Navarra Center for International Development

Working Paper nº 06/2010

TOURISM IN SOUTH AFRICA. TIME SERIES PERSISTENCE AND THE NATURE OF THE SHOCKS. ARE THEY TRANSITORY OR PERMANENT?

Luis A. Gil-Alana
University of Navarra, Faculty of Economics, Pamplona, Spain

Navarra Center for International Development
WP-06/2010

Tourism in South Africa. Time series persistence and the nature of the shocks. Are they transitory or permanent?
Luis A. Gil-Alana
Working Paper No. 06/2010
October 2010

-

ABSTRACT

This paper deals with the analysis of several time series related with the tourism sector in South Africa. We examine the orders of integration of the series in order to determine if they are stationary or nonstationary, and, more importantly, if the shocks affecting them are of a permanent or a transitory nature. Policy implications based on the results obtained are derived.

Keywords: South Africa; tourism; fractional integration.
JEL Classification: C22.

Corresponding author: Prof. Luis A. Gil-Alana
University of Navarra
Faculty of Economics
E-31080 Pamplona, Spain

Phone: 00 34 948 425 625
Fax: 00 34 948 425 626
Email: alana@unav.es

1. Introduction

International travel to South Africa has surged since the end of apartheid. In 1994, the year of South Africa's first democratic elections, the country has improved its tourism position from the 52^{nd} most visited destination in the world to the 17^{th} most visited in 2005. In 1994, only 3.9 million foreign visitors arrived in the country. In 2004, international arrivals had more than doubled to 6.7 million, and in 2007, a total of 9.07 million foreigners visited South Africa _ an 8.3% increase over 2006. Moreover, during the last ten years the country has expanded its tourism plant significantly. This includes growth in the number of hotels, guest houses, game farms, lodges and even the number of airlines servicing the country. From a demand side, the country has experienced an increase of more than 100% in tourist arrivals over the same period. This growth has been despite many problems and threats experienced on the African continent such as political instability, poverty, disease and low level of development (Saayman and Saayman, 2008), and has positioned South Africa as Africa's leading tourist destination.

Directly and indirectly, tourism constitutes approximately 7% of employment in South Africa (TBCSA, Tourism Business Council of South Africa). Ideally placed to create new jobs and to add value to the country's many natural and cultural resources, tourism has been earmarked by the government as one of South Africa's growth sectors.

The challenge for a country like South Africa, as for most other tourist destinations is to sustain tourism growth and to improve the country's position among the most visited countries (Saayman, 2006). In order to address the sustainability issue it is important to understand the stochastic properties of the time series related with the tourism sector since the past history of the data can give us information about the stationary/nonstationary and the mean-reversion/non-mean-reversion behaviour of the series. This is crucial to determine the nature of the shocks, to know if they have permanent or transitory effects. Clearly, different policy measures should be implemented depending on the nature of the shocks, and this is clearly connected with the mathematical models associated to the tourism data. In a shock we mean an event which takes place at a particular point in the series, and it is not confined to the point at which it occurs. A shock is known to have a temporary or short term effect, if after a number of periods the series returns back to its original performance level. On the other hand, a shock is known to have a persistent or long term impact if its short run impact is carried over forward to set a new trend in performance. In the context of the I(d) processes examined in this work, if a series is stationary I(0) shocks disappear relatively fast and there is no need for strong policy actions to recover the original levels. If the series is fractionally integrated (or I(d)) with d in the range (0, 1) shocks generally tend to take longer to disappear than in the previous case, with an hyperbolic rate of decay of the autocorrelations as opposed to the exponential decay associated to the stationary I(0) AR(MA) processes. Finally, if the series possesses a unit root (i.e., it is I(1)), shocks will be permanent and will persist forever. In this latter case, strong policy measures should be implemented if we want to recover the original mean of the process.

In this paper we will examine the orders of integration in eight South African series related to the tourism sector. In doing so, we will be able to determine if shocks associated to the series are mean reverting or not, which is important in order to assist in future policy formulation towards improving or revitalising the tourism sector.

The article is structured as follows: Section 2 presents a brief literature review on tourism modelling and forecasting. Section 3 describes the methodology employed. Section 4 presents the dataset. Section 5 displays the empirical results, while Section 6 contains some concluding comments and extensions.

