http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (20): 5447-5453, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.5447.5453

Construction and Implementation of Demand Forecasting Management model of Group Steel Enterprise

[1]Yu Hu and [2]Wen-jie Li
[1]School of Computer Science and Technology, Tianjin University, Tianjin, 300072, China
[2]School of Computer and Communication Engineering, Tianjin University of Technology, Tianjin, 300384, China

Abstract: By analyzing the demand for iron and steel enterprise group differences, this study presents the basic ideas to achieve demand forecast and the key of the ideas is to identify the various groups affiliated enterprises demand forecasting modeling and implementation. Introduction of applying stochastic processes of non-determine seasonal ARIMA model is to forecast the demand for the group subsidiary company's products or market segments. In the model solution process, due to capacity constraints of iron and steel enterprises, the use of a certain confidence interval is solved. By testing and verifying the Iron and Steel Group's sales data, we know using seasonal ARIMA model for a particular area of a product demand forecasting modeling and forecasting is feasible when we deal with non-stationary time series for the steel and iron industry requests.

Key words: Iron and steel enterprise, ARIMA model, time series, demand forecast

INTRODUCTION

According to the current group of steel company's actual operations and the company's overall sales, production, purchasing, logistics, the key to complete steel product task is to have an accurate overall demand as the basis of input conditions. This study achieved demand forecasting management mainly through the seasonal ARIMA modeling to ensure the accuracy of demand forecasts.

IDEAS TO ACHIEVE GROUP OF STEEL COMPANY'S DEMAND MANAGEMENT

The main idea to achieve demand forecasting is: by the establishment of a multi-dimensional product and market structure, determine the demand forecasting model. By using market segment product historical sales data and model operation, generate forecast data at different stages. By manual review and adjustment, superimpose the forecast data of each product or segment market to produce group forecast data.

Multi-dimensional product and market structure: Demand forecasting implementation should firstly define a hierarchy of markets and products and the hierarchy should be very flexible and users can easily adjust it if necessary (Lin and Guo, 1999).

Market and product hierarchy allows users to browse and modify all levels of sales history, sales forecasting, sales planning of any information and when modify once, all relevant data synchronizatedly update (Jiang *et al*., 2003).

Demand forecast based on historical sales data: Demand forecasting model will use a sales history database to record the history of the factory product sales, aggregate historical data according to user-specified products and market segments structures and time units. According to this aggregated history sales data, do the demand forecast by using models. Sales forecast data includes the demand distribution of each segment-market and representative product (maybe produce group) (Zhang and Zhang, 2009) and prediction results can be compared and analyzed with the sales history data.

Based on the structural features of HTML web information extraction technology (Hu and Wang, 2009), this study extracts data to the current sale system. Sales history information stored inside the database should include the actual number of orders, the actual number of shipments and average selling prices and so on.

Multi-level planning cycle prediction: Demand forecast supports annual, quarterly, monthly plans and other multi-level planning cycle forecasts. Based on historical sales records, different plans can be predicted, such as

Corresponding Author: Wen-jie Li, School of Computer and Communication Engineering, Tianjin University of Technology, Tianjin 300384, China

product categories, market segmentation and the time period of detail. Do request forecast for the group subsidiary company's products or market segments, based on historical sales data.

MATHEMATICAL MODEL FOR FORECAST DEMAND

ARIMA model profile: ARIMA (p, d, q) is the differential autoregressive moving average model (Box *et al*., 1994), AR refers to the autoregressive, MA refers to the moving average. Where, p is the number of items from the regression; d is the integration order (processing times after smoothly differencing the original sequence); q is the moving average number of items. ARIMA model for the steel industry needs analyze historical data, input differential calculated value, output demand forecast for the next time.

Hypothesis $\{x_t, t = 0, \pm,...\}$ is a random sequence, B is the post-shift operator, then $Bx_t = x_{t-1}$. Mark $\nabla = 1\text{-}B$, say ∇ as a different operator.

