



Hierarchical Forecasting of the Zimbabwe International Tourist Arrivals

Tendai Makoni ^{1*}, Delson Chikobvu ¹ and Caston Sigauke ²

¹*Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, South Africa*

²*Department of Statistics, University of Venda, South Africa*

Abstract The objectives of the paper is to: (1) adopt the hierarchical forecasting methods in modelling and forecasting international tourist arrivals in Zimbabwe; and (2) coming up with Zimbabwe international tourist arrivals Prediction Intervals (PIs) in Quantile Regression Averaging (QRA) to hierarchical tourism forecasts. The unavailability of statistical models for Zimbabwe international tourist arrivals that cater for disaggregated tourism data and account for uncertainty due to parameter estimation methods, has resulted in poor marketing strategies, infrastructure and policies targeting wrong tourism groups. Furthermore, the country is failing to attract significant Foreign Direct Investment for particular tourist arrivals. Zimbabwe's monthly international tourist arrivals data from January 2002 to December 2018 was used. The data set was disaggregated according to the purpose of the visit. Three hierarchical forecasting approaches, namely top-down, bottom-up and optimal combination approaches were applied to the data. The results showed the superiority of the bottom-up approach over both the top-down and optimal combination approaches. Forecasts indicate a general increase in aggregate series. The combined methods provide a new insight into modelling tourist arrivals. The approach is useful to the government, tourism stakeholders, and investors among others, for decision-making, resource mobilisation and allocation. The Zimbabwe Tourism Authority (ZTA) could adopt the forecasting techniques to produce informative and precise tourism forecasts. The data set used is before the COVID-19 pandemic and the models indicate what could happen outside the pandemic. During the pandemic the country was under lockdown with no tourist arrivals to report on. The models are useful for planning purposes beyond the COVID-19 pandemic.

Keywords Aggregation, bottom-up, disaggregation, forecasting, international tourist arrivals, optimal combination, top-down

DOI:10.19139/soic-2310-5070-959

1. Introduction

The tourism sector is an important economic sector of every economy. It helps alleviate poverty and promotes economic growth ([1]). Therefore the sector needs to be properly managed. Accurate tourism demand forecasts lead to proper tourism management as they help in decision-making and planning. Tourism managers, the government and investors are interested in tourism purpose of visit as well as forecasts at the national level, provincial level and each tourist attraction centre. Accurate international tourism demand forecasts are beneficial to investors in particular, in determining how much to invest and in which tourism sector. The government needs accurate forecasts for resource allocation and policy implementation. The tourism managers need accurate forecasts in policy formulation and amendment. Accurate tourism forecasts can be generated through the use of statistical methods and other techniques. Statistical forecasting methods have better forecasting performance as compared to judgemental methods ([2]). Pure time series forecasting methods deliver more accurate tourism forecasts as compared to econometric models ([3]). However, the general agreement in terms of forecasting

*Correspondence to: Tendai Makoni (Email: tpmakoni@gmail.com). Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, South Africa.

performance is that, no single model performs best in all situations.

Forecasting international tourist arrivals is often done using statistical methods with aggregated series ([4]). However, it is also important and feasible to use disaggregated data in certain instances. The national level and various different sub-national level tourism forecasts are fundamental to tourism managers, government, investors and policy-makers for efficient and effective planning. [5] noted the importance of having forecasts for various levels of aggregation when modelling electricity demand for the United Kingdom (UK). The authors developed an algorithm for the purpose of producing coherent probabilistic density forecast for an individual series in the hierarchy. Electricity data was disaggregated according to geographic and demographic information. The algorithm was applied to the data and the results were better than those of benchmark models.

Tourism aggregated data can be decomposed according to various aspects. Decomposition and disaggregation are progressively used because of the availability of data and this allows knowledgeable recommendations on specific aspects of tourism ([6]).

The arrival of international tourists in Zimbabwe is influenced by several factors, such as the purpose of visit (PoV), season and weather conditions. Therefore, hierarchical forecasting is appropriate for this type of data. Hierarchical forecasting will improve the decision-making processes, as particular characteristics and dynamics in the tourism data can be viewed and explained ([2]).

In this paper, the Zimbabwe international tourist data are categorised according to PoV that is: In-transit (I), Business (B), Education (E), Holiday (H) and Shopping (S). This implies total international arrivals can be disaggregated into these five groups, resulting in a one level hierarchy. This decomposition is chosen based on the availability of data, as highlighted by [6]. Political instability and a deteriorating economy can be attributed to tourist arrivals decline ([7]). With hierarchical forecasting, one can identify the declining tourist arrivals category as well as other factors contributing to the decline, so that the necessary corrective measures can be taken.

Prediction intervals (PIs) are vital in forecasting; they provide supplementary information that can be used in decision-making. According to [8], PIs allow better assessment of future uncertainty and help in planning. The purpose of this paper is to adopt both the hierarchical forecasting methods proposed by [3] and the Quantile Regression Averaging (QRA) PIs in forecasting international tourist arrivals in Zimbabwe, for both aggregated and disaggregated data by PoV. Hierarchical forecasting methods observe the grouped structure of time series data, and in so doing improve forecasting accuracy ([9]). PIs take into account uncertainty from the parameter estimation method as well as the noise in the input data and they improve prediction accuracy ([10]).

To the best of our knowledge, the application of the QRA PIs to hierarchical forecasting tourism forecasts has not been considered on African tourism, particularly Zimbabwe, where international tourist arrivals are disaggregated by PoV. Using the best hierarchical forecast approach estimates to come up with the QRA PIs in the tourism industry provides new insight into modelling tourist arrivals in Africa and Zimbabwe in particular.

The rest of the paper is organised as follows: Section 1 presented the introduction, Section 2 reviews the literature. The modelling and individual series prediction approaches are presented in Section 3, empirical results and discussions are in Section 4. Section 5 concludes the paper.

2. Literature Review

In the tourism industry, tourism forecasting plays a significant role in decision-making. Because of its importance, several and frequently used models namely: Autoregressive (AR), Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA), Autoregressive Distributed Lag Model (ADLM), Error Correction

Model (ECM), Artificial Neural Networks (ANN), Vector Autoregressive (VAR), and the Na?ve model among others, have been developed. These models offer various advantages over others.

Several casual methods (see: [11]; [12]; [13]; [14]; [15]; [16]) and non-casual tourism studies for Africa exists (see: [17]; [18]). Casual methods analyse the causal relationships between the tourism demand variable and its influencing factors ([19]). In casual methods, data gathering and analysing is cost and time consuming ([20]) and the approach is not especially accurate in forecasting tourism demand ([21]), hence are not employed in this paper. Non-casual methods (univariate time series) such as the Autoregressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA) and Exponential Smoothing (ES) are less time consuming and costly than regression-based or econometric models ([22]). Univariate time series are incorporated in hierarchical forecasting techniques.

In Zimbabwe, [23], [24], [25], [26] and [27] are among those researchers who focused tourism determinants. The current study differs in that it applies non-casual methods because of their ability to capture tourism dynamics, quantifies future uncertainty and generates more informative forecasts. [28] focused at specific tourist destination sites (the Victoria Falls Rainforest, Zimbabwe) and not at country level as considered in this paper. Tourism forecasts for national level are more important for policy formulation.

The current paper employ the hierarchical forecasting technique, Quantile Regression Averaging (QRA) and PIs which capture tourism dynamics and quantify the uncertainties in future values due to parameter estimation unlike the hierarchical forecasting model developed by [29] which only captures tourism seasonality. QRA PIs provide supplementary information useful for effective marketing strategies and decision-making ([30]; [8]).

[31], employed a SARIMA(0,0,0)(1,1,0)₁₂ model in forecasting Bhutan's tourism demand. [32], modelled tourist demand for Macedonia and an ARIMA(1,1,1) model was found as an appropriate model. The model projected an increase in tourist arrivals. In Sri Lanka, [33] fitted a SARIMA model using monthly data. The HEGY test was used to identify seasonality. The SARIMA(1,0,16)(36,0,24)₁₂ model with a seasonal component was found to be a better model for the data.

[34] fitted a SARIMA model to Taiwan's outbound tourists. They projected an increase in outbound tourists to Japan, Hong Kong and the USA, which are the major tourist destinations for Taiwanese. [35] used a SARIMA model in estimating South Africa's tourism demand from its major overseas tourism markets namely: the USA, Germany, France, Great Britain and the Netherlands. The SARIMA(1,0,1)(0,1,1)₁₂ model and SARIMA(1,1,1)(0,1,1)₁₂ model fitted well to tourist arrivals from France and Netherlands, respectively. Tourist arrivals from the USA, Germany and Great Britain were found to follow a SARIMA(0,1,1)(0,1,1)₁₂ model. SARIMA models have stood the test of time since their existence from the 1970s to date.

