Prediction of South Africa's Tourism Hotel Accommodation Monthly Income: Challenges in an Environment Characterised by a World Recession and a World Cup

C. Sigauke

School of Statistics and Actuarial Science, University of the Witwatersrand, Johannesburg,
Private Bag 3, Wits, 2050, South Africa
Email: csigauke@gmail.com

T.B. Darikwa

Department of Statistics and Operations Research, School of Mathematical and
Computer Sciences, University of Limpopo, P. Bag X1106,
Sovenga, 0727, South Africa

M.I. Masemola

Department of Statistics and Operations Research, School of Mathematical and
Computer Sciences, University of Limpopo, P. Bag X1106,
Sovenga, 0727, South Africa

Abstract

This paper uses official data to develop exponential smoothing models for predicting monthly total income in tourism hotel accommodation in South Africa for the period 2004 to 2011. The performance of the developed models is evaluated against a neural network model for in and out of sample predictions. Empirical results from the study show that the triple exponential smoothing model produces better forecast accuracy compared to the single and double exponential smoothing models. The neural network model outperforms all the three exponential smoothing models prior to the onset of recession in 1998. The triple exponential smoothing model produces better forecast accuracy after the onset of the world recession. This suggests that a time series undergoing structural changes may require different forecasting approaches. This study is meant to provide insight into the dynamics of the South African hotel accommodation income time series thereby stimulating future definitive work on methods of forecasting.

Keywords: Tourism accommodation; lodging industry; exponential smoothing; income.

1. Introduction

The need for tourism forecasting in destination countries, especially as a tool in management processes cannot be overemphasised (see Cummings and Busser, 1994; Archer, 1994). Accurate forecasting results have the capability of shaping public or private tourism-related policies (Chen et al., 2008). South Africa increasingly needs accurate forecasts in the tourism sector as the sector has grown in stature in recent years. This paper explores some of the challenges posed by the 1998 recession and the 2010 World Cup in obtaining accurate tourism demand forecasts and gives suggestions as to how researchers in the field can possibly overcome these challenges.

The importance of the tourism industry to South Africa can be viewed in the context of it outstripping gold in terms of income generation for the country (Tourism in South Africa website). Given the available plethora of services that range from transport, hunting and accommodation to catering, it is not a surprise that tourism is deemed by the government as a growth sector that is best placed to draw previously disadvantaged people into the mainstream of South Africa's economy. Currently one job is created for every ten foreign tourists who visit South Africa (Whitepapers, 1996) and approximately 7% of the employed people in the country are in the tourism sector or its auxiliary industries (Gil-Alana, 2011 : after, (Tourism Business Council of South Africa (TBCSA). The highest contributor to South Africa's tourism receipts is the accommodation services.

Tourist accommodation establishments in South Africa are assessed according to the nine types of

accommodation they provide, namely: (1) bed-and-breakfast; (2) guesthouse; (3) hotel; (4) self-catering; (5) backpacker and hostelling; (6) caravan and camping; (7) country house; (8) restaurants and; (9) meetings, exhibitions and special events. The hotel industry generates the most income (see Figure 1) and will be at the centre of this study. There is currently no known research that has specifically focused on forecasting hotel accommodation income in South Africa.

Lundberg et al. (1995) indicate that hotels, like most other business activities, are subject to whims of the economy with particular variables such as occupancy rates rising during prosperity and falling in recession. There has been a global recession since 2008 and a World Cup in South Africa in 2010, which events have opposing effects on economy in general and the hotel business, in particular.

Forecasting of tourist accommodation demand and forecasting guest nights in the lodging industry is discussed in literature (Choi et al., 1999; Sorenson, 1999; Krakover, 2000; Choi, 2003; Koenig and Bischoff, 2004a, b; Athanasopoulos and Hyndman, 2008; Lim et al., 2009; among others). Various statistical methods ranging from classical time series, regression, exponential smoothing, SARIMA and neural networks have been used in tourist accommodation demand forecasting. Neural networks have been particularly used to predict tourism expenditure (Law, 1998; Palmer et al., 2006; among others).

Burger et al. (2001) provide a forecasting guideline for researchers who lack information to apply structural models for South Africa's tourism demand data. Their study reveals that neural network methods give the best forecasts when applied to tourism demand data at a metropolitan level in Durban in comparison with such models as naïve, moving average, single exponential smoothing, ARIMA, multiple regression and genetic regression.

