African Journal of Hospitality, Tourism and Leisure, Volume 7 (4) - (2018) ISSN: 2223-814X

Hierarchical forecasting of tourist arrivals at the Victoria Falls Rainforest, Zimbabwe

Tendai Makoni* and Delson Chikobvu
Department of Mathematical Statistics and Actuarial Science
University of the Free State, South Africa
tpmakoni@gmail.com *
Chikobvu@ufs.ac.za

Corresponding author*

Abstract

In this paper, the aim was to model and forecast the Victoria Falls Rainforest tourism demand using hierarchical forecasting methods. The methods are capable of capturing tourism data dynamics, producing precise, coherent and sensible short-term tourism forecasts. The monthly tourism data from the Zimbabwe Parks and Wildlife Management Authority (ZPWMA) is used. The data are disaggregated according to tourism source (international, regional and locals) and further disaggregated by age (adult and children). The ZPWMA needs future tourism forecasts for all tourism sources, adult and children tourists for planning purposes. Accurate and coherent forecasts for both adult and children tourists are needed for hotel accommodation including family rooms, child minders, recreational facilities, food types at the site to mention a few. Time series plots of the disaggregated tourism data indicated more foreign adult tourists and fewer children tourists at the site. To model and forecast the tourism demand, the top-down, bottom-up, and the optimal combination approaches are adopted. The exponential smoothing techniques (EST) and the autoregressive integrated moving average (ARIMA) methods are the forecasting methods considered under the mentioned approaches. Accuracy measures indicated the bottom-up approach under ARIMA models as the best approach to the data and produced sensible future tourism forecasts. An overall slow increase in future tourist arrivals is suggested by the models, with significant numbers expected from foreign adult tourists and less children tourists. Effective marketing strategies targeting the younger generation from abroad need to be implemented as this will impact positively to our economy in the near future. The educational issues around the rainforest have the potential to attract school tours from the rest of the world. The methods used are of significant benefit to the ZPWMA, tourism managers, investors, government and policy makers. These methods are strongly recommended as they give economically feasible solutions.

Keywords: Tourist arrivals; bottom-up; top-down; optimal-combination

Introduction

Future accurate tourist forecasts are important to tourism business practitioners and policy makers (Chatziantoniou., 2016). Both the private and public sectors are in need of timely, accurate tourism forecasts considering social and economic tourism impacts (Khadivi and Ramakrishnan, 2016). Accurate tourism forecasts are needed at national level as well as at particular tourist destination sites and they necessitate effective destination tourism site management. Future planning is essential in the tourism industry, especially for managers so as to minimize failure risk.

Being able to predict future values is one of the important properties of time series models (Jonathan and Kung-Sik, 2008). Tourism destination planning through better forecasting may result in sustainable tourism development and this can be achieved through making use of time series or statistical methods that are capable of giving accurate tourism projections. Both short-term and long-term investment opportunities either in tourism equipment and infrastructure all depends on accurate tourism forecasts. Accurate tourism forecasts at different levels provide useful feedback to tourism managers as they support decision making, cost saving as well as identification and capitalization on opportunities (Athanasopoulos et al., 2017).

Athanasopoulos et al., (2011) reports the tourism industry as the most speedily growing industries. Furthermore, Armindo, João and Álvaro (2015) indicated that tourism activities have high benefits to many regions. The tourism industry contributes immensely to the growth of various countries in the world (Yao et al., 2014) and Zimbabwe is not an exception. According to Jones and Ohsawa (2016) and the United Nations World Tourism Organisation (2015), nature tourism is among the fastest growing tourism sectors, and the growth is also being noticed in Zimbabwe. Agriculture is the backbone of the Zimbabwean economy, although mining and tourism contribute significantly (Zhou, 2018). The Victoria Falls rainforest, Lake Kariba, Gonarezhou national park, Hwange national park, Chinhoyi Caves and Great Zimbabwe monuments are among the tourist attractions in Zimbabwe. These tourist attraction sites contribute significantly to the growth of the country's tourism industry.

The Victoria Falls is one of the Seven Wonders of the World, is the epicenter of Zimbabwean tourism and is the most visited area (Zhou, 2018). The town hosts the Victoria Falls Rainforest that is managed by the Zimbabwe Parks and Wildlife Management Authority (ZPWMA). The Zimbabwean tourism industry is customer driven (Zhou, 2018), therefore effective tourist destination or attraction site branding and management is important for all tourism sources and age groups including family tourists and school tours. The Annual Tourism Trends and Statistics Report (2015) indicated that the Victoria Falls Rainforest tourism site as the most popular tourist attraction site for all tourism sources and age groups. The same document reported that the 70% of the tourists are foreigners and the other 30% are locals. The Victoria Falls Rainforest tourism site retained top position in 2015 and was visited by more than half of the tourists who visited the national parks and the country (Annual Tourism Trends and Statistics Report, 2015). According to the Annual Tourism Trends and Statistics Report (2015), the Victoria Falls Rainforest is frequented by local group visitors (families, churches and schools) with schools visiting seasonally, especially during the summer school term.

In Kenya, Job and Paesler (2013) concluded that the most protected tourist attraction sites in developing countries are frequented by wildlife tourists since they are characterized by various attractive animals and plant species. The Victoria Falls Rainforest is one of the protected tourist attraction sites in Zimbabwe that hosts uncommon natural attractiveness species in the region. Furthermore, protected tourist attraction sites offer coolness and attractive natural environments which is also educational in nature and can be enjoyed by tourists (Balmford et al., 2015) and the Victoria Falls Rainforest is one of a kind.