According to a Denmark research done by [36], a Time-Varying Parameter (TVP) model fitted well to the tourism demand of Denmark. The results were achieved after the evaluation of the forecasting performance of both univariate and econometric models. Exponential Smoothing (ES), ARIMA, and ANN models were used by [37] in modelling Hong Kong's tourism demand. The ANN models were found to be the best forecasting approach for Hong Kong's tourism demand. [4] compared forecasting performance of ECM, AR, VAR and ARIMA models while projecting UK's outbound tourism demand. Research findings indicated the superiority of the ECM over other models.

Much has been achieved in modelling tourism demand using current techniques, but the use of hierarchical forecasting techniques has received little attention. Credit is given to [2] in hierarchical forecasting. They proposed the approach while forecasting tourist arrivals in Australia. The major advantage of the hierarchical forecasting approach is that it generates more accurate and coherent disaggregated forecasts than other corresponding

forecasting approaches applied to the same data set.

In Australia, [2] disaggregated the tourism data by states, zones and regions and used hierarchical forecasting methods. [2] used two different top-down, two mixed bottom-up and the optimal combination approaches in forecasting Australian domestic tourism demand. They noted that the optimal combination and the top-down approaches based on forecast of proportions performed better in their paper and this was based on the accuracy measures used.

The new top-down and bottom-up approaches proposed by [2], together with the optimal combination approach, can capture various characteristics of tourism data. The optimal combination approach makes use of a regression model to come up with unbiased combination forecasts as well as minimum variances ([2]). Results from their study were used for decision-making in Australia and they indicated areas that needed special attention.

When forecasting these types of hierarchical series, the aggregated forecasts should exactly correspond to disaggregated estimates ([38]). However, according to [9], estimates of higher-level may not automatically tally with forecasts from lower levels. The incoherence of the series is applicable to tourism data as base forecasts of PoV (using disaggregated series) may not tally with national level forecasts (using aggregated series). Therefore, reconciling international tourist arrivals using hierarchical forecasting methods proposed by [2] is suggested as a solution to the problem articulated. The bottom-up, top-down (average historical proportions, the proportion of historical averages and forecast proportions) and optimal combination approaches developed by [3] are adopted. These approaches are capable of summing up the point forecasts of PoV, so that they are identical to the point forecasts of the aggregated data.

Different views exist in determining the best approach in hierarchical time series analysis. [39] concluded that no method could outperform the other and further indicated that performance varies according to the situation at hand in hierarchical models. [40] noted no difference in terms of performance between the top-down and bottom-up approaches during forecasting the aggregate demand in production planning. During an investigation of aggregated variable time series forecast strategies with specific sub-aggregated time series statistical correlation, [41] concluded the top-down approach outperforming other approaches. Similar findings were also obtained by [42], while applying hierarchical forecasting methods in the production planning and inventory control cases. Their top-down approach produced better results than the results produced by the bottom-up approach.

Conversely, [43], indicated that the bottom-up approach performed better and also showed that it was the best method when forecasts are statistically generated. Whilst forecasting items in a product line using 15000 aggregated series, [44] noted that the bottom-up approach produced more accurate forecasts than the top-down approach. They compared predictions from the bottom-up and top-down approaches. Another study by [45] indicated the bottom-up approach as better than the top-down approach while estimating future total-entity earnings of 26 diversified companies using segment and consolidated revenue and profitability data. However, [2], [3], [39] and [5] all noted that: for disaggregated series, the bottom level series are volatile, hence tricky to model and the bottom-up approach can be outdone by other methods.

PIs based on the Quantile Regression Averaging (QRA) technique are more accurate than those from the best performing individual model ([46]). PIs perform well when the sample size considered is large ([47]). The sample size considered in this study allows for the construction of PIs.

According to [48], PIs generate a range of future values; thus, the interval width and coverage rate can be used to evaluate the accuracy of PIs. A good PI model is one with a narrow width and coverage rate closer to the nominal coverage rate ([49]; [50]). [51] developed an instrument useful for PI forecasting assessment called Winkler score. During combining tourism interval forecasts for tourist arrivals in Hong Kong, [49] used Winkler scores to measure interval forecasting performance.

3. Methods

In this section, statistical methods in the form of the hierarchical forecasting approach, QRA and PIs and other methods used in this paper will be discussed.

3.1. Hierarchical structure

According to [52], a grouped time series is a collection of time series that can be grouped together in a number of non-hierarchical ways, while a hierarchical time series is a collection of several time series are linked together in a hierarchical structure. Tourist arrivals by state and PoV are an example of grouped time series, while tourist arrivals by state and region represent a hierarchical time series. The hierarchical structure adapted to the Zimbabwe international tourist arrivals is the total series and PoV series as shown in Figure 1.

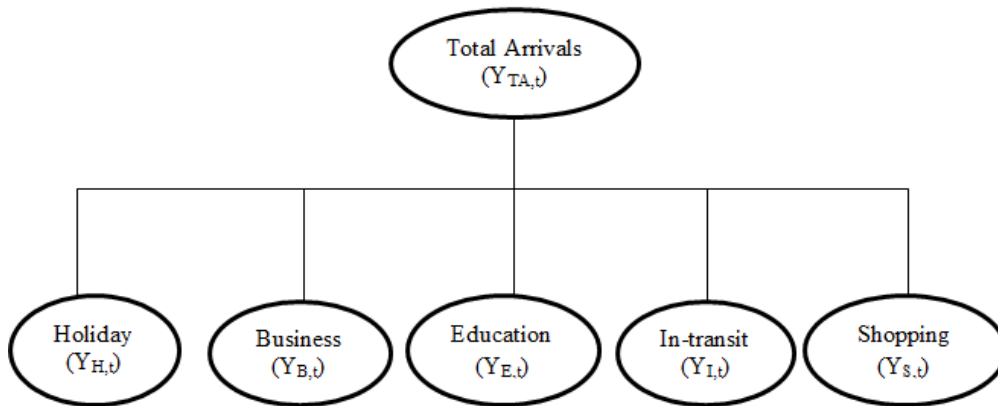


Figure 1. Hierarchical tree diagram for Zimbabwean international tourist arrivals.

This is a Level 1 hierarchy. Completely aggregated series is named total arrivals’ and is under Level 0, whereas Level 1 is constituted by disaggregation series (decomposed by PoV).

In hierarchical time series, individual forecast at each level will make up forecast of the upper level when summed. For grouped time series, each group’s forecasts will be equivalent to the forecasts of individual series constituting that group. Both hierarchical and grouped time series are capable of producing coherent forecasts. Hierarchical time series is less complex as compared to grouped time series with disaggregation constraints. For this study, the authors considered grouped time series, since total international tourist arrivals are being disaggregated according to PoV. Hierarchical forecasting approach deals with both grouped and hierarchical time series.

Adopting the notation of [52], \mathbf{Y}_t represents an m -vector of all tourism observations at time t , \mathbf{B}_t becomes an n -vector ($n = 5$) of disaggregated tourism observations at the bottom level (level 1) and $\mathbf{Y}_{(X,t)}$ being the vector of observation on series X at time t . From these, we have Equation (1),

$$\mathbf{Y}_t = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} Y_{H,t} \\ Y_{B,t} \\ Y_{E,t} \\ Y_{I,t} \\ Y_{S,t} \end{pmatrix}. \tag{1}$$

Equation (1) can be reduced to Equation (2),

$$Y_t = SB_t, \tag{2}$$

where matrix $S_{m \times m_k}$ ($S_{6 \times 5}$) is the summing matrix that stores the structure of the hierarchy ($m=6$ is the total number of series in the hierarchy and $m_k = 5$ is the number of bottom level series). There are several methods used in hierarchical time series modelling, some of which are discussed in the following sections.

3.2. The bottom-up method

With this approach, lowest level forecasts are produced, which are later on aggregated to the higher level through the use of the summing matrix ([2]). The dynamics of specific series will be captured since information is not lost with this approach but fitting a forecasting model becomes difficult if the bottom level series is unstructured and noisy ([2]). More bottom level series imply more series to be forecasted at the lower levels.

3.3. The top-down method

According to [2], top-level forecasts are produced and are disaggregated to the lower levels of the hierarchy through the use of proportions. The approach works well in the presence of low count data ([2]). With this method, it is easy to come up with single forecasting models and reliable forecasts for aggregated levels. [53], highlighted three methods (average historical proportions, proportions of historical averages and forecasted proportions) of coming up with the proportions.

Average historical proportions

For each series at the lower level of the hierarchy, proportions are found, $p_i, i = 1, 2, \dots, m_k,$

$$p_i = \frac{1}{n} \sum_{t=1}^n \frac{y_{i,t}}{y_t}. \tag{3}$$

Under this method, the individual proportion p_i is mirroring the average of the historical proportions of the bottom level series (level 1 series) ($Y_{(X,t)}$) from $t = 1, \dots, 204,$ relative to the total aggregate Y_t , which is the total international tourist arrivals.

Proportions of historical averages

The formula is given by:

$$p_i = \frac{(\sum_{t=1}^n \frac{y_{i,t}}{n})}{(\sum_{t=1}^n \frac{y_t}{n})}. \tag{4}$$

Under this method, the individual proportion p_i uses the average historical value of the bottom level series (level 1 series) ($Y_{(X,t)}$) from $t = 1, \dots, 204,$ relative to the total aggregate (Y_t), which is the total international tourist arrivals.

