The 9th International Conference on
Electrical and Control Technologies
May 8–9, 2014, Kaunas, Lithuania

ORDERS PREDICTION FOR SMALL IT COMPANY

Rūta UŽUPYTĖ*, Tomas KRILAVIČIUS**
*Vytautas Magnus University, Baltic Institute of Advanced Technology, Lithuania
** Baltic Institute of Advanced Technology, Lithuania

***Abstract*:** Reliable methodology for service orders prediction can significantly improve the quality of business strategy. It is very important to identify the seasonal behavior in order data to correctly predict customer demand and make appropriate business decisions. There are several methods to model and forecast time series with seasonal pattern. This paper compares seasonal naive, Holt – Winters seasonal, SARIMA and neural networks methods in order to evaluate their performance in prediction of the future values of time series that consist of the monthly orders in a small IT company.

***Keywords*:** orders prediction, time series.

1. Introduction

Orders forecasting is an essential tool for business management which enables to make more efficient decisions in allocation of company's resources, planning employment needs, targeting and satisfying customer needs and etc. Usually, such information is analyzed as time series that often exhibit seasonality. Different approaches, such as regression analysis [1, 2], exponential smoothing [2, 3], moving average [4, 5], Box – Jenkins methodology [1, 6] and more sophisticated modeling techniques such as artificial [2, 6], fuzzy logic [3, 7] or evolving [8] neural networks, are used for orders prediction. Our goal is to analyze different methods for modeling and forecasting time series with seasonal behavior.

This research focuses on the monthly orders forecasting of a small IT company and applies seasonal naive, exponential smoothing, SARIMA methods and also simple neural network in order to find the most appropriate methodology for orders prediction.

The remainder of the paper is organized as follows. In section 2 data collection, problem formulation and used methods are presented. Section 3 discusses experimental results. In section 4 concluding remarks are provided.

2. Methods and materials

2.1. Orders data

In this research we have used a small IT company's monthly service orders data from the period July, 2010 - March, 2013. Visualization shows growth trend (Fig. 1) and seasonality (Fig. 2, Fig. 3): orders volume increases in July, August, and November. Seasonality can be due:

- Heating season: the end of the heating season reduces consumer's expenses on heating and as a result customers bring computers to perform preventive maintenance or to repair older failures.
- Holidays: during holidays a number of different accidents, such as dropped or wet computers, increases.
- Summer season: more failures due to the computers overheating as well as computers left outside during the rain.

Of course, other factors can influence it as well.

All models were developed using data from the period July, 2010 - February, 2013. The observation from March 2013 was used to measure performance.

2.2. Problem formulation

Predict future monthly orders y_{T+h} (h - time horizon) by using the previous monthly orders data: $y_1, y_2, \dots, y_T$. Three different time horizons are chosen: 1, 3 and 12 months.

