International Journal of Current Trends in Engineering & Technology
Volume : I, Issue : II (Jan.- Feb. 2015)

Seasonal Interval Time Series Models in Comparative Study of Industrial Forecasting

Mehdi Khashei*
Department of Industrial and systems Engineering, Isfahan University of Technology Isfahan, Isfahan 84156-83111, Iran.

Farimah Mokhatab Rafiei
Department of Industrial Engineering, Faculty of Engineering, Tarbiat Modares University, Tehran 14115-111, Iran.

Akram Mirahmadi
Department of Industrial and systems Engineering, Isfahan University of Technology Isfahan, Isfahan 84156-83111, Iran.

Abstract — Time series forecasting is an active research area that has drawn considerable attention for applications in a variety of fields. In recent years various types of seasonal time series models have been developed in both industrial and financial markets. The accuracy of forecasting is known to be one of the most important factors to value the models. Therefore, many researches have been made to enhance the efficiency of seasonal forecasting models. In this paper performance of four seasonal interval time series models including seasonal autoregressive integrated moving average (SARIMA), fuzzy seasonal autoregressive integrated moving average (FSARIMA), fuzzy seasonal multi-layer perceptron (FSMLP), and Watada models are compared together and they are applied to forecast two seasonal time series data, the total production value of the Taiwan machinery industry and the sales volume of soft drinks. Empirical results show that the obtained interval of the FSMLP is narrower than the ones of other those used models, so that the FMSLP would be the most satisfactory among all mentioned ones.

Keyword — Seasonal interval forecasting; Multi Layer perceptrons (MLPs); Seasonal Auto-Regressive Integrated Moving Average (SARIMA); Fuzzy logic and Fuzzy models; Industrial and financial time series.

1. Introduction

With the fast development of new technologies and the stiff competition among the various enterprises, the whole business environment has become more dynamic and unstable. It is crucial for the enterprises to arrive at accurate and quick-response decisions. Hence, effective forecasting is essential for the enterprises to identify future technological trends and customer demands [1]. Many industrial and financial time series exhibit seasonal and trend variations. The seasonal time series is a sequence of seasonal data points recorded sequentially in time. Although seasonal variations are perhaps the most significant component in a seasonal time series, a stochastic trend is often accompanied with the seasonal variations and can have a significant impact on various forecasting methods. A time series with trend is considered to be nonstationary and often needs to be made stationary before modeling and forecasting processes take place. Accurate forecasting of seasonal and trend time series is very important for effective decisions. Thus, how to model and forecast seasonal and trend time series has long been a major research topic that has significant practical implications. [2].

In the literature of seasonal time series forecasting, many works have been devoted to develop and improve seasonal time series models over the past several decades. These models can be generally classified in linear and nonlinear models. The SARIMA model is one of the most popular approaches in linear time series models. The SARIMA model has been successfully utilized in many fields, such as in forecasting social, economic, medical, industrial, financial problems [3]. The popularity of the SARIMA models is due to their statistical properties as well as to the well-known Box–Jenkins methodology [4] in the model building process. According to the Box–Jenkins methodology, SARIMA models require that the data be seasonally differenced to achieve stationarity condition [4].

Second class is nonlinear models. The recent up-surging research activities in artificial neural networks (ANNs) as well as their numerous successful forecasting applications suggest that they can also be an important candidate for seasonal and trend time series forecasting. Several distinguishing features of ANNs make them valuable and attractive for a forecasting task. The major advantage of neural networks is their flexible nonlinear modeling capability, which can approximate any continuous measurable function with arbitrarily desired accuracy [5]. No prior assumption of the model form is required in the model building process. Instead, the network model is largely determined by the characteristics of the data [6]. Commonly used neural networks include multi-layer perceptrons (MLPs), radial basis functions (RBFs), probabilistic neural networks (PNNs), and general regression neural networks (GRNNs).

International Journal of Current Trends in Engineering & Technology
Volume : I, Issue : II (Jan.- Feb. 2015)

Improving forecasting especially time series forecasting accuracy is an important yet often difficult task facing decision-makers in many areas; therefore, several researchers have proposed different seasonal time series models and they have used hybrid models or combined several models in order to improve the accuracy of forecasting. Reference [7] constructed a hybrid methodology that exploits the unique strength of the seasonal autoregressive integrated moving average (SARIMA) model and the support vector machines (SVM) model in forecasting seasonal time series. Reference [8] proposed using a hybrid model called SARIMABP that combines the seasonal autoregressive integrated moving average (SARIMA) model and the back-propagation (BP) neural network model to predict seasonal time series data. Reference [9] proposed a fuzzy seasonal ARIMA (FSARIMA) forecasting model, which combines the advantages of the seasonal time series ARIMA (SARIMA) model and the fuzzy regression model. Reference [10] proposed a hybrid approach based on the partial high order bivariate fuzzy time series models in order to analyse the seasonal fuzzy time series. Reference [11] developed a seasonal support vector regression (SSVR) model to forecast seasonal time series data. Reference [12] proposed a hybrid model that combines the seasonal autoregressive integrated moving average (SARIMA) model and grey system theory to forecast MSW generation at multiple time scales without needing to consider other variables such as demographics and socioeconomic factors.