The ZPWMA groups the Victoria Falls Rainforest tourist arrivals according to tourism sources. Tourists from the Southern African Development Community (SADC) excluding Zimbabweans are referred as to regional tourists. Tourists from outside SADC are referred to as international tourists. The data are further disaggregated according to age (Adult tourists and Children tourists). According to the ZPWMA criteria, those aged 6 to 12 years are considered as children tourists and attract specific tariff fees at a discount. Those above 12 years are considered as adults for the purposes of paying entry fees. Those under 6 may be too young to enjoy the beauty around them and entry is free. The young minds may still be intrigued by nature. The entry of these young children into a park may not be recorded. For the Victoria Falls Rainforest, adult and children tourists who visit the site are a profitable source of revenue since they pay site viewing fees among other things. In Nepal, Dhakal (2014) analyzed tourism trends and among the tourism variables considered, age was among the variables. The study indicated adult and youth as popular tourists in the country and recommended the improvement and increase in recreational facilities for the groups. Underlying tourism trends for both adults and children tourists are also significant to transport and accommodation management at any tourism site. School children often share accommodation and transport as a group. Buses and bunk beds prove to be very economical for such groups.

The ZPWMA will need future tourism forecasts for all tourism sources as well as for adult tourists and children tourists for planning purposes. Forecasts for both adult and children tourists are very informative to the government, hotel accommodation and infrastructure development managers. Enough recreational facilities, accommodation facilities and food types at the site, particularly for children tourists depend on precise forecasts as they need special attention to suit the child’s needs. Currently there is no statistical model being used by the ZPWMA in projecting future tourist arrivals at its sites. They are using judgmental methods which are not good enough to produce accurate forecasts.

Accommodation and recreation facilities at the site need to be proportional to the nature or type of tourists that arrive. Meaningful information for marketing and development purposes is gathered if tourism age is considered (Perera, 2017). The ZPWMA and other responsible authorities need to be aware about future sources of tourist arrivals and socio-economic factors (gender and age) in order to provide adequate accommodation and ablution facilities among other things. This paper is aimed at finding a suitable hierarchical forecasting approach to forecast tourist arrivals to the Victoria Falls Rainforest tourism site. The approach will play a pivotal role in decision making purposes. According to Athanasopoulos et al., (2017) and Wickramasuriya et al., (2018), hierarchical forecasting improves forecasting accuracy and at the same time reconciles forecasts for easy decision making purposes.

A hierarchical forecasting approach is adopted to make sure the forecasts balance across the hierarchy. Traditional time series methods like ARIMA and exponential smoothing methods are the commonly used in modelling and forecasting tourism demand worldwide. Based on the authors’ knowledge, no African country (or developing country) has adopted the hierarchical forecasting approaches in the tourism industry while in developed countries, only two tourism researches used the hierarchical forecasting approach (Hyndman et al., 2011 and Athanasopoulos et al., 2008). Sensible tourism forecasts were produced by these approaches. This research will contribute to the board of knowledge on the adoption of coherent and precise hierarchical forecasting approaches in African and developing countries like Zimbabwe.

It is beneficial to have precise tourism demand at the Victoria Falls Rainforest in order to make the right amount of touring vehicles and tourist viewer assistants. Knowing tourist statistics, like the carrying capacity and future tourist arrivals for the Victoria Falls Rainforest is vital for resource allocation, infrastructure development, monitoring ecosystem and policy formulation. Apart from this, necessary interventions at the Victoria Falls Rainforest are needed in order to be able to safeguard the environmental assets at the site as they are the pillars tourist attraction. Informed decisions and interventions can be achieved through the use of scientific and statistical methods.

Mutanga et al., (2017) used line and bar graphs in analyzing tourism trends at Gonarezhou National Park in Zimbabwe and noted an increase between the years 1991 and 1998, 2008 and 2014 with a major decline being recorded in the years 1999 and 2007. The research concluded that tourism is volatile; therefore managing the destination’s image will result in wildlife tourism success. Good management of the Victoria Falls Rainforest will ensure the success of the tourism industry.

Athanasopoulos, Ahmed, and Hyndman (2009) indicated that hierarchical forecasting is applicable to a vast number of scenarios, including but not limited to production, tourism, mortality and national economic accounts. Shang (2017) used the bottom-up and optimal combination methods in forecasting the regional infant mortality in Australia and obtained informative coherent forecasts. Shang and Haberman (2017) adopted grouped forecasting

African Journal of Hospitality, Tourism and Leisure, Volume 7 (4) - (2018) ISSN: 2223-814X

methods and used Japanese mortality data in pricing annuity. The methods improved forecast accuracy and helped estimating annuity prices in insurance and pensions industries. However, Athanasopoulos and Hyndman (2008) disaggregated the Australian tourism data according to geographical regions and adopted the hierarchical forecasting approach. Accurate short-term forecasts were obtained using the optimal combination and top-down approach (forecast proportions). This approach can be adopted by the ZPWMA in order to come up with precise and coherent tourism forecast that will help in planning and decision making purposes.

Hyndman et al., (2011) proposed a new hierarchical forecasting approach and applied it in forecasting Australian tourism demand after data was aggregated according to geographical region and purpose of travel. The proposed approach makes use of a regression model to combine and reconcile forecasts produced individually at each and every hierarchical level. The conclusion indicated that the forecasts from the new approach were better than that obtained using the bottom-up and top-down approaches.

Hierarchical forecasting methods

In this paper, the Victoria Falls Rainforest tourism data is disaggregated by tourism source (international, regional and locals). The data are further disaggregated according to age (adults and children), hence a two level hierarchical model is appropriate. A two level hierarchical tree diagram for this data set is shown in Figure 1 below.

