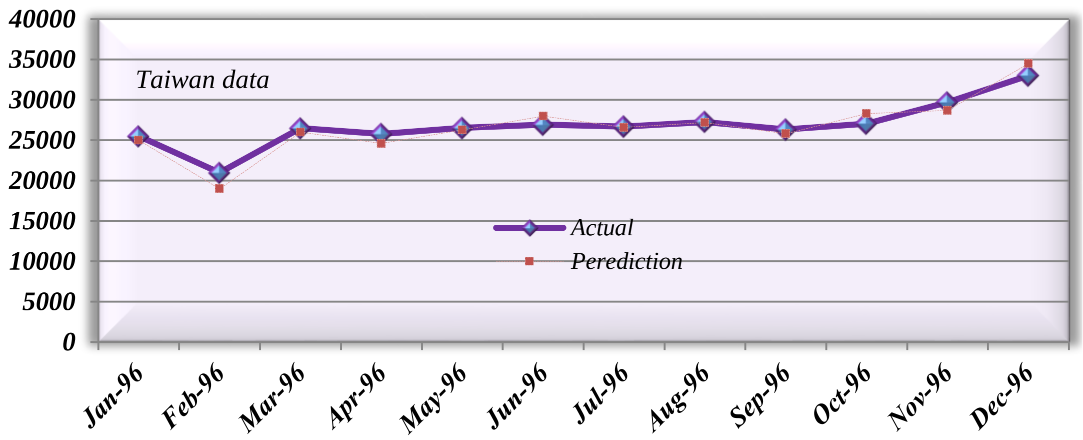

Figure 4. Proposed model prediction of production value data set (test sample)

4.2 The sales volume of soft drinks

The second seasonal time-series data, which is considered in this investigation, contains the sales volume of soft drinks from Montgomery's book "Forecasting and Time Series Analysis" [31]. The time series demonstrates the growth trend and seasonality, as is shown in Figure 5. This data set has also been extensively applied in the linear and nonlinear seasonal time series literature. This data set has 48 observations, corresponding to the period of January 1972 to December 1975, which is divided into two samples of training and testing. The training data set are 36 observations, and the testing data set has 12 observations.

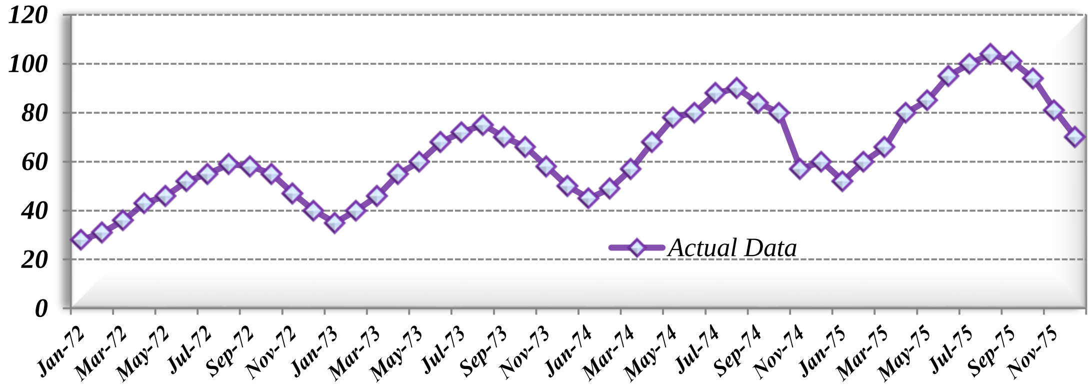

Figure 5. The monthly sales volume of soft drinks (January 1972 to December 1975)

The different modeling steps based on the proposed structure using SARIMA and SMLP models for sales volume of soft drinks data are as follows:

Stage I: (Linear modeling) In the first stage of the linear/nonlinear hybrid models using the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) and using the Eviews package software, the best-fitted model is SARIMA (1, 1, 0) $(0, 1, 0)_{12}$.