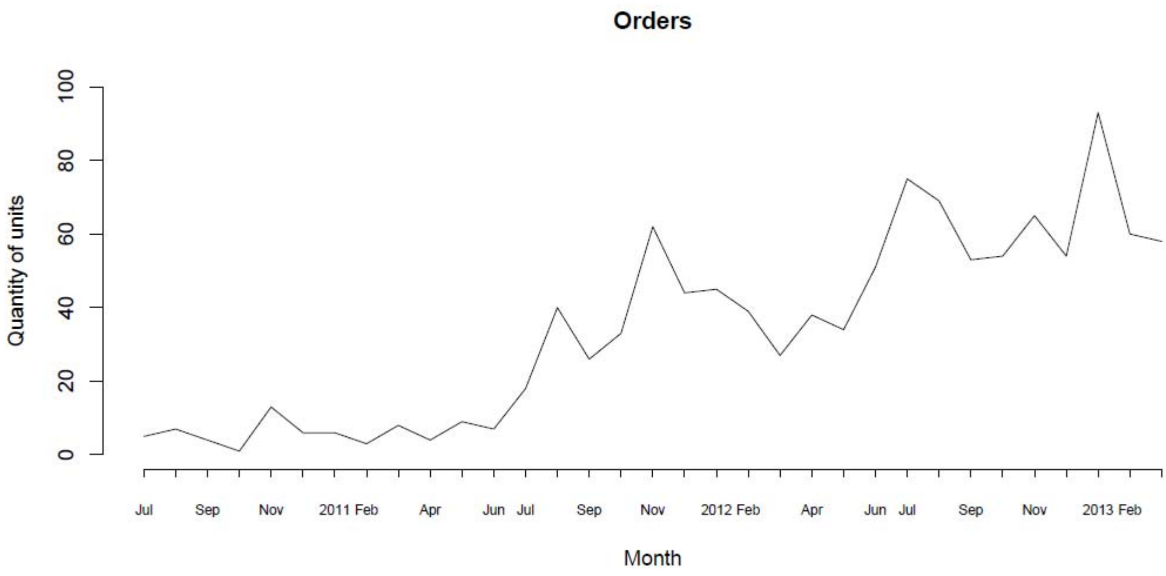

Fig. 1. Monthly orders

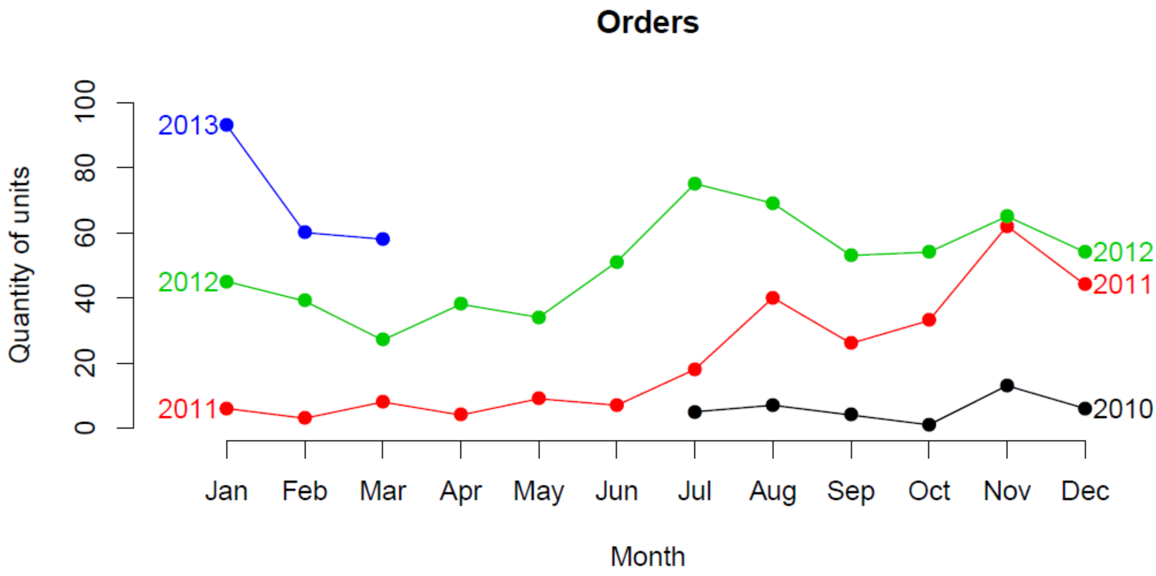

Fig. 2. Seasonal plot

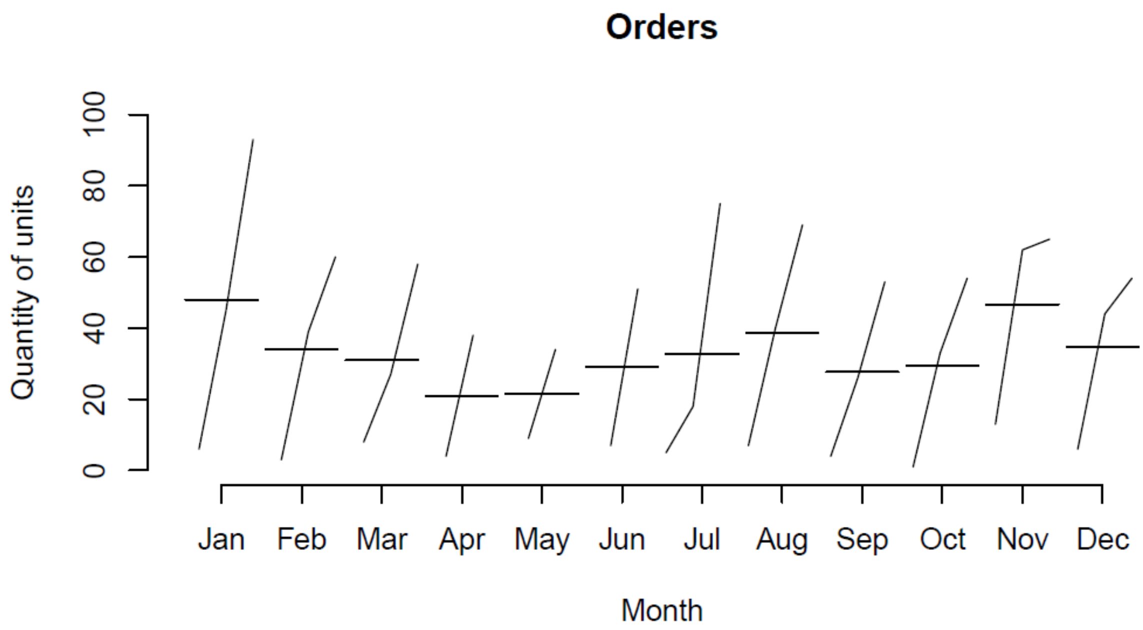

Fig. 3. Seasonal subseries plot

2.3. Loss function

In order to evaluate performance of different models we decided to use the mean absolute error (MAE) [9] and mean absolute percentage error (MAPE) [9]. MAE was selected due to its simplicity – easy to understand and calculate and MAPE was selected because this measure is not inclined to respond to the increase of time series values. Let y_t denote the tth observation and $\hat{y}_t$ denote predicted value. Then the forecast error is equal to

$$e_t = y_t - \hat{y}_t. \quad (1)$$

Mean absolute error can be found using formula

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |e_t| \quad (2)$$

and mean absolute percentage error

$$MAPE = \frac{1}{T}\sum_{t=1}^{T} \left|\frac{100e_t}{y_t}\right|. \quad (3)$$

As mentioned above we used the observation from March 2013 as one of evaluation criteria:

$$e_{March} = y_{March} - \hat{y}_{March}; \quad (4)$$

where y_{March} is a real value of March, 2013 monthly orders and $\hat{y}_{March}$ is forecasted value.

2.4. Seasonal naive method

A simple method is to forecast (seasonal) data. In this case, each forecast is equal to the last observed value from the same season of the year, e.g. the same month of the previous year.

$$\hat{y}_{T+h} = y_{T+h-km}; \quad (5)$$

