African Journal of Hospitality, Tourism and Leisure, Volume 6 (4) - (2017) ISSN: 2223-814X
Copyright: © 2017 AJHTL - Open Access- Online @ http//: www.ajhtl.com

Modelling international tourist arrivals and volatility to the Victoria Falls Rainforest, Zimbabwe: Application of the GARCH family of models

Makoni Tendai*
Great Zimbabwe University, Zimbabwe
Email address: tpmakoni@gmail.com

and

Delson Chikobvu
University of the Free State, South Africa
Email address: Chikobvu@ufs.ac.za

Corresponding author*

Abstract

An increase in international tourist arrivals in Zimbabwe improves the country's foreign currency reserves. The changes in international tourist arrivals flows in Zimbabwe are affected by the vagaries of weather, climate change, and various economic factors. These changes affect the growth of the tourism industry. Accurate forecasting methods are helpful to tourism officials, government officials, tourism stakeholders and all tourism managers and planners, especially for effective future marketing and investment planning. In this paper, we fit a SARIMA-GARCH model to the monthly international tourist arrivals and forecast volatility. Monthly international tourist arrivals for 10 years (2006-2016) from the Zimbabwe Tourism Authority (ZTA) and the Zimbabwe Parks and Wildlife Management Authority (ZPWMA) is used and gives a good fit. The data are logarithmic transformed and the growth rate series is stationary after a seasonal difference. A SARIMA(0,0,1)(0,1,1)$_{12}$ model proved to be the best for forecasting monthly international tourist arrivals and the forecasts from this model indicated a slow increase in international tourist arrivals in the first half of 2017 followed by a decrease in the second half. Tourism officials could introduce new marketing strategies, policies and re-brand the destination so as to keep the tourism industry growing. Squared residuals from the SARIMA model indicated volatility clustering prompting the fitting of a SARIMA-GARCH model. The SARIMA(1,0,0)(0,1,1)$_{12}$-GARCH(1,0) model under normal distribution of errors proved to be the best for modelling international tourist arrival volatility and the model quantifies uncertainty in the future, hence tourism and government officials need to come up with stern measures to effectively and efficiently deal with the future uncertainty.

Keywords: SARIMA, SARIMA-GARCH model, volatility, uncertainty, international tourist arrivals.

Introduction

Tourism is a leading industry with international tourist arrivals and receipts increasing 35 times and 425 times respectively in 2009 since the 1950s worldwide (UNWTO, 2010). The tourism industry is a booster for the economy and international tourism is considered as one of the key source of export receipts in various countries worldwide (Chang et al, 2009**)** and Zimbabwe is no exception. In 2013, international tourist arrivals in the world were above 1 billion (UNWTO, 2014). Over 101 million jobs were being created through the tourism industry; US$2.15 trillion was contributed to the global Gross Domestic Product (GDP) by the tourism sector by 2013, (WTTC, 2014a). This highlight the importance of the tourism industry to the world and Zimbabwe is no exception as the tourism industry is one of the country's key economic sectors. Most of the country's economic sectors (transport sector, infrastructure and

Copyright: © 2017 AJHTL - Open Access- Online @ http//: www.ajhtl.com

development sector, etc) are linked to the tourism sector and benefit from international tourist arrivals.

The Zimbabwean tourism industry is one of the economic revival catalysts as it provides the foreign currency that is needed for sustainable development in many African countries. International tourist arrivals are important to the country as well as the tourist destination town. Balance of payment is directly linked to tourism earnings since positive shocks to international tourist arrivals increase financial reserves, government revenues as well as current account balance (Shareef & McAleer, 2005, Ozer, 2011). This means exchanging rate will be strengthened, resulting in cheaper imports and better lives for local citizens. Improvements in the country's tourism facilities positively affect the country's economic and employment status. Considering that Zimbabwe depends mostly on agriculture, the tourism industry offers diversification and can help in sustaining the country since the agriculture sector is sometimes affected by the vagaries of weather.

Zimbabwe is a tourist destination to several countries from Africa and Europe. It enjoys a fair share of tourist arrivals from all over the world, whose main interest is to visit the country's wildlife areas for photographic and hunting safaris, while others visit cultural sites and mystical areas whose aesthetic value have no parallels anywhere in the world. Zimbabwe is gifted with minerals, geographical, historical and cultural environments that attract tourists from all over the world. Events which attract tourists like, the Zimbabwe International Trade Fair (ZITF), Sanganai Festivals and Carnival among others results in tourism growth (Zimbabwe Statistics Report, 2014).

Victoria Falls Rainforest, Great Zimbabwe Monuments, Chinhoyi curves, Gonarezhou National Park and Hwange National Park are among the popular tourist attractions in Zimbabwe. According to the Zimbabwe Parks and Wildlife Management Authority Statistics Overview (2008), the Zimbabwean Parks estate receives most of its tourist arrivals from European, Chinese, other Asian, Australian and NewZeland markets. It is also noted in the same report that the Victoria Falls Rainforest continues to dominate on contribution to visitors to the estate. It has one of the 7 Natural Wonders of the world.

The Victoria Falls Rainforest is the largest rainforest in Zimbabwe and is endowed with most species, fauna and flora that attract most tourists, scientists and medical researchers. It is housed in the Rainforest National Park and there is the Zambezi National Park, which is within a 7 km radius of the Rainforest. Private game parks, many hotels and lodge facilities are within the range and the Victoria Falls Bridge (over 100yrs old) is nearby. Local art and craft ware, an African Cultural village where different Zimbabwean dances are demonstrated for the enjoyment of tourists are also close to the Victoria Falls Rainforest. Dense forests with very tall trees are some of the features of the Victoria Falls Rainforest. The nature of tourists who visit the Victoria Falls Rainforest are photographing and site viewing visitors, whose main interest is viewing the mystical Rainforest that occupies a piece of land measuring 436 hectares facing the falls (Parks and Wildlife Act 20:14). The majority of visitors (82.4%) who visited Victoria Falls come to the country, mainly for holiday and leisure (Zimbabwe-Visitor Exit Survey Report, 2015/16).