This paper develops forecasting models which will be used for predictions of monthly income in hotels in tourism accommodation in South Africa in the absence of data for structural modelling. The rest of the paper is organized as follows. In Section 2 we describe the data, while the models are discussed in Section 3 and the empirical results are presented in Section 4. A comparative analysis of the models is given in Section 5 and Section 6 concludes.

2. Data

Official monthly income data for tourism accommodation in South Africa for the period 2004 to 2011was obtained from the South African Tourism website. The data is for hotels, caravan parks and camping, guest houses and guest farms, and other accommodation. The breakdown of income generated by tourism accommodation in South Africa for the period September 2004 to September 2011 is shown in Figure 1. The variable H stands for hotels, CPCS for caravan parks and camping sites, GHGF for guest houses and guest farms while OA is for other accommodation.
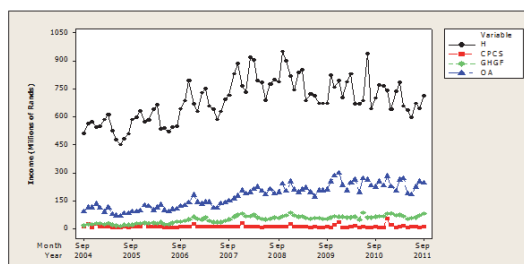
Figure 1. Breakdown of income generated by tourism accommodation in South Africa.

Figure 1 shows that hotels have the largest income while caravan parks and camping sites had the lowest monthly income during the sampling period. The bold solid lines indicate an upward moving trend of the income generated by hotel accommodation since 2004, up until the end of 2008 when the recession set in. This was, thereafter, followed by a general downward moving trend as the recession took hold. The "tweak" in June of 2010 is due to the mega event which is the World Cup. Thus, there is a general structural change in the time series data of the total monthly income of the tourist hotel accommodation induced by the onset of the recession in 1998 and, possibly, by the World Cup of 2010.

3. The Models

3.1 Historical note on exponential smoothing

Exponential smoothing can be traced back to Second World War by Robert Brown, who worked as an operations research analyst for the United States. After the war, Brown (1959) applied his work on simple exponential smoothing to inventory management problems. Brown's work was extended by Charles Holt who came up with an exponential smoothing model that could handle an additive trend (Holt, 1957). Winters (1960) implemented and tested Holt's models on empirical data, and since then, this early form of exponential smoothing is referred to as the Holt-Winters method. Hyndman et al. (2002) further extended this framework to cater for multiplicative and damped additive trends. It was observed that "exponential smoothing methods are optimal for a very general class of state-space models that is in fact broader than the ARIMA class" (Gardner, 2006).

3.2 Single exponential smoothing (SES): an adaptive approach

The single exponential smoothing (SES) model developed by Brown (1959) can be written as:

$$L_t = \alpha X_t + (1 - \alpha)L_{t-1} \tag{1}$$

where $\alpha\ (0 \leq \alpha \leq 1)$ is the smoothing parameter, X_t is the time series observation at time t. Equation (1) can also be written as

$$L_t = L_{t-1} + \alpha(X_{t-1} - L_{t-1}) \tag{2}$$

where $(X_{t-1} - L_{t-1})$ is the forecast error. The forecast L_{t+1} is given as

$$L_{t+1} = \alpha X_t + (1 - \alpha)L_t = L_t + \alpha(X_t - L_t) \tag{3}$$

3.3 Holt's linear method (double exponential smoothing model)

Holt (1957) extended single exponential smoothing to linear exponential smoothing to allow forecasting of data with trends. The forecast for Holt's linear exponential smoothing is found using two smoothing constants alpha, beta (with values between 0 and 1) and two equations:

$$\left. \begin{array}{l} L_t = \alpha X_{t-1} + (1 - \alpha)(L_{t-1} + T_{t-1}), 0 \leq \alpha \leq 1 \\ T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, 0 \leq \beta \leq 1 \end{array} \right\} \tag{4}$$

where T_t represents the trend term, β represents the trend coefficient and the other terms are as defined in equation (1). The $h - \text{step}$ ahead forecast, $\hat{X}_t(h)$ at time t is given as