2. Literature review

There is a wide variety of articles on modelling and forecasting tourist data. These articles can be grouped into two large categories: those using time series techniques and those using panel data studies. Within the time series framework we find articles using log-linear and cointegration analysis (Kulendran and Witt, 2001; Lim and McAleer, 2002; Dritsakis, 2004; Lim, 2004; Algieri, 2006); unit root testing procedures (Narayan, 2005; Bhattacharya and Narayan, 2005); persistence in volatility models (Hoti, León and McAleer, 2006; Hoti, McAleer and Shareef, 2007 and Kim and Wong, 2006), etc.

Using multivariate techniques, Syriopoulos (1995), Kulendran (1996), Kulendran and King (1997), Seddighi and Shearing (1997), Kim and Song (1998), Vogt and Wittayakorn (1998) and others documented high persistence in tourism and arrivals time series. Most of these authors argue that the tourism series are nonstationary I(1) processes, implying thus the existence of a unit root and permanent effects of shocks. On the other hand, García-Ferrer and Queralt (1997), Chu (1998), Kim (1999), Lim and McAleer (2001, 2002), Goh and Law (2002), Gustavsson and Nordström (2001) and Brännäs et al. (2002) employed pure time series analytical models, some of these authors arguing that the univariate time series approach may be preferable to the multivariate models from a forecasting viewpoint.[1]

In this paper we use methodologies based on fractional integration. This approach enables the identification of the level of persistence of a series in a continuous way and therefore overcomes the restrictive view of traditional econometric methods which identifies a series as either persistent (i.e. I(1)) or non-persistent (i.e., I(0)), but is unable to evaluate the middle term of the persistence level. Fractionally integrated techniques have been recently applied to tourism time series data in a number of papers. Thus, for example, Chu (2008) uses an AutoRegressive Fractionally Integrated Moving Average (ARFIMA) model to forecast monthly international tourist arrivals in Singapore. Seasonal long memory models have been applied to tourist arrivals in Gil-Alana et al. (2004) and Gil-Alana (2005) and other recent papers with fractional integration in tourism data are Cunado et al. (2008a,b). Though this paper does not include any new methodological contribution, it is important to examine the degree of persistence in tourism-related series in South Africa so that we could determine the need of the implementation of policies to recover the level of the series in the event of negative shocks.

In the context of South African tourism data, Burger et al. (2001) present a survey of time series models to predict the US demand for travel to Durban. They used a variety of specifications including naïve, moving average, decomposition, single exponential smoothing, ARIMA and neural networks, and conclude that the latter performs the best. In a different context, Briedenhann and Wickens (2003) investigated

[1] Alternatively, artificial intelligence, as a group of emerging tourism forecasting techniques (including genetic algorithms, fuzzy logic, artificial neural networks and support vector machines, see e.g., Wang, 2004; Kon and Turner, 2005) has also emerged in recent years.

the development of rural tourism routes in South Africa, while Naude and Saayman (2005) examined the determinants of tourist arrivals in a panel of 43 African countries including South Africa. It is concluded in the paper that attention should be given to improving the overall stability of the continent, and the availability and quantity of tourism infrastructure. We should finally note that we were unable to find papers dealing with the persistence of South African tourism data, and, in this context, the use of I(d) models in the present paper seems overdue.

3. Materials and methods

Two well known features commonly present in tourism series are the dependence across time of the observations and the seasonality of the data. With respect to the time dependence various approaches have been adopted. Until the 80s, the most common one was to assume a deterministic function of time, generally under the assumption that the detrended series was stationary I(0). Later, and especially after the seminal work by Nelson and Plosser (1982), many series were found to be I(1), under the presumption that the first differences of the data were stationary I(0). These two approaches, usually named as "*trend stationarity*" and "*stochastic difference*" prevailed in economics until quite recently. During the last twenty years or so, a new approach has emerged that models many series in terms of fractionally integrated processes. That means that the number of differences required to render a series stationary I(0) may not necessarily be an integer value (1 in most of the cases) but a real one that can be between 0 or 1, or even above 1. This approach is clearly more general than the previous ones in the sense for example that it considers the I(1) or unit root model (the "*stochastic difference*" approach) as a particular case, and thus allows a much richer degree of flexibility in the dynamic specification of the series.

On the other hand, seasonality is another issue that should be taken into account when modelling tourist data, and depending on the nature of the seasonal component of the series different approaches should be adopted.