The Victoria Falls Rainforest recorded the highest number of visitations among all the parks in the years 2008, 2009 and 2010 (Tourism Trends and Statistics Report, 2008-2010). According to the Tourism Trends and Statistics Report (2012), the Victoria Falls Rainforest is the most popular Park and Wildlife station. 65% and 44% of the tourist arrivals that visits the top 10 national parks in Zimbabwe during the year 2011 and 2012 were accounted for by the Rainforest (Tourism Trends and Statistics Report, 2011, 2012). During the years 2011 and 2012, 72% and 69% of the Victoria Falls Rainforest visitors were foreigners (Tourism Trends and Statistics Report, 2012). The Rainforest has one of the wonders of the 7 natural Wonders of the world, the Victoria Falls (*Mosi-Oa-Tunya*). An increase in the number of international tourist arrivals at the Victoria Falls Rainforest corresponds to an increase in hotels and lodges in Zimbabwe.

The Zimbabwean tourism industry has a potential for expansion as well as to benefit from the Victoria Falls Rainforest though international tourism is still underdeveloped in Victoria Falls town. Tourism activities in the town are creating employment and generating income for the region. There are many new investment opportunities in African tourism sectors, including Zimbabwe, hence modelling international tourist arrivals and volatility in this rich Victoria Falls Rainforest will guide investors in deciding the types of investment and how much to invest towards the tourism sector. Thoplan (2014) noted the importance of forecasting in modern business when it comes to planning and decision making processes, therefore marketing strategies of the Victoria Falls Rainforest as a tourist destination for the international community will be based on the forecasting model results. Academics and tourism managers have the opportunity to benefit from this research.

The modelling aspect of tourism volatility is important for future planning because tourism volatility is normally due to different advertising campaigns, changes in income earned by tourists, exchange rate in the destination country as well as random events among others. The effects of the economic meltdown and different policy reforms in Zimbabwe resulted in dramatic changes in international tourist arrivals, especially at the Victoria Falls Rainforest tourist destination in particular. Variability in tourist arrivals influenced various researchers (Akuno et al., 2015, Singh, 2013, Loganathan, 2010 and Lim, 2001) use Autoregressive Moving Average (ARIMA) models and Seasonal Autoregressive Moving Average (SARIMA) models in modelling the variations. These models ignored the tourism volatility aspect yet tourism volatility has an impact on, occupancy tare of accommodation, transport, Information Communication Technology (ICT) facilities as well as the country's economy. Knowing the tourism volatility pattern is vital in the formulation of macroeconomic policies and strategic planning in all economic sectors, hence Chang et al., (2009), McAleer et al., (2011), Shareef and McAleer (2005) considered the volatility aspect of tourism demand modelling. It is important for tourism planners have statistical models that explain and forecast international tourist arrival volatility.

The identification of trends and seasonal components of historical time series data is vital for future predictions (Rangika & Peiris, 2016). Tourist arrival series are generally seasonal and volatile hence stochastic models like the SARIMA by Box-Jenkins (1946) and Autoregressive Conditional Heteroscedasticity (ARCH) and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models developed by Engle and Bollerslev, (1982) and (1986) respectively are more appropriate in handling a time-varying series with clustering. McAleer and Divino (2008) concluded that tourism data series can be analysed through the ARCH model proposed by Engle (1982). They concluded this when they were modelling tourism demand to Peru. According to Amos, (2009) forecasting accuracy of a model is affected by heteroscedasticity therefore the adoption of ARCH and GARCH models improves model's forecasting accuracy. This was discovered during modelling South African inflation data.
There is very little literature on the economic implications and importance of tourism in particular destination such as the Victoria Falls Rainforest exist. The linkages between economic performance and tourism development of particular tourist destinations is little known. Wang & Lim (2005) noted that a tourist destination needs accurate forecasts as long as satisfactory tourism services are to be provided, especially in terms of accommodation arrangements, transport arrangements among other services. The purpose of this paper is to adopt the SARIMA-GARCH model in forecasting international tourist arrivals and volatility to Victoria Falls Rainforest using monthly tourist arrival data from January 2006 to December 2016 obtained from the Zimbabwe Parks and Wildlife Management Authority (ZPWMA). This model has the ability to come up with accurate forecasts since it captures seasonality in tourist arrivals as well as volatility in international tourist arrivals. Sigauke et al, (2011) noted that the SARIMA-GARCH model is a model in which a GARCH process is fitted to the variance of the

SARIMA model residuals. The same model was used by Huynh et al, (2015) in forecasting Taiwan tourism demand and the model gave good forecasts.

Saayman and Botha (2015) acknowledged that, the need for accurate tourist arrival predictions are the source of different modelling techniques, hence the adoption of the SARIMA-GARCH model is in line with this. Sigauke et al, (2011) used a SARIMA, RARIMA-GARCH model and Reg-SARIMA-GARCH model in modelling demand for electricity in South Africa. The SARIMA-GARCH model fitted well to the data. The Reg-SARIMA-GARCH model produced better results. These models were used as the data exhibited seasonal patterns.

Volatility clustering analysis leads to better strategic planning and decision making processes (Ozer, Turkyilmaz, 2004) as it accounts better for uncertainty in the series. Monthly international tourist arrival series were used in modelling international tourism volatility by various researchers (McAleer & Divino, 2008, Shareef & McAleer, 2008, Hoti, McAleer & Shareef, 2007, Chan, Lim & McAleer, 2005 and Chan et al., 2004). The estimated volatility models like the (Glosten, Jagannathan and Runkle) GJR(1,1) and GARCH(1,1) models fitted well with the monthly series. This paper will adopt these univariate volatility models because of their ability to capture volatility and to the best of our knowledge, is the first modelling paper on international tourist arrivals to Victoria Falls Rainforest tourist destination.

The paper is structured as follows: From the introduction, literature review of various quantitative methods used in analysing tourism data follow in section 2. Discussion of various steps, univariate seasonal and volatility models used in this study follow in section 3. Data analysis, empirical results and interpretations follow in section 4 and the final section 5 concludes and gives proposed recommendations to the responsible tourism stakeholders.

Literature Review

According to Goh and Law (2002), tourism products and services are perishable hence forecasting and accurate forecasts are important for the tourism industry. Forecasting methods will guide, tourism decision makers. Investors make use of these models because of the dynamic nature of international tourist arrivals as well as uncertainty in the Zimbabwean economic environment.

Gil-Alana et al., (2010) noted that seasonality is an influential factor in tourism demand, therefore models that capture this aspect like the SARIMA model need to be adopted when modelling tourist arrivals. SARIMA models fit well to time series data with seasonality and trends than ARIMA models (Liang, Y-H, 2014). He later adopted the SARIMA-GARCH model in modelling Taiwan tourism demand that surpassed other models (regression, Holt-winter exponential smoothing). SARIMA model can capture the seasonality in the data and GARCH model identifies dynamic changes in international tourist arrivals. Osarumwense and Waziri (2013) fitted an ARMA(1,0)-GARCH(1,0) model during the modelling of the monthly inflation rate of Nigeria. The model indicated less volatility in the future.