Forecasted proportions

This method makes use of h -step ahead base forecasts that can be denoted by $\hat{Y}_n^l(h)$. Independent lowest level forecasts for all the series in the hierarchy are produced. Proportions of individual base forecasts at each level against the cumulative base forecasts at that level are computed. The general formula for M level hierarchy is given by:

$$p_a = \prod_{l=0}^{M-1} \frac{\hat{Y}_{a,t}^l}{\hat{S}_{a,t}^{l+1}}, \tag{5}$$

where a ($a = 1, 2, \dots, m_k$) denotes a particular bottom level series, $\hat{Y}_{a,t}^l$ represents the base forecasts of the series corresponding to the node which is l levels above a whereas the sum of base forecasts below the series that are l

levels above node a and are directly linked with that series. As for the bottom level series H in Figure (1), we have:

$$p_H = \begin{pmatrix} \hat{Y}_{H,t} \\ \hat{S}_{H,t} \end{pmatrix} \begin{pmatrix} \hat{Y}_{TA,t} \\ \hat{S}_{Total,t} \end{pmatrix}, \tag{6}$$

where

$$\hat{S}_{Total,t} = \hat{S}_{Total\ arrivals,t} \text{ and } \hat{S}_{H,t} = \hat{y}_{(H,t)} + \hat{y}_{(B,t)} + \hat{y}_{(I,t)} + \hat{y}_{(E,t)} + \hat{y}_{(S,t)}.$$

3.4. Optimal combination approach

[54] introduced this approach that makes use of a regression model for the purpose of coming up with optimally combined and reconciled forecasts. With this method, interactions and relationships between series at every level in the hierarchy, as well as special tourism events and seasonality, are considered. The h -step-ahead base forecasts in a hierarchy are represented by:

$$\hat{Y}_n(h) = \mathbf{S}\beta_h + \varepsilon_h, \tag{7}$$

where

$$\beta_h = E [Y_{(M,n+h)} | Y_1, Y_2, \dots, Y_n], \tag{8}$$

is the vector of the unknown means of base forecasts of bottom level M , $\hat{Y}_n(h)$ and ε_h is the error term with a mean of zero and a covariance matrix \sum_h . Equation (7) then reduces to:

$$\hat{Y}_n(h) = \mathbf{S} \times E [Y_{(M,n+h)} | Y_1, Y_2, \dots, Y_n] + \varepsilon_h. \tag{9}$$

Due to the unavailability of information and complexity of computing \sum_h , [54] assumed that the error is predicted using the forecast error of the bottom level, hence $\varepsilon_h \approx \mathbf{S}\varepsilon_{(M,h)}$. With this assumption, the error terms will meet the aggregation constraints as the data in the hierarchy. The unbiased estimator for β_h becomes:

$$\hat{\beta}_h = (\mathbf{S}'\mathbf{S})^{-1} \mathbf{S}'\hat{Y}_n(h) \tag{10}$$

and revised forecasts will be given by:

$$\bar{Y}_n(h) = \mathbf{S}\hat{\beta}_h. \tag{11}$$

The revised forecasts are unbiased because $\mathbf{SPS} = \mathbf{S}$ since $\mathbf{P} = (\mathbf{S}'\mathbf{S})^{-1} \mathbf{S}'$.

3.5. Individual series prediction method

There are various models (Exponential Smoothing (ES), Random Walk (RW) and ARIMA, etc.) that are used in the prediction of individual series at various levels and hierarchies. The ES and ARIMA were adopted by [55] while [2], adopted the ES alone. In this paper, the modelling and forecasting of individual hierarchical series at all the two levels is done using the ES methods through a default R software algorithm incorporated in forecast package by [54]. According to [56], this is the commonly used methods when it comes to generating individual forecasts and in time series analysis. They further acknowledged the inexpensiveness of ES and their ability to produce accurate forecasts in various fields.

According to [57], the best three ES methods are the simple exponential smoothing, Holt-Winters' seasonal method and the Holt's linear trend method. The simple exponential smoothing (SES) is most appropriate to non-seasonal and de-trended time series data ([58]); [56] and [49]). [59] indicated that the Holt-Winters smoothing techniques are more applicable to seasonal business data. Tourism data usually exhibit seasonality and trend hence the individual series cannot be predicted using the SES. As a way of dealing with the tourism trends and seasonal components, the paper will use the Holt-Winters' methods.

3.6. Forecast combination and prediction Intervals

Quantile Regression Averaging (QRA)

The forecasts are combined using Quantile Regression Averaging (QRA) ([46]; [60]. [61] noted that combining or averaging forecasts of two or more models improves accuracy while also reducing the variance of forecasting errors. In this paper we will also use QRA to compute the PIs. The QRA is given as:

$$\hat{y}_{t,\tau}^{QRA} = \beta_0 + \sum_{i=1}^k \beta_i \hat{y}_{ti} + \varepsilon_{t,\tau}, \tag{12}$$

where \hat{y}_{ti} represents the forecasts from method i , $\hat{y}_{t,\tau}^{QRA}$ denotes the combined forecasts and $\varepsilon_{t,\tau}$ is the error term. The parameters of Equation (12) are estimated by minimizing the loss function for a particular τ^{th} quantile,

$$\arg \min_{\beta \in \mathbb{R}^p} \sum_{t: \hat{y}_t^{QRA} > x_t^T \beta} \tau(\hat{y}_t^{QRA} - x_t^T \beta) + \sum_{t: \hat{y}_t^{QRA} < x_t^T \beta} (1 - \tau)(\hat{y}_t^{QRA} - x_t^T \beta). \tag{13}$$

3.7. Performance measures

The Winkler score is used to measure the accuracy of PIs. Coverage rate and interval widths are useful in evaluating performance of PIs, but this study will make use of Winkler score. Winkler’s score accounts for merits of both coverage rates and interval widths ([49]). The Winkler score measures both the interval width and the coverage rate. According to [49], a smaller Winkler score shows better performance of PIs. Given that y_t denotes future predicted values, Winkler score can be expressed as:

$$\text{Winkler} = \begin{cases} w_t, & U_t < y_t \leq L_t \\ w_t + \frac{2(L_t - y_t)}{1 - p}, & L_t \geq y_t \\ \frac{2(y_t - U_t)}{1 - p}, & U_t < y_t. \end{cases} \tag{14}$$

where y_t denotes exact future values, p is the PI percentage, w_t denotes the width of the interval, L_t and U_t are the lower and upper limits, respectively.

The Mean Absolute Percentage Error (MAPE) is used to measure accuracy of hierarchical forecasting approaches. The MAPE can be expressed as:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{\hat{Y}_t} \right|, \tag{15}$$

where \hat{Y}_t and Y_t are the forecasted and actual values, respectively, and n represents total observations.

4. Empirical results and discussion

4.1. Exploratory Data Analysis

Monthly Zimbabwe international tourist arrivals data provided by both the Zimbabwe Tourism Authority (ZTA) and the Zimbabwe National Statistics Agency (ZIMSTAT) from January 2002 to December 2018 are used. The data set used was before the COVID-19 period. The models derived from the data indicate what could happen outside of the pandemic period. The data are disaggregated according to PoV to come up with a hierarchy with Level 0 and Level 1. The descriptive statistics of the tourist arrivals are summarised in Table 1

Table 1. **Summary statistics for tourist arrivals.**

PoV	Mean	Median	Max	Min	St. Dev.	Skewness	Kurtosis
Business	22991.75	18042.50	93166.00	2990.00	16614.55	1.6372	2.4672
Education	2214.97	1369.00	17116.00	84.00	2278.59	2.7502	1.0839
Holiday	82805.38	75267.50	355448.00	14592.00	49188.34	2.1078	7.2440
In-transit	55847.13	53838.00	13618.50	14283.00	22159.74	0.7346	0.7621
Shopping	10937.66	5511.00	95304.00	0.00000	15472.46	2.8064	9.3987

Table 1 shows the distribution of tourist arrivals by PoV. Tourist arrivals for holiday are the highest. The skewness values of international tourist arrivals under different categories are positive, meaning their distributions are non-normal with extremes. The kurtoses for all the cases of PoV are positive which suggests heavy-tailed distributions.

[62] concluded that tourist arrival pattern can be recognised using box-plots. The box-plots for the monthly international tourist arrivals are illustrated in Figure 7 (Annexure 1).

The box-plots in Figure 7 (Annexure 1) show summaries of international tourist arrivals for each category. The box-plots indicate that the international tourist arrivals to Zimbabwe are dominated by those who come for holiday, followed by in-transit, shopping, business and lastly education. All these international tourist arrivals may visit various tourist attractions in various provinces in the country.