Stage II: (Nonlinear modeling) In order to obtain the optimum network architecture, based on the concepts of artificial neural networks design and using pruning algorithms in MATLAB 7 package software, different network architectures are evaluated to compare the SANNs performance. The best-fitted network which is selected, and therefore, the architecture which presented the best forecasting accuracy with the test data, is composed of seven inputs, THREE hidden and one output neuron (in abbreviated form, $N^{(5\text{-}3\text{-}1)}$). The structure of the network is shown in Figure 6.

Stage III: (Combination) In this stage, obtained results from stages I and II are combined. The estimated values of the SARIMA-SMLP model with test data are plotted in Figure 7.

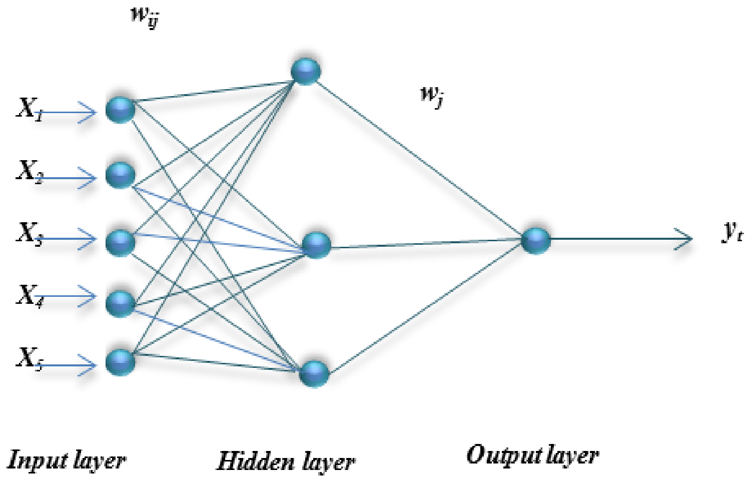

Figure 6. Structure of the best-fitted network (soft drinks case), $N^{(5\text{-}3\text{-}1)}$.

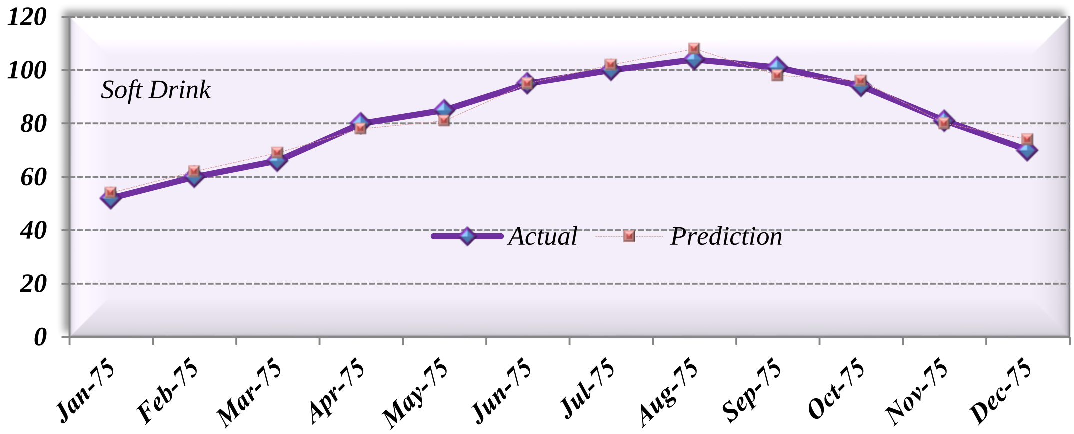

Figure 7. Proposed model prediction of soft drinks data (test sample)

4.3 Evaluation criteria

The MAE (Mean Absolute Error), MSE (Mean Squared Error), and MAPE (Mean Absolute Percentage Error) criteria were used to compare the performance of the proposed model (SARIMA-SMLP) with the single model (SARIMA and SANN) using the two known real datasets above.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}[e_i] \quad (9)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 \quad (10)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}|e_i/y_i| \quad (11)$$