Then for d-order single whole sequence xt~I(d), have:

$$W_t = \nabla^d x_t = (1\text{-}B)^d{}_{xt} \quad (1)$$

Claimed w_t is a stationary sequence and thus we can establish its ARMA (p, q) model, then the resulting model is called x_t~ARIMA(p, d, q), the model form is:

$$\phi(B)\nabla^d{}_{xt} = \theta(B)\alpha_t \quad (2)$$

in which:

$$\phi(B) = 1\text{-}\phi_1 B\text{-}\phi_2 B_2\text{-}\ldots\text{-}\phi_p B^p,$$
$$\phi(B) = 1\text{-}\theta_1 B\text{-}\theta_2 B^2\text{-}\ldots\text{-}\theta_q B^q$$

$|B|\leq 1$ and $\phi(B)$ and $\theta(B)$ comprise,$\phi_p\theta_q\neq 0$, $\{a_t\}$ is white noise, $E\alpha_t = 0$, $E\alpha_t^2 = \sigma^2<\infty$, put $\{x_t\}$ as ARIMA(p, d, q), called autoregressive integrated moving average sequence, then $\{x_t\}$ obey ARIMA(p, d, q) model, d indicates the order of differencing, p represents the regression coefficients, q represents the average coefficient of sliding.

Seasonal demand forecasting model for the steel industry: In practice, the general demand for steel products is based on seasonal changes. Namely, it produces cyclical changed random series according to time. Although these sequences fluctuate in practice, however, at a specific time in each cycle, the corresponding value remains substantially consistent. Various cycle data is not stationary but the peak value of each cycle tends to a linear horizontal. Differentially, making a subtraction between the corresponding data in the two cycles at a certain time can eliminate the periodic variation and the sequence after subtraction can be considered close to the stationary sequence.

Mark $\nabla_s = (1\text{-}B_s)$, known as seasonal differencing operator, then:

$$\nabla_s x_t = (1\text{-}B^s)x_t = x_t\text{-}x_{t\text{-}s} \quad (3)$$

if necessary, consider seasonal differencing D times, then:

$$\nabla_S^d x_t = (1\text{-}B^S)^D = \nabla_S^{D\text{-}1}x_t\text{-}\nabla_S^{D\text{-}1}x_{t\text{-}s} \quad (4)$$

let's be a positive integer, d and D non-negative integer, if the random sequence $\{x_t, t = 0, \pm 1,...\}$ satisfy the differential equation:

$$\phi(B^S)\nabla_S^D x_t = \Theta(B^S)\varpi_t \quad (5)$$

in the equation:

$$\Phi(B^S) = 1\text{-}\Phi_1 B^S\text{-}\Phi_2 B^{2S}\text{-}\ldots\text{-}\Phi_P B^{PS} \quad (6)$$

$$\theta(B^S) = 1\text{-}\theta_1 B^S\text{-}\theta_2 B^{2S}\text{-}\ldots\text{-}\Phi_Q \theta^{QS} \quad (7)$$

$\{\varpi_t\}$ as the ARIMA(p, d, q) sequence, namely:

$$\phi(B)\nabla^d\varpi_t = \theta(B)\alpha_t \quad (8)$$

Among this equation above, $\phi(B)$, $\theta(B)$ satisfy the conditions of the previous section and $\phi_0\neq 1$, $\phi_p\neq 0$, $\theta_0\neq 1$, $\theta_q\neq 0$, $|B|\leq 1$, $\phi_p\theta_q\neq 0\{a_t\}$ is the white noise, then you can put $\{x_t\}$ as seasonal ARIMA sequences, also know $\{x_t\}$ as obey seasonal ARIMA model, the order use (p, d, q)X(P, D, Q)$_S$ as a representation.

Take the Eq. 5 into 8, we obtain the general multiplicative seasonal models:

$$\phi_p(B)\Phi_p(B^S)\nabla_d\nabla^D{}_S x_t = \theta_q\Theta_Q(B^S)\alpha_t \quad (9)$$

wherein, p, d, q, P, D, Q is the order of a variety of operator, so that, the resulting product can be called seasonal model of order (p, d, q)X(P, D, Q)$_S$.