Throughout this paper we will consider the following process:

$$y_t \;=\; \alpha \;+\; \beta t \;+\; x_t, \qquad t \;=\; 1, 2, \ldots, \tag{1}$$

$$(1 - L)^d\, x_t \;=\; u_t\,, \qquad t \;=\; 1, 2, \ldots, \tag{2}$$

where y_t corresponds to the original time series; α and β are the coefficients corresponding to the intercept and the time trend respectively; d may be a real value, and u_t is supposed to be I(0), defined, for the purpose of the present work, as a covariance stationary process with a spectral density function, which is positive and finite at the zero frequency. Thus, u_t in (2) may be a white noise process but it may also allow for weak autocorrelation of the ARMA-form. Note that this is a very general specification in the sense that it includes the "*trend stationary*" representation in case of d = 0, and the unit root advocated by many authors in case of d = 1. Seasonal AR processes will be incorporated throughout the error term.

If $d > 0$ in (2), x_t (and thus y_t) it is said to be long memory, so-named because of the strong association between observations widely separated in time. Note that the polynomial $(1–L)^d$ in (2) can be expressed in terms of its Binomial expansion, such that, for all real d,

$$(1-L)^d = \sum_{j=0}^{\infty} \psi_j L^j = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^j = 1 - d\,L + \frac{d\,(d-1)}{2} L^2 - \ldots ,$$

and thus

$$(1-L)^d x_t = x_t - d\,x_{t-1} + \frac{d\,(d-1)}{2} x_{t-2} - \ldots .$$

In this context, the parameter d plays a crucial role since it will be an indicator of the degree of dependence of the time series. Thus, the higher the value of d is, the higher the level of association will be between the observations. These processes are characterized because the spectral density is unbounded at the zero frequency. This was first noticed in the 60s by Granger (1966) and Adelman (1965), who pointed out that most aggregate economic time series have a typical shape where the spectral density increases dramatically as the frequency approaches zero. However, differencing the data frequently leads to overdifferencing at the zero frequency. Fifteen years later, Robinson (1978) and Granger (1980) showed that aggregation could be a source of fractional integration[2]. Since then, fractional processes have been widely employed to describe the dynamics of many time series (see, e.g. Diebold and Rudebusch, 1989; 1991; Sowell, 1992; Baillie, 1996; Gil-Alana and Robinson, 1997; etc.).

[2] These authors showed that fractionally integrated data could arise as a result of aggregation when data are aggregated across heterogeneous autoregressive (AR) processes; data involving heterogeneous dynamic relationships at the individual level are then aggregated to form the time series.

The methodology employed in this work is based on the Whittle function in the frequency domain (Dahlhaus, 1989) along with a testing procedure developed by Robinson (1994) that permits us to test different hypotheses under the null. In fact, the latter is a Lagrange Multiplier (LM) procedure that is supposed to be the most efficient method in the context of fractional integration. It tests the null hypothesis H_o: $d = d_o$ for any real value d_o, in (1) and (2) with a standard (normal) null limit distribution. Moreover, this standard behaviour holds independently of the inclusion or not of deterministic terms and the way of modelling the I(0) disturbances. The functional form of Robinson's (1994) tests can be found in any of the numerous empirical applications of his tests (e.g. Gil-Alana and Robinson, 1997; Gil-Alana, 2000; etc.)

4. Data

The data examined in this paper correspond to the following variables all referring to the total tourism industry: "Stay units available" (SUA); "Stay units nights sold" (SUNS); "Occupancy rate" (OR); "Income per stay unit nights sold " (IPSUNS); "Income from accommodation" (IFA); "Income from restaurant and bar sales" (IFRBS); "Other income" (OI); and "Total income" (TI), monthly, for the time period September 2004 – August 2009. These data were obtained from the online service "Statistics South Africa" (http://www.statssa.gov.za). (See Table 1 for each time series numbering).

[Insert Table 1 and Figure 1 about here]

Figure 1 displays each of the time series plots. We observe a slight increase across time, particularly in the income-related series (IPSUNS, IFA, IFSRBS, OI and

TI). On the other hand, the three time series related to accommodation issues (SUA, SUNS and OR) may have a stationary appearance.

[Insert Figure 2 about here]

Figure 2 displays the first 25 sample autocorrelation values for each series. We observe significant values in all cases with a slow decay that may be consistent with a simple autoregressive model but also with a fractionally integrated structure. Seasonal lags appear significant in the cases of SUNS, OR and IFRBS.

5. Empirical results

Table 2 displays the estimates of d (and the 95% confidence bands) in the model given by (1) and (2) under the assumption that the error term u_t in (2) is white noise. We report the results for the three cases of no regressors in the undifferenced regression model (1) (i.e, $\alpha = \beta = 0$ a priori); an intercept (α unknown and $\beta = 0$ a priori), and an intercept with a linear time trend (α and β unknown).