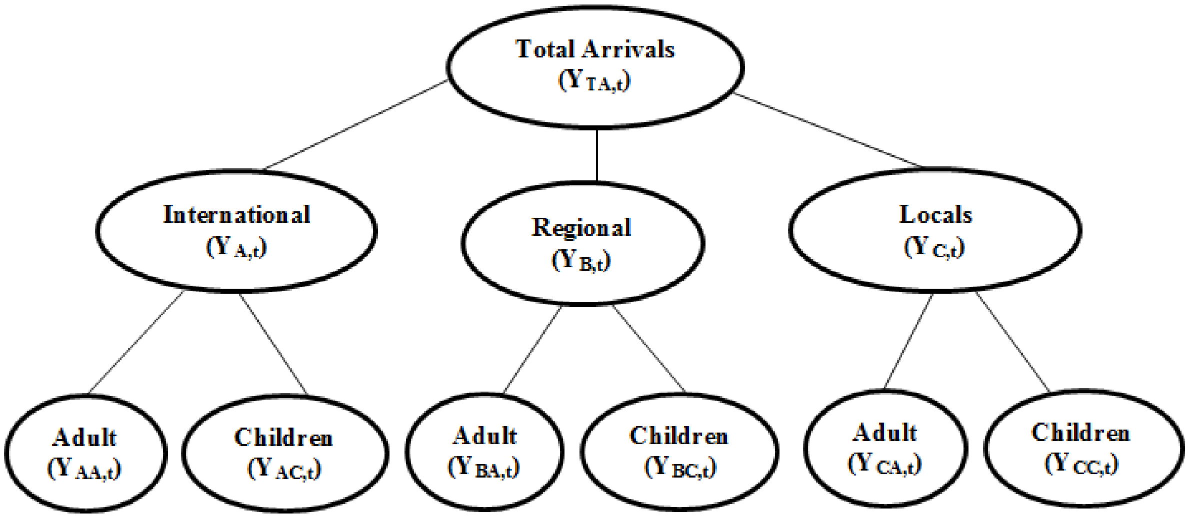

Figure 1: A two level hierarchical tree diagram for Victoria Falls Rainforest tourist arrivals.

From Figure 1, level 0 represents the completely aggregated tourism data named "Total Arrivals" series are denoted by $Y_{TA,t}$ where $t = 1, 2, 3, \ldots, 84$ and are obtained by adding all the series at level 1 and level 2 as indicated by Hyndman et al. (2011). Level 1 represents tourism data disaggregated according to tourism source and level 2 represents tourism data disaggregated according to age (adult and children). Level 1 and level 2 series can be denoted by $Y_{i,t}$, where i denotes the node in the hierarchical tree diagram. The tourism data consists of 84 monthly observations (t=1, 2,…, 84). Forecasts for each level will be produced based on the bottom-up, top-down and optimal combination approaches and then evaluated using accuracy measures such as the MASE, MAPE and RMSE. The approach with lower accuracy measure will be used to produce future forecasts for the site.

According to Morgan (2015), matrix and vector notation are most useful when dealing with hierarchical time series, hence the approach is adopted in this paper. Le $\boldsymbol{Y}_t$ and $\boldsymbol{S}_{10X6}$ be the vector of the tourism data and a summing matrix storing the hierarchical structure shown in Figure 1 respectively.

$$\boldsymbol{Y}_t = [Y_{TA,t}, Y_{A,t}, Y_{B,t}, Y_{C,t}\ Y_{AA,t}, Y_{AC,t}, Y_{BA,t}, Y_{BC,t}, Y_{CA,t}, Y_{CC,t}]' \quad (1)$$

$$\boldsymbol{S} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

Making use of the summing matrix (S), equation 1 can be expressed as

$$\boldsymbol{Y}_t = \boldsymbol{S}\boldsymbol{Y}_{2,t} \quad (3)$$

The bottom-up method

The approach is concerned with producing individual base forecasts at the lower level of the hierarchy and combining the forecasts upwards through $\boldsymbol{S}$. With this approach all dynamics of the series will be captured and no loss of information is encountered, although the bottom level series may be noisy (Morgan, 2015). Applying this approach to the hierarchy shown in Figure 1, it means the approach starts by producing h-step-ahead individual base forecasts for, $\ddot{Y}_{AA,h}$, $\ddot{Y}_{AC,h,}$ $\ddot{Y}_{BA,h}$, $\ddot{Y}_{BC,h}$, $\ddot{Y}_{CA,h}$, and $\ddot{Y}_{CC,h}$ which are for level 2. These forecasts will be aggregated in order to obtain the h-step-ahead forecasts for the higher level. This mean level 1 h-step-ahead forecasts ($\tilde{Y}_{A,h}$, $\tilde{Y}_{B,h}$ and $\tilde{Y}_{C,h}$) will be given by:

$$\tilde{Y}_{A,h} = \ddot{Y}_{AA,h} + \ddot{Y}_{AC,h} \quad (4)$$

$$\tilde{Y}_{B,h} = \ddot{Y}_{BA,h} + \ddot{Y}_{BC,h} \quad (5)$$

$$\tilde{Y}_{C,h} = \ddot{Y}_{CA,h} + \ddot{Y}_{CC,h} \quad (6)$$

Generally, making use of the summing matrix, the h-step-ahead forecasts of the hierarchy exhibited in Figure 1 using the bottom-up approach will be given by:

$$\tilde{Y}_h = S\ddot{Y}_{K,h} \quad (7)$$

where $k = 0, 1, 2$.

The top-down method

The approach is concerned with producing individual base forecasts at the top level of the hierarchy and disaggregating them downwards through the use of proportions (Athanasopoulos et al., 2009). Historical data is used in the calculation of the proportions and the approach has the ability to yield reliable forecasts for the aggregate levels (Morgan, 2015). The three approaches in calculating the proportions as explained by Hyndman and Athanasopoulos (2014) are the average historical proportions, proportions of historical averages and forecasted proportions.

Average historical proportions

The proportions are given by the formula:

$$p_i = \frac{1}{N}\sum_{t=1}^{N}\frac{Y_{i,t}}{Y_t} \tag{9}$$

where $i = 1,2, \ldots, m_k$. According to Morgan (2015), every proportion reveals the average of the historical proportions of the bottom level series over time relative to the aggregated series (Y_t) for $t = 1, 2, 3, \ldots, N\ (N = 84)$.