During forecasting Australian tourism demand, Smeral and Wüger (2005) noted that the Naïve I outperformed both the SARIMA and ARIMA models. Goh and Law (2002) used different models in modelling Hong Kong's tourism demand and the SARIMA model outperformed several models (Naïve I, Naïve II, Holt exponential smoothing, ARIMA, etc).
Despite this, Song and Li, (2008) concluded that the exponential smoothing, Naïve I, Naïve II and the moving average models are benchmark models though they are good in forecasting purposes.

Osarumwense and Waziri (2013) noted that Autoregressive Moving Average (ARIMA) removes trends on the series; this is the same with SARIMA models, they remove trend and captures seasonality on the series and assumes constant variance which cannot apply to volatile series like tourism series. This indicates the need for ARCH and GARCH models that deals with volatile series. Athanasopoulos et al (2009) modelled and forecasted domestic tourism demand for Australia using GARCH models. Gil-Alana et al., (2009) applied different seasonal statistical models to forecast tourist arrivals to Spain and Canary Islands and obtained useful results.

Chan et al., (2005) noticed that conditional variances affected tourism demand when they used three multivariate GARCH models in examining volatility of Australian tourism demand and various shock effects. ARCH models explain time-varying conditional variances for tourism and financial data better, according to McAleer and Divino (2008), hence these models can be applied to the Zimbabwean case in modelling international tourist arrivals. Chang et al., (2009) showed that the GARCH, GJR and EGARCH models are good enough to model the international tourist arrivals to Taiwan. These models are capable of accounting for the higher volatility persistence. Chang et al., (2009) mentioned that symmetric and asymmetric conditional volatility models like GARCH, GJR and EGARCH models all fit tourism data very well and account for the higher volatility persistence. Furthermore, there will be little clarity on leptokurtosis on the data from the normal GARCH model if conditional variance is not normally distributed since Student's t and EGARCH model captures higher conditional moments (Alexander & Lazar, 2006). Chan et al., (2004) used constant conditional correlation (CCC) volatility models (symmetric CCC-MGARCH model, symmetric VARMA-GARCH model and asymmetric VARMA-AGARCH model) in modelling Conditional Correlations in International Tourism Demand in Australia. They found interdependent effects in the conditional variances among several leading countries. Shareef and McAleer (2007) used GARCH models in examining tourism uncertainty to the Maldives while Hoti et al (2007) used GARCH models in comparing tourism growth and volatility for Cyprus and Malta.

The time varying parameter (TVP) model is one of the techniques that was introduced by Li et al., (2006) and is much more applicable in modelling tourist arrivals. During modelling tourist arrivals to Hong Kong, Song et al., (2011) combined the time-varying parameter (TVP) regression model and structural time series model (STSM) to come up with the TVP-STSM model that outperforms other different models

Reviews of methods used in this study.

A time series approach is used in modelling international tourist arrival data. The approach consists of identifying tourism pattern through the use of time series plot, unit root testing, identifying the order of the model and fitting the appropriate SARIMA and SARIMA-GARCH models. A SARIMA model is used to determine the mean equation whose errors is then modelled by the GARCH family of models due to presence of ARCH effects. Characteristics of the monthly international tourist arrival data favours SARIMA-GARCH family of models because of their ability to quantify volatility. R and E-views are the major statistical packages used in the data analysis process.

Stationarity test

ARCH and GARCH models make use of stationary series, thus the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) unit root test are used to test for stationarity. The tests are conducted on logarithm transformed series. The ADF formula can be expressed as:

$$\Delta X_t = \alpha + \beta_t + \gamma X_{t-1} + \delta_1 \Delta X_{t-2} + \cdots + \delta_{p-1} \Delta X_{t-p+1} + \epsilon_t \qquad (1)$$

Where the difference operator is represented by Δ, α is constant, β is the trend coefficient and γ is the autoregressive process lag.

Seasonal ARIMA model

Most international and domestic tourism data seem to be seasonal and can be modelled better by SARIMA models. A multiplicative seasonal ARIMA model X_t can be denoted by SARIMA(p,d,q)(P,D,Q)$_s$ where *p*, *d*, and *q* are non-seasonal orders and *P,D* and *Q* are seasonal orders and *s* is the seasonal period. The SARIMA model can be expressed as;

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D X_t = \theta_q(B)\Theta_Q(B^s)e_t,\ \ e_t \sim N(0,\sigma_t^2) \quad (2)$$

where, X_t is the monthly international tourist arrival series, B is the backward shift operator that can be defined by $BX_t = BX_{t-1}$, $\nabla_s^D X_t = (1-B^s)^D X_t$ and $\nabla^d X_t = (1-B)^D X_t$.

GARCH Models

Bollerslev (1986) proposed the GARCH(p,q) model which is an econometric term that describes volatility estimation approach using volatile series. This model is normally used to model residuals of various models showing unequal variance as well as exhbit ARCH effect. The GARCH(p,q) model is fitted under three distributional assumptions, namely normal distribution, Student's *t* distribution and Generalized Error Distribution (GED). GARCH(1,1) model can be expressed as:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha e_{t-1}^2 \quad (3)$$

Where σ_t^2 is the constant variance, $\omega > 0, \alpha \geq 0$, $\beta \geq 0$ and $(\alpha + \beta < 1)$. An ARMA(p,q) process with a GARCH noise denoted by X_t that is used to model linear dependency on the series and ARCH effect on residuals can be expressed as:

$$X_t = \mu + \sum_{j=1}^{q}\beta_j\mathcal{E}_{t-j} + \sum_{i=1}^{p}\alpha_i X_{t-i} + \mathcal{E}_t$$

$$\mathcal{E}_t = Z_t\sigma_t, \quad Z_t \sim \text{i.i.d N}(0,1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1\mathcal{E}_{t-1}^2 + \alpha_2\mathcal{E}_{t-2}^2 + \cdots + \alpha_p\mathcal{E}_{t-p}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \cdots + \beta_q\sigma_{t-q}^2 \quad (4)$$

Where α_i, β_j and μ are model parameters and $\mathcal{E}_t$ are the model residuals.