Figure 8 (Annexure 2) shows the map of Zimbabwe indicating the country's provinces. Zimbabwe has ten (10) provinces. Each of the ten provinces houses at least one tourist attraction site. In Masvingo there is the famous Great Zimbabwe Monuments. In the Midlands province there is Sebakwe Mountain. There is Matobo National Park, Hwange National Park, Chinhoyi Caves, Goshu Park, Nyanga National Park and Harare International Festival of the Arts in Matabeleland South Province, Matabeleland North Province, Mashonaland West Province, Mashonaland East Province, Manicaland Province and Harare province, respectively.

4.2. Zimbabwe international tourist arrivals.

International tourist arrivals, according to PoV are shown in Figure 2.

From Level 1, the blue line represents holiday (H), the red line represents in-transit (I), the purple line represents shopping (S), the light green represents business (B) and the green line represents education (E). A mixed pattern of international tourist arrivals is shown at both levels, with high numbers experienced between years 2000 and 2008, then a slight increase around 2016. At Level 1, it is noted that most international tourists are for holiday purposes, followed by in-transit, then shopping, and finally business and education visitors. In-transit tourist arrivals are high, probably due to the fact that Zimbabwe connects several Southern African countries (South Africa, Mozambique, Zambia, Botswana and Namibia), hence a lot of foreign tourists pass through Zimbabwe on their way to various neighbouring countries.

4.3. Forecasting the grouped tourism series

[2] used ES based on innovations state space model in forecasting the Austrian tourism demand. The same method is used in this study in modelling and forecasting all Level 0 and Level 1 tourism series. The methods include additive models that result in better forecast accuracy, according to [2].

4.4. Forecasting accuracy of the models

Forecasting accuracy performance at various levels of the hierarchy is done using the mean absolute percentage error (MAPE) based on out-of-sample forecasts. The data from January 2002 to September 2017 are used for

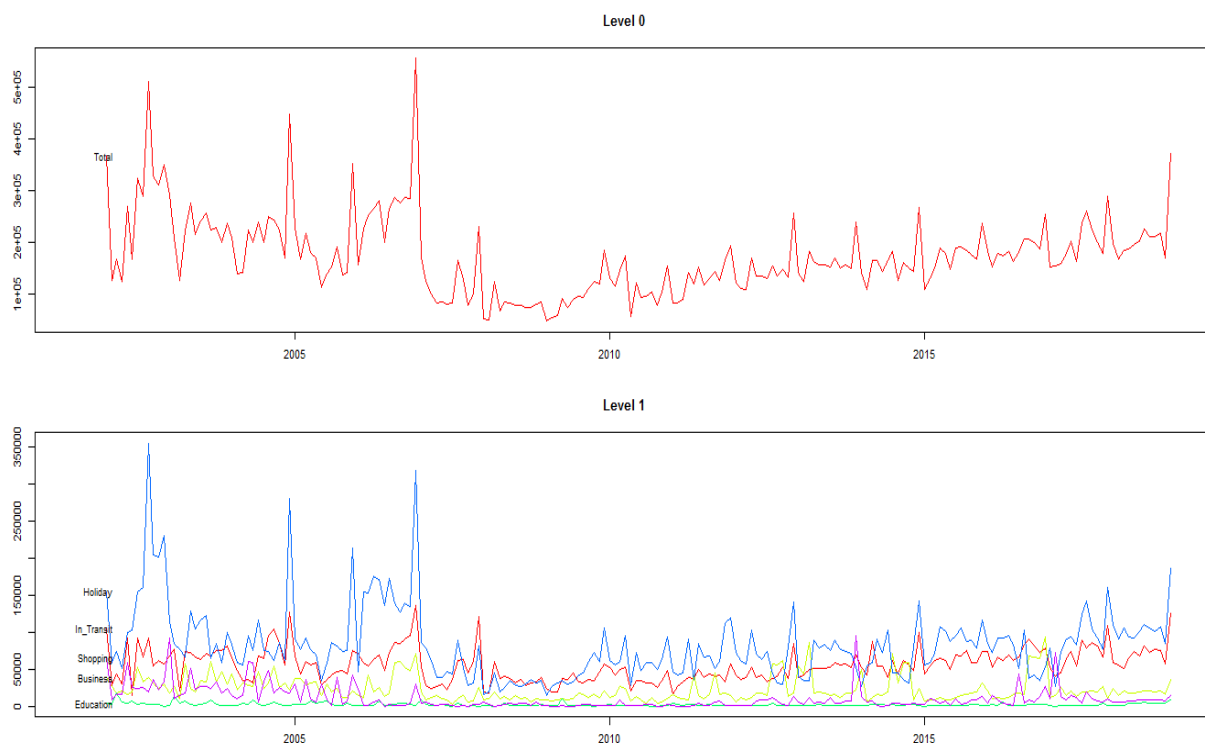


Figure 2. Time series plot of Zimbabwe international tourist arrivals.

training the models. Tourism observations from October 2017 to March 2018 are used for validation. One to six-step-ahead tourism forecasts are done using the bottom-up, top-down (average historical proportions, the proportion of historical averages and forecast proportions) and the optimal combination approaches. These approaches reconcile and optimally combine the international tourist forecasts. The one to six-step-ahead tourism forecasts are done by increasing the sample size by one month (observation) and re-doing the forecasts. Table 2 summarises the results.

The last column labelled 'Average MAPE' in Table 2 indicates the MAPE value for each approach across the entire forecasted horizons (1, 2, 3, 4, 5 and 6). The best performing method is the one with a lower MAPE value and is identified by bold entries in Table 2. For both the upper level series (total tourist arrivals) and lower level series (purpose of visit), it is noted that the bottom-up approach is the best performing method according to the average MAPE values displayed in Table 2. The method is capable of generating more accurate tourism forecasts.

The bottom-up approach is the best performing for Level 0 and Level 1 series, as exhibited in Table 2, since the average MAPE values are lower. The bottom-up approach outperforms the other approaches when forecasting total arrivals as well as educational and shopping tourists, as supported by the MAPE values. However, the optimal combination approach seems to be the best for holiday tourists. The top-down approach based on average historical proportions and forecasted proportions are the best for modelling and forecasting business and in-transit international tourists respectively. [3] showed that any top-down method introduces bias into the reconciled forecasts at each disaggregation level, even if the base forecasts are unbiased. To overcome this short-coming, we are going to combine the forecasts using QRA.

Table 2. Forecasting accuracy results.

Forecast horizon (h)	1	2	3	4	5	6	Average MAPE
Top Level: Zimbabwe							
Bottom-up	0.83	2.55	8.16	6.03	7.63	15.04	6.71
^a Top-down AHP*	3.88	4.87	6.36	5.90	8.17	18.57	7.46
Top-down PHA*	3.88	4.87	6.36	5.90	8.17	15.57	7.46
Top-down FP*	3.88	4.87	6.36	5.90	8.17	15.57	7.46
Optimal Combination	2.17	3.46	6.84	6.03	7.86	16.59	7.16
Bottom level 1: Purpose of visit							
Bottom-up	30.63	28.57	26.94	25.34	29.53	31.35	28.73
Top-down AHP*	32.51	22.83	30.76	27.68	36.27	45.48	32.59
Top-down PHA*	26.78	18.25	27.81	25.75	33.42	41.78	28.97
Top-down FP*	36.12	31.7	26.02	25.61	31.01	33.83	30.72
Optimal Combination	34.22	29.32	26.09	24.51	31.67	33.75	29.93

^aAHP*—average historical proportions, PHA*—proportion of historical averages, FP*—forecast proportions.

It is concluded that the bottom-up approach gives the overall best fit for both hierarchies based on the lower average MAPE values associated with the approach. This may be due to the nature of tourism data and the way it is disaggregated, that is according to PoV. Furthermore, the short-term forecasts used for evaluation purposes may have favoured these approaches. This is consistent with part of the results of [2] study, when they concluded that the bottom-up approach forecasts outperformed top-down forecasts based on forecasted proportions. Furthermore, [43] noted the good performance of the bottom-up approach over the top-down approach.

4.5. International tourist forecasts

The bottom-up approach is applied to predict future 60-months out-of-sample Zimbabwean international tourist arrivals across all levels of the hierarchy. Tables 4 and 5 (Annexure 3 and 4) summarises the forecasts produced using the bottom-up approach, while Figure 3 is the graphical representation of the forecasts and the original series data.

From Figure 3, the solid line(s) represent(s) the historical international arrivals tourism data, while the dashed line(s) represent(s) the projected international arrivals tourism forecasts. Of course these forecasts did not anticipate the Covid-19 pandemic. However, the forecasts give the potential for each grouping beyond the pandemic. The colours represent Level 1 forecasts, the blue line represents holiday (H), red line represents in-transit (I), the purple line represents shopping (S), light green represents business (B) and green line represents education (E).

According to the bottom-up approach, international tourist arrivals are seasonal, as high peaks are observed towards the end of every year. The plot for aggregated series, holiday tourists and in-transit also indicated seasonality with a slightly trend. Holiday tourists have the highest frequency, probably due to the wonderful geographical nature and warm temperatures in Zimbabwe. Harare is the sunshine city of the world climate wise. Investors can invest in the accommodation and transport sectors, targeting these holiday makers and in-transit tourists. The government could channel more resources to the accommodation and transport sectors, so as to accommodate all international tourist arrivals. Tourism managers could also keep on marketing the available tourist attraction centres and be creative to introduce new cultural and other entertaining activities when the COVID-19 pandemic is over or manageable.