In this paper, the performance of four different seasonal interval time series models, Seasonal Auto-Regressive Integrated Moving Average model (SARIMA), Fuzzy Seasonal Auto-Regressive Integrated Moving Average (FSARIMA), Fuzzy Seasonal Multi-Layer Perceptron (FSMLP) and Watada models are compared for industrial and financial markets forecasting. The rest of the paper is organized as follows. In the next section, basic concepts of four used seasonal time-series models are briefly reviewed. Empirical results of machinery industry forecasting are presented in Section 3. The performance of each model is compared together in section 4, and finally the conclusions are discussed.

2. Seasonal Interval Time Series Models

There are several different approaches for time series modelling. Interval models are a special class of the quantitative forecasting models, in which an interval is calculated as optimum forecast of independent variable. In this section, four used seasonal interval models are briefly reviewed.

2.1. Seasonal Auto-Regressive Integrated Moving Average model (SARIMA)

In a seasonal time series, there are two types of variations: The first type is between consecutive observations, while the second one is between pairs of corresponding observations belonging to consecutive seasons. ARIMA (p,d,q) models can be constructed to depict the relationship between consecutive observation values, whereas ARIMA $(P,D,Q)_s$ models can be formed to show the relationship between corresponding observation values of consecutive seasons. A time series $\{ y_t \mid t=1,2,...,k \}$ is generated by SARIMA (p,d,q) $(P,D,Q)_s$ process of Box-Jenkins time series model with the mean μ if:

$$\phi_p(B)\,\Phi_P(B^s)\,\nabla^d \nabla_s^D (y_t-\mu) = \Theta_Q(B^s)\,\theta_q(B)\,a_t \qquad (1)$$

Where y_t and a_t are the observed values and random errors at time period t, $t=1,2,...,k$,respectively, $\phi_p(B) = 1-\sum_{i=1}^{p}\phi_i B^i$, and $\theta_q(B) = 1-\sum_{i=1}^{q}\theta_i B^i$, are the nonseasonal autoregressive operator and moving average (MA) operator, respectively, $\Phi_P(B^s) = 1-\sum_{i=1}^{P}\Phi_{is} B^{is}$, and $\Theta_Q(B^s) = 1-\sum_{i=1}^{Q}\Theta_{is} B^{is}$, are the seasonal AR operator and MA operator, respectively, B is the backshift operator, $\nabla^d = (1-B)^d$ is the nonseasonal d*th* differencing, $\nabla_s^D = (1-B^s)^D$ is the seasonal D*th* differencing at s number of lags, *s* is equals 12 months, *p* is the order of nonseasonal AR process, *P* the order of seasonal AR process, *q* the order of nonseasonal MA process, and *Q* the order of seasonal MA process. The Box–Jenkins [3] methodology for fitting a seasonal autoregressive integrated moving average model to data involves the following four-step iterative cycles:

(a) Identify the SARIMA (p,d,q) (P,D,Q)s structure;
(b) Estimate unknown parameters;
(c) Perform goodness-of-fit tests on the estimated residuals;
(d) Forecast future outcomes based on the known data.

The basic idea of model identification is that if a time series is generated from an ARIMA process, it should have some theoretical autocorrelation properties. By matching the empirical autocorrelation patterns with the theoretical ones, it is often possible to identify one or several potential models for the given time series. Reference [4] proposed to use the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data as the basic tools to identify the order of the ARIMA model. Some other order selection methods have been proposed based on validity criteria, the information-theoretic approaches such as the Akaike's information criterion (AIC) [13] and the minimum description length (MDL) [14]-[16]. In addition, in recent years different approaches based on intelligent paradigms, such as neural networks [17], genetic algorithms [18], [19] or fuzzy system [20] have

International Journal of Current Trends in Engineering & Technology
Volume : I, Issue : II (Jan.- Feb. 2015)

been proposed to improve the accuracy of order selection of ARIMA models.

2.2. Fuzzy Seasonal Auto-Regressive Integrated Moving Average (FSARIMA)

The parameter of SARIMA (p, d, q) (P, D, Q)s, $\phi_1,\phi_2,....,\phi_p$, $\Phi_1,\Phi_2,....,\Phi_P$; $\theta_1,\theta_2,....,\theta_q$ and $\Theta_1,\Theta_2,....,\Theta_Q$ are all crisp values. The SARIMA model is a precise forecasting model for short time periods, although it is limited by the large amount of historical data required. However, we usually have to forecast future situations using limited amounts of data in a short span of time. So this model addresses the limitations of real world applications [21]. Instead of using crisp, fuzzy parameters, $\tilde{\phi}_1,\tilde{\phi}_2,....,\tilde{\phi}_p$, $\tilde{\Phi}_1,\tilde{\Phi}_2,....,\tilde{\Phi}_P$, $\tilde{\theta}_1,\tilde{\theta}_2,....,\tilde{\theta}_q$, and $\tilde{\Theta}_1,\tilde{\Theta}_2,....,\tilde{\Theta}_Q$ in the form of triangular fuzzy numbers are used. A FSARIMA (p, d, q)(P, D,Q)s model is described by a fuzzy function with fuzzy parameter

$$\tilde{\phi}\ B\ \tilde{\Phi}\ B^s\ W_t = \tilde{\theta}_q\ B\ \tilde{\Theta}\ B^s\ a_t \qquad (2)$$