where T is the index of the last observation, h is time horizon, $k = (h-1)/m + 1$ and m is the period.

2.5. Holt - Winters seasonal method

Holt-Winters seasonal method [9] can be used for forecasting data with trend and seasonal components. There are two variations to this method: additive and multiplicative. Model form for the additive method is:

$$\hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m}, \quad (6)$$

$$l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}), \quad (7)$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}, \quad (8)$$

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}; \quad (9)$$

while for the multiplicative:

$$\hat{y}_{t+h|t} = (l_t + hb_t)s_{t+h-m}, \quad (10)$$

$$l_t = \alpha\frac{y_t}{s_{t-m}} + (1-\alpha)(l_{t-1} + b_{t-1}), \quad (11)$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1}, \quad (12)$$

$$s_t = \gamma\frac{y_t}{(l_{t-1} - b_{t-1})} + (1-\gamma)s_{t-m}; \quad (13)$$

where $\hat{y}_{t+h|t}$ is predicted value at the time $t+h$, l_t is the smoothed value at the time t, b_t denotes trend component at t, s_t denotes seasonal component at t, $0 \le \alpha \le 1$ is the smoothing parameter, $0 \le \beta \le 1$ is the smoothing parameter for the trend, $0 \le \gamma \le 1$ is the smoothing parameter for the seasonal component and m is the period of the seasonality. Values for the smoothing parameters can be estimated by minimizing the sum of the squared errors:

$$\sum_{t=1}^{T} (y_t - \hat{y}_{t|t-1})^2 \rightarrow min. \quad (14)$$

2.6. Seasonal ARIMA model

A seasonal ARIMA model (SARIMA) [9] is an extension of ARIMA model and is denoted as an $ARIMA\ (p,d,q)(P,D,Q)_m$, where p is the order of the autoregressive part, d is the order of the differencing, q is the order of the moving-average process, P is the order of the seasonal autoregressive part, D is the order of the seasonal differencing, Q is the order of the seasonal moving-average process and m is the period of the seasonality. SARIMA model in general form is:

$$(1-B)^d(1-B^m)^D y_t = \mu + \frac{\theta(B)\theta_m(B^m)}{\phi(B)\phi_m(B^m)} e_t; \quad (15)$$

where

- B is backshift operator: $By_t = y_{t-1}$;
- μ is average of series $\{y_t\}_{t=1}^{T}$;
- $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$;
- $\theta_m(B^m) = 1 - \theta_{m,1}B^m - \cdots - \theta_{m,Q}B^{mQ}$;
- $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$;
- $\phi_m(B^m) = 1 - \phi_{m,1}B^m - \cdots - \phi_{m,P}B^{mP}$.

Akaike's Information Criterion (AIC) [9] was used in order to select the most appropriate order of an ARIMA model. In the general case, the AIC is

$$AIC = 2l - 2\ln(L); \quad (16)$$

where $\boldsymbol{l}$ is the number of model parameters and $\boldsymbol{L}$ is the maximized value of the likelihood function. The preferred model is the one with the lowest AIC value.

2.7. Neural networks

Neural network [10] is a structure composed from an input layer, one or more intermediate layers and an output layer, each consisting of several neurons. Each neuron inputs are combined using a weighted linear combination:

$$net = \sum_{t=1}^{T} w_t x_t + w_0; \quad (17)$$

where $x_t, t = 1, \ldots, T$ are inputs, $w_t, t = 1, \ldots, T$ are corresponding weights and w_0 is a threshold value. Transfer function is used to calculate output value:

$$y = f(net). \quad (18)$$

The most common transfer function is a sigmoid [10]:

$$f(net) = \frac{1}{1 + e^{-net}}. \quad (19)$$

The parameters $w_t, t = 0, \ldots, T$ are evaluated using training data: in the beginning all weights take random values which are then updated using the observed data.

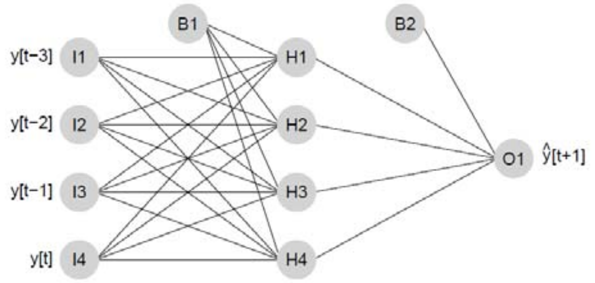

Fig. 4. Scheme of a constructed neural network

In this case we have constructed neural network with 4 inputs and one intermediate layer consisting of 4 neurons (see Fig. 4). Prediction of monthly orders is based on the previous four months data:

Input				**Output**
July	Aug.	Sept.	Oct.	Nov.
Aug.	Sept.	Oct.	Nov.	Dec.
…	…	…	…	…

Neural network was trained using 80% of the data (26 observations) and tested using the rest.

3. Experimental results

Experiments can be divided into two parts: appropriate model selection and forecasting. See Table 1 for models performance.

Table 1. Results of model fitting

Model	**Accuracy measures**		
	MAE	MAPE	e_{March}
Seasonal naive	19.36	42.55	31/54%
Holt - Winters (A) $\alpha = 0.4, \beta = 0.3, \gamma = 1$	10.42	23.35	**1/0.02%**
Holt - Winters (M) $\alpha = 0.8, \beta = 0.1, \gamma = 0.5$	22.64	47.39	50/86%
$ARIMA(2,0,7)(0,2,2)_{12}$	**1.26**	**1.89**	3/0.05%
Neural network	2.21	4.6	7/0.12%