Figure 3. Bottom-up approach forecasts.

The plot also indicated that the overall educational tourism flow in Zimbabwe will be low and stable over time, probably due to the deteriorating economy and industry, as well as poor educational facilities. Unavailability and relatively high costs of some basic goods in shops may be the reasons for the relatively constant number of shopping tourists. The education sector may introduce some role model scholarships, with the help of the government, as a way of luring more educational tourists. Zimbabwe has a good reputation in education in the region, but has not harvested any fruits in the form of large numbers of educational tourists, like Australia. Graduates will be ambassadors in their respective countries. Business visitor arrivals were increasing slightly, probably due to the on-going campaign by the president of Zimbabwe, telling the world that Zimbabwe is open for business. There is a need for better policies from the government so that the country will receive more business tourist arrivals and this will improve the country's foreign currency reserves. Improving the political environment in Zimbabwe will certainly help in this endeavour. However the COVID-19 pandemic needs to be brought under control first.

4.6. Forecast combination results and prediction intervals

Forecasts from the bottom-up approach are used in the QRA, particularly the aggregated series denoted by BUtotal, since the rest were not giving good prediction intervals. Plots of the forecasts, density, normal Q-Q plot and the box plot are done and displayed in Figure 4.

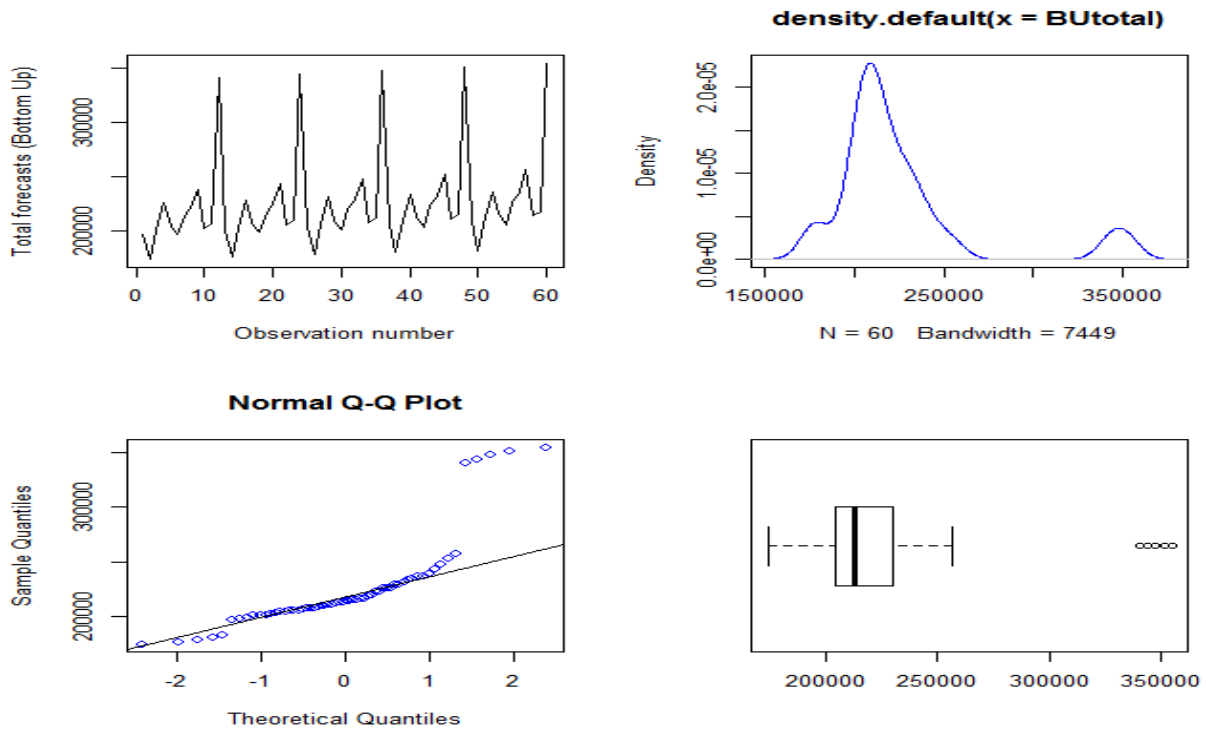


Figure 4. Bottom-up total forecasts plots.

Figure 4 shows some variations in international tourist arrivals and the forecasts are a little bit diverting from a normal distribution, according to the density and normal Q-Q plot. The forecasts (πt) are smoothed using the penalised cubic smoothing spline function. The function is given by:

$$\pi(t) = \sum_{t=1}^n (y_t - f(t))^2 + \alpha \int (f''(t))^2 dx, \tag{16}$$

where α is the smoothing parameter, $f(t)$ is a smoothing spline function at time t obtained from a noisy observation denoted by y_t and $f''(t)$ is the second derivative of the smoothing spline function at time t . According to [63], the Generalised Cross Validation (GCV) method gives an optimal regularisation parameter based on leave-one-out cross validation procedure. The α value based on the GCV ($\alpha = 1:49995$) was not used, as it gave a straight line. It was then selected based on the trial and error method until we got a fairly good fit as shown in Figure 5 ($\alpha = 0:049995$).

The fitted values were then extracted and used as the independent variable in the quantile regression model. The quantile(s) to be estimated is denoted by τ , BUtotal is regressed against BUfitted values for the $\tau = 0 : 5$ for median forecasts and $\tau = 0 : 025$ and $\tau = 0 : 975$ for the lower and upper 95% PIs. 99% PIs were also constructed using a similar procedure. The performance of 99% and 95% PIs were evaluated using Winkler score and the results are in Table 3.

From the results in Table 3, Winkler score values for a 99% PIs are lower than that of 95% PIs for all forecasting periods. This implies 99% PIs outperform the 95% PIs. Because of this, the 99% PI is considered as the best.

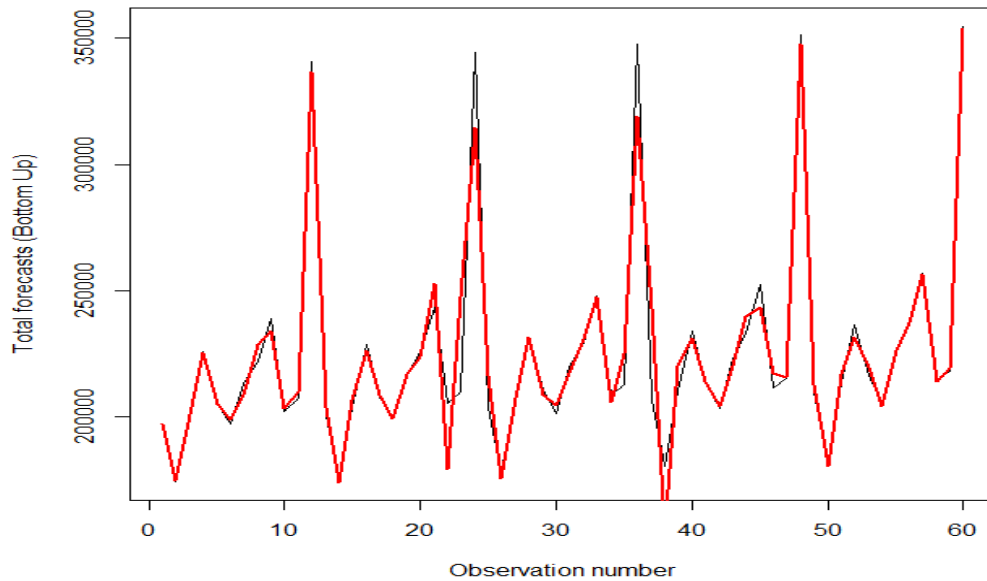


Figure 5. The penalised cubic smoothing spline function.

Table 3. Winkler score.

Forecasting horizon	99% Prediction interval			95% Prediction interval		
	1-step ahead	6-step ahead	12-step ahead	1-step ahead	6-step ahead	12-step ahead
BU	0.416	0.402	0.359	0.419	0.412	0.374

For a full probability distribution of the forecast, we find quantiles in the range $\tau \in (0; 1)$. We further regressed BUtotal against BUfitted values for the $\tau = 0.01$ for median forecasts and $\tau = 0.01$ and $\tau = 0.99$ for the lower and upper 99% prediction intervals. Figure 6 displays forecasts from these values.

Figure 6 indicates forecasts with small deviations from the PIs. The obtained operational forecasts from the QRA are important to decision-makers in the tourism sector.

5. Conclusion

The hierarchical forecasting methodology is applied to Zimbabwe's international tourist arrivals. The bottom-up, optimal combination and three versions of the top-down approaches are considered. The bottom-up approach is considered the best performing approach for Zimbabwe's international tourist arrivals according to the forecasting performance evaluation method. These results are consistent with what was concluded by [64], when their bottom-up approach outperformed other approaches. Our aggregate Zimbabwean international tourist arrival forecasts exhibited a mixed pattern with seasonality over the 60 months forecasted.