SARIMA-GARCH model

If the variance of the disturbance term of a SARIMA model can be modelled by a GARCH process therefore, that model can be termed a SARIMA-GARCH model. SARIMA-GARCH model can be expressed as follows;

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D X_t = \theta_q(B)\Theta_Q(B^s)e_t,\ \ e_t = z_t\sigma_t,\ z_t \sim N(0,1) \quad (5)$$

where

$$\sigma_t^2 = \mu + \sum_{j=1}^{q}\beta_j\,\mathcal{E}_{t-j}^2 + \sum_{i=1}^{p}\alpha_i\sigma_{t-i}^2 \quad (6)$$

where X_t is the monthly international tourist arrivals growth rate, *p* and *q* are the ARCH and GARCH orders of the model, μ is the constant term, β_j and α_i are the positive model coefficients whose sum should be less than one.

3.5 Maximum Likelihood Estimation (MLE) method

MLE method is adopted in parameter estimation under three distributional assumptions.
The likelihood function of a GARCH(1,1) model with residuals normally distributed is given by

$$\log L_{normal} = -0.5\sum_{t=1}^{n}[\log(2\pi) + \log(\sigma_t^2) + v_t] \quad (7)$$

The likelihood function of a GARCH(1,1) model with residuals that follow Student's t distributed is given by

$$\log L_{Student} = n\left\{\log \Gamma\left(\frac{u+1}{2}\right) - \log \Gamma\left(\frac{u}{2}\right) - 0.5\log[\pi(u-2)]\right\} - 0.5\sum_{t=1}^{n}\left[\log(\sigma_t^2) + (1+u)\log\left(1+\frac{v_t^2}{u-2}\right)\right] \qquad (8)$$

where u represents the degrees of freedom.
Then as for the the log-likelihood function of the GED, the formula will be:

$$logL_{GED} = \sum_{t=1}^{n}\left[log\left(\frac{\mu}{\lambda_\mu}\right) - 0.5\left|\frac{v_t}{\lambda_\mu}\right|^{\mu} - (1+\mu^{-1})\log(2) - log\Gamma\left(\frac{1}{\mu}\right) - 0.5\log(\sigma_t^2)\right] \qquad (9)$$

where $\lambda_\mu = \sqrt{\frac{\Gamma\left(\frac{1}{\mu}\right)2^{\frac{-2}{\mu}}}{\Gamma\left(\frac{3}{\mu}\right)}}$ and $\mu > 0$.

Model selection

The model selection is based on the Akaike information criterion(AIC) proposed by Akaike (1973). It approximates the quality of each model in a given collection of model for data and best model is considered as the one with a smaller AIC value. The AIC is given by the formula:

$$AIC = 2K - 2\ln(L) \qquad (10)$$

where K represents the number of estimated model parameters and L is the maximum value of the likelihood function for the model.

Model diagnostics

Model diagnostics are done on the model residuals where normality and independence of the residuals of the fitted model will be checked. Normality test is done using the Jaque Bera test under the null hypothesis of residuals being normally distributed. Histogram, density plot as well as the normal Q-Q plot is also used in testing the normality of residuals. The independence of the residuals is tested using the Ljung Box test under the null hypothesis of residuals being independently distributed versus the alternative hypothesis of residuals not independently distributed. The Ljung-Box test statistic is given by the formula:

$$Q_{calculated} = n(n+2)\sum_{k=1}^{m}\frac{\hat{r}_k}{n-k} \qquad (11)$$

Where n is the sample size, $\hat{r}_k$ is the autocorrelation at lag k and m is the number of the tested lags.

Forecasting performance evaluation

The Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE) are the measures used in the evaluation of the forecasting performance of the models. The MAPE is given by the formula:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|F_t - A_t|}{A_t} \qquad (12)$$

Where n is the number of observations, F_t are the forecasted values and A_t are the actual values.

Data collection and description

Copyright: © 2017 AJHTL - Open Access- Online @ http//: www.ajhtl.com

The monthly international tourist arrivals series for Victoria Falls Rainforest used in this study are from the ZTA and ZPWMA. The data set is for the period January 2006 to December 2016. The sample size is divided into two parts, the in-sample period (January 2006-December 2015) that is used in the estimation of a SARIMA and SARIMA-GARCH models. The out-of-sample years (January 2016-December 2016) is used in the prediction of out-of sample forecasts and determination of forecasting accuracy performance. The monthly international tourist arrivals series is denoted by X_t and t represents the month. The monthly international tourist arrivals growth rate Y_t is given by the formula:

$$Y_t = \log\left(\frac{X_t}{X_{t-1}}\right) * 100 \qquad (13)$$

where t is the time period in months, X_t and X_{t-1} are the current and previous monthly international tourist arrivals respectively.

The main characteristics of a data set are summarised by the exploratory data analysis (Tukey, 1961), therefore it is done. The summary comprises of sample means, standard deviations, maximums, minimums, medians, skewness, kurtosis, Jarque-Bera (JB) test as well as the p-value. It helps in the preliminary model selection; examine certain distributional assumptions as well as possible data transformations that are needed.

Statistics	Original (X_t)	Growth rate (Y_t)
Number of Observations	132	131
Mean	11654.72	0.646705
Median	10808.50	-2.346629
Maximum	27019.00	57.62033
Minimum	4521.000	-46.72356
Std. Dev.	4849.444	23.25319
Skewness	0.768244	0.315047
Kurtosis	3.035192	2.412438
Jarque-Bera	12.99118	4.051426
Probability	0.001510	0.131900

Table 4.1: Descriptive statistics of X_t and Y_t

Table 4.1 shows the descriptive statistics of the monthly international tourist arrivals that consists of 132 observations and the international monthly tourist arrival growth rate with 131 observations. The growth rate has a mean of 0.646705, a median of -1.544103, maximum of 66.31760 and a minimum of -103.9542. The Jarque-Bera test probability value (0.011190) in Table 4.1 suggests the acceptance of the null hypothesis of normal distribution of the data at the 1% significance level. The tail of the data on the left side is longer than that on the right side because of the negative skewness coefficient; therefore the data is positively skewed. The data have fat tails and is also peaked as indicated by the kurtosis value which is above 3.

Time series plots

The evaluation of the behaviour, patterns and seasonal effects of the log-transformed monthly international tourist arrivals and the monthly international tourist arrivals growth rate over time is done using time series plots.