Results show that Seasonal naive method has high MAE and MAPE values and forecast results are very inaccurate (31 orders mismatch). It leads to the conclusion that Seasonal naive method is not suitable for orders prediction, so the method will not be further analyzed. In order to find the most appropriate parameters values for both Holt – Winters method variations we have used a grid-based optimization [11]. Results show that additive Holt – Winters method is far superior to the multiplicative in terms of the accuracy measures. Also, it is easy to see that this method most accurately predicts orders for March while $ARIMA(2,0,7)(0,2,2)_{12}$ model has the lowest MAE and MAPE values. The order of SARIMA model was identified using the following procedure:

1. Finding a list of models with minimal AIC value.
2. Selection of the model with the lowest values of accuracy measures (MAE and MAPE).
3. Model validation: apply Box-Ljung test [12] to determine whether residuals are random and Dickey – Fuller test [12] to ascertain whether residues are stationary.

Neural network also has small loss functions values but the prediction of March is not very accurate in comparison with other methods.
Selected models were used to forecast monthly orders for 1, 3 and 12 months ahead, see Table 2 for results.

Table 2. Forecast for orders

Model	**Forecasts values (month)**		
	1	3	12
Holt - Winters (A) $\alpha = 0.4, \beta = 0.3, \gamma = 1$	69	81	101
$ARIMA(2,0,7)(0,2,2)_{12}$	93	105	109
Neural network	64	64	88

Results show, that seasonal ARIMA model forecast is the most optimistic while neural network is the most pessimistic. Furthermore, prediction obtained from the neural network does not have the growth trend (see Fig. 7). Therefore, neural network shows very high accuracy.