The first thing we observe in this table is that if we do not include regressors the unit root hypothesis (i.e. d = 1) cannot be rejected in any of the series. Including an intercept or an intercept with a linear time trend, the unit root cannot be rejected in any of the income series with the exception of IFRBS where d is found to be strictly smaller than 1. The same mean reverting behaviour is observed for the three accommodation series (SUA, SUNS and OR) where d is strictly below 1.

[Insert Tables 2 and 3 about here]

Table 3 displays the parameter estimates of the selected model for each time series according to the specification of the deterministic terms. It is observed that only for IRFBS the time trend is required, while for the remaining series an intercept seems to be sufficient to describe the deterministic part. We see that the estimates of d widely vary from one series to another. The lowest degrees of integration are obtained in the cases of SUNS (d = 0.475) and OR (d = 0.493), followed by IFRBS (d = 0.512) and SUA (d = 0.758). In all these series the estimates are strictly smaller than 1 implying long memory and mean reverting behaviour. Estimates of d below 1 are also observed in the cases of OI (d = 0.782), IFA (d = 0.800) and TI (d = 0.812), though in these three series the unit root null cannot be rejected. The highest degree of integration is obtained at IPSUNS (d = 1.013) and the unit root cannot be rejected at the 5% level. Thus, according to this simple specification, all series are I(d), being d strictly smaller than 1 (and thus showing mean reversion) in case of the accommodation series, and with values of d around 1 (the unit root case) for the income series.

The results presented so far assume that the time dependence between the observations is fully captured through the fractional differencing parameter d. In what follows, we include another source of dependence throughout the error term. The standard way of modelling such I(0) dependence is by using stationary ARMA models. In this paper, however, we employ an alternative approach that is based on the exponential spectral model of Bloomfield (1973). This is a non-parametric method that produces autocorrelations decaying exponentially as in the ARMA case. In this approach, the spectral density function is given by:

$$f(\lambda;\tau) = \frac{\sigma^2}{2\pi} \exp\left(2 \sum_{r=1}^{m} \tau_r \cos(\lambda r)\right), \qquad (3)$$

where m is the number of parameters required to describe the short run dynamics of the series. Bloomfield (1973) showed that the logarithm of an estimated spectral density function is often found to be a fairly well-behaved function and can thus be approximated by a truncated Fourier series. He showed that the spectral density of an ARMA process can be well approximated by (3). Moreover, this model is stationary across all values of τ, and the model accommodates extremely well in the context of fractionally integrated models.[3] The results using this model (with m = 1) for the three cases of no regressors, an intercept, and an intercept with a linear trend are displayed in Table 4.

We notice that most of the estimates are in the range (0, 1) though we also observe some negative values in the case of the inclusion of a linear time trend. Negative values of d imply anti-persistence (Mandelbrot, 1977)[4] and suggest a kind of competition with the error term in describing the time dependence.

[Insert Tables 4 and 5 about here]

Table 5 displays the estimates for each series. The time trend is now required in all cases except in one series (SUA). Once more the estimates of d widely vary across

[3] See Gil-Alana (2004) for a paper relating fractional integration with the exponential spectral model of Bloomfield (1973) in the context of Robinson's (1994) tests.

[4] A process is said to be anti-persistent if it reverses itself more often than a random series would.

the series from -0.317 (IFRBS) to 0.726 (SUA), and the 95% confidence intervals are now very wide, including the I(0) hypothesis in all except one series (SUA). For the latter, the I(1) null cannot be rejected.

Finally, seasonality is also taken into account, and a seasonal AR(1) process is assumed for the error term.[5] The results are displayed across Tables 6 and 7. We first observe that most of the estimates are in the range (0.5, 1) suggesting that according to this specification the series are nonstationary though mean reverting.

[Insert Tables 6 and 7 about here]

In Table 7 we see that only for IPSUNS and IFRBS the time trends are required. In all the other cases, the intercept is sufficient to describe the deterministic part. If we focus now on the estimates of d we see that the lowest value is obtained for IFRBS (with d = 0.555) followed by OR (d = 0.666) and SUNS (d = 0.668). Then, come SUA (d = 0.775), OI (d = 0.784), IPSUNS (d = 0.792) and IFA (d = 0.796), and the highest degree of dependence is obtained for TI, with d = 0.853. Also, note that there are three income series where the unit root cannot be rejected; they are OI, IPSUSN and TI, while mean reversion is detected in the remaining five series. If we look at the seasonal AR coefficients we notice that the highest values correspond to OR, SUNS and IFRBS, which are precisely the series with the lowest values of d, indicating that the two parameters might be competing in describing the time dependence.