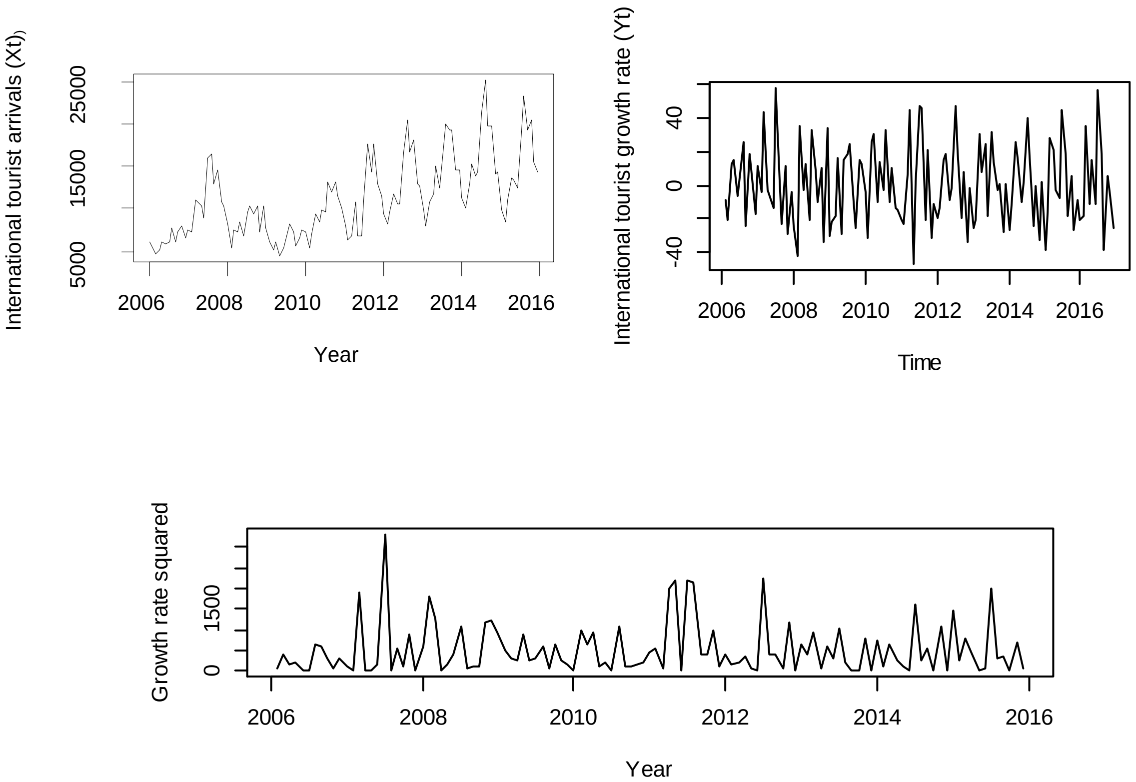

Figure 4.1 Time series plot of X_t, Y_t and the growth rate squared

The visual inspection of the time series plot in Figure 4.1 indicates that the plots of the monthly international tourist arrivals X_t are not stationary. There is an upward surge around 2007 followed by a decrease and later an upward trend on the larger part of the graph. An upward trend, especially starting from 2009 is due to the introduction of multiple currency and improved tourism policies in the country. The international monthly tourist arrivals growth rate displays volatility clustering and the series are fluctuating around the constant mean. Volatility persistence is being noticed in the growth rate squared.

Unit root test

Unit root test based on both the ADF and PP test statistic is done using E-views 9.0 software package. The test is done on the growth rate (Y_t) and results are summarised in Table 4.2.

Monthly international tourist arrival growth rate (Y_t)		
ADF test statistic: -3.184041	1% Critical value: -2.584539 5% Critical value: -1.943540 10% Critical value: -1.614941	P-value: 0.0017
PP test statistic: -12.60172	1% Critical value: -2.582872 5% Critical value: -1.943304 10% Critical value: -1.615087	P-value: 0.0000

Table 4.2: ADF unit root test results

Both the ADF and PP test results in Table 4.2 lead to the rejection of the null hypothesis of presence of a unit root as supported by the ADF test value of -3.184041 with p-value of 0.0017 and the PP test value of -12.60172 with p-value of 0.0000. We can safely conclude that the monthly international tourist arrival growth rate is stationary.

Copyright: © 2017 AJHTL - Open Access- Online @ http//: www.ajhtl.com

SARIMA model identification and estimation

The ACF and PACF of the stationary series indicate strong seasonality hence a seasonal differencing leads to a series that can be modelled parsimoniously. The ACF and PACF after a seasonal difference of the monthly international tourist arrival growth rate is used in the identification of the rough order of the mean equation.

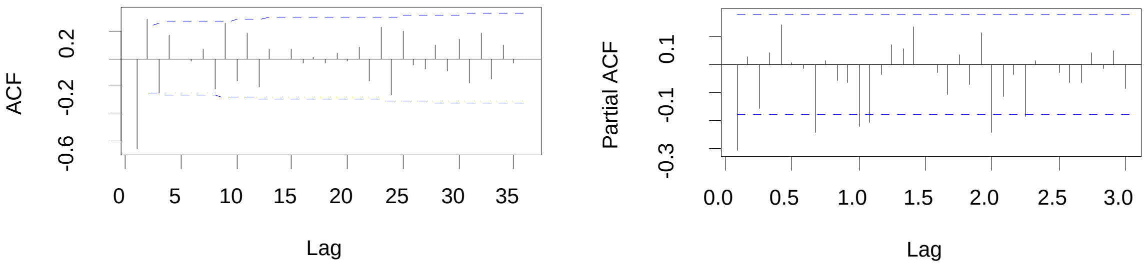

Figure 4.2: ACF and PACF of stationary series

Figure 4.2 confirms very little autocorrelation in the series after the seasonal difference and suggests a simple model which incorporates lag 1and lag 2 seen in the ACF as well as significant lags in the PACF. A SARIMA$(1,0,0)\times(0,1,1)_{12}$ model is fitted first as suggested by the ACF and PACF though other models are being fitted.

Model	AIC
SARIMA$(1,0,0)\times(0,1,1)_{12}$	1010.63
SARIMA$(0,0,1)\times(0,1,1)_{12}$	1012.27
SARIMA$(1,0,1)\times(0,1,1)_{12}$	1012.51
SARIMA$(1,0,1)\times(1,1,1)_{12}$	1014.18
SARIMA$(1,0,0)\times(1,1,0)_{12}$	1033.99
SARIMA$(1,1,1)\times(1,1,1)_{12}$	1014.60

Table 4.3: AIC of SARIMA models

The SARIMA$(0,0,1)\times(0,1,1)_{12}$ model proved to be a better model for the monthly international tourist arrival growth rate because of lowest AIC (1082.23). Parameters of this model are summarised in Table.4.4.