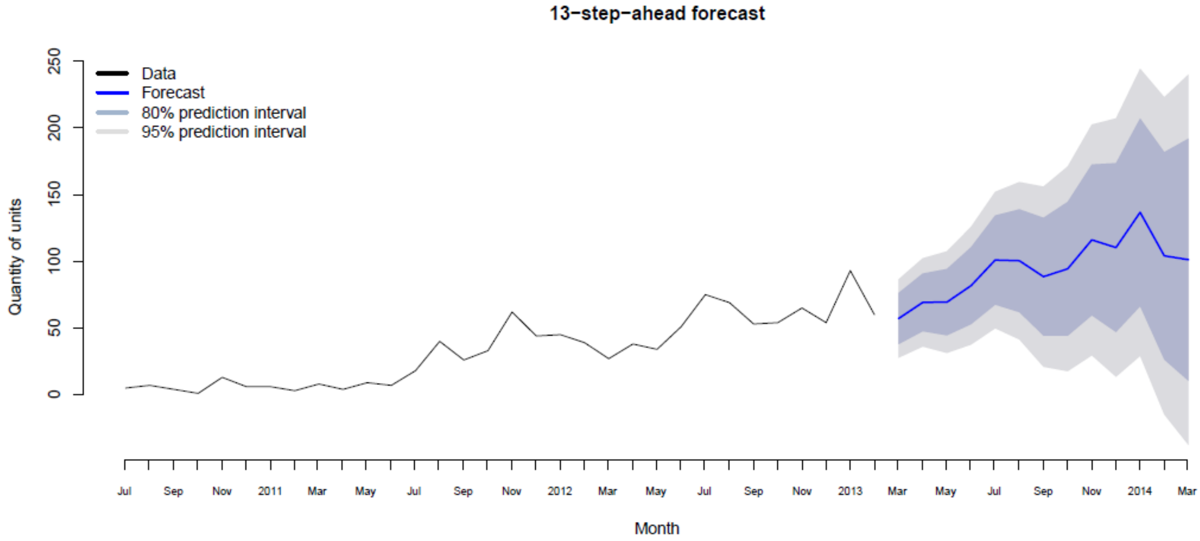

Fig. 5. Holt - Winters forecast

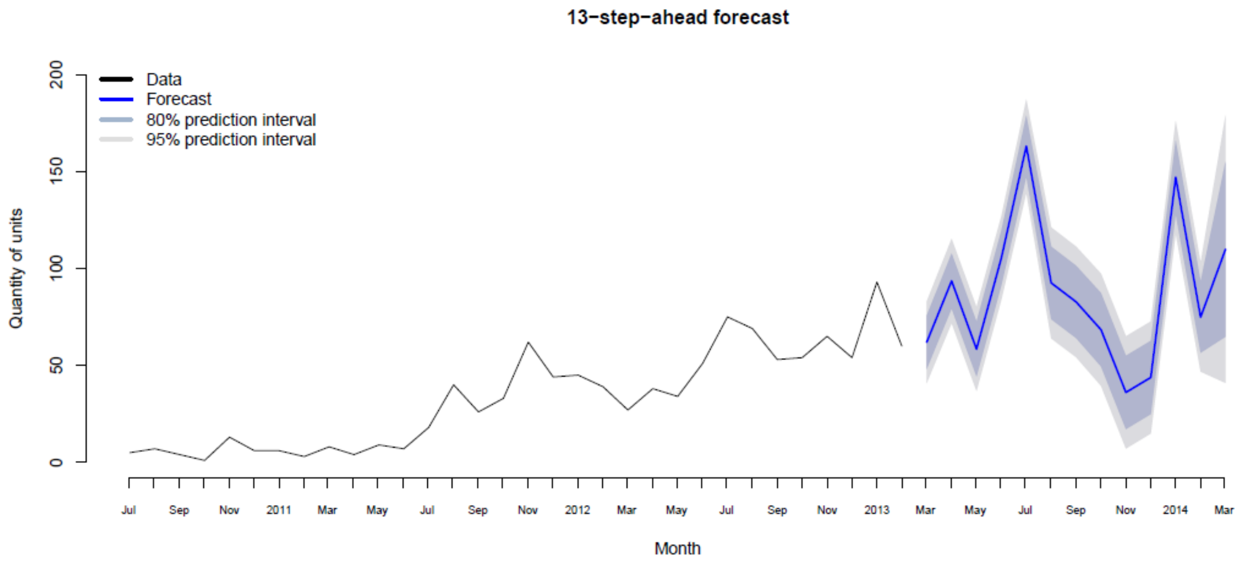

Fig. 6. $\boldsymbol{ARIMA(2,0,7)(0,2,2)_{12}}$ forecast

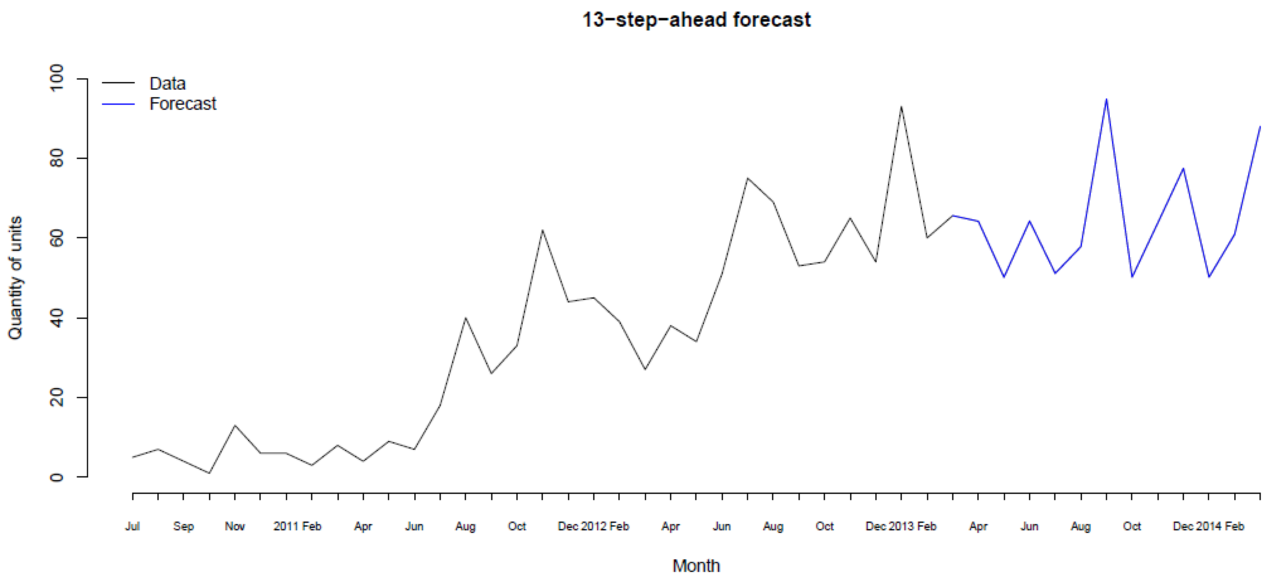

Fig. 7. Neural network forecast

Analyzing the result from Holt - Winters forecasting (see Fig. 5) it is easy to see that prediction has a growing trend but the range of prediction interval is wide, so this model is more appropriate for the first steps prediction.