$$\hat{X}_t(h) = L_t + hT_t \tag{5}$$

3.4 Holt- Winters' trend and seasonality method (triple exponential smoothing model)

Holt's method was extended by Winters (1960) to capture seasonality. The Holt-Winters' method is based on three smoothing equations-one for level, one for trend and one for seasonality. It is similar to Holt's method with one additional equation to deal with seasonality. Seasonality is included in the model either multiplicatively or additively. We present the multiplicative Holt-Winter's exponential smoothing model. It is given as:

$$\left. \begin{array}{l} L_t = \alpha \frac{X_t}{S_t} + (1 - \alpha)(L_{t-1} - T_{t-1}) \\ T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \\ S_t = \gamma \frac{X_t}{L_t} + (1 - \gamma)S_{t-s} \end{array} \right\} \tag{6}$$

where s is the seasonal period, $\frac{X_t}{L_t}$ captures the seasonal effects, α is the smoothing parameter, β is the trend coefficient and γ is the seasonality coefficient. The $h - \text{step}$ ahead forecast, $\hat{X}_t(h)$ at time t is given as:

$$\hat{X}_t(h) = (L_t + hT_t)S_{t+h-s} \tag{7}$$

4. Empirical Results and Discussion

The component analysis for income from hotels is illustrated in Fig. 2. Figure 2(c) and 2(d) shows that there is a sudden surge in income in June 2010 for hotels. This is due to the

FIFA 2010 world cup which was hosted by South Africa. This figure is removed and replaced by the mean income for hotels for the sampling period 2004 to 2011, excluding the outlier.
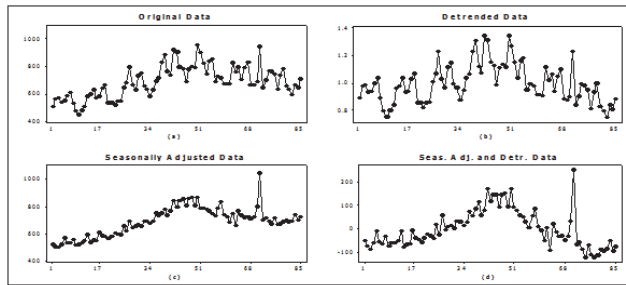
Figure 2. Component analysis for income from hotels.

The mean absolute error (MAE) and the mean absolute percentage error (MAPE) are used to assess the accuracy of the models under discussion. A summary of the estimates of the parameters of the exponential smoothing models, i.e. single, double and triple together with some important accuracy measures are given in Table 1. The triple exponential smoothing model produces better out of sample forecast accuracy with a MAPE of 4.9% and a MAE of 34.6. This is not surprising given the strong seasonal variation inherent in the data. The multiplicative triple exponential model gives better results as it is able to capture the interaction of the seasonal, secular, and cyclical trends.

Table 1. Summary results of exponential smoothing models for hotels.

	Single	Double	Triple
α	0.400	0.300	0.500
β		0.100	0.100
γ			0.100
MAE	61.83	64.70	34.60
MAPE	8.887%	9.387%	4.956%

4.1 Comparative analysis of the Holt-Winters method and neural network model

Both the exponential smoothing and neural network models are developed using data for the period 2004 up to the onset of the recession (September 2004 to March 2008). This procedure is repeated using the data for the period September 2004 to December 2010 to include data after the recession. A 12-step forecast is done for each of the models and an accuracy evaluation is done using MAPE and MAE. Table 2 shows a comparative analysis of the triple exponential smoothing model with results from using a neural network model.

Table 2. Comparison of the Triple Exponential Smoothing and Neural Network Models.

Performance Criteria	Prior Recession		Post-Recession	
	Triple Exponential Smoothing	Neural Network	Triple Exponential Smoothing	Neural Network
MAPE	8.98%	6.15%	3.67%	11.89%
MAE	72.15	49.52	24.77	78.03

Table 2 shows that prior to the onset of the recession the neural network model outperforms the triple exponential smoothing. The neural network model has the least values of 49.52 for MAE and 6.15% for MAPE compared to a MAE of 72.15 and a MAPE of 8.98 for exponential smoothing, respectively. But, after the onset of the recession and the World Cup mega event neural network models are outperformed by the triple exponential smoothing model since the latter had the least MAE (24.77) and the least MAPE (3.67%). Thus the neural network performs better than the exponential smoothing method before the onset of the recession but is outperformed by the triple exponential technique after the cyclical effects of the recession.