Forecasts using disaggregated tourism series by the PoV (holiday and in-transit) indicated a slight increase, while forecasts for educational tourists indicated a low and stable decrease. Educational infrastructure and resources

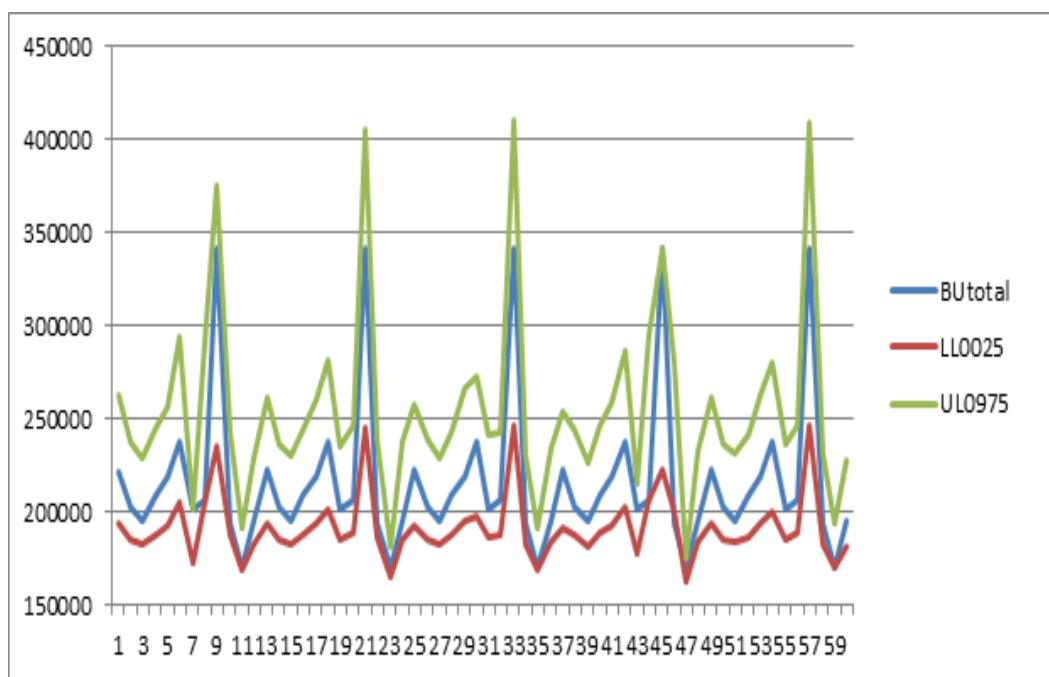


Figure 6. **BUtotal with the 99% prediction interval.**

need to be improved so that the educational sector attracts more foreigners. Zimbabwe has the potential to boost its education tourism in the region of Southern Africa after South Africa. The government, through the Reserve Bank of Zimbabwe (RBZ), could alleviate foreign currency shortages experienced by shops and manufacturing companies so that they are able to restock and produce more goods. This will bring in the much-needed foreign currency into the country, through shopping tourist, thus creating a virtuous cycle.

Since business tourists are few, the president of Zimbabwe and other interested parties have been campaigning and advertising that Zimbabwe is open for business. If a success, this will lead to foreign direct investments and improve the economy and livelihoods of Zimbabweans. Further, business policies need to be amended so that they are favourable to the international investors in the tourism sector. Holiday tourists are the only group with high figures on the hierarchy, motivated mainly by the geographical nature of tourist attraction centres in the country (Victoria Fall Rainforest, Chinhoyi Caves, Great Zimbabwe Monuments, etc.). Disaggregated series allowed us to recognise the nature of international tourists received in Zimbabwe. The hierarchical approach helped us in coming up with demand for specific categories and this helps in the identification of the needed attention and interventions in the particular categories. This is a wealth of information in preparation for the post Covid-19 period.

QRA techniques are applied to the aggregated forecasts from the bottom-up approach. Under the QRA, smoothing techniques are applied to the forecasts and points forecasts with PIs are generated. The aggregated QRA forecasts using 99% PIs indicated a good fit as a small variance is observed. The QRA forecasts can be used to guide business people and other investors. Hence the ZTA could adopt the QRA approach in coming up with improved tourism forecasts with prediction intervals, for investors and government to be guided accordingly. PIs are helpful in planning purposes. The mentioned approaches apply outside the COVID-19 period, assuming the pandemic is eventually contained.

Annexure 1

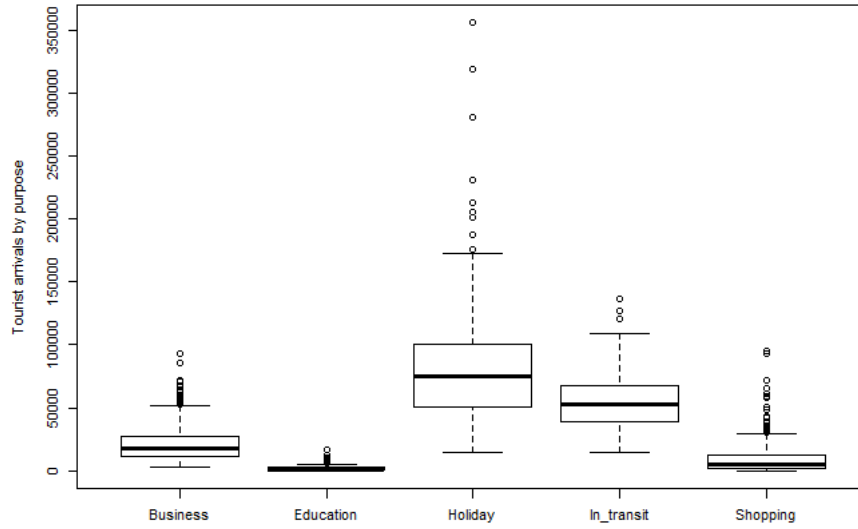


Figure 7. International tourist arrivals from January 2002 to December 2018.

Annexure 2

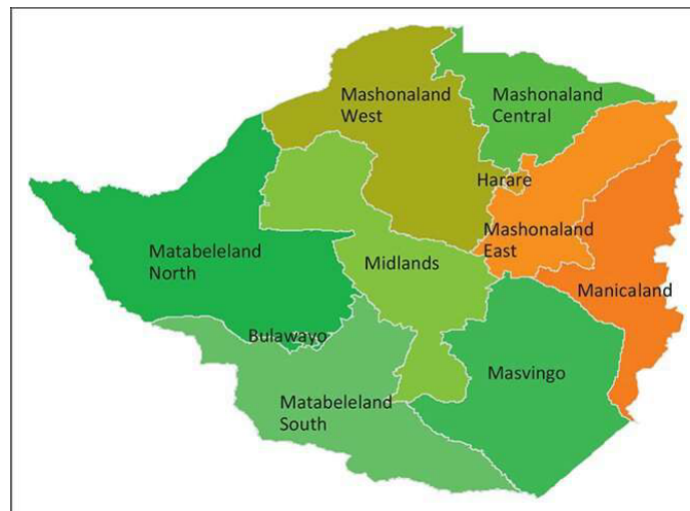


Figure 8. Map of Zimbabwe's provinces.[†]

Annexure 3

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

Date	Total	In-Transit	Business	Education	Holiday	Shopping
Jan-19	197081	68463	26661	4420	87984	9650
Feb-19	174164	60136	17050	3895	83530	9600
Mar-19	200634	67269	26906	4152	92752	9582
Apr-19	225833	69386	22908	7566	116419	9504
May-19	204866	72910	24392	5310	92701	9653
Jun-19	196933	65403	18139	3237	100601	9554
Jul-19	213108	75269	28711	3154	96422	9536
Aug-19	222049	75389	32053	4467	100588	9648
Sep-19	238713	78923	40931	3815	105491	9503
Oct-19	202103	71236	28324	3151	89839	9650
Nov-19	207194	68312	24128	3773	101428	9753
Dec-19	340767	109533	32152	3054	186474	9663
Jan-20	200299	68463	29878	4420	87984	9546
Feb-20	176202	60136	19087	3895	83531	9501
Mar-20	203817	67270	30089	4152	92753	9624
Apr-20	228517	69386	25592	7566	116419	9536
May-20	207696	72910	27221	5310	92702	9589
Jun-20	199018	65403	20223	3237	100602	9572
Jul-20	216376	75269	31978	3154	96422	9503
Aug-20	225663	75389	35667	4467	100588	9533
Sep-20	243285	78923	45502	3815	105492	9663
Oct-20	205237	71236	31458	3151	89839	9702
Nov-20	209840	68312	26773	3773	101429	9576
Dec-20	344261	109533	35646	3054	186475	9656
Jan-21	203517	68463	33095	4420	87985	9544
Feb-21	178239	60136	21124	3895	83531	9460
Mar-21	207000	67270	33272	4152	92753	9506
Apr-21	231201	69387	28275	7566	116420	9554
May-21	210526	72910	30051	5310	92702	9551
Jun-21	201102	65403	22307	3237	100602	9579

Annexure 4

Table 5. Out-of-sample future tourism forecasts.

