Time Series Forecasting for a Call Centre

Bahar Sennaroglu
Department of Industrial Engineering
Marmara University
Istanbul, Turkey
Department of Engineering and the Built Environment
Anglia Ruskin University
Chelmsford Essex, UK
bsennar@marmara.edu.tr, Bahar.Sennaroglu@anglia.ac.uk

Gokcag Polat
Department of Engineering Management
Marmara University
Istanbul, Turkey
gokcagpolat@gmail.com

Abstract

Forecasting is an important tool in operations planning. Determining resource requirements and scheduling of personnel at a call centre require accurate forecasts. This study includes analysis and forecasting of the number of calls of general type of customers at a private bank's call centre. These customers are referred to as mass customers and their calls constitute more than 60% of all calls at the call centre. Time series data contain double seasonality with intraday pattern for hourly calls within day and interday pattern within week. The data set of hourly calls of ten-week period is separated into a training set consisting of the first eight weeks and test set consisting of the last two weeks. The methods considered include seasonal Autoregressive Integrated Moving Average (ARIMA) and regression with ARIMA errors. Two models out of different candidate models are identified through statistical inferences of the model parameters, diagnostic checking and fit indices by using the training set for the original series and logged series. Forecasting performances of the models are evaluated with one-step ahead forecasts by using the test set and it is concluded that suggested models adequately predict hourly calls of the call centre.

Keywords
Time Series Forecasting, Call Centre, Seasonality, ARIMA, Regression

1. Introduction

It is very important for banks to provide customer-satisfying services by their call centres while running profitable call centre operations. Gans et al. (2003) explain the economic importance and complexity of telephone call centres. They state that most call centres capacity costs in general, and human resource costs in particular, account for 60-70% of operating expenses. The cost of staff members handling phone calls at call centres typically comprises 60-80% of the overall operating budget (Aksin, et al., 2007).

Forecasting is an important tool in operations planning. Determining resource requirements and scheduling of personnel at a call centre require accurate forecasts. Accurate forecasts of call arrivals are key for the achievement of optimal operational efficiency as under-forecasting leads to under-staffing and long customer waits, while over-forecasting results in money wasted on over-staffing (Ibrahim et al., 2016). Steckley et al. (2009) point out that it is worthwhile to investigate techniques for choosing staffing levels while accounting for the trade-off between service quality and staffing costs in the context of arrival rate variability as its ignorance can have a significant negative impact on system performance. There is an increasing interest in describing the call arrival patterns, modelling and

forecasting call arrivals for the call centres. Ibrahim et al. (2016) review the existing literature on modeling and forecasting call arrivals. Ibrahim and L'Ecuyer (2013) model interday and intraday dependence structures of call centre arrivals via a Gaussian linear mixed model after applying a squared root transformation to the data and extend the mixed model into a bivariate mixed model to generate forecasts that are more accurate. Based on the results from comparing univariate time series methods, Taylor (2008) suggests the use of seasonal Autoregressive Integrated Moving Average (ARIMA) or exponential smoothing for double seasonality for forecasting half-hourly arrivals at a call centre and applies a squared root transformation to the series containing periods with zero arrivals, whereas a logarithmic transformation to the other series.

This study examines forecasting performances of models developed using seasonal ARIMA and regression with ARIMA errors methods to predict hourly calls of the call centre of a bank. The detailed information on the methods used can be found in Box et al. (1994) and Hyndman and Athanasopoulos (2014).

2. Data Set

The time series data used in this study belong to the number of calls of general type of customers at a private bank's call centre in Turkey. These customers are referred to as mass customers by the bank and their calls constitute more than 60% of all calls at the call centre. The data set of hourly calls of ten-week period is separated into a training set consisting of the first eight weeks and a test set consisting of the last two weeks. The training set is used for model fitting and the test set for forecasting. The ten-week time period does not include any special day.

Figure 1 shows hourly calls at the call centre, which serves 24 hours and 7 days, the time 0:00 indicating (23:00,0:00], 1:00 indicating (0:00,1:00], and so on. There is no trend pattern, but obvious seasonal patterns. Time series data contain double seasonality with intraday pattern for hourly calls within day and interday pattern within week. Therefore, the data have two seasonal cycles: a daily cycle of length 24 and a weekly cycle of length 168 (24 hours x 7 days).
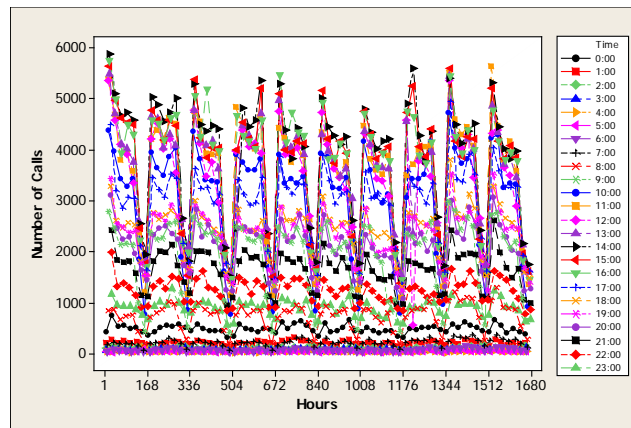