$$W_t = \ 1-B\ ^d\ \ 1-B^s\ ^D\ \ Z_t - \mu \qquad (3)$$

$$\begin{aligned}\tilde{W}_t &= \sum_{i=1}^{p}\tilde{\phi}_i W_{t-i} + \sum_{i=1}^{P}\tilde{\Phi}_i W_{t-is} \\ &\quad - \sum_{i=1}^{p}\sum_{j=1}^{P}\tilde{\phi}_i\tilde{\Phi}_j W_{t-i-js} \\ &\quad - \sum_{i=1}^{q}\tilde{\theta}_i a_{t-i} - \sum_{i=1}^{Q}\tilde{\Theta}_i a_{t-is} \\ &\quad + \sum_{i=1}^{q}\sum_{j=1}^{Q}\tilde{\theta}_i\tilde{\Theta}_j a_{t-i-js}\end{aligned} \qquad (4)$$

Where Z_t are observations, $\tilde{\phi}_1,\tilde{\phi}_2,....,\tilde{\phi}_p$, $\tilde{\Phi}_1,\tilde{\Phi}_2,....,\tilde{\Phi}_P$, $\tilde{\theta}_1,\tilde{\theta}_2,....,\tilde{\theta}_q$, and $\tilde{\Theta}_1,\tilde{\Theta}_2,....,\tilde{\Theta}_Q$ are fuzzy numbers. Now (4) can be modified as follows:

$$\begin{aligned}\tilde{W}_t &= \sum_{i=1}^{p}\tilde{\beta}_i W_{t-i} + \sum_{i=1}^{P}\beta_{p+i} W_{t-is} \\ &\quad - \sum_{j=1}^{P}\sum_{i=1}^{p}\tilde{\beta}_i\beta_{p+j} W_{t-i-js} \\ &\quad - \sum_{i=1}^{q}\tilde{\beta}_{p+P+i} a_{t-i} - \sum_{i=1}^{Q}\tilde{\beta}_{p+P+q+i} a_{t-is} \\ &\quad + \sum_{j=1}^{Q}\sum_{i=1}^{q}\tilde{\beta}_{p+P+i}\tilde{\beta}_{p+P+q+j} a_{t-i-js}\end{aligned} \qquad (5)$$

Fuzzy parameters in the form of triangular fuzzy numbers are used

$$\mu_{\beta_i}\ \beta_i\ = \begin{cases} 1-\dfrac{|\alpha_i-\beta_i|}{c_i} & if \quad \alpha_i - c_i \le \beta_i \le \alpha_i + c_i, \\ 0 & otherwise, \end{cases} \qquad (6)$$

Where $\mu_{\tilde{\beta}}\ \beta_i$ is the membership function of the fuzzy set that represents parameter β_i, α_i is the center of the fuzzy number, and c_i is the width or spread around the centre of the fuzzy number. Using the extension principle, and approximation formula as follows to get fuzzy multiplication

$$A_i \otimes A_j \cong\ c_i c_j, a_i a_j, b_i b_j \qquad (7)$$

Where $A_i =\ c_i, a_i, b_i$ and $A_j =\ c_j, a_j, b_j$ are triangular fuzzy numbers. The fuzzy multiplication of $\tilde{\beta}_i\tilde{\beta}_{p+j}$ will be given as:

$$\tilde{\beta}_i\tilde{\beta}_{p+j} \cong\ c_i c_{p+j}, a_i a_{p+j}, b_i b_{p+j} \qquad (8)$$

Now, applying fuzzy parameters $\tilde{\beta}_i$ in the form of triangular fuzzy numbers the membership of W in (5) is given as:

$$\mu_{\tilde{W}}\ W_t\ = \begin{cases} 1-{|W_t - E_t|}\big/{F_t} & for \ \ W_t \ne 0,\ \ a_t \ne 0, \\ 0 & Otherwise \end{cases} \qquad (9)$$

Where

$$\begin{aligned}E_t &= \alpha_0 + \sum_{i=1}^{p}\alpha_i W_{t-i} + \sum_{i=1}^{P}\alpha_{p+i} W_{t-is} \\ &\quad - \sum_{j=1}^{P}\sum_{i=1}^{p}\alpha_i\alpha_{p+j} W_{t-js-i} + a_t \\ &\quad - \sum_{i=1}^{q}\alpha_{p+P+i} a_{t-i} + \sum_{i=1}^{Q}\alpha_{p+P+q+i} a_{t-is} \\ &\quad + \sum_{j=1}^{Q}\sum_{i=1}^{q}\alpha_{p+P+i}\alpha_{p+P+q+j} a_{t-js-i},\end{aligned} \qquad (10)$$

$$\begin{aligned}F_t &= c_0 + \sum_{i=1}^{p} c_i\,|W_{t-i}| + \sum_{i=1}^{P} c_{p+i}\,|W_{t-is}| \\ &\quad + \sum_{j=1}^{P}\sum_{i=1}^{p} c_i c_{p+j}\,|W_{t-js-i}| \\ &\quad + \sum_{i=1}^{q} c_{p+P+i}\,|a_{t-i}| + \sum_{i=1}^{Q}\alpha_{p+P+q+i}\,|a_{t-is}| \\ &\quad + \sum_{j=1}^{Q}\sum_{i=1}^{q} c_{p+P+i}c_{p+P+q+j}\,|a_{t-js-i}|.\end{aligned} \qquad (11)$$