5. Analysis of results

5.1 Total production revenues of the Taiwan machinery industry

In this section, comparison of the performance of the hybrid model (SARIMABP) model with single models Seasonal Autoregressive Integrated Moving Average (SARIMA), Artificial Neural Networks (ANNs) (Deseasonalized data), Artificial Neural Networks (ANNs) (Differenced data), Seasonal Artificial Neural Network (SANN) using MSE, MAE, MAPE criteria is presented in Table 2. The data in this section is the total production revenues of the Taiwan machinery industry. Table 3 shows the percentage performance improvement of the proposed compound structure with its single models.

Table 2. Comparison of the performance of the proposed model with those of other forecasting models

Model	Taiwan machinery data set		
	MSE	**MAE**	**MAPE**
Seasonal Autoregressive Integrated Moving Average (SARIMA)	1197387	816.72	2.88
Artificial Neural Networks (ANNs) (Deseasonalized data)	3046768	1477.67	5.72
Artificial Neural Networks (ANNs) (Differenced data)	1028880	899.28	3.45
Seasonal Artificial Neural Network (SANN)	864520	772.05	2.94
Our proposed model (SARIMABP)	535639	591.89	2.25

Table 3. Improvement of the proposed model in comparison with those of other forecasting models

Model	Taiwan machinery data set		
	MSE	**MAE**	**MAPE**
Seasonal Autoregressive Integrated Moving Average (SARIMA)	55.27%	27.53%	21.87%
Artificial Neural Networks (ANNs) (Deseasonalized data)	82.42%	59.94%	60.66%
Artificial Neural Networks (ANNs) (Differenced data)	47.94%	34.18%	34.78%
Seasonal Artificial Neural Network (SANN)	38.04%	23.34%	23.47%

As can be seen in the Tables above, considering the total production revenue data of the Taiwan machinery, the proposed model with hybrid structures is more accurate than the single-component model. For example, it can be said that considering the MAE criterion, the percentage of proposed structure improvement compared to SARIMA and SANN models is 27.53% and 23.34%, respectively, also the percentage of improvement compared to single model Deseasonalized data is 59.94% and compared to the Differenced data model is 34.18%. Composition structure SARIMA-SANN compared to the SARIMA model considering the MAPE criterion, has 21.87% improvement and compared to the SANN model, it has improved 23.47%. as shown in Table 3, the accuracy of the Seasonal Artificial Neural Network (SANN) model over the models Artificial Neural Networks (Differentiated data) and Artificial Neural Networks (Deseasonalized data) is better. According to the Tables, the SANN model is found to be more accurate than the SARIMA model.

5.2 The sales volume of soft drinks forecasting results

The performance of the proposed composite structure is compared with the single models of Seasonal Autoregressive Integrated Moving Average (SARIMA) and Artificial Neural Networks (ANNs) (Deseasonalized data) and artificial Neural Networks (ANNs) (Differenced data) and Seasonal Artificial Neural Network (SANN) in Table 4. Table 5 shows the percentage of performance improvement of the proposed model compared to the individual models.

Table 4. Comparison of the performance of the proposed model with those of other forecasting models.

Model	Soft drinks data set		
	MSE	**MAE**	**MAPE**
Seasonal Autoregressive Integrated Moving Average (SARIMA)	172.96	12.20	14.33
Artificial Neural Networks (ANNs) (Deseasonalized data)	114.23	8.81	10.48
Artificial Neural Networks (ANNs) (Differenced data)	108.84	9.12	12.54
Seasonal Artificial Neural Network (SANN)	8.27	2.25	3.08
Our proposed model (SARIMABP) (SARIMABP)	14.49	2.32	5.74

Table 5. Improvement of the proposed model in comparison with those of other forecasting models