	Coefficients	Std. Error	t-Statistic	Prob.
ar(1)	-0.3649	0.0864	-4.2234	0.0000
sma(12)	-0.8329	0.1257	-6.6261	0.0000

Table 4.4: Parameter estimates of the SARIMA$(1,0,0)\times(0,1,1)_{12}$ model

Model parameters in Table 4.4 are all significant at the 1% level of significance as indicated by the probabilities.

SARIMA model validation

We further examine the model residuals starting with constructing the ACF of the model residuals.

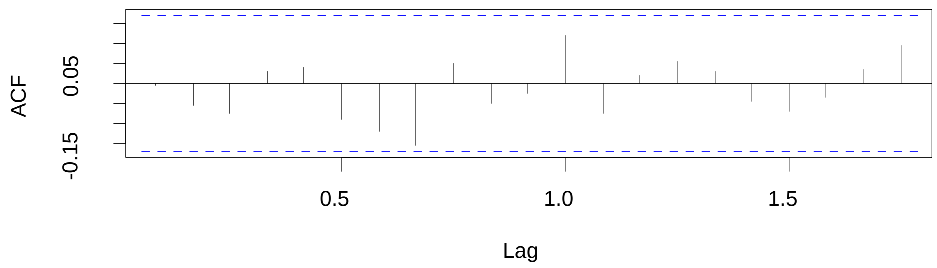

Figure 4.2: Residuals of SARIMA$(1,0,0)\times(0,1,1)_{12}$ model

Figure 4.2 shows that the model residuals are uncorrelated as all the spikes are within the desired bounds. Again, the Box-Ljung statistic 13.8014 and the probability 0.8404 supports the assumption of no autocorrelation on the residuals. The Jarque-Bera test for normality suggested that the model residuals are normally distributed at 1% level of significance because of the Jarque-Bera test statistic of 7.2371 and the probability of 0.02682. The forecasting performance of this SARIMA(1,0,0)(0,1,1)$_{12}$ model together with other various SARIMA models is determined.

Model	RMSE	MAE	MAPE
SARIMA(0,0,1)(1,1,0)$_{12}$	17.15400	13.19631	213.5209
SARIMA(1,0,0)(0,1,1)$_{12}$	15.76115	12.35983	175.6433
SARIMA(1,0,0)(1,1,0)$_{12}$	18.36120	14.20347	225.1933

Table 4.5: Forecasting accuracy measures

When it comes to forecasting international tourist arrivals, SARIMA(1,0,0)(0,1,1)$_{12}$ model is the best according to RMSE (15.76115), MAE (12.35983) and MAPE (175.6433).

SARIMA forecasts

The SARIMA(1,0,0)(0,1,1)$_{12}$ model observed all the assumptions and used to come up with the future international tourist arrivals out of sample forecasts for the next 24 months.

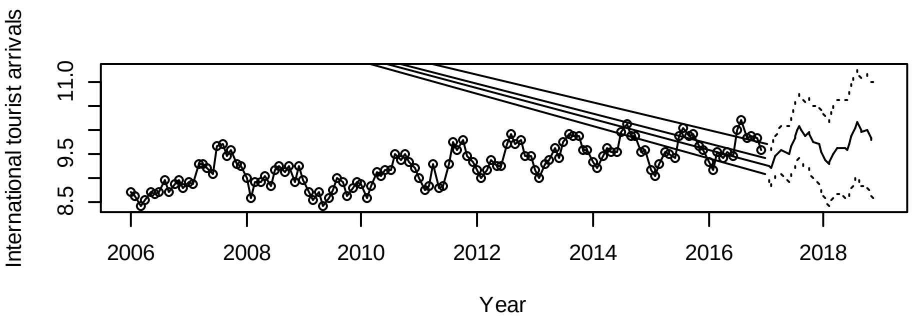

Figure 4.3: Future monthly international tourist arrivals out-of-sample forecasts.

It is noted from Figure 4.3 that international tourist arrivals will increase in the first half of each year (2017 and 2018) though at a slower rate and later decrease in the second half of each of the years (2017 and 2018). From these results, the Ministry of Tourism in conjunction with the government must come up with revived policies and marketing strategies that lure more international tourists, so that Zimbabwe will keep on receiving more tourists, especially from abroad as this have an impact to our foreign currency reserves as well as economy.

Heteroskedasticity and ARCH test on residuals

Heteroskedasticity and ARCH effect on the residuals of the SARIMA(1,0,0)(0,1,1)$_{12}$ model are examined in order to see if the GARCH family of models can be adopted to model volatility. The ARCH LM statistic 18.842 and probability of 0.09242 suggest the rejection of the null hypothesis of no ARCH effect on the model residuals. Presence of ARCH effect is also supported by the Box-Ljung test statistic of 19.5445 and the probability of 0.04867 on squared residuals. A plot of residuals squared is done to visualisation heteroskedasticity.

African Journal of Hospitality, Tourism and Leisure, Volume 6 (4) - (2017) ISSN: 2223-814X
Copyright: © 2017 AJHTL - Open Access- Online @ http//: www.ajhtl.com

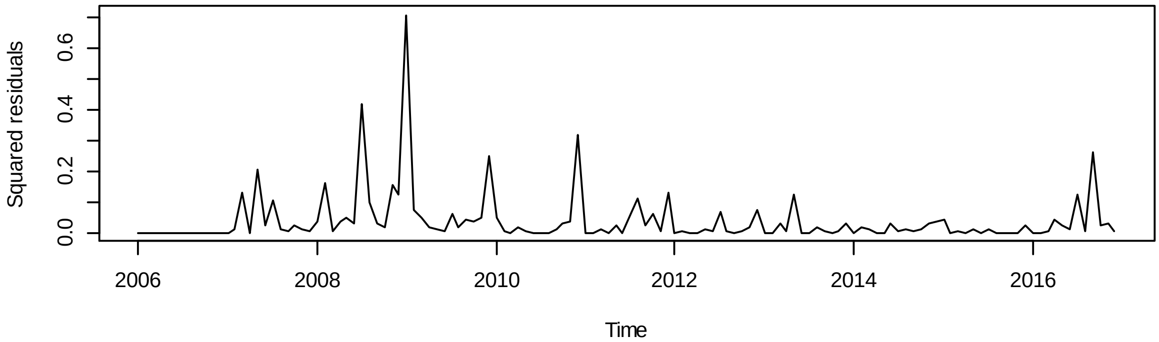

Figure 4.4: Squared Residuals of SARIMA(1,0,0)×$(0,1,1)_{12}$ model

Figure 4.4 suggests the presence of heteroskedasticity on the squared residuals. There is volatility clustering in the squared residuals prompting the adoption of GARCH family of models. Because of this, different SARIMA-GARCH models are fitted and the AIC is used in selecting the best model.