Proportions of historical averages

The proportions are given by the formula:

$$p_i = \left(\sum_{t=1}^{N}\frac{Y_{i,t}}{N}\right) / \left(\sum_{t=1}^{N}\frac{Y_t}{N}\right) \tag{11}$$

where $i = 1,2, \ldots, m_k$. The average historical bottom level series $(Y_{i,t})$value relative to the average value of the top level aggregated series (Y_t) over time, $t = 1, 2, 3, \ldots, N\ (N = 84)$ is used.

Forecasted proportions

Individual base forecasts for all the series in the hierarchy are generated. The proportions of the h-step-ahead forecast to the aggregate of all the h-step-ahead base forecasts for each level (top level to bottom level) are the calculated. The process is repeated for every node in the hierarchy starting from the top moving downwards. For a K level hierarchy, the forecasted proportions are given by:

$$p_i = \prod_{l=0}^{K-1}\frac{\hat{y}_{i,t}^{(l)}}{\hat{S}_{i,t}^{(l+1)}} \tag{12}$$

where $i = 1,2, \ldots, m_k$. $\hat{y}_{i,t}^{(l)}$ represents the h-step-ahead base forecast of the series corresponding to the node which is l levels above i. The sum of the h-step-ahead base forecasts under the node that is l levels above i and are directly linked to that node is $\hat{S}_{i,t}^{(l+1)}$. Applying this approach to one of the nodes in Figure 1 and the bottom level series CC,

$$p_{ZC} = \left(\frac{\hat{y}_{CC,t}}{\hat{S}_{C,t}}\right)\left(\frac{\hat{y}_{C,t}}{\hat{S}_{Total,t}}\right) \tag{13}$$

where $\ddot{S}_{Total,t} = \ddot{Y}_{A,t} + \ddot{Y}_{B,t} + \ddot{Y}_{C,t}$ and $\ddot{S}_{C,t} = \hat{y}_{CA,t} + \hat{y}_{CC,t}$

2.3 Optimal combination method

The optimal combination approach was proposed by Hyndman et al., (2011) as part of the hierarchical forecasting approaches. In this approach, a regression model is used in combining and reconciling individual base forecasts optimally. With this technique, correlations and interactions between series at every level are allowed and taken into consideration. The approach is aimed at estimating unknown future expectation values of the bottom level of the dataset *K*. According to Hyndman et al., (2011), let there be a vector of the unknown means $(\boldsymbol{\beta}_n(h))$.

$$\boldsymbol{\beta}_n(h) = E[\boldsymbol{Y}_{k,n+h} | \boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n] \tag{14}$$

African Journal of Hospitality, Tourism and Leisure, Volume 7 (4) - (2018) ISSN: 2223-814X

For recall here $\boldsymbol{Y_t}$ represents the vector of all observations at time t while and $\boldsymbol{Y}_{k,n+h}$ represents the vector of observations in the bottom level K. The base forecasts $(\ddot{\boldsymbol{Y}}_n(h))$ will be presented in a regression format to give:

$$\ddot{\boldsymbol{Y}}_n(h) = \boldsymbol{S}\boldsymbol{\beta}_n(h) = +\boldsymbol{\varepsilon}_h \tag{15}$$

for $\boldsymbol{\varepsilon}_h$ denoting a white noise term with covariance matrix $\sum h$ which is difficult to find in large hierarchies (Morgan, 2015). In an attempt to solve this challenge, Hyndman et al., (2011) hypothesised estimating the white noise term by the forecast error in the bottom level, which is $\boldsymbol{\varepsilon}_h \approx \boldsymbol{S}\boldsymbol{\varepsilon}_{k,h}$. With this hypothesis, errors satisfy the same aggregation constraint as the dataset, resulting in

$$\sum h = \boldsymbol{S}\ \mathrm{Var}(\boldsymbol{\varepsilon}_{k,h})\boldsymbol{S}' \tag{16}$$

Forecasting individual series method

The commonly used methods in forecasting individual series in hierarchical time series are the exponential smoothing (ETS) and the autoregressive integrated moving average (ARIMA). Mahkya et al., (2017) forecasted individual series of export value in Central Java using the ARIMA, Radial Basis Function Neural Network (RBFNN) and Hybrid ARIMA-RBFNN. The findings indicated that the bottom-up approach with Hybrid ARIMA-RBFNN produced best long-term forecasts. Wickramasuriya et al., (2018) proposed reconciliation approach which incorporates full covariance matrix information of forecast errors in an attempt to produce coherent forecasts used the ARIMA and ETS. The proposed approach is said to outperform other existing approaches. The general ARIMA model can be expressed as:

$$Y_t - \Phi_1 Y_{t-1} - \Phi_2 Y_{t-2} - \cdots - \Phi_p Y_{t-p} = a_t - \Theta_1 a_{t-1} - \Theta_2 a_{t-2} - \cdots - \Theta_q a_{t-q} \tag{17}$$

where $\Phi's$ and $\Theta's$ are model parameters.

According to Sharpe et al., (2010) the ETS allocates exponentially decaying heavier weights to recent values than the old ones. Holt-Winters are part of ETS designed for a dataset that exhibit permanent, constant, linear trend and multiplicative seasonal variations indicated by Saayman and Saayman (2010). The general forms of the Holt-Winters with permanent constant (equation 18), linear trend (equation 19) and multiplicative seasonal variations (equation 20) are will be used to generate base forecasts.

$$\tilde{Y}_t = \alpha\left(\frac{Y_t}{S_{t-s}}\right) + (1-\alpha)(\tilde{Y}_{t-1} + \tilde{B}_{t-1}) \tag{18}$$

$$\tilde{B}_t = \beta\left(\tilde{Y}_t - \tilde{Y}_{t-1}\right) + (1-\beta)\tilde{B}_{t-1} \tag{19}$$

$$S_t = \gamma\left(\frac{Y_t}{\tilde{Y}_t}\right) + (1-\gamma)S_{t-s} \tag{20}$$

where the smoothing parameters (γ, α and β) take values between 0 and 1. The smoothed series and seasonality period is denoted $\tilde{Y}_t$ and s respectively.