Model	Soft drinks data set		
	MSE	**MAE**	**MAPE**
Seasonal Autoregressive Integrated Moving Average (SARIMA)	91.62%	80.98%	59.94%
Artificial Neural Networks (ANNs) (Deseasonalized data)	87.32%	73.67%	45.23%
Artificial Neural Networks (ANNs) (Differenced data)	86.67%	74.56%	54.23%
Seasonal Artificial Neural Network (SANN)	-42.93%	-3.11%	-46.34%

Considering the soft drinks data, as shown in the tables above, the proposed composition structure performs better than single models. For example, considering the MAPE criterion, the proposed structure performed better than the SARIMA models with 59.94%. Performance improvement of the proposed model in terms of MAE is 73.67% compared to the deseasonalized data model and 74.56% compared to the Differenced data model. As can be seen, the Seasonal Artificial Neural Network (SANN) model performs better than the Seasonal Autoregressive Integrated Moving Average (SARIMA). In short, it can be said that the combination structure of the linear and non-linear models performs better than the individual models. It can also be noted that the use of the seasonal Artificial Neural Network model was more accurate than the moving average autoregressive model in predictions.

6. Conclusion

The supply chain is subject to wide uncertainties, especially in the area of demand that makes decision-making a challenge for managers. To this end, the purpose of this study is to model the structure of the compound to improve the performance of demand forecasts, which uses high-precision models for the first time .The combination structure used consists of a linear model of classical statistical models and a nonlinear model of neural network models. The models used for the first time are SARIMA and SANN. At first, in this structure, the data are modeled by SARIMA, and then the residuals are considered as inputs to the SANN model, and finally, the two functions will be aggregate. In this study are used two datasets: the volume of soft drinks and the total production revenue of the Taiwan machinery industry for seasonal time series forecasting. The performance of the proposed compound structure is compared and compared with its single-component models. The results show that the prediction accuracy using the proposed compound structure is more than the single-component model. MAE, MSE, and MAPE criteria were used to compare the performance of the models. Suggestions for future research are as follow:

1- Using new compound structures for modeling and forecasting in different fields.

2- Using other single models of classical statistical and intelligent domains.

3. Using other data for forecasting.

7. References

[1] Shen., Z.M. and Daskin, M.S. 2005. Trade-offs between customer service and cost supply chain design. Manufacturing and service operations Management. 7(3):188-207.

[2] Armstrong, J. 2001. Principles of forecasting. Kluwer Academic Publisher, Dordrecht.

[3] Cassettari, L., Bendato, I., Mosca, M. and Mosca, R. 2017. A new stochastic multi source approach to improve the accuracy of the sales forecasts, Emerald Insight. 19(1): 48–64 .

[4] Khashei, M., and Bijari, M. 2011. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting, Applied Soft Computing. 11:2664– 2675.

[5] Duerr, I., Merrill, H.R., Wang, Ch., Bai, R., Boyer, M., Dukes, M. D. and Bliznyuk. N. 2018. Forecasting urban household water demand with statistical and machine learning methods using large space-time data: a comparative study. Environmental Modelling & Software. 102:29-38.

[6] Ahmad, T. and Chen, H. 2018. Potential of three variant machine-learning models for forecasting district level medium-term and long-term energy demand in smart grid environment. Energy. 160:1008-1020.

[7] Ahmad, T. and Chen, H. 2018. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches. journal of Energy & Buildings. 166: 460–476.

[8] Candelieri, A., Giordani, I., Archetti, F., Barkalov, K., Meyerov, I., Polovinkin, A., Sysoyev, A. and Zolotykh, N. 2019. Tuning hyperparameters of a SVM-based water demand forecasting system through parallel global optimization. Computers and Operations Researc. 106: 202-209.

[9] Boutselis, P. and Mc Naught, k. 2019. Using Bayesian Networks to forecast spares demand from equipment failures in a changing service logistics context. International Journal of Production Economics. 209: 325-333.

[10] Ahmad, T. and Chen, H. 2018. Utility companies strategy for short-term energy demand forecasting using machine learning based models. Sustainable Cities and Society. 39: 401–417.