In fact, the multiplicative seasonal ARIMA model is sparse coefficient ARIMA(p+P, d+D, q+Q) model, in that many of the parameters of the model is zero (Xiang, *et al*., 1988).

We can think the time series consist of different frequency sine wave, cosine wave and spectral analysis is an important way to analyze time series. After spectral analysis, we can draw a time series of the cycle as basic modeling conditions for the season model.

For the ARMA series, it is relatively easy to deal with and like such a non-stationary series ARIMA modeling, we use a differential approach to convert the ARIMA sequence to the corresponding ARMA series to solve. Then the Eq. 9 can be rewritten as:

$$\omega_t = \nabla^d \nabla^D_S x_t \tag{10}$$

in which,

$$\omega_t = \phi_p^{-1}(B)\Phi_p^{-1}(B^S)\theta_q(B)\Theta_Q(B^S)a_t \tag{11}$$

If the time series is known, then it should estimate the values of d and D and by the equation (10) obtain the corresponding season ARMA series $\{\omega_t\}$. And we can also get values of order d and D from the basis statistical properties of order d and D values, then difference $\{x_t\}$ to get $\{\omega_t\}$. Seasonal product ARMA sequence $\{\omega_t\}$ can be seen as a special form of ARMA series. Then according to the $\phi(B)$, $\theta(B)$, $\Phi(B^S)$, $\theta(B^S)$, in Eq. 11, we can get the estimated value of all those parameters: $\phi_1, \phi_2, \ldots \phi_p, \theta_1, \theta_2, \ldots \theta_q, \sigma^2, \phi_{P1}, \phi_{P2}, \ldots \phi_{PP}, \theta_{Q1}, \theta_{Q2}, \ldots, \theta_{QQ}$ (Gihman and Skorohod, 1975; Dickey and Fuller, 1981).

Through the above analysis, seasonal ARIMA model can be given to establish the basic steps.

Step 1:

- The time period of the sequence S (spectral analysis)
- To obtain estimates of d and D from trend analysis, determine d and D values through the ADF test. In actual production, for the seasonal time series, the difference does not exceed the first order
- Do the differential operation as shown in Eq. 10 for the sequence $\{x_t\}$ and we can get the corresponding stationary sequence after the operation
- **Pattern recognition:** determine the order of ARIMA model. Although the hybrid model order is usually very unobvious but in the practical application for the demand forecast in the steel industry, the value of the general order p, q, P, Q is usually taken as 0, 1 and 2. Then the selection range of (p, q)×(P, Q) is combinations of 0, 1 or 2 and p, q, P, Q zero simultaneously. Then through the AIC criterion or BIC criterion, to find a suitable value of (p, d)×(P, Q) and to definite different orders of the model
- The all parameters of the model can use the following estimated value of maximum likelihood parameter: $\phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q, \sigma^2, \phi_{P1}, \phi_{P2} \ldots, \phi_{PP}, \theta_{Q1}, \theta_{Q2}, \ldots \theta_{QQ}$
- Set the order and assessment: compare and judge the order of Eq. 9, we can get the time series fitting model

Solving for demand forecasting model: Using the time series model above, we can predict future corresponding observed values after inputting historical observed values. As follows:

- Use the seasonal ARIMA model to forecast time series. Assuming the observed value is $\{x_t, t = 0, \pm 1, \ldots\}$, then:

$$\alpha_t = \sum_{j=0}^{\infty} v_j x_{t-j} \tag{12}$$

among:

$$x_t = \sum_{j=0}^{\infty} \psi_j \alpha_{t-j}, \ \psi_0 = 1 \tag{13}$$

Then:

$$\sum_{j=0}^{\infty} \psi_j B^j = \theta_q(B)\Theta_Q(B^S)\,\phi_p^{-1}(B)\Phi_P^{-1}(B^S)\nabla^{-d}\nabla_S^{-d} \tag{14}$$

$$\sum_{j=0}^{\infty} v_j B^j = \phi_p(B)\Phi_P(B^S)\,\theta_q^{-1}(B)\Theta_Q^{-1}(B^S)\nabla^d\nabla_S^d \tag{15}$$

Equation from 12 to 15 is the general form of prediction algorithm. Therefore, you can use the MMSE forecast (Minimum Mean Square Error) to forecast such a linear model like seasonal ARIMA.