SARIMA-GARCH model identification.

The E-views and R software packages are used in the model identification and parameter estimation since they supports a range of univariate distributions. AIC and BIC of the estimated models are noted.

Models	**AIC**	**BIC**
SARIMA(1,0,0)$(1,1,0)_{12}$-GARCH(1,1) ged	8.412077	8.574415
SARIMA(1,0,0)$(1,1,0)_{12}$-GARCH(1,0) std	8.404934	8.567272
SARIMA(1,0,0)$(0,1,1)_{12}$-GARCH(1,0) norm	8.037506	8.138013
SARIMA(1,0,1)$(1,1,1)_{12}$-GARCH(1,1) std	8.250077	8.466528
SARIMA(1,0,0)$(0,1,0)_{12}$-GARCH(1,1) std	8.628607	8.754241
SARIMA(0,0,1)$(0,1,0)_{12}$-GARCH(1,1) ged	8.753681	8.878580

Table 4.6: AIC and BIC of SARIMA-GARCH models

According to Table 4.6, the SARIMA(1,0,0)$(0,1,1)_{12}$-GARCH(1,0) model under normal distribution proved to be the model with the smallest AIC (8.037506) and BIC (8.138013) and is preferred among the other models. The optimal parameters of this model are summarised in Table 4.7.

	Estimate	Std. Error	z-Statistic	Prob
ar(1)	-0.413203	0.110550	-3.737698	0.0002
sma(12)	-0.959783	0.013248	-72.44786	0.0000
mu	114.7081	17.83469	6.431737	0.0000
alpha1	0.438834	0.214318	2.047578	0.0406

Table 4.7: SARIMA(1,0,0)$(0,1,1)_{12}$-GARCH(1,1) parameters.

All the estimated coefficients of SARIMA(1,0,0)$(0,1,1)_{12}$-GARCH(1,0) under the normal distribution assumption in Table 4.7 are statistically significant at the 5 % level of significance and their signs are appropriate to ensure that the conditional variance is positive. The significance of α indicates that past period's volatility news impacts current volatility. The value of alpha is small (0.4) meaning that volatility is not spiky and reacts slowly to tourism shocks according to Dowd, (2002). This indicates minimum uncertainty level in the long-run and short-run effects of shocks in the monthly international tourist arrival growth rate.

SARIMA(1,0,0)(0,1,1)$_{12}$-GARCH(1,0) model validation

Serial correlation on model residuals is examined through the correlogram of the standardised residuals squared. The null hypothesis of no serial correlation on residuals is accepted as evidenced by the p-values which are above 5% (see Annexure 1). Presence of ARCH effect on model residuals is also examined.

Heteroskedasticity Test: ARCH			
F-statistic	0.385271	Prob.F (1,103)	0.5376
Obs*R-squared	0.388556	Prob.Chi-square(1)	0.5331

Table 4.8: Heteroskedasticity Test

The null hypothesis of no ARCH effect on model residuals is being accepted as supported by the p-values in Table 4.8. The Jarque-Bera test for normality suggests that the residuals are normally distributed at 1% because of the p-value of 0.0153. From all the diagnostic check-up, the model residuals behave as a white noise, hence the model specification is good and fitted well to the data. Forecasting performance of the SARIMA-GARCH models are examined so as to find the best model that can forecast international tourist arrivals as well as modelling international tourist arrival volatility.

Model	RMSE	MAE	MAPE
SARIMA(0,0,1)(0,1,0)$_{12}$-GARCH(1,1) ged	19.65878	14.47363	156.0056
SARIMA(1,0,0)(0,1,1)$_{12}$-GARCH(1,0) norm	13.52516	10.01009	126.6030
SARIMA(1,0,1)(1,1,1)$_{12}$-GARCH(1,1) std	13.94712	10.22037	129.0571

Table 4.9: Forecasting performance measures

The SARIMA(1,0,0)(0,1,1)$_{12}$-GARCH(1,1) model under normal distribution is the best in modelling and forecasting international tourist arrivals volatility as suggested by small RMSE, MAE and MAPE in Table 4.9. The model is used to forecast conditional variance.

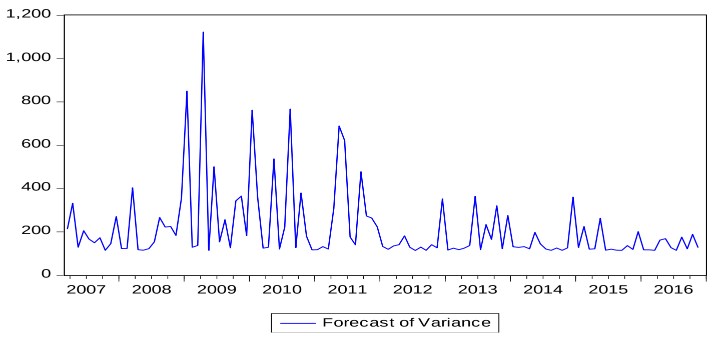

Figure 4.5: Forecast of variance

The forecast of conditional variance in Figure 4.5 is non-constant over the study period. This suggests uncertainty in the monthly international tourist arrivals to Victoria Falls Rainforest. The Presidential elections held in 2008 and the inclusive government formed in 2009 probably resulted in high volatility between 2008 and 2010 while the stabilising of the economy and the end of the inclusive government resulted in low volatility visualised from around 2012 and afterwards. The SARIMA(1,0,0)(0,1,1)$_{12}$-GARCH(1,1) model is used to predict future volatility.

Month	Jan 2017	Feb 2017	Mar 2017	Apr 2017	May 2017	Jun 2017
Volatility	115.1509	165.2397	187.2201	196.8658	201.0986	202.9561
Month	Jul 2017	Aug 2017	Sep 2017	Oct 2017	Nov 2017	Dec 2017
Volatility	203.7712	204.1289	204.2859	204.3548	204.3850	204.3983

Table 4.10: Volatility values

Volatility values in Table 4.10 are increasing at a slow rate implying minimum uncertainty in the future. Similar conclusions were made in Taiwan by Huynh et al, (2015) during modelling

Taiwan's tourism demand. The result calls for the best policies from the tourism industry officials in conjunction with the government to take advantage of the reduced uncertainty in future international tourist arrivals. It becomes easier to plan for future tourist arrivals in the face of reduced uncertainty.