Both the ETS and The ARIMA default algorithm of Hyndman and Khandakar (2007) effected in the forecast package for R are used in generating forecasts for all levels in the hierarchy. The ARIMA(p, d, q) model will take 0, 1 and 2 as p and q values, with equal probability while d take either 0 or 1 with equal probability according to Wickramasuriya et al., (2018)

Data analysis and discussion of findings

Data and descriptive statistics

The Victoria Falls Rainforest monthly tourism data are obtained from the ZPWMA, the custodians of the rainforest. The available data spans from January 2010 to December 2016, giving a total of 84 observations (see Figure 1). Table 1 gives the structure of the hierarchy.

Table 1: Hierarchy of Victoria Falls Rainforest tourism data

Level	Number of series
Total tourist arrivals	1
Tourism source(International, Regional, Locals)	3
Age category X Tourism source	6

The descriptive statistics of the tourism data are constructed and are shown in Table 2.

Table 2: Descriptive statistics

Series	Mean	Median	Max	Min	St. Dev.	Skewness	Kurtosis
AA	10578.30	9798.50	20971	4380	4106.52	0.56	2.56
AC	199.48	126.50	760	24	171.89	1.57	4.71
BA	2632.71	2414.50	5324	836	1094.54	0.76	2.95
BC	202.70	154.50	751	16	170.72	1.42	4.53
CA	4069.55	3554.50	10673	1339	2168.14	1.61	4.89
CC	1451.38	1052.50	5959	90	1236.54	1.02	3.67

Mean arrivals are ranging from 199.48 to 10578.30. Minimum tourist arrivals are 16 regional children tourists and the maximum number of tourist arrivals is 20971 for international adult tourists. The tourism data for all the series is skewed to the right as indicated the skewness value. Kurtosis values indicated that most of the data are leptokurtic.

The characteristics of the disaggregated tourist arrivals series are done. Figure 2 shows the behaviour of tourist arrivals at the site. Level 0 shows the total tourist arrivals through time. Level 1 shows tourist arrivals from the three tourism sources through time. Level 2 shows tourists arrivals by age through time.

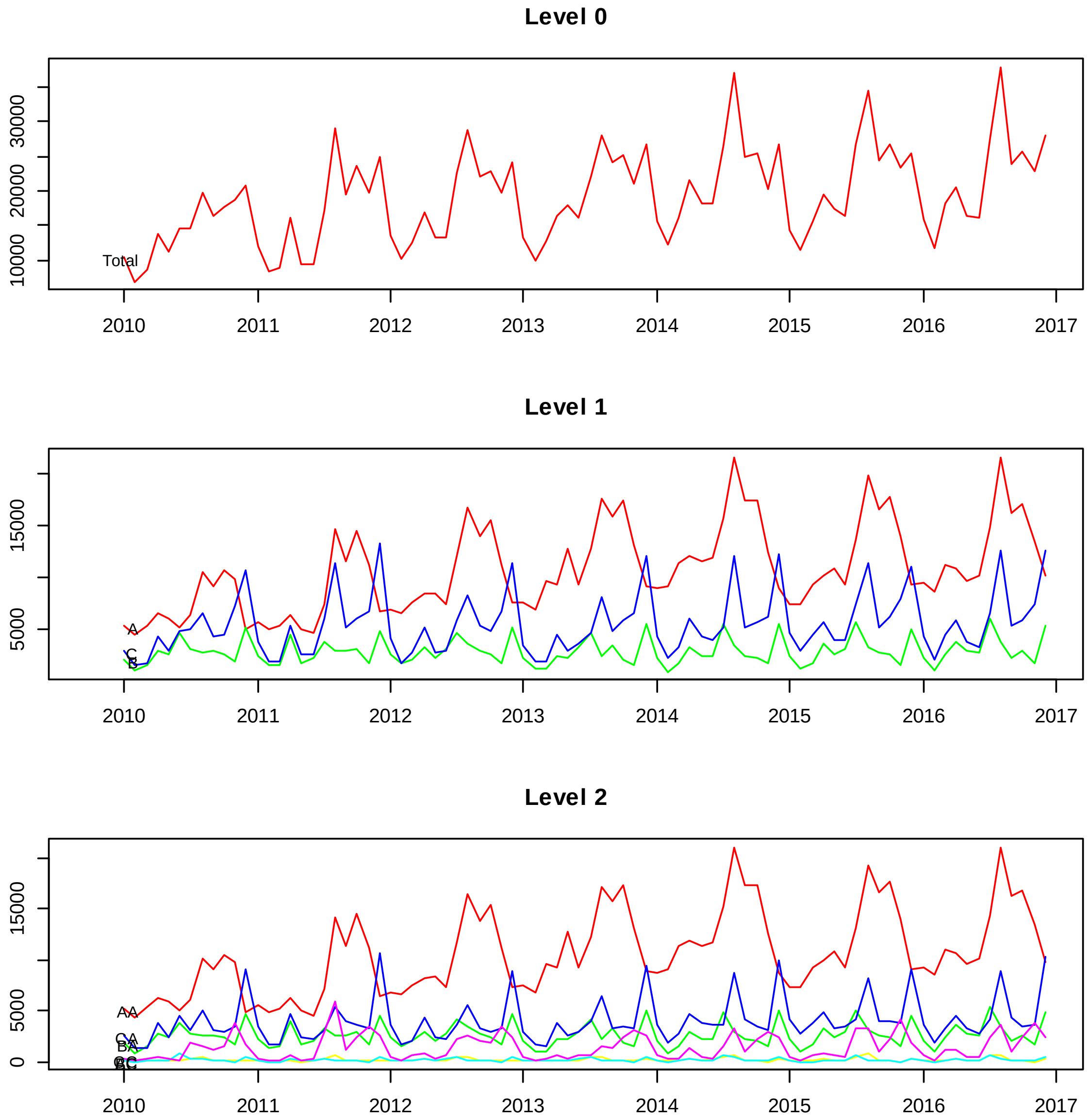

Figure 2: Hierarchical time series of the Victoria Falls Rainforest tourist arrivals.