Suppose to do MMSE forecast for x_{t+h}, let t (current time) is the origin, lead or step is h, forecasting value is $\hat{x}_t(h)$, $\hat{x}_t(h)$ is a linear function for current and historical observations $x_t, x_{t-1}, x_{t-2} \ldots$. After calculation:

$$\hat{x}_t(h) = -\sum_{j=0}^{\infty} v_j^{(h)} \hat{x}_{t+h-j} \tag{16}$$

if there is $v_j^{(1)}$-v_j:

$$v_j^{(h)} = v_{j+h-1} - \sum_{i=1}^{h-1}, \ h > 1$$

then:

$$\hat{x}_t(h) = \sum_{j=0}^{\infty} \psi_j \alpha_{t+h-j} \tag{17}$$

then $\hat{\sigma}_t^2(h)$ of MMSE is: (h is increments)

$$\hat{\sigma}_t^2(h) = E(x_{t+h} - \hat{x}_t(h))^2 = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2) \tag{18}$$

- To do forecasts based on iron and steel enterprise production capacity constraints

In the group of iron and steel enterprises, the size of predicted values is affected by internal production status. Because effects such as business capacity are limited, it usually uses confidence intervals to forecast.

Set α as a given level of confidence, $\hat{x}_t^{\alpha}(h)\pm$ is the adjusted value of h-step prediction, that is, the level of the probability confidence limits, the equation is:

$$\hat{x}_t^{\alpha}(h)\pm = \hat{x}_t(h)\pm\xi_{\alpha} \quad (19)$$

In the equation, $\hat{x}_t(h)$ is the predicted values for the MMSE, ξ_{α} is the deviation. Suppose $e_t(h)$ is the prediction error, then it follows a normal distribution. Then, ξ_{α} and α satisfies the Eq. 20:

$$p[e_t(h)\leq\xi_{\alpha}] = \alpha,\ 0.5<\alpha<1 \quad (20)$$

- Demand forecasting process and methods of implementation

Step 2:

- According to the actual plant capacity situation, set α, usually confidence level α set to 0.95
- Through α determine expressions of ξ_{α}, such as $\alpha = 2\sigma_t(h)$;
- Based on actual data situation and product requirements, determine the time interval and the h-step prediction parameters h
- Using step one in the above section to do time series modeling
- Use MMSE method to do h-step prediction
- By the Eq. 20 obtain the prediction deviation and give the probability confidence limits

EMPIRICAL ANALYSIS

This section will use a set of real data to do research modeling and forecasting of seasonal ARIMA model. Actual data comes from a steel group's historical sales data from May 1, 2012 to November 16, 2012 and accumulate the accepted actual volume of sales orders at 1-hour interval.

Express the accepted all types of steel sales order data within the time interval of 1 hour as TDDATA; based on historical data of 200 days, do forecast about the next 20 days after November 16, 2012; then, compare the actual sales data with the predicted values. In the forecasting process, calculating and drawing tools are using s-plus software package (Venables and Ripley, 1997).