Figure 3 further illustrates the extent to which the forecasted values for both the exponential and neural network models fit the actual data. This is done for the model using data before the recession (panel A) and the data after the recession (panel B). Generally forecasting accuracy worsens with time as would be expected. Overall, on one hand, it is clear that the neural network model gives better forecast results before the recession than the exponential model. This is in line with the findings of Burger *et al.* (2001). On the other hand, after the onset of the recession the triple exponential smoothing model produces better forecast accuracy in forecasting hotel accommodation monthly income. The exponential smoothing model gives a better estimate after the onset of the recession as it is able to identify structural changes in the data as it gives weight to the most recent events. On the other hand, the failure by the neural network to read structural changes to the data can be attributed to the bias within the model due to its training in the previous months. Predictions are based on previous historical patterns only and does not take care or give more weight to recent changes that may have occurred in the data. It can be seen from panel B of Figure 4 that the neural network model tends to overestimate the actual values that have been depressed by the recession.
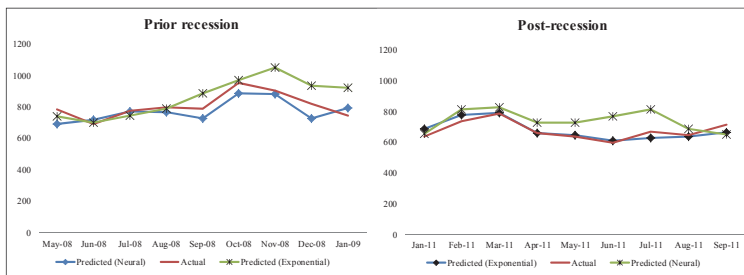

Figure 3. Actual and predicted for hotels using triple exponential smoothing and neural network models.

5. Conclusion

This paper explored the classical time series techniques to forecast monthly tourism accommodation income in South Africa. The exponential smoothing models were developed for predicting income from tourism accommodation. The out of sample performance of the developed models was evaluated by comparing them with a neural network model for the period prior to the onset of the recession in 2008 and for the months after the recession. Empirical results from the study shows that the triple exponential smoothing model produces better forecast accuracy compared to the single and double exponential smoothing models. This would be expected as the former is designed to capture seasonality that is present in the data.

This study has shown that there are situations in which neural network models are not the best models. Thus no one model can be expected to outperform all the other models in all situations. This is not a new discovery, but nonetheless a very important finding. It was found that the neural network proved to be a better model when compared with the triple exponential smoothing technique before the beginning of the recession but after the recession the reverse was true. This can be attributed to the inability of the neural network model to quickly adapt to new changes in the structure of the time series as it is trained on historical patterns whereas the triple exponential smoothing method can easily and quickly adapt to the new changes as it gives more weight to the most recent observations.

The models used in this study were so chosen as to be able to capture the seasonal variations evident in the data and the structural changes since the recession of 2008. It is not easy to ascertain the impact of the 2010 World Cup and the 2008 recession but future research can be designed to capture these dynamics in the South African tourism hotel accommodation monthly income with a view to give improved forecasting approximations.

Additionally, improved forecasts may be obtained through the use of advanced cause and effect structural models that may provide better capability of capturing the dramatic seasonal variations and the structural changes induced by the recession of 2008 and the 2010 World Cup. This may imply application of adaptive models such as multivariate adaptive regression splines that can be used to capture piece-wise changes in the trends inherent in this series. Combining forecasts is also a viable alternative as work by Palm and Zellner (1992) showed that a simple average of individual forecasts may work well in practice, where equal weights are assigned to individual forecasts. The work by Athanasopoulos and Hyndman (2008) in which they predict tourism demand using state space models may be worth pursuing as the models capture shocks in the time series such as the impact of the Bali bombings on the Australian

domestic tourism industry.

It may be desirable to make use of complex models but the difficulties in handling such models coupled with the scarcity of data on the drivers of income generation in the hotel accommodation industry in South Africa deem the results of the triple exponential smoothing model of this study good first approximations to forecast tourist hotel accommodation monthly income. These short to medium term forecasts can be used for planning and management purposes. Subsequent forecasts can be obtained by regularly updating the model from time to time as actual realisations are obtained as suggested by Burger *et al.* (2001).