Simultaneously, Z_t represents the t*th* observation, and *h*-level is the threshold value representing the degree to which the model should satisfy all the data points $Z_1, Z_2,....,Z_t$. The problem of finding the fuzzy seasonal ARIMA parameters using threshold value is formulated as the following linear programming problem:

International Journal of Current Trends in Engineering & Technology
Volume : I, Issue : II (Jan.- Feb. 2015)

$$Minimize \quad S=\sum_{t=1}^{k} F_t$$

$$subject.to \begin{cases} E_t + (1-h)\, F_t \geq W_t \quad t=1,2,....,k, \\ E_t - (1-h)\, F_t \geq W_t \quad t=1,2,...,k, \\ c \geq 0, \end{cases} \tag{12}$$

2.3. Fuzzy Seasonal Multi-Layer Perceptron (FSMLP)

In fuzzy seasonal multi-layer perceptron model, a seasonal autoregressive integrated moving average model is initially fitted in order to model the linear component (L_t) of time series y_t . Results are the estimation of actual values of time series ($\hat{L}_t$) and model parameters as follows:

$$z_t=\sum_{i=1}^{p}\phi_i z_{t-i}+\sum_{j=1}^{P}\Phi_{js} z_{t-js}+\sum_{i=1}^{p}\sum_{j=1}^{P}\phi_i\Phi_{js} z_{t-i-js}+a_t$$
$$+\sum_{i=1}^{q}\theta_i a_{t-i}+\sum_{j=1}^{Q}\Theta_{js}\cdot a_{t-js}+\sum_{i=1}^{q}\sum_{j=1}^{Q}\theta_i\Theta_{js} a_{t-i-js}, \tag{13}$$

where, $\phi_p \ (p=1,2,...,p)$, and $\theta_q \ (q=1,2,...,q)$ are the seasonal ARIMA nonseasonal parameters, and $\Phi_P \ (P=1,2,...,P)$, and $\Theta_Q \ (Q=1,2,...,Q)$ are the SARIMA seasonal parameters, $z_t=\nabla^d\nabla_s^D (y_t-\mu) = (1-B^s)^D (1-B)^d (y_t-\mu)$ in which B is the backward shift operator, D and d are integers and often referred to as order of seasonal and nonseasonal differencing, respectively. Considering a seasonal time series to be composed of a linear and a nonlinear component ($y_t = N_t + L_t$), the SARIMA model cannot model the nonlinear seasonal patterns (N_t); hence, the residual of SARIMA model will contain only the nonlinear seasonal patterns ($e_t = y_t - \hat{L}_t$). Therefore, in the second phase, a neural network is used to model the SARIMA residuals. By modelling residuals and using artificial neural networks, nonlinear seasonal relationships can be discovered. With M input nodes, the ANN model for the residuals will be [5]:

$$\begin{aligned} e_t &= f(e_{t-1},...,e_{t-M})+\varepsilon_t \\ &= w_0+\sum\nolimits_{j=1}^{N} w_j g(w_{0j}+\sum\nolimits_{i=1}^{M} w_{i,j}\, e_{t-i})+\varepsilon_t \\ &= \sum\nolimits_{j=0}^{N} w_j\, u_{t,j}+\varepsilon_t, \end{aligned} \tag{14}$$

Where, f is a nonlinear function determined by the neural network, $u_{t,j}=g(w_{0j}+\sum\nolimits_{i=1}^{M} w_{i,j}\, e_{t-i})$, and ε_t is the random error. Note that if the model f is not an appropriate one, the error term is not necessarily random. Therefore, the correct model identification is critical. The forecast from (13) denoted as N_t , the combined forecast will be:

$$\begin{aligned} \hat{z}_t = \hat{L}_t+\hat{N}_t = (&\sum_{i=1}^{p}\phi_i z_{t-i}+\sum_{j=1}^{P}\Phi_{js}\cdot z_{t-js}+ \\ &\sum_{i=1}^{p}\sum_{j=1}^{P}\phi_i\Phi_{js} z_{t-i-js}++\sum_{i=1}^{q}\theta_i\cdot a_{t-i}+ \\ &\sum_{j=1}^{Q}\Theta_{js}\cdot a_{t-js}+\sum_{i=1}^{q}\sum_{j=1}^{Q}\theta_i\Theta_{js} z_{t-i-js})+ \\ &(w_0+\sum\nolimits_{j=1}^{N} w_j\cdot u_{t,j}). \end{aligned} \tag{13}$$