For level 2 series, the red line is for international adult tourists, the green line is for regional adult tourists and the blue line is for Zimbabwean adult tourists. The purple line is for Zimbabwean children tourists. The yellow and light-blue lines are for international and regional children tourists respectively. According to Figure 2, more foreign adult tourists visited the Victoria Falls Rainforest than local adults during the study period. Very few children tourists are coming to the Victoria Falls Rainforest, in particular international and regional children. Generally, most of the tourists who visited the site are from abroad as shown in Figure 2.

Forecasting accuracy evaluation

Forecasting accuracy of the models will be evaluated using the mean absolute percentage error (MAPE). The accuracy measure is calculated using equation 10.

$$MAPE = \frac{100\%}{N}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{C}\right| \qquad (21)$$

African Journal of Hospitality, Tourism and Leisure, Volume 7 (4) - (2018) ISSN: 2223-814X

N denotes total observations while A_t and F_t denotes actual and forecasted values, respectively.

An out-of-sample forecasting accuracy measure is done. 12 step-ahead tourism forecasts are produced prior to setting up the period January 2010 to December 2015 as the training sample and the 2016 period as the testing sample. Table 3 is the summary of forecasting accuracy measures done using both the ETS and ARIMA as forecasting methods.

Table 3: Forecast error measures (MAPE)

	ETS					ARIMA				
	BU	**TDHP**	**TDHA**	**TDFP**	**OC**	**BU**	**TDHP**	**TDHA**	**TDFP**	**OC**
Total	20.03	27.91	27.91	27.91	22.6	26.67	26.42	26.42	26.42	26.03
A	18.9	18.27	18.03	16.52	16.69	24.6	17.82	17.51	29.26	24.96
B	53.33	58.11	54.88	64.68	57.23	52.84	54.68	51.54	53.92	53.24
C	45.14	50.75	53.97	61.41	52.86	40.13	48.46	51.6	37.29	37.46
AA	18.68	18.16	17.86	16.2	16.42	24.53	18.06	17.57	29.51	25.08
AC	95.79	116.96	124.67	119.81	110.32	124.72	108.38	115.49	129.4	126.97
BA	49.08	56.25	52.57	59.48	52.27	50.52	52.43	49.07	51.5	50.82
BC	130.18	111.8	114.21	158.5	154.56	100.35	112.72	115.15	103.33	105.06
CA	52.59	54.71	55.41	67.73	56.04	48.08	51.29	51.96	47.54	46.55
CC	98.51	145.86	169.66	117.57	120.5	78	157.05	182.61	74.36	74.67
Average	58.223	65.878	68.917	70.981	65.949	**57.044**	64.731	67.892	58.253	57.084

BU-Bottom-up approach, TDHP-Top-down approach with average historical proportions, TDHA- Top-down approach with proportion of historical averages, TDFP- Top-down approach with forecast proportions and OC-optimal combination approach.

First and last four columns under the heading ETS and ARIMA are accurate measures for individual series in the hierarchy. The average accuracy measures from each model are under the row named “Average”. It is shown in Table 3 that the bottom-up approach produces small MAPE values under the two forecasting methods (ETS and ARIMA). However the bottom-up approach under ARIMA with MAPE error (57.044) is the best hence is used in forecasting future tourist arrivals for the site. Table 4 and Figure 3 show forecasted future values and the graphical display.

African Journal of Hospitality, Tourism and Leisure, Volume 7 (4) - (2018) ISSN: 2223-814X

Table 4: Out-of-sample future tourism forecasts.

Month	Total	A	B	C	AA	AC	BA	BC	CA	CC
Jan-17	16183	9121	2251	4810	8961	161	2102	149	4194	616
Feb-17	12465	8760	1028	2677	8643	117	1002	27	2447	230
Mar-17	17497	11044	2170	4282	10859	186	2069	101	3569	713
Apr-17	21211	11322	3619	6270	10984	338	3370	249	4900	1370
May-17	17837	11103	2686	4048	10987	117	2581	105	3560	487
Jun-17	17075	10600	2750	3724	10366	234	2607	143	3383	342
Jul-17	27042	15064	5767	6211	14451	613	5087	679	4412	1799
Aug-17	37351	21544	3543	12264	20805	739	3212	331	8815	3448
Sep-17	25108	17271	2396	5442	17145	126	2299	97	4441	1001
Oct-17	27140	18274	2660	6206	18032	242	2488	172	3917	2289
Nov-17	23421	14585	1682	7154	14523	62	1608	74	3906	3248
Dec-17	28389	10548	5240	12601	10238	310	4827	413	10161	2440
Jan-18	17160	10084	2244	4833	9905	179	2102	142	4321	512
Feb-18	13307	9465	1028	2814	9333	132	1002	27	2627	187
Mar-18	18473	11890	2193	4389	11669	221	2069	124	3748	641
Apr-18	21640	11879	3628	6133	11532	348	3370	258	5080	1053
May-18	18186	11196	2690	4300	11064	132	2581	109	3740	560
Jun-18	17983	11221	2759	4003	10968	253	2607	152	3562	440
Jul-18	28801	15737	5771	7293	15104	634	5087	683	4592	2702
Aug-18	38195	22391	3516	12288	21652	739	3212	303	8995	3293
Sep-18	25659	17586	2399	5674	17447	138	2299	100	4620	1054
Oct-18	27461	18448	2665	6348	18189	259	2488	177	4097	2251
Nov-18	24351	14839	1679	7832	14765	74	1608	71	4086	3746
Dec-18	28799	11127	5239	12433	10799	328	4827	412	10341	2092