Figure 1. Hourly calls at the call centre

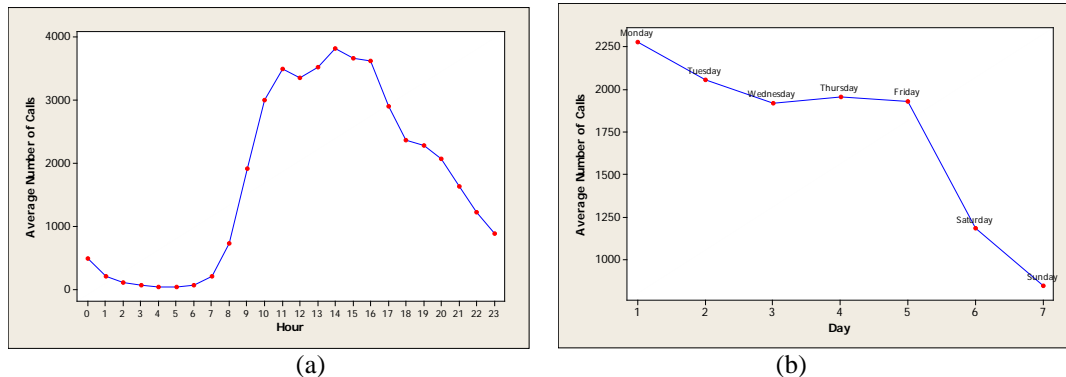(a)   (b)

Figure 2. Average number of calls (a) hourly (b) daily

Figure 2 shows seasonal patterns clearly by the average number of calls. It is observed in Figure 2 (a) that peaks occur between 10 pm and 17 pm and less number of calls arrive at the night times than the day times. Figure 2 (b) shows that peaks occur on Mondays, valleys occur on Sundays and less number of calls arrive on weekends than the working days.

3. Forecasting and Comparison

Two forecasting methods, which are seasonal ARIMA and regression with ARIMA errors, are considered. The models are fitted to both the original series and the natural-logged series using the training set. Calendar-related dummy variables for time of day (24 hours) and day of week (7 days) are defined to use for the regression part of the models as follows:

$$H_i = \begin{cases} 1 & \text{if hour is } i, i = 1, \dots, 2 3 \\ 0 & \text{otherwise} \end{cases}$$

where hour 1 corresponds to (23:00,0:00], 2 to (0:00,1:00],…, 23 to (21:00-22:00] and when all H_i's are zero, hour corresponds to (22:00,23:00] and

$$D_j = \begin{cases} 1 & \text{if day is } j, j = 1, \dots, 6 \\ 0 & \text{otherwise} \end{cases}$$

where day 1 corresponds to Monday, 2 to Tuesday,…, 6 to Saturday and when all D_j's are zero, day corresponds to Sunday.

Two models out of different candidate models are identified through statistical inferences of the model parameters, diagnostic checking and fit indices (Table 1). The residuals of the models fitted to logged series are not serially correlated but there is serial correlation with some lags of the residuals of the models fitted to original series. It can be assumed that the residuals are approximately normally distributed. Akaike information criterion (AIC) and Schwarz information criterion (SIC) are used as fit indices; their smaller values indicate the better model fits.

Table 1. Forecasting models

Data	Method	Forecasting Model	AIC	SIC
Original series	Seasonal ARIMA	$(1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3)(1 - \emptyset_{24} B^{24})(1 - \emptyset_{168} B^{168} - \emptyset_{336} B^{336} - \emptyset_{504} B^{504} - \emptyset_{672} B^{672})(1 - B^{24}) y_t = e_t$	14.08	14.12
	Regression with ARIMA errors	$y_t = c + \sum_{i=1}^{23} \beta_i H_i + \sum_{j=1}^{6} \beta_j D_j + n_t$ $(1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3)(1 - \emptyset_{168} B^{168}) n_t = e_t$	13.71	13.86
Logged series	Seasonal ARIMA	$(1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3)(1 - \emptyset_{24} B^{24})(1 - \emptyset_{168} B^{168} - \emptyset_{336} B^{336})(1 - \emptyset_{672} B^{672})(log y_t - c) = e_t$	-0.89	-0.83
	Regression with ARIMA errors	$log y_t = c + \sum_{i=1}^{23} \beta_i H_i + \sum_{j=1}^{6} \beta_j D_j + n_t$ $(1 - \emptyset_1 B - \emptyset_2 B^2 - \emptyset_3 B^3)(1 - \emptyset_{168} B^{168}) n_t = e_t$	-0.76	-0.61

Forecasts are generated using the test set consisting of the last two weeks of time series data (Figure 3). Forecasting performances of the models are evaluated with one-step ahead forecasts and the accuracy of the forecasts is measured by Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). It is observed in Table 2 that both seasonal ARIMA and regression with ARIMA errors produce better results with logged series. It is also observed that when original data is used, the regression model with ARIMA errors results in negative values for some forecasts.