In next section, we will consider the parameters of the two mentioned models ($\phi_p \ \ p=1,2,...,p$, $\theta_q \ \ q=1,2,...,q$, $\Phi_P \ \ P=1,2,...,P$, $\Theta_Q \ \ Q=1,2,...,Q$, and $\tilde{w}_j \ \ j=0,1,2,...,N$), in the form of triangular fuzzy numbers ($\tilde{\phi}_p \ \ p=1,2,...,p$, $\tilde{\theta}_q \ \ q=1,2,...,q$, $\tilde{\Phi}_P \ \ P=1,2,...,P$, $\tilde{\Theta}_Q \ \ Q=1,2,...,Q$, and $\tilde{w}_j \ \ j=0,1,2,...,N$. Then the fuzzy regression will be used to calculate the fuzzy parameters. In addition, this study adapts the methodology formulated by [22] for condition which includes a wide spread of the forecasted interval. A FSMLP model is described by a fuzzy parameter as follows:

$$\begin{aligned} \tilde{z}_t = &\sum_{i=1}^{p}\tilde{\phi}_i z_{t-i}+\sum_{j=1}^{P}\tilde{\Phi}_{js}\cdot z_{t-js}+\sum_{i=1}^{p}\sum_{j=1}^{P}\tilde{\phi}_i\tilde{\Phi}_{js} z_{t-i-js}+ \\ &\sum_{i=1}^{q}\tilde{\theta}_i\cdot a_{t-i}+\sum_{j=1}^{Q}\tilde{\Theta}_{js}\cdot a_{t-js}+ \\ &\sum_{i=1}^{q}\sum_{j=1}^{Q}\tilde{\theta}_i\tilde{\Theta}_{js} z_{t-i-js}+\sum_{j=0}^{N}\tilde{w}_j\cdot u_{t,j}. \end{aligned} \tag{16}$$

Now, (15) is modified as:

$$\begin{aligned} \tilde{z}_t = &\sum\nolimits_{i=0}^{p}\sum\nolimits_{j=0}^{P}\tilde{\beta}_{ij}\cdot z_{t-i-js}+ \\ &\sum\nolimits_{i=l_1}^{l_1+q}\sum\nolimits_{j=l_1}^{l_1+Q}\tilde{\beta}_{ij}\cdot a_{t-i_1-j_1 s}+ \\ &\sum\nolimits_{i=l_2}^{l_2+N}\tilde{\beta}_i\cdot u_{t,i_2}. \end{aligned} \tag{17}$$

International Journal of Current Trends in Engineering & Technology
Volume : I, Issue : II (Jan.- Feb. 2015)

where, $\tilde{\beta}_{0,0}=0$, $\tilde{\beta}_{i,0}=\tilde{\phi}_i$, $\tilde{\beta}_{0,j}=\tilde{\Phi}_{js}$, $\tilde{\beta}_{i,j}=\tilde{\phi}_i\tilde{\Phi}_{js}$ for $i=0,1,2,...,p$, $j=0,1,2,...,P$; and $\tilde{\beta}_{l_l,l_l}=0$, $\tilde{\beta}_{i_1,l_l}=\tilde{\theta}_i$, $\tilde{\beta}_{0,j_1}=\tilde{\Theta}_{js}$, $\tilde{\beta}_{i_1,j}=\tilde{\theta}_i\tilde{\Theta}_{js}$ for $i_1=l_1,l_1+1,...,l_1+q$, $j_1=l_1,l_1+1,...,l_1+Q$, $l_1=P\times p$, $i_1=i-l_1$; and $\tilde{\beta}_i=\tilde{w}_i$ for $i_2=l_2,l_2+1,...,l_2+N$, $l_2=P\times p+q\times Q$, and $i_2=i-l_2$. Then, fuzzy parameters in the form of triangular fuzzy numbers as (6) are used. Using fuzzy parameters β_i in the form of triangular fuzzy numbers and applying the extension principle, it becomes clear that the membership of Z_t in (17) is given as (18).

Simultaneously, z_t represents the *t*th observation and h-level is the threshold value representing the degree to which the model should satisfy all the data points $z_1,z_2,...,z_k$ [23].

$$\mu_{\tilde{z}}\left(z_t\right)\geq h \qquad for\quad t=1,2,....,k. \tag{19}$$

The index t refers to the number of nonfuzzy data used for constructing the model. On the other hand, the fuzziness S included in the model is defined by:

$$\begin{aligned} S=&\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|z_{t-i-js}\right|+\\ &\sum_{t=1}^{k}\sum_{i=l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}c_{ij}\cdot\left|a_{t-i_1-j_1s}\right|+\\ &\sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N}c_i\cdot\left|u_{t,i_2}\right|, \end{aligned} \tag{20}$$

where, ϕ_i is the nonseasonal autoregressive coefficient of the time lag *i*, Φ_{is} is the seasonal autoregressive coefficient of the time lag is, θ_i is the nonseasonal moving average coefficient of the time lag *i*, Θ_{is} is the seasonal moving average coefficient of the time lag is, and w_i is the connection weight between output neuron and *ith* hidden neuron.

3. Application of seasonal models for forecasting

In this section, the appropriateness and effectiveness of four aforementioned seasonal time series models are compared together in applications of the total production value of the Taiwan machinery industry and the sales volume of soft drinks. These data sets show strong seasonality, and growth trends, as shown in Fig. 1 and Fig. 2, respectively. In addition, these data sets have been extensively applied in the fuzzy and nonfuzzy, and linear and nonlinear seasonal time series literature with a focus on the linear fuzzy modelling in incomplete data situations. The Taiwan machinery industry data set totally has 48 observations, corresponding to the period of the January 1994 to December 1997, which is divided into two samples of training and testing in order to assess the forecasting performance of proposed model.