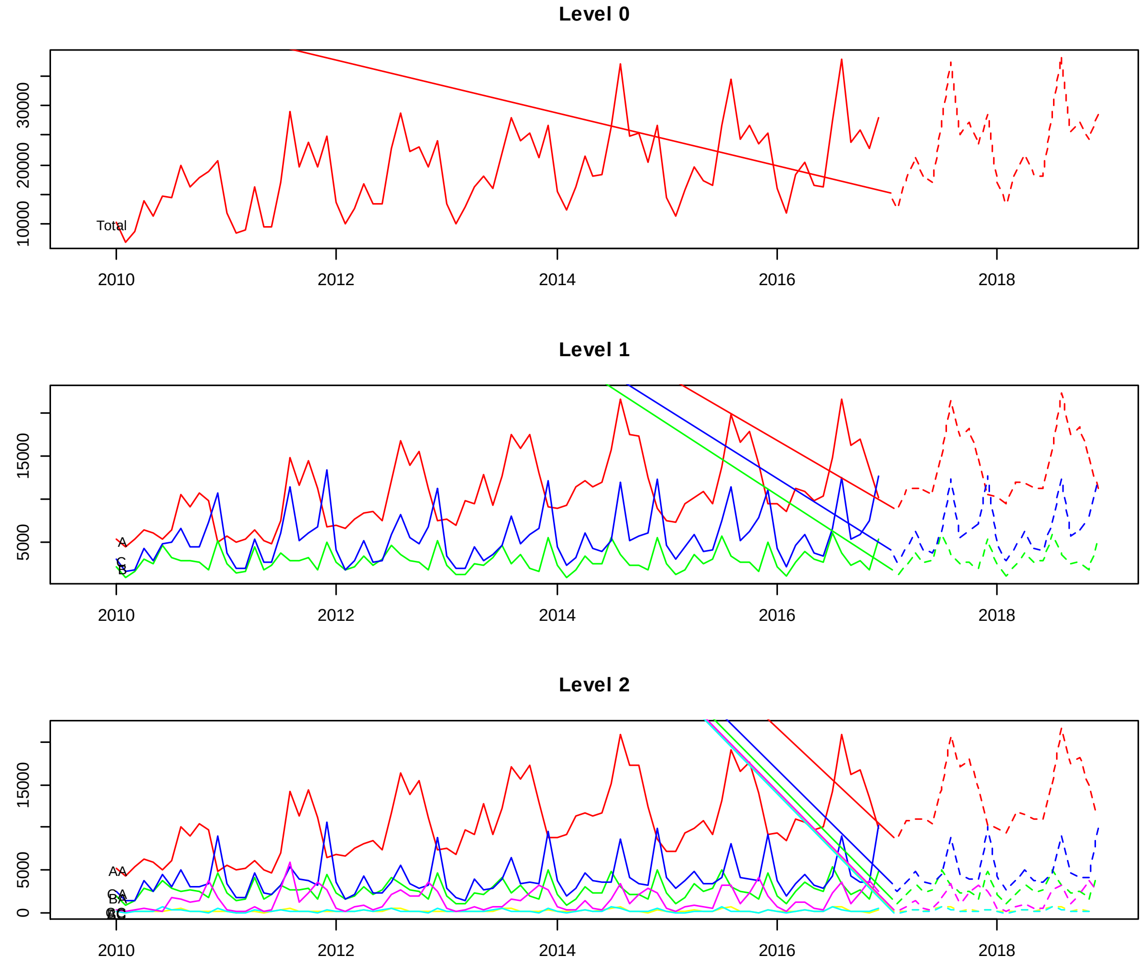

Figure 3: Future out-of-sample tourism forecasts.

Future tourist arrival forecasts are indicated by the dashed line(s). Historical, tourist arrival data is indicated by a solid line(s). For level 2 colours, the red line and the blue line represent international adult tourists (AA) and Zimbabwean adult tourists (BA) respectively. The green line represents regional adult tourists (CA). The purple line represents Zimbabwean children tourists (CC). The light-blue and yellow lines represent regional children tourists (BC) and international children tourists (AC).

Figure 3 indicates some seasonality in future tourist arrivals at the site. There is an overall slow increase in tourist arrivals for the year 2017 and 2018. Regional children tourists and international children tourists remain very low. Effective marketing strategies are needed so as to motivate the young generation to visit the site. Foreign currency can be earned from this age group if managed well. The Victoria falls rainforest environment, apart from entertaining, is also educational. Zimbabweans are willing to share the knowledge.

Conclusions

The Victoria Falls Rainforest tourist arrivals are modelled and forecasted using the hierarchical forecasting methods. The monthly tourism data are disaggregated according to tourism source (international tourists, regional tourists, locals) and further by age (adult tourist and children tourists). Families need decent accommodation where they are all together at all tourism sites. Moreover, adults and children have different demands and expectation that need to be considered by transport, food and accommodation managers at tourism sites. School tours need places that are good and considered safe by parents and tour operators. The bottom-up, top-down and optimal combination approaches are used, with the ETS and ARIMA models adopted as forecasting methods. Forecasting accuracy of the approaches under ETS and ARIMA models are evaluated using MAPE and the bottom-up approach under ARIMA models outperform all other approaches. Forecasts produced from the bottom up approach under ARIMA model indicate seasonality and an overall increase in future tourist arrivals.