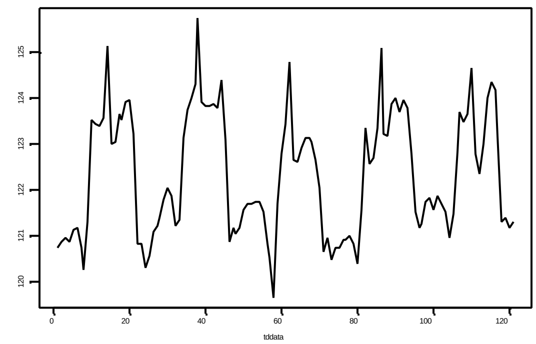

Fig. 1: Actual data diagram

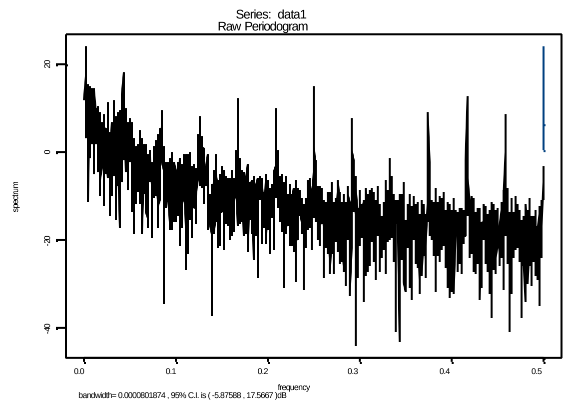

Fig. 2: Spectrum analytical diagram

Analyze data TDDATA: Figure 1 shows part information of the intercepted historical sales data and the data uses 24 (h) as a cycle. Through its spectrum analysis, choose method of suppressing the digital filtering in ECG to correct waveform which is obtained after filtration and organize clutter (Hu and Chen, 2011).

Seen from Fig. 2, the first peak appears near 0.042 and the space for each peak is 0.042, so the data has a determined cycle and is 1/0.042≈24, the same as in Fig. 1. After spectral analysis, we can get the cycle is 24 h (1 day).

In the actual forecasts, the study takes $\xi_{\alpha} = 2\sigma_t(24)$.

Establish multiplicative seasonal mode for TDDATA: By TDDATA analysis results, we can see the multiplicative seasonal model of $(p, d, q)\times(P, D, Q)_{24}$ is applicable. This study uses step one to select models of different p, d, q and P, D, Q to diagnose (Akaike, 1978), after comparing its AIC value and diagnosing the function data series as $(1, 0, 0)X(0, 1, 1)_{24}$, it's diagnose is shown as Fig. 3. The convergence and confidence level of this model is good and is line with the actual requirements.