$$\mu_{\tilde{z}_t}\left(z_t\right)=\begin{cases} 1-\dfrac{\left|z_t-\left[\sum\limits_{i=0}^{p}\sum\limits_{j=0}^{P}\alpha_{ij}\cdot z_{t-i-js}+\sum\limits_{i=l_1}^{l_1+q}\sum\limits_{j=l_1}^{l_1+Q}\alpha_{ij}\cdot a_{t-i_1-j_1s}+\sum\limits_{i=l_2}^{l_2+N}\alpha_i\cdot u_{t,i_2}\right]\right|}{\sum\limits_{i=0}^{p}\sum\limits_{j=0}^{P}c_{ij}\cdot\left|z_{t-i-js}\right|+\sum\limits_{i=l_1}^{l_1+q}\sum\limits_{j=l_1}^{l_1+Q}c_{ij}\cdot\left|a_{t-i_1-j_1s}\right|+\sum\limits_{i=l_2}^{l_2+N}c_i\cdot\left|u_{t,i_2}\right|}, & z_t\neq 0,\ a_t\neq 0,\ u_{t,i}\neq 0 \\ 0 & Otherwise. \end{cases} \tag{18}$$

International Journal of Current Trends in Engineering & Technology
Volume : XX, Issue : XXX

$$Minimize\quad S=\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|z_{t-i-js}\right|+\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|a_{t-i_1-j_1s}\right|+\sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N}c_i\cdot\left|u_{t,i_2}\right|$$

$$subject.to\begin{cases}\sum_{i=0}^{p}\sum_{j=0}^{P}\alpha_{ij}\cdot z_{t-i-js}+\sum_{i-l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}\alpha_{ij}\cdot a_{t-i_1-j_1s}+\sum_{i=l_2}^{l_2+N}\alpha_i\cdot u_{t,i_2}+(1-h)\cdot\left(\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|z_{t-i-js}\right|+\ldots\right.\\ \left.\ldots+\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|a_{t-i_1-j_1s}\right|+\sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N}c_i\cdot\left|u_{t,i_2}\right|\right)\geq z_t\qquad t=1,2,..,k,\\ \sum_{i=0}^{p}\sum_{j=0}^{P}\alpha_{ij}\cdot z_{t-i-js}+\sum_{i-l_1}^{l_1+q}\sum_{j=l_1}^{l_1+Q}\alpha_{ij}\cdot a_{t-i_1-j_1s}+\sum_{i=l_2}^{l_2+N}\alpha_i\cdot u_{t,i_2}-(1-h)\left(\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|z_{t-i-js}\right|+\ldots\right.\\ \left.\ldots+\sum_{t=1}^{k}\sum_{i=0}^{p}\sum_{j=0}^{P}c_{ij}\cdot\left|a_{t-i_1-j_1s}\right|+\sum_{t=1}^{k}\sum_{i=l_2}^{l_2+N}c_i\cdot\left|u_{t,i_2}\right|\right)\geq z_t\qquad t=1,2,..,k,\\ c_i\geq 0\quad for\quad i=1,2,\ldots,l_2+N.\end{cases}\tag{21}$$

The training data set, 36 observations (January 1994 - December 1996), is exclusively used in order to formulate the model and then the test sample, the last 12 observations (January 1997 - December 1997), is used in order to evaluate the performance of the established model. The soft drinks data set totally has 48 observations, corresponding to the period of the January 1972 to December 1975, which is also divided into two samples of training and testing. The training data set, 36 observations (January 1972 - December 1974), is used to formulate the model and then the test sample, the last 12 observations (January 1975 - December 1975), is used to evaluate the performance of the established model [21]. The procedure of the aforementioned seasonal models, as example, is illustrated for the Taiwan machinery industry data set forecasting.

3.1. Seasonal Autoregressive Integrated Moving Average model (SARIMA)

Using the *Eviews* package software, the best-fitted model for the production value, is SARIMA (1, 1, 0) (0, 1, 1)$_{12}$, as (22). The results are given in Table1.

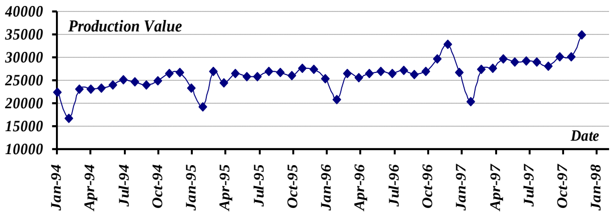

Fig. 1: The production value of the Taiwan machinery industry time series.

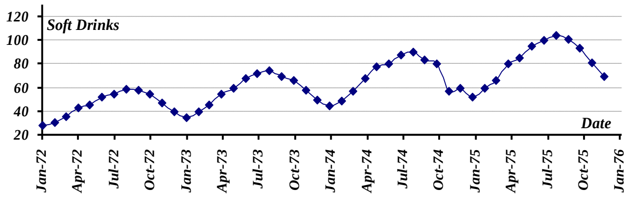

Fig. 2: The monthly sales volume of soft drinks time series.

$$z_t=\nabla^1\nabla_{12}^1(y_t-\mu)=(1-B^{12})^1(1-B)^1(y_t-\mu).$$
$$\hat{z}_t=-0.2588\,z_{t-1}-0.73997\,a_{t-12}\,,\tag{22}$$

3.2. Fuzzy Seasonal Autoregressive Integrated Moving Average model (FSARIMA)

Setting $(\alpha_1,\alpha_2)=(-0.2588,-0.73997)$, the fuzzy parameters are obtained by (10) (with $h=0$). The results after deleting the outlier data are given in Table 1.