The forecasts from the approach suggest the need for effective marketing strategies targeting the younger generation as the average number of children tourists is very low particularly for international and regional tourists. The marketing needs to all target parents and tour operators. If the marketing is successful, it will have a positive impact in terms of foreign currency earnings.

The hierarchical forecasting methods used are more informative and of significant benefit to the ZPWMA, accommodation managers, transport managers, food managers, tourism managers, investors, government and policy makers. The research findings strongly recommend the use of hierarchical forecasting methods in the tourism industry since they capture tourism data dynamics, produce precise and sensible forecasts and give economically feasible solutions.

Acknowledgements
The authors would like to acknowledge the Zimbabwe Parks and Wildlife Management Authority (ZPWMA) for providing the tourism data.

References

Annual Tourism Trends and Statistics Report. (2015) Government of Zimbabwe. Printflow: Harare. http://www.zimstat.co.zw/

Athanasopoulos, G. & Hyndman, R. J. (2008) Modelling and forecasting Australian domestic tourism. *Tourism Management,* 29(1): 19-31.

Athanasopoulos, G., Ahmed, R. A. & Hyndman, R. J. (2009) Hierarchical forecasts for Australian domestic tourism. *International Journal of Forecasting,* 25(1): 146-166.

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. & Petropoulos, F. (2017) Forecasting with temporal hierarchies. *European Journal of Operational Research,* 262(1): 60-74.

Athanasopoulos, G., Hyndman, R. J., Song, H. & Wu, D. C. (2011) The tourism forecasting competition. *International Journal of Forecasting,* 27 (3): 822-844.

Balmford, A., Green, J.M., Anderson, M., Beresford, J., Huang, C., Naidoo, R., Walpole, M. & Manica, A. (2015) Walk on the wild side: Estimating the global magnitude of visits to protected areas. *PLoS Biology,* 13(2): e1002074.

Chatziantoniou, I., Degiannakis, S., Eeckels, B. & Filis, G. (2016) Forecasting tourist arrivals using origin country macroeconomics. *Applied Economics,* 48(27): 2571-2585.

Dhakal, K. R. (2014) An Analysis of the Tourism in Nepal Trend of Tourist Arrivals. *The Third Pole: Journal of Geography Education*, 13: 46-53.

Frias, A., Cabral, J. & Costa, A. (2015) Logistic optimization in tourism networks. *European Regional Science Association (ERSA)*.

Hyndman, R. J. & Athanasopoulos, G. (2014) Optimally reconciling forecasts in a hierarchy. Foresight. *The International Journal of Applied Forecasting,* 35: 42-48.

Hyndman, R.J. & Khandakar, Y. (2007) Automatic time series for forecasting: the forecast package for R. Monash University, Department of Econometrics and Business Statistics. 6/07.

Hyndman, R.J., Ahmed, R. A., Athanasopoulos, G. & Shang, H. L. (2011) Optimal combination forecasts for hierarchical time series. *Computational Statistics and Data Analysis,* 55(9): 2579-2589.

Job, H. & Paesler, F. (2013) Links between nature-based tourism, protected areas, poverty alleviation and crises. The example of Wasini Island (Kenya). *Journal of Outdoor Recreation and Tourism,* 1(2): 18-28.

Jonathan, D. C. & Kung-Sik, C. (2008) Time Series Analysis with Applications in R (Lowa City, Lowa: Springer)

Jones, T. & Ohsawa, T. (2016) Monitoring nature-based tourism trends in Japan's national parks: Mixed messages from domestic and inbound visitors. *Parks,* 22(1): 25-36.

Khadivi, P. & Ramakrishnan, N. (2016) Wikipedia in the tourism industry: Forecasting demand and modeling usage behaviour. *In AAAI* 4016-4021.

Mahkya, D., Ulama, B. S. & Suhartono, B. (2017) Hierarchical time series bottom-up approach for forecast the export value in Central Java. *J. Phys,* 893 (1): 012033.

Morgan, L. (2015) Forecasting in Hierarchical Models

Mutanga, C. N., Gandiwa, E. & Muboko, N. (2017).] An analysis of tourist trends in northern Gonarezhou national park, Zimbabwe, 1991-2014. *Cogent Social Sciences,* 3(1): 1392921.

Perera, K. K. E. (2017) An Analysis of Recent Trends in Tourist Arrivals in Sri Lanka. *Athens Journal of Tourism*, 4(1): 51-79.

Saayman, A. & Saayman, M. (2010) Forecasting tourist arrivals in South Africa. *Acta Commercii,* 10(1): 281-293.

Shang. H. L. (2017) Reconciling forecasts of infant mortality rates at national and sub-national levels: Grouped time-series methods. *Population Research and Policy Review,* 36(1): 55-84.

Shang, H. L. & Haberman, S. (2017) Grouped multivariate and functional time series forecasting: An application to annuity pricing. *Insurance: Mathematics and Economics,* 75: 166-79.

Sharpe, R. N., De Vaux, R. & Velleman, P. F. (2010) Business Statistics, 2nd Edition, Boston, Addison Vesley - Pearson Education.

United Nations World Tourism Organisation. (2015) *Towards measuring the economic value of wildlife watching tourism in Africa: Briefing Paper*.

Wickramasuriya, S. L., Athanasopoulos, G. & Hyndman, R. J. (2018) Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 71(353): 68-71.

Yao L, Ma, R., Jin, M., Ge, P. & Ren, P. (2014) Forecasting real-time tourist arrivals using Hierarchical Cluster and Gaussian fitting algorithm: A Case Study of Jiuzhai Valley. *Applied Mathematics and Information Sciences,* 8(3): 1425-1434.

Zhou, Z. (2018) A survey of visitor satisfaction at Victoria Falls Rainforest. *African Journal of Hospitality, Tourism and Leisure,* 7 (1).