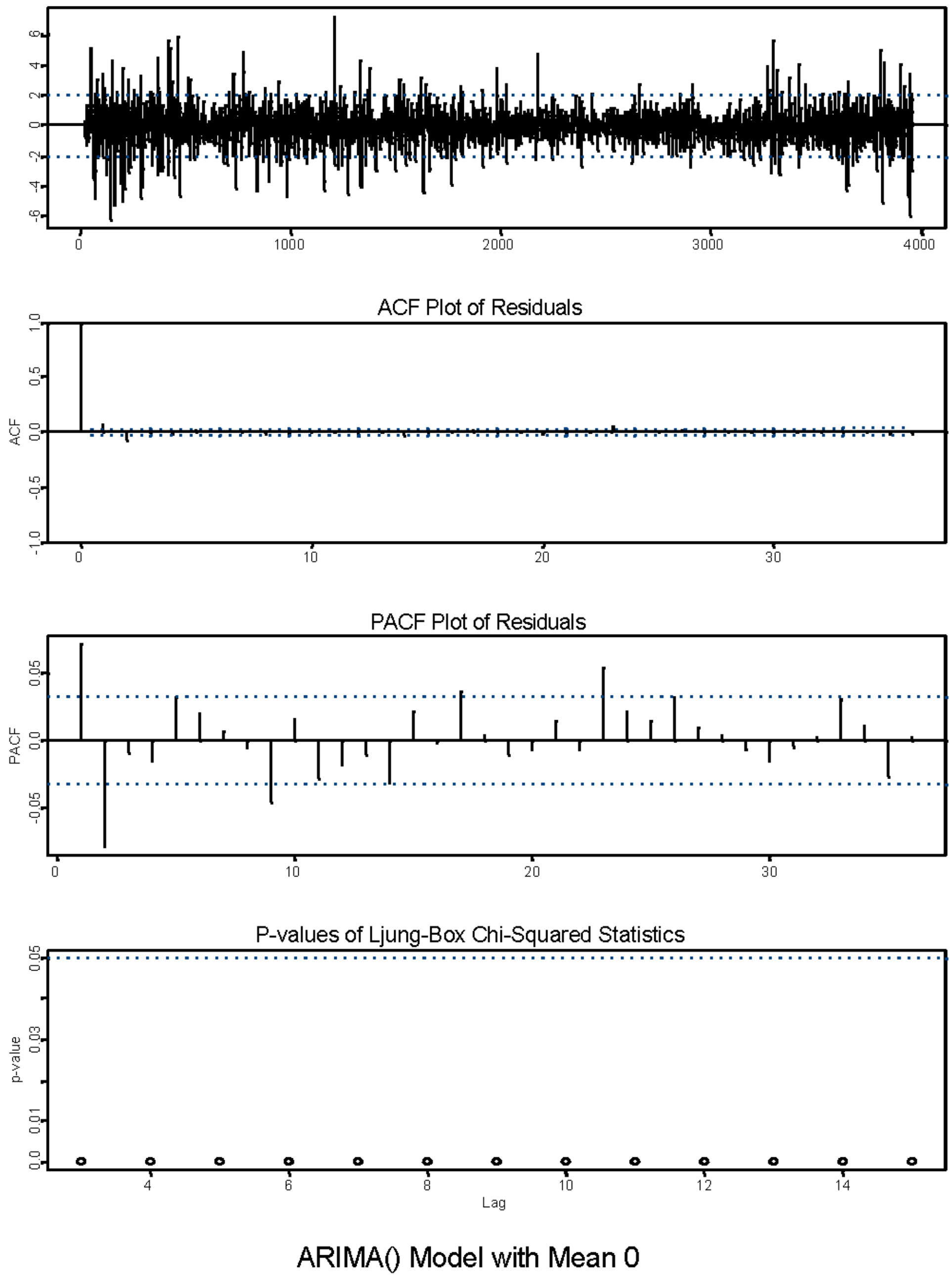

Fig. 3: Diagnostic diagram

Predict TDDATA by using the model: This study again using MMSE prediction to do a 24-step prediction for TDDATA which uses values of the 4800 (200 days) historical sales data to forecast the sales data in the next 24 h (one day).This study will use 95% confidence limit region of $\xi_{\alpha} = 2\sigma_t(24)$ and the actual and predicted forecast value at different time periods are shown, respectively in Fig. 4 (prediction of 2012.11.17) and Fig. 5 (forecasts of 2012.11.21). After a comparative analysis, we can see the predicted values are basically consistent with the true value.

Similarly, we can predict the first 20 days and will aggregate forecast data into data DATA1 (interval in days). To verify its validity, we make a comparison between the actual sales values and forecast data DATA1 of the first 20 days. It's shown in Fig. 6.