3.3. The Fuzzy Seasonal Multi-Layer perceptron (FSMLP)

By Setting obtained results from the designed a crisp MLP model $(\alpha_1^*,\alpha_2^*)=(-0.2588,-0.73997)$ and $(\alpha_3^*,\alpha_4^*,\alpha_5^*)=(5.8673,0.34745,0.59424)$, the fuzzy parameters are obtained using (21) (with $h=0$) as (23). The obtained results of FSMLP model after deleting the outlier data are given in Table 1.

$$\tilde{z}_t=\langle -0.2588,0.0124\rangle z_{t-1}-\langle 0.73997,0.0000\rangle a_{t-12}+\langle 5.8673,0.0000\rangle u_{t,1}+\langle 0.34745,0.04862\rangle u_{t,2}+\langle 0.59424,0.05641\rangle u_{t,3}.\tag{23}$$

International Journal of Current Trends in Engineering & Technology
Volume : XX, Issue : XXX

4. Comparison the Performance of Models

In this section, based on the empirical results of these examples, the predictive capabilities of the aforementioned seasonal models are compared together. The information of forecasted lower and upper bounds of each model for the Taiwan machinery industry and Soft drinks cases is given in Table 1 and Table 2, respectively. The Watada model achieves the lowest and unacceptable performance; so its results are not reported the lowest performance. The seasonal autoregressive integrated moving average (SARIMA) and fuzzy seasonal autoregressive integrated moving average (FSARIMA) take second and third places, respectively. The fuzzy seasonal multi-layer perceptron has the best performance among other those models in both data sets. The obtained results of the fuzzy seasonal multi-layer perceptron (FSMLP) for the Taiwan machinery industry and Soft drinks data sets are shown in Fig. 3 and Fig. 4; respectively.

Table (1). Forecasted interval width and related performance of each model (Taiwan machinery industry).

Date	Actual	FSMLP		FSARIMA		SARIMA	
		Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Jan-97	26910	25082	28893	24690	28731	15133	38288
Feb-97	20489	18901	21626	17930	24943	9558	33316
Mar-97	27489	23868	28329	26453	28800	15284	39969
Apr-97	27669	25585	29796	25672	29959	15072	405586
May-97	29737	26103	30576	27160	28705	14789	41076
Jun-97	29053	27302	31871	27110	30233	15143	42200
Jul-97	29279	26784	31293	26166	33232	15795	43603
Aug-97	29020	26580	31230	27622	31270	15177	43715
Sep-97	28251	25567	29992	26336	31155	14120	43371
Oct-97	30288	27657	32259	27338	31949	14671	44617
Nov-97	30188	28909	34143	30128	31981	15742	46368
Dec-97	35099	33087	39115	25194	37570	15736	470275

Table (2)*. Forecasted interval width and related performance of each model (Soft drinks).

Date	Actual	FSMLP		SARIMA		FSARIMA	
		Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
Jan-75	52	47.69	65.57	39.12	55.68	35.09	91.62
Feb-75	60	51.30	69.76	49.32	60.54	39.54	106.38
Mar-75	66	58.54	80.50	55.502	71.99	42.6	126.94
Apr-75	80	68.28	94.58	69.002	83.26	50.74	157.46
May-75	85	76.96	108.22	69.59	100.55	51.62	172.64
Jun-75	95	78.77	111.25	76.19	95.73	57.6	201.46
Jul-75	100	85.25	121.61	74.39	107.78	58.69	218.15
Aug-75	104	86.00	123.58	81.74	113.94	61.86	240.64
Sep-75	101	81.54	117.30	61.84	117.78	58.99	242.01
Oct-75	94	76.90	110.91	58.56	116.666	54.86	235.31
Nov-75	81	61.52	86.38	32.38	108.636	45.68	205.55
Dec-75	70	58.66	83.64	42.68	99.36	38.1	179

*Note: The upper and lower bound values are rounded.

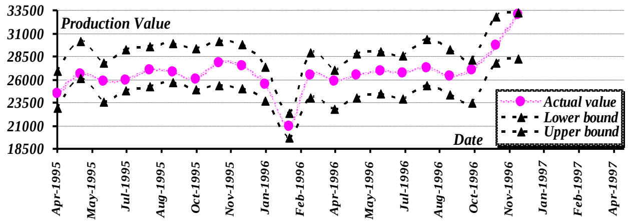

Fig. 3: Results obtained from the FSMLP model (Production Value case).

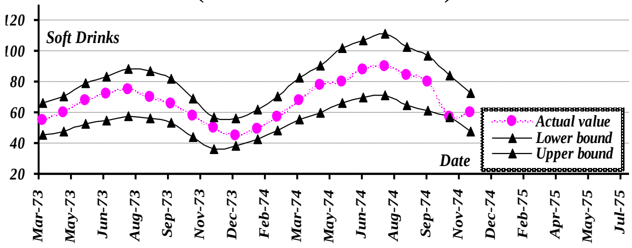

Fig. 4: Results obtained from the FSMLP model (sales volume case).