Date	Total	In-Transit	Business	Education	Holiday	Shopping
Jul-21	219644	75269	35245	3154	96423	9515
Aug-21	229278	75389	39280	4467	100588	9586
Sep-21	247857	78923	50074	3815	105492	9501
Oct-21	208372	71236	34592	3151	89840	9603
Nov-21	212486	68312	29419	3773	101429	9511
Dec-21	347756	109534	39139	3054	186476	9599
Jan-22	206734	68464	36312	4420	87985	9548
Feb-22	180277	60136	23161	3895	83531	9552
Mar-22	210184	67270	36455	4152	92754	9577
Apr-22	233885	69387	30958	7566	116421	9508
May-22	213356	72910	32880	5310	92703	9553
Jun-22	203187	65403	24392	3237	100602	9566
Jul-22	222912	75269	38513	3154	96423	9541
Aug-22	232892	75389	42894	4467	100589	9545
Sep-22	252429	78923	54645	3815	105493	9584
Oct-22	211507	71236	37727	3151	89840	9596
Nov-22	215133	68312	32064	3773	101430	9571
Dec-22	351250	109534	42633	3054	186477	9543
Jan-23	209952	68464	39529	4420	87986	9557
Feb-23	182314	60136	25198	3895	83532	9508
Mar-23	213367	67270	39637	4152	92754	9553
Apr-23	236569	69387	33642	7566	116421	9578
May-23	216187	72910	35710	5310	92703	9566
Jun-23	205272	65403	26476	3237	100603	9726
Jul-23	226180	75269	41780	3154	96423	9591
Aug-23	236506	75389	46508	4467	100589	9576
Sep-23	257001	78923	59217	3815	105493	9588
Oct-23	214642	71236	40861	3151	89840	9555
Nov-23	217779	68312	34710	3773	101430	9601
Dec-23	354745	109534	46126	3054	186478	9546

REFERENCES

1. World Bank, *Ethiopia - Towards a strategy for pro-poor tourism development (English)*, Washington, DC: World Bank, viewed 2 April 2019, 2006. <http://documents.worldbank.org/curated/en/463631468256450819/Ethiopia-Towards-a-strategy-for-pro-poor-tourism-development>.
2. G. Athanasopoulos, R.A. Ahmed, and R.J. Hyndman, *Hierarchical forecasts for Australian domestic tourism*, *International Journal of Forecasting*, vol. 25, no. 2, pp. 146–166, 2009. <https://doi.org/10.1016/j.tourman.2007.04.009>.
3. R.J. Hyndman, R.A. Ahmed, G. Athanasopoulos, and H.L. Shang, *Optimal Combination Forecasts for Hierarchical Time Series*, *Computational Statistics and Data Analysis*, vol. 55, no. 9, pp. 2579–2589, 2011. <https://doi.org/10.1016/j.csda.2011.03.006>.
4. H. Song, and S.F. Witt, *Forecasting international tourist flows to Macau*, *Tourism management*, vol. 27, no. 2, pp. 214–224, 2006.
5. S.B. Taieb, J.W. Taylor, and R.J. Hyndman, *Hierarchical probabilistic forecasting of electricity demand with smart meter data*, viewed 10 March 2019, 2017. http://www.souhaib-bentaieb.com/pdf/jasa_probhts.pdf.

6. H. Booth, *Demographic forecasting: 1980 to 2005 in review*, International Journal of Forecasting, vol. 22, no. 3, pp. 547–581, 2006. <https://doi.org/10.1016/j.ijforecast.2006.04.001>.
7. J. Ndlovu, and E. Heath, *Re-branding of Zimbabwe to enhance sustainable tourism development: Panacea or Villain*, African Journal of Business Management, vol. 1, no. 12, pp. 947–955, 2013.
8. C. Chatfield, *Time series Forecasting*, Chapman & Hall/CRC, Boca Raton, Florida, 2000.
9. H.L. Shang, *Mortality and life expectancy forecasting for a group of populations in developed countries: a multilevel functional data method*, The Annals of Applied Statistics, vol. 10, no. 3, pp. 1639–1672, 2016. <https://doi.org/10.1214/16-aos953>.
10. A. Khosravi, E. Mazloumi, S. Nahavandi, D. Creighton, and J.W.C. Van Lint, *Prediction intervals to account for uncertainties in travel time prediction*, IEEE Transactions on Intelligent Transportation Systems, vol. 12, no. 2, pp. 537–547, 2011. <https://doi.org/10.1109/TITS.2011.2106209>.
11. H. Constantino, J.P. Teixeira, and P.O. Fernandes, *Tourism demand modelling and forecasting with artificial neural network models: The Mozambique case study*, IV Congresso Internacional de Turismo da ESG/IPCA, Tourism for the 21st Century, 3rd and 4th of December 2015.
12. M.A.M.A. Ibrahim, *The determinants of international tourism demand for Egypt: Panel Data Evidence*, European Journal of Economics, Finance and Administrative Sciences, vol. 3, pp. 50-58, 2011.
13. A. Muhammad, and D. Andrews, *Determining tourist arrivals in Uganda: The impact of distance, trade and origin-specific factors*, African Journal of Accounting, Economics, Finance and Banking Research, vol. 2, no. 2, pp. 51–62, 2008.
14. P.K. Chabari, *Determinants of inbound tourism to Kenya*, newblock 2013. <https://pdfs.semanticscholar.org/dd62/a31a707409934801d307b9dc50ff50b09a4d.pdf>.
15. S. Hathroubi, *Modelling and forecasting international tourism demand for Tunisia: A time varying parameter approach*, Interdisciplinary journal of contemporary research in business, vol. 3, no. 7, pp. 261–279, 2011.
16. H. Choyakh, *A model of tourism demand for Tunisia: inclusion of the tourism investment variable*, Tourism Economics, vol. 14, no. 4, pp. 819–838, 2008.
17. A.O. Akuno, M.O. Otieno, C.M. Mwangi, and L.A. Bichanga, *Statistical models for forecasting tourists arrival in Kenya*, Open Journal of Statistics, vol. 5, pp. 60–65, 2015. <http://dx.doi.org/10.4236/ojs.2015.51008>.
18. A. Zakhary, N. El Gayar, and A.F. Atiya, *A comparative study of the pickup method and its variations using a simulated hotel reservation data*, ICGST International Journal on Artificial Intelligence and Machine Learning, vol. 8, pp. 15–21, 2009.
19. H. Song, and G. Li, *Tourism demand modelling and forecasting. A review of recent research*, Tourism Management, vol. 29, no. 2, pp. 203–C220, 2008. <https://doi.org/10.1016/j.tourman.2007.07.016>.
20. D.C. Frechtling, *Practical Tourism Forecasting*, Oxford, UK: Butterworth-Heinemann, 1996.
21. S.F. Witt, and C.A. Witt, *Forecasting tourism demand: a review of empirical research*, International Journal of Forecasting, vol. 11, no. 3, pp. 447–475, 1992.
22. C. Petropoulos, K. Nikolopoulos, A. Patelis, and V. Assimakopoulos, *A technical analysis approach to tourism demand forecasting*, Applied Economics Letters, vol. 12, no. 6, pp. 327–333, 2005. doi:10.1080/13504850500065745.
23. C.N. Mutanga, E. Gandiwa, and N. Muboko, *An analysis of tourist trends in northern Gonarezhou National Park, Zimbabwe, 1991-2014*, Cogent Social Sciences, vol. 3, no. 1, 1392921, 2017.
24. F. Kabote, B. Mashiri, and S. Vengesayi, *Pricing and domestic tourism performance in Zimbabwe*, African Journal of Hospitality, Tourism and Leisure, vol. 3, no. 2, pp. 1–C12, 2014.
25. C. Mutsena, and F. Kabote, *Zimbabwe policy environment and domestic tourism performance*, International Journal of Safety and Security in Tourism and Hospitality, vol. 1, no. 13, pp. 1–14.
26. F. Choga, *Comprehending determinants of demand: Zimbabwe tourism destination scenario*, International Economics and Business, vol. 1, no. 1, pp. 17–23, 2015.
27. E. Muchapondwa, and O. Pimhidzai, *Modelling International Tourism Demand for Zimbabwe*, International Journal of Business and Social Science, vol. 2, no. 2, pp. 71–81, 2011.
28. T. Makoni, and D. Chikobvu, *Hierarchical forecasting of tourist arrivals at the Victoria Rainforest, Zimbabwe*, African Journal of Hospitality, Tourism and Leisure, vol. 7, no. 4, pp. 1–15, 2018.
29. T. Makoni, and D. Chikobvu, *Modelling and Forecasting Zimbabwe's Tourist Arrivals Using Time Series Method: A Case Study of Victoria Falls Rainforest*, Southern African Business Review, vol. 22, pp. 1–22, 2018. <https://doi.org/10.25159/1998-8125/3791>.
30. L. Pascual, J. Romo, and E. Ruiz, *Bootstrap prediction for returns and volatilities in GARCH models*, Computational Statistics & Data Analysis, vol. 50, no. 9, pp. 2293–2312, 2006.
31. C. Choden, and S. Unhapipat, *ARIMA model to forecast international tourist visit in Bumthang, Bhutan*, Journal of Physics: Conference Series, vol. 1039, 8th International Conference on Applied Physics and Mathematics (ICAPM 2018) 27–29 January 2018, Phuket, Thailand. doi:10.1088/1742-6596/1039/1/012023.
32. B. Petrevska, *Predicting tourism demand by A.R.I.M.A. models*, Economic Research-Ekonomska Istraživanja, vol. 30, no. 1, pp. 939–950, 2017. doi:10.1080/1331677x.2017.1314822.
33. H.R.I. Peiris, *A Seasonal ARIMA Model of Tourism Forecasting: The Case of Sri Lanka*, Journal of Tourism, Hospitality and Sports, vol. 22, pp. 98–109, 2016.
34. Y.-W. Chang, and M.-Y. Liao, *A Seasonal ARIMA Model of Tourism Forecasting: The Case of Taiwan*, Asia Pacific Journal of Tourism Research, vol. 15, no. 2, pp. 215–221, 2010. doi:10.1080/10941661003630001.
35. A. Saayman, and M. Saayman, *Forecasting tourist arrivals in South Africa*, Acta Commercii, vol. 10, no. 1, pp. 281–293, 2010. <https://doi.org/10.4102/ac.v10i1.141>.
36. F.S. Witt, H. Song, and P. Louvieris, *Statistical testing in forecasting model selection*, Journal of Travel Research, vol. 42, no. 2, pp. 151–158, 2003.
37. V. Cho, *A comparison of three different approaches to tourist arrival forecasting*, Tourism Management, vol. 24, no. 3, pp. 323C–330, 2003.