[11] Cam, M., Zmeureanu, R. and Daoud, A. 2017. Cascade-based short-term forecasting method of the electric demand of HVAC system. Energy. 119: 1098-1107.

[12] Bikcora, C., Verheijen, L. and Weiland., S. 2018. Density forecasting of daily electricity demand with ARMA-GARCH, CAViaR, and CARE econometric models. Sustainable Energy, Grids and Networks. 13:148–156.

[13] Han,Y., Wang, L., Gao, J., Xing, Z. and Tao, T. 2017. Combination Forecasting Based on SVM and Neural Network for Urban Rail Vehicle Spare parts Demand. Proceedings of the 36th Chinese Control Conference. July 26-28. 4660-4665.

[14] Arunraj, N. and Ahrens, D. 2015. A Hybrid Seasonal Autoregressive Integrated Moving Average and Quantile Regression for Daily Food Sales Forecasting. International Journal of Production Economics. 170(Part A):321-335.

[15] Mohammadi,R., Fatemi Ghomi, S.M.T. and Zeinali. F. 2014. A new hybrid evolutionary based RBF networks method for forecasting time series: A case study of forecasting emergency supply demand time series. Engineering Applications of Artificial Intelligence. 36: 204–214.

[16] Aburto, l. and Weber, R. 2007. Improved supply chain management based on hybrid demand forecasts. journal of Applied Soft Computing. 7: 136–144.

[17] Al-Fattah, S. M. and Aramco, S. 2021. Application of the artificial intelligence GANNATS model in forecasting crude oil demand for Saudi Arabia and China. Journal of Petroleum Science and Engineering. 200:108368.

[18] Sharma, S. K. and George, S., A. 2021. A vendor recommendation model using neural networks. International Journal of Process Management and Benchmarking. 11(2): 254 – 270.

[19] Hess, A., Spinler, S. and Winkenbach, M., 2020. Real-time demand forecasting for an urban delivery platform. 145: 1-20.

[20] Pérez-Chacón, R., Asencio-Cortés, G., Martínez-Álvarez, F. and Troncoso, A. 2020. Big data time series forecasting based on pattern sequence similarity and its application to the electricity demand. Information Sciences. 540: 160-174.

[21] Van Steenbergen, R., M. and Mes, M., R., K. 2020. Forecasting demand profiles of new products. Decision Support Systems. 139(113401): 1-15, 2020.

[22] Oey, E., Wijaya, W. A. and Hansopaheluwakan, S. 2020. Forecasting and aggregate planning application – a case study of a small enterprise in Indonesia. International Journal of Process Management and Benchmarking. 10(1): 1 – 21.

[23] Xu, S., Chan, H, K. and Zhang, T. 2019. Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. Transportation Research. Part E, 122:169–180.

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

[25] Cybenko, G. 1989. Approximation by superposition of a sigmoidal function, Mathematical Control Signal Systems. 2:303–314.

[26] Hornik, K., Stinchcombe, M. and White, H. 1989. Multilayer feed-forward networks are universal approximators. Neural Networks. 2:359–366.

[27] Funahashi, K. 1989. On the approximate realization of continuous mappings by neural networks. Neural Networks. 2:183–192.

[28] Zhang, G., P. 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing. 50:159 – 175.

[29] Tseng, F.M., Yu, H.C. and Tzeng, G.H. 2002. Combining neural network model with seasonal time series ARIMA model. Technological Forecasting & Social Change. 69: 71–87.

[30] Tseng, M., Tzeng, G.H. and Yu, HC. 1999. Fuzzy Seasonal Time Series for Forecasting the Production Value of the Mechanical Industry in Taiwan. Technological Forecasting and Social Change. 60: 263–273.

[31] Montgomery, D.C., Johnson, L.A. and Gardiner, J.S. 1990. Forecasting and Time Series Analysis. McGraw-Hill, New York.