6. Acknowledgements

References

Archer, B.H. 1994. *Demand forecasting and estimation*. In J.R.B. Ritchie & C.R. Goeldner (Eds.), Travel, tourism and hospitality research (4th ed., pp. 105-114). New York: Wiley.

Athanasopoulos, G. & Hyndman, R.J. 2008. Modelling and forecasting Australian domestic tourism, *Tourism Management*, 29: 19-31.

Burger, C.J.S.C., Dohnal, M., Kathrada, M. & Law, R. 2001. A practioners' guide to time-series methods for tourism demand forecasting– a case study of Durban, South Africa, *Tourism Management*, 22(4): 403-409.

Brown, R.G. 1959. *Statistical forecasting for inventory control*. New York: McGraw-Hill.

Chen, R.J.C., Bloomfield, P. & Cubbage, F.W. 2008. Comparing forecasting models in tourism, *Journal of Hospitality and Tourism Research*, 32(1): 3-21.

Choi, J.G. 2003. Developing an economic indicator system (a forecasting technique) for the hotel industry, International *Journal of Hospitality Management*, 22: 147-159.

Choi, J.G., Olsen, M.D., Kwansa, F.A. & Tse, E.C. 1999. 'Forecasting industry turning points: the US hotel industry cycle model', *International Journal of Hospitality Management*, 18: 159-170.

Cummings, L.E., Busser, J.A. 1994. Forecasting in recreation and park management: Need, substance, and reasonableness, *Journal of Park and Recreation Administration*, 12(1): 35-50.

Gardner, E.S. 2006. Exponential smoothing: The state of the art – Part II, *International Journal of Forecasting*, 22: 637-666.

Gil-Alana, L.A. 2011. Tourism in South Africa- time series persistence and the nature of shocks: are they transitory or permanent?, *African Journal of Business Management*, 5: 4686-4698.

Holt, C.C. 1957. *Forecasting seasonals and trends by exponentially weighted moving averages*. ONR Memorandum, vol. 52. Pittsburgh, PA7 Carnegie Institute of Technology.

Hyndman, R.J., Koehler, A.B., Snyder, R.D. & Grose, S. 2002. 'A state space framework for automatic forecasting using exponential smoothing methods'. *International Journal of Forecasting*, 18: 439-454.

Hyndman, R.J., Koehler, A.B., Ord, J.K. & Snyder, R.D. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*. Springer-Verlag, Berlin, Heidelberg

Koenig, N. & Bischoff, E.E. 2004a. Tourism demand patterns in turbulent times: analysing Welsh accommodation occupancy rate for 1998–2001, *International Journal of Tourism Research*, 6,:205-220.

Koenig, N. & Bischoff, E.E. 2004b. Analyzing seasonality in Welsh room occupancy data. *Annals of Tourism Research*, 31: 374-392.

Krakover, S. 2000. Partitioning seasonal employment in the hospitality industry, *Tourism Management*, 21: 461-471.

Law, R. (1998). Room occupancy rate forecasting: a neural network approach, *International Journal of Contemporary Hospitality Management*, 10: 234-239.

Lim, C., Chang, C. & McAleer, M. 2009. 'Forecasting h(m)otel guest nights in New Zealand', International *Journal of Hospitality Management*, 28: 228-235.

Lundberg, D., Drishnamoorthy, M. & Stavenga, M. 1995. *Tourism Economics*, Wiley, New York.

Palmer, A., Montano, J.J. & Sese, A. 2006. Designing an Artificial neural network for Forecasting Tourism Time Series. *Tourism Management*, 27: 781-790.

Sorenson, N. 1999. Modelling the seasonality of hotel nights in Denmark by county and Nationality, *Tourism Economics*, 5: 9-23.

Winters, P.R. 1960. Forecasting sales by exponentially weighted moving averages, *Management Science*, 6: 324-342.

http://www.info.gov.za/whitepapers/1996/tourism.htm Online Accessed 3-12-2011.

http://www.tourismrsa.com Online Accessed 4-12-2011.

http://en.wikipedia.org/wiki/Tourism_in_South_Africa Online Accessed 8-01-2012.