38. S.L. Wickramasuriya, G. Athanasopoulos, and R.J. Hyndman, *Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization*, *Journal of the American Statistical Association*, vol. 71, no. 353, pp. 68–71, 2018.
39. L. Morgan, *Forecasting in Hierarchical Models*, viewed 02 January 2019, 2015. <https://www.lancaster.ac.uk/morganle/images/HierarchicalModels.pdf>
40. H. Widiarta, S. Viswanathan, and R. Piplani, *Forecasting aggregate demand: an analytical evaluation of top-down versus bottom-up forecasting in a production planning framework*, *International Journal of Production Economics*, vol. 118, no. 1, pp. 87–94, 2009. <https://doi.org/10.1016/j.ijpe.2008.08.013>.
41. G. Fliedner, *An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation*, *Computers and Operations Research*, vol. 26, no. 10-11, pp. 1133–1149, 1992. [https://doi.org/10.1016/s0305-0548\(99\)00017-9](https://doi.org/10.1016/s0305-0548(99)00017-9).
42. D.W. McLeavey, L. Seetharama, and S.L. Narasimhan, *Production planning and inventory control*, *The Journal of the Operational Research Society*, vol. 36, no. 6, pp. 536–545, 1995. <https://doi.org/10.2307/2582831>.
43. P. Wanke, and E. Saliby, *Top-down or bottom-up forecasting?*, *Pesquisa Operacional*, vol. 27, no. 3, pp. 591–605, 2007. <https://doi.org/10.1590/s0101-74382007000300010>.
44. B.J. Dangerfield, and J.S. Morris, *Top-down or bottom-up: Aggregate versus disaggregate extrapolations*, *International Journal of Forecasting*, vol. 8, no. 2, pp. 233–241, 1992. [https://doi.org/10.1016/0169-2070\(92\)90121-o](https://doi.org/10.1016/0169-2070(92)90121-o).
45. W.R. Kinney Jr., *Predicting earnings: entity versus subentity data*, *Journal of Accounting Research*, vol. 9, no. 1, pp. 127–136, 1971.
46. J. Nowotarski, and R. Weron, *Computing electricity spot price prediction intervals using quantile regression and forecast averaging*, *Computational Statistics*, vol. 30, pp. 791–803, 2015. <https://doi.org/10.1007/s00180-014-0523-0>.
47. E. Torsen, and L.I. Seknewna, *Bootstrapping nonparametric prediction intervals for conditional value-at-risk with heteroscedasticity*, *Journal of Probability and Statistics*, vol. 2019, pp.1–6, 2019. <https://doi.org/10.1155/2019/7691841>.
48. G. Athanasopoulos, J.H. Rob, H. Song, and D. Wu, *The tourism forecasting competition*, *International Journal of Forecasting*, vol. 27, no. 3, pp. 822–845, 2011. <https://doi.org/10.1016/j.ijforecast.2010.04.009>.
49. Z. Li, H. Yu, Y. Liu, and F. Liu, *An improved adaptive exponential smoothing model for short-term travel time forecasting of urban arterial street*, *Acta automatica sinica*, vol. 34, no. 11, pp. 1404–1409, 2008. [https://doi.org/10.1016/s1874-1029\(08\)60062-2](https://doi.org/10.1016/s1874-1029(08)60062-2).
50. J.H. Kim, K. Wong, G. Athanasopoulos, and S. Liu, *Beyond point forecasting: Evaluation of alternative prediction intervals for tourist arrivals*, *International Journal of Forecasting*, vol. 27, no. 3, pp. 887–901, 2011. doi:10.1016/j.ijforecast.2010.02.014.
51. R.L. Winkler, *A Decision-Theoretic Approach to Interval Estimation*, *Journal of the American Statistical Association*, vol. 67, pp. 187–191, 1972.
52. R.J. Hyndman, A.J. Lee, and E. Wang, *Fast computation of reconciled forecasts for hierarchical and grouped time series*, *Computational Statistics and Data Analysis*, vol. 97, pp. 16–32, 2016. <https://doi.org/10.1016/j.csda.2015.11.007>.
53. R.J. Hyndman, and G. Athanasopoulos, *Forecasting: principles and practice*, viewed 12 January 2019, 2014. https://robjhyndman.com/uwa_les/fpp-notes.pdf.
54. R.J. Hyndman, and Y. Khandakar, *Automatic Time Series Forecasting: The Forecast Package for R*, *Journal of Statistical Software*, vol. 27, no. , pp. 1–22, 2007. <https://doi.org/10.18637/jss.v027.i03>.
55. G. Athanasopoulos, R.J. Hyndman, N. Kourentzes, and F. Petropoulos, *Forecasting with temporal hierarchies*, *European Journal of Operational Research*, vol. 262, no. 1, pp. 60–74, 2017. <https://doi.org/10.1016/j.ejor.2017.02.046>.
56. E. Ostertagova, and O. Ostertag, *Forecasting using simple exponential smoothing method*, *Acta Electrotechnica et Informatica*, vol. 12, no. 3, pp.62, 2012. <https://doi.org/10.2478/v10198-012-0034-2>.
57. S. Wang, *Exponential smoothing for forecasting and Bayesian validation of computer models*, viewed 5 May 2019, 2006. <https://smartech.gatech.edu/handle/1853/19753>.
58. M. Ruiter, *Using Exponential Smoothing Methods for Modelling and Forecasting Short-Term Electricity Demand (Bachelor Thesis, Erasmus University Rotterdam, Rotterdam, Netherlands)*, viewed 20 March 2019, 2017. <https://thesis.eur.nl/pub/38564/Ruiter-de.pdf>.
59. S. Gelper, R. Fried, and C. Croux, *Robust forecasting with exponential and Holt-Winters smoothing*, *Journal of forecasting*, vol. 29, no. 3, pp. 285–300, 2010. <https://doi.org/10.2139/ssrn.1089403>.
60. B. Liu, J. Nowotarski, T. Hong, and R. Weron, *Probabilistic load forecasting via Quantile Regression Averaging on sister forecasts*, *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 730–737, 2017. <https://doi.org/10.1109/tsg.2015.2437877>.
61. S. Makridakis, R.M. Hogarth, and A. Gaba, *Forecasting and uncertainty in the economic and business world*, *International Journal of Forecasting*, vol. 25, no. 4, pp. 794–812, 2009. <https://doi.org/10.1016/j.ijforecast.2009.05.012>.
62. W.G.S. Konarasinghe, and N.R. Abeynayake, *Modelling stock returns of individual companies of Colombo Stock Exchange*, *Conference Proceedings of the 1st International Forum for Mathematical Modelling*, pp. 111–115, 2014.
63. G. Wahba, *The approximate solution of linear operator equations when the data are noisy*, *SIAM Journal on Numerical Analysis*, vol. 14, no. 4, pp. 651–667, 1977.
64. A. Zellner, J. Tobias, *A note on aggregation, disaggregation and forecasting performance*, *Journal of Forecasting*, vol. 19, no. 5, pp. 457–469, 2000. [https://doi.org/10.1002/1099-131x\(200009\)19:5](https://doi.org/10.1002/1099-131x(200009)19:5)