The study shows that the data transformation is necessary to stabilize variance and hence satisfy model assumptions and produce better forecasts. The logarithmic transformation (log base *e*) works well for the time series data used in this study. The results suggest the use of seasonal ARIMA method and regression with ARIMA errors method for forecasting call arrivals. Regression with ARIMA errors method also has its advantage if time period includes special days; the effects of special days can be observed by adding dummy variables to the regression terms of the model. For the seasonal ARIMA method, the solution is to smooth out the special days using an adjustment method or filtering method prior to model fitting.
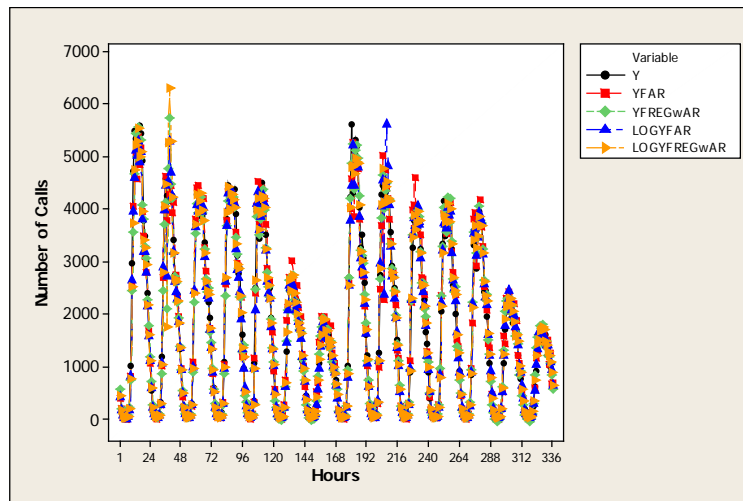
Figure 3. One-step ahead forecasts

Table 2. Performance comparison of forecasting models

Data	Forecasting Method	Label for Model	RMSE	MAE	MAPE
Original series	Seasonal ARIMA	YFAR	268.96	158.72	15.12
	Regression with ARIMA errors	YFREGwAR	243.89	141.02	19.28
Logged series	Seasonal ARIMA	LOGYFAR	251.40	128.50	10.31
	Regression with ARIMA errors	LOGYFREGwAR	255.87	124.84	11.09

4. Conclusion

It is vital to forecast arrivals accurately at call centres to be able to manage effectively the operations in terms of staffing, scheduling and quality of service. This study includes analysis and forecasting of the number of calls of general type of customers at a private bank's call centre. The forecasting models obtained using seasonal ARIMA method and regression with ARIMA errors method adequately predict hourly calls of the call centre. Since the time series data of call centres have both a daily cycle and a weekly cycle, the forecasting methods chosen should capture these seasonal patterns. As a further study, artificial neural networks can be used with the same data by adding trigonometric dummy variables for seasonal cycles and the results can be compared with the results of this study.

References

Aksin, Z., Armony, M., and Mehrotra, V., The modern call center: A multi-disciplinary perspective on operations management research, *Production and Operations Management*, vol. 16, no. 6, pp. 665-688, 2007.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C., *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1994.

Gans, N., Koole, G., and Mandelbaum, A., Telephone call centers: Tutorial, review, and research prospects, *Manufacturing & Service Operations Management*, vol. 5, no. 2, pp. 79-141, 2003.

Hyndman, R. J., and Athanasopoulos, G., *Forecasting: principles and practice*, Available: http://www.otexts.org/book/fpp, OTexts, 2014.

Ibrahim, R., and L'Ecuyer, P., Forecasting call center arrivals: Fixed-effects, mixed-effects, and bivariate models, *Manufacturing & Service Operations Management*, vol. 15, no. 1, pp. 72-85, 2013.

Ibrahim, R., Ye, H., L'Ecuyer, P., and Shen, H., Modeling and forecasting call center arrivals: A literature survey and a case study, *International Journal of Forecasting*, vol. 32, no. 3, pp. 865-874, 2016.

Steckley, S. G., Henderson, S. G., and Mehrotra, V., Forecast errors in service systems, *Probability in the Engineering and Informational Sciences*, vol. 23, no. 2, pp. 305-332, 2009.

Taylor, J. W., A comparison of univariate time series methods for forecasting intraday arrivals at a call center, *Management Science*, vol. 54, no. 2, pp. 253-265, 2008.

Biography

Bahar Sennaroglu is an Associate Professor in the Industrial Engineering Department of Faculty of Engineering of Marmara University. Dr. Sennaroglu received her MSc and PhD degrees in Industrial Engineering from Marmara University. Her research interests include design and analysis of experiments, decision techniques, forecasting techniques, project management, production and quality systems. She is an academic visitor for the academic year 2016-2017 in the Department of Engineering and the Built Environment of Faculty of Science and Technology of Anglia Ruskin University. Her ORCID ID is orcid.org/0000-0002-6809-634X.

Gokcag Polat is PhD Candidate in the Engineering Management Department of Institute of Pure and Applied Sciences of Marmara University. Gokcag Polat received his MBA degree in Business Administration from Bahcesehir University. He is a Finance Professional with 6 years of experience in working with closely banking and insurance companies.