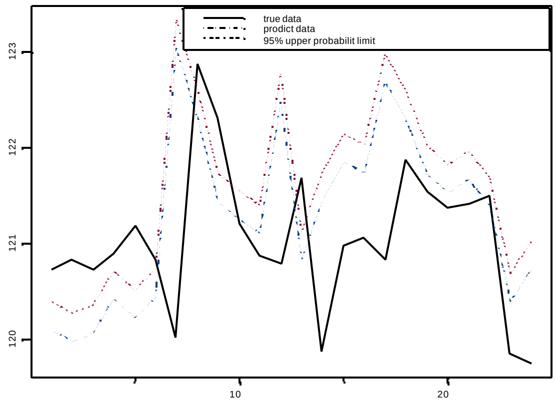

Fig. 4: Forecast comparison diagram on 11.17

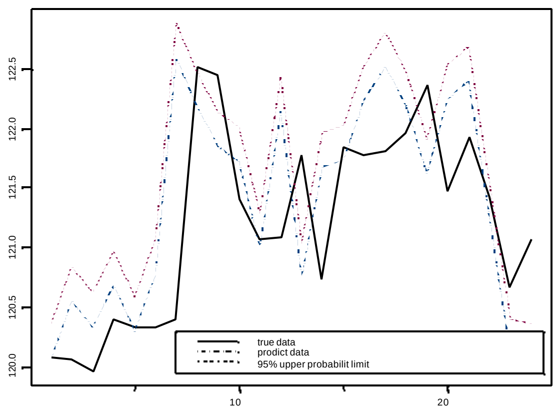

Fig. 5: Forecast comparison diagram on 11.21

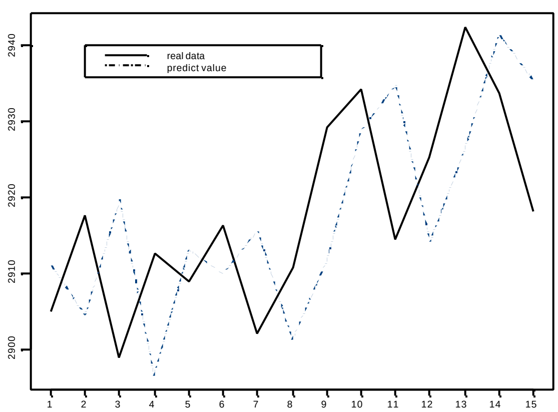

Fig. 6: Forecast comparison diagram after 20 days

After a comprehensive analysis of each demand forecast and the actual value in Fig. 4-6, we can calculate that the error between the predicted value and the actual value is within 2.5%. So, we can assert that it's effective to use multiplicative seasonal model for demand forecast modeling and forecasting. Through empirical analysis, we can prove: seasonal ARIMA model is an effective method to handle this non-stationary time series.

CONCLUSION

The introduction of seasonal ARIMA model with non-application of stochastic processes is used to do demand forecasting for the group subsidiary company's products or market segments. According to the steel products seasonal cyclical, the observed data at a time corresponding to the period minus the next time observation data can eliminate cyclical changes. In the model solution process, due to capacity constraints of iron and steel enterprises, the use of a certain confidence interval is solved. The iron and steel group's sales data validation shows that seasonal ARIMA model using a particular area of a produce demand forecast modeling and forecasting is feasible.

ACKNOWLEDGMENT

The study was sponsored by the National Natural Science Foundation of China (Grant No. 61202169 and No. 61170173).

REFERENCES

Akaike, H., 1978. Time Series Analysis and Control through Parametric Models. In: Applied Time Series Analysis, Findley, D.F. (Ed.) Academic Press, New York, pp: 1-23.

Box, G., G. Jenkins and C. Reinsel, 1994. Time Series Analysis: Forecasting and Control. 3rd Edn., Prentice Hall, New Jersey, ISBN-13: 978-0130607744.

Dickey, D.A. and W.A. Fuller, 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49: 1057-1072.

Gihman, I.I. and A.V. Skorohod, 1975. The Theory of Stochastic Processes II. Springer Verlag, Berlin.

Hu, Y. and L.Z. Wang, 2009.. Page information extraction based on the structure of the HTML. J. Liaoning Univ. Petroleum Chem. Technol., 29: 65-69.

Hu, Y. and T. Chen, 2011. Restrain the power frequency interference in ECG by digital filtering. Electr. Measure. Technol., 34: 22-26.

Jiang, S.P., Y.R. Zou and G. Liang, 2003. Study on production planning in iron and steel enterprise CIMS. Metall. Ind. Automa., 6: 1-6.

Lin, J. and Y.H. Guo, 1999. An object-oriented approach to mode generation and coupling. J. Manage. Sci., 2: 93-102.

Venables, W.N. and B.D. Ripley, 1997. Modern Applied Statistics with S-PLUS. Springer-Verlag, Berlin.

Xiang, J.T., J.G. Du and J.E. Shi, 1988. Dynamic Data Processing: Time Series Analysis. 1st Edn., Meteorology Press, Beijing, (In Chinese).

Zhang, J.X. and L.P. Zhang, 2009. Forecast of commodities demand based on association rule of time series. Comput. Eng., 35: 65-67.