Conclusion

The paper modelled monthly international tourist arrivals to Victoria Falls Rainforest and their volatility. It was noted that the SARIMA$(1,0,0)(0,1,1)_{12}$ model with RMSE (15.76115), MAE (12.35983) and MAPE (175.6433) is the best for monthly international tourist arrivals. The forecasts from this model indicated a growth in international tourist arrivals. Accurate prediction of monthly international tourist arrivals is vital to decision and policy makers in the tourism industry. Marketing strategies, policy reforms and destination re-branding depends on these accurate forecasts with minimum error or uncertainty. Presence of ARCH effect on the SARIMA $(1,0,0)(0,1,1)_{12}$ model residuals lead to the fitting of the SARIMA$(1,0,0)(0,1,1)_{12}$-GARCH(1,0) model under normal distribution of errors. The model was shown to be best as measured by lower RMSE (13.52516), MAE (10.01009) and MAPE (126.6030) and outperformed other various models in modelling monthly international tourist arrival volatility. Tourism shocks will have a short-term impact on international tourist arrivals according to the variance equation. Modelling and forecasting international tourist arrival volatility gives a rough picture of the uncertainty ahead. Tourism authorities will make future plans based on these results. Further studies will be to analyse the effectiveness of current advertising strategies as well as the optimal advertising time for international tourist arrivals. Another further study will be determining the major factors affecting international tourist arrivals.

References

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedacity. *Journal of Econometrics*. 31(3), 307–327.

Chan, F., Lim, C. & McAleer, M. (2005), Modelling multivariate international tourism demand and volatility, *Tourism Management*, 26(3), 459-471.

Coskun, I.O. & Ozer, M. (2011). MGARCH modeling of inbound tourism demand volatility in Turkey. *MIBES Transactions*, 5(1), 24-40.

Goh, C., & Law, R. (2002). Modelling and forecasting tourism demand for arrivals with stochastic non-stationary seasonality and intervention. Tourism Management, 23(5), 499-510.

Hoti, S., McAleer, M. & Shareef, R. (2007), Modelling international tourism and country risk spillovers for Cyprus and Malta, *Tourism Management*, 28(6), 1472-84.

Liang, Y. H. (2014). Forecasting models for Taiwanese tourism demand after allowance for Mainland China tourists visiting Taiwan. Computers & Industrial Engineering, 74, 111-119

McAleer, M. & Divino, J. A. (2008). Modelling the growth and volatility in daily international mass tourism to Peru, unpublished paper, Department of Economics, Catholic University of Brasilia, Brazil.

Osarumwense, O, I. & Waziri, E. I., (2013). Modeling Monthly Inflation Rate Volatility, using Generalized Autoregressive Conditionally Heteroscedastic (GARCH) models: Evidence from Nigeria. *Australian Journal of Basic and Applied Sciences*, 7(7), 991-998.

Shareef, R. & McAleer, M (2007), “Modeling the Uncertainty in Monthly international Tourist Arrivals to the Maldives”, *Tourism Management*, 28(1), 23-45.

Shareef, R. & McAleer, M. (2005), "Modeling International Tourism Demand and Volatility in Small Island Tourism Economies", *International Journal of Tourism Research*, 7(6), 313-333.

Shareef, R. & McAleer, M. (2008), "Modeling International Tourism Demand and Uncertainty in Maldives and Seychelles: A Portfolio Approach", *Mathematics and Computers in Simulation*, 78(2-3), 459-468.

Shareef, R. & McAleer, M. (2008), Modelling international tourism demand and uncertainty in Maldives and Seychelles: a portfolio approach, *Mathematics and Computers in Simulation*, 78, 459-68.

Song H, Li G, Witt SF, Athanasopoulos, G. (2011), Forecasting Tourist Arrivals Using Time-Varying Parameter Structural Time Series Models. *International Journal of Forecasting*, 27 (3), 855-869.

Thoplan, R. (2014) Qualitative v/s Quantitative Forecasting of Yearly Tourist Arrivals in Mauritius, *International Journal of Statistics and Applications*, 4(4), 198.

UNWTO (2014). Tourism Highlights. Retrieved from http://mkt.unwto.org/publication/unwtotourism-highlights-2014-edition.

UNWTO, 2010, World tourism barometer 8(1), (http://www.unwto.org).

Wang, Y. & Lim, C. (2005). Using Time Series Models to Forecast Tourist Flows, Proceedings of the 2005 International Conference on Simulation and Modelling.

WTTC (2014a). Travel and Tourism Economic Impact 2014: World.

Copyright: © 2017 AJHTL - Open Access- Online @ http//: www.ajhtl.com

Annexure 1: Serial correlation test results

Correlogram of Standardized Residuals Squared

Included observations: 100

Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.061	-0.061	0.4001	
		2	0.090	0.086	1.2879	
		3	0.250	0.263	8.2228	0.004
		4	-0.029	-0.003	8.3180	0.016
		5	0.019	-0.035	8.3590	0.039
		6	-0.025	-0.096	8.4315	0.077
		7	0.126	0.139	10.263	0.068
		8	0.038	0.082	10.430	0.108
		9	0.076	0.095	11.103	0.134
		10	0.135	0.069	13.277	0.103
		11	0.046	0.023	13.534	0.140
		12	-0.098	-0.174	14.706	0.143
		13	0.200	0.161	19.651	0.050
		14	-0.092	-0.068	20.715	0.055
		15	0.030	0.077	20.828	0.076
		16	0.103	0.012	22.175	0.075
		17	-0.062	-0.047	22.664	0.092
		18	0.051	-0.046	23.006	0.114
		19	0.030	0.059	23.127	0.145
		20	0.046	0.026	23.412	0.175
		21	0.026	0.065	23.500	0.216
		22	0.001	-0.040	23.500	0.265
		23	-0.079	-0.167	24.360	0.276
		24	0.144	0.131	27.237	0.202
		25	-0.028	0.073	27.350	0.241
		26	0.094	0.110	28.618	0.235
		27	0.002	-0.050	28.619	0.280
		28	0.004	-0.071	28.621	0.329
		29	0.195	0.116	34.300	0.157
		30	-0.102	0.013	35.878	0.146
		31	0.006	-0.060	35.883	0.177
		32	0.019	-0.033	35.939	0.210
		33	-0.020	-0.013	36.001	0.246
		34	-0.005	-0.030	36.005	0.286
		35	-0.029	-0.053	36.144	0.324
		36	-0.051	-0.060	36.577	0.350