5. Conclusions

Improving forecasting especially time series forecasting accuracy is an important yet often difficult task facing many decision makers in a wide range of areas. Accurate forecasting of seasonal and trend time series is very important for effective decisions in many areas. Thus, how to model and forecast seasonal and trend time series has long been a major research topic that has significant practical implications. However, predicting seasonal movements has always been a problematic task for academic researchers and despite the paramount modelling effort registered in the last three decades, it is widely recognized that seasonal time series are extremely difficult to forecast.

Many empirical studies including several large-scale forecasting competitions with a large number of commonly used time series forecasting models also conclude that combining forecasts obtained from more than one model often leads to improved performance. That is the reason why research on improving the effectiveness of seasonal time series models has been never witnessed a halt. In this paper the performance of four different seasonal interval time series models (Seasonal Autoregressive Integrated Moving Average (SARIMA), Fuzzy Seasonal Autoregressive Integrated Moving Average (FSARIMA), Watada fuzzy time series, and Hybrid Seasonal Multi-Layer Perceptrons and Fuzzy logic (FSMLP) are compared together in the Taiwan machinery industry and Soft drinks data sets forecasting. Empirical results indicate that among all models FSMLP, FSARIMA, SARIMA, Watada model has the best to the worst performance, respectively.

In addition, the obtained interval of the FSMLP is narrower than other those used models. However; all of the four models have the capacity to handle growth trends and seasonal cycles, with the unit cost of forecasting being relatively low. These evidences indicate that the FSMLP model can be an effective way to improve forecasting accuracy; therefore, it can be used as an appropriate alternative tool for seasonal time series forecasting, especially in incomplete data situations.

Acknowledgment

The authors wish to express their gratitude to Dr. Gh. A. Raisi Ardali, Industrial Engineering Department, Isfahan University of Technology, for their insightful and constructive comments, which helped to improve the paper greatly.

References

[1] H.T. Liu, M.L. Wei, "An improved fuzzy forecasting method for seasonal time series", Expert Systems with Applications 37, 6310–6318, 2010.

[2] G. P. Zhang, M. Qi, "Neural network forecasting for seasonal and trend time series", European Journal of Operational Research 160, 501–514, 2005.

[3] M. Khashei, M. Bijari, "An artificial neural network (p, d, q) model for time series forecasting", Expert Systems with Applications 37, 479–489, 2010.

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

[5] Khashei, M., Bijari, M., Raissi, G. A., "Improvement of Auto-Regressive Integrated Moving Average Models Using Fuzzy Logic and Artificial Neural Networks", Neurocomputing, Volume 72, Pages 956- 967, 2009.

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

[7] K. Chen, C. Wang, "A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan, Expert Systems with Applications 32, 254–264, 2007.

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

[9] F.M. Tseng, G.H. Tzeng, "A fuzzy seasonal ARIMA model for forecasting", Fuzzy Sets and Systems 126, 367–376, 2002.

[10] E. Egrioglu, C. H. Aladag, U. Yolcu, M. A. Basaran, V. R. Uslu, "A new hybrid approach based on SARIMA and partial high order bivariate fuzzy time series forecasting model", Expert Systems with Applications 36, 7424–7434, 2009.

[11] P. F. Pai, K. P. Lin, C. S. Lin, P. T. Chang, "Time series forecasting by a seasonal support vector regression model", Expert Systems with Applications 37, 4261–4265, 2010.

[12] L. Xu, P. Gao, S. Cui, C. Liu, "A hybrid procedure for MSW generation forecasting at multiple time scales in Xiamen City, China", Waste Management 33, 1324−1331,2013.

[13] R. Shibata, "Selection of the order of an autoregressive model by Akaike's information criterion", Biometrika AC-63 (1), pp. 117–126, 1976.

[14] R.H. Jones, "Fitting autoregressions", J. Amer. Statist. Assoc. 70 (351), pp. 590–592, 1975.

[15] C. M. Hurvich, C. L. Tsai, "Regression and time series model selection in small samples", Biometrica 76 (2), pp. 297–307, 1989.

[16] L. Ljung, "System Identification Theory for the User", Prentice-Hall, Englewood Cliffs, NJ, 1987.

[17] H.B. Hwang, "Insights into neural-network forecasting time series corresponding to ARMA *(p; q)* structures", Omega 29, pp. 273–289, 2001.

[18] T. Minerva, I. Poli, "Building ARMA models with genetic algorithms", in: Lecture Notes in Computer Science, Vol. 2037, pp. 335–342, 2001.

[19] C. S. Ong, J. J. Huang, G. H. Tzeng, "Model identification of ARIMA family using genetic algorithms", Appl. Math. Comput. 164 (3) , pp. 885–912, 2005.

[20] M. Haseyama, H. Kitajima, "An ARMA order selection method with fuzzy reasoning", Signal Process 81, pp. 1331–1335, 2001.

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

[22] H. Ishibuchi, H. Tanaka, "Interval regression analysis based on mixed 0-1 integer programming problem", J. Japan Soc. Ind. Eng 40 (5), 312-319, 1988.

[23] M. Khashei, S. R. Hejazi, M. Bijari, "A new hybrid artificial neural networks and fuzzy regression model for time series forecasting", Fuzzy Sets and Systems 159, pp. 769 – 786, 2008.