4. Conclusions

Data analysis shows that orders data has trend and seasonal components: orders volume increase in July, August, and November. Based on this assumption we chose to use seasonal naïve, Holt - Winters seasonal, SARIMA and neural network methods. Results show that seasonal naive method is not suitable for orders prediction. Holt - Winters method with the set of parameters $\alpha = 0.4,\ \beta = 0.3,\ \gamma = 1$ is the most accurately predicts orders for March, 2013 but the range of prediction interval is wide, so this model is more appropriate for the short term prediction. $ARIMA(2,0,7)(0,2,2)_{12}$ model has the lowest MAE and MAPE values, and its forecast has the strongest growing trend. Neural network also has small loss functions values but prediction obtained from this method does not include the growth.

A more detailed investigation should be performed in order to estimate the application of neural networks in orders prediction, e.g. development of a more sophisticated model with different number of inputs or hidden layers. Furthermore, we are planning to test models on new data and apply different methodology: de-seasonalize the data before applying forecasting methods and then to re-seasonalize the forecasts. Moreover, we plan to perform separate analysis of business and consumer service orders, because business orders may show sudden jumps due to an agreement with bigger company (it is not really common and predictable), while consumer dynamics should be more regular.

Acknowledgment: The authors would like to thank to Dainora Kuliešienė (UAB Ekostream) for the data and cooperation and European Social Fund and the Republic of Lithuania (grant number VP1-3.2-ŠMM-01-K-02-002) for financial support.

5. References

1. Forst F. G. Forecasting restaurant sales using multiple regression and box-jenkins analysis. Journal of Applied Business Research, 8(2), 1992. p. 15–19, ISSN 0892-7626.
2. Sadowski R. J., Alon I. and Min Q. Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. Journal of Retailing and Consumer Services, 8(3), 2001. p. 147–156, ISSN 0969-6989.
3. Wang Y. W. and Changa P.C. Fuzzy delphi and back-propagation model for sales forecasting in pcb industry. Expert Systems with Application, 30(4), 2006. p. 715-726, ISSN 0957-4174.
4. Vornberger O. and Thiesing F. M. Forecasting sales using neural networks. Computational Intelligence Theory and Applications, 1226, 1997. p. 321–328. ISSN 0302- 9743.
5. Robb E. A. and Silver D. J. Using composite moving averages to forecast sales. Journal of the Operational Research Society, 53(1), 2002. p. 1281-1285. ISSN 0160-5682.
6. Tang Z., Ch. De Almeida and Fishwick P. A. Time series forecasting using neural networks vs. Box Jenkins methodology. Transactions of The Society for Modeling and Simulation International, 57(5), 1991. p. 303–310. ISSN 0037-5497.
7. Chen T. A fuzzy back propagation network for output time prediction in a wafer fab. Applied Soft Computing Journal, 2, 2003. p. 211-222. ISSN 1568-4946.
8. Chang P. C., Wang Y. W. and Tsai C. Y. Evolving neural network for printed circuit board sales. Expert Systems with Applications, 29(5), 2005. ISSN 0957-4174.
9. Athanasopoulos G., Hyndman R. J. Forecasting: principles and practice. OTexts, 2012.
10. Gelžinis A., Verikas A. Neuroniniai tinklai ir neuroniniai skaiciavimai. KTU, Kaunas, 2008. ISBN 978-9955-591-53-5.
11. Coope I. D. and Price C. J. On the convergence of grid-based methods for unconstrained optimization. SIAM Journal on Optimization, 11(4), 2000. p. 859 – 869. ISSN: 1052-6234
12. Shumway R. H. and Stoffer D. S. Time Series Analysis and Its Applications: With R Examples, Third Edition, Springer, 2011. 604 p. ISBN 978-1-4419-7864-6.