[5] Higher seasonal AR processes lead essentially the same results.

Looking at the results displayed in this work it seems reasonable to argue that the model that combines fractional integration with seasonal autoregressions, is the one that should be taken into consideration for these series given the significancy of the two parameters that describe the time dependence, i.e., the fractional differencing parameter d, and the seasonal AR coefficient.[6] Figure 3 displays the first 120 impulse responses for the selected model for each series according to the parameter estimates reported in Table 7.

[Insert Figure 3 about here]

We observe that the responses decay to zero in all cases though at a very slow rate. We also notice that seasonality is important in the majority of the series, the two exceptions being the "Stay Units Available" series (SUA) and "Other Income" (OI). The fastest processes of convergence seem to take place for SUA, IFRBS and OI. Table 8 displays the numerical values of the first 12 impulse responses, while Table 9 focuses on the first 10 multiple of 12 responses.

[Insert Tables 8 and 9 about here]

Starting with the short run responses (in Table 8) we see that for IFRBS, SUNS and OR, the responses of a shock are smaller than 0.5 after three periods, while for the remaining series, half of the effect remains even after one year. If we look now at the seasonal long run effects (in Table 9) we observe that the most persistent effects take

[6] Moreover, we conducted several diagnostic tests on the residuals of the differenced models and the results indicate that the model with seasonal AR(1) error produces the best results.

place in the series OR and IF, while the less persistent ones are IFRBS and especially OI and SUA.

6. Final comments and conclusions

In this paper we have examined several time series related with the tourism sector in South Africa. The aim was to determine the nature of the shocks in the series and for this purpose we employed fractionally integrated techniques. The series examined were: the number of stay units available, the number of stay unit nights sold, the occupancy rate, and five more series related with income in the tourism sector: income per stay units sold, income from accommodation, income from restaurants and bar sales, other income and total income.

The results indicate that the eight series examined are mean reverting, implying that shocks have transitory though long-lasting effects. Moreover, the level of persistence substantially changes from one series to another and the highest degrees of persistence are obtained in the short run in the cases of Total Income, Income From Accommodation, and Income Per Stay Units Night Sold (see Table 8), while the seasonal effects seem particularly important in the cases of Occupancy Rate and Income From Accommodation (see Table 9). Therefore, in the event of a negative exogenous shock stronger measures must be adopted in relation with these series to recover the original levels. Other issues such as the presence of asymmetric effects of the shock, the presence of structural breaks or even non-linearities in tourism data will be examined in future papers.

References

Adelman, I. (1965). Long cycles: Fact or artifacts. American Economic Review, 55, 444-463.

Algieri, B. (2006). An econometric estimation of the demand for tourism. The case of Russia. Tourism Economics, 12, 1, 5-20.

Baillie, R.T. (1996). Long memory processes and fractional integration in econometrics. Journal of Econometrics, 73, 5-59.

Bhattacharya, M., & Narayan, P.K. (2005). ´Testing for the random walk hypothesis in the case of visitor arrivals: evidence for Indian tourism. Applied Economics, 37, 1485-1490.

Bloomfield, P. (1973). An exponential model in the spectrum of a scalar time series. Biometrika, 60, 217-226.

Brännäs, K., Hellström, J. & Nordström, J. (2002). A new approach to modelling and forecasting monthly guest nights in hotels. International Journal of Forecasting, 18, 19-30.

Briedenhann, J., & Wickens, E. (2003). Tourism routes as a tool for the economic development of rural areas _ Vibrant hope or impossible dream?. Tourism Management, 25, 71-79.

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

Chu, F.L. (1998). Forecasting tourist arrivals. Journal of Travel Research, 36, 79-84.

Chu, F.L. (2008). A Fractional Integrated Autoregressive Moving Average approach to Forecasting Tourism Demand. Tourism Management, 29, 79-88.

Cunado, J., Gil-Alana, L.A., & Perez de Gracia, F. (2008a). Fractional integration and structural breaks. Evidence from the international monthly arrivals in the US. Tourism Economics, 14, 13-23.

Cunado, J., Gil-Alana, L.A., & Perez de Gracia, F. (2008b). Persistence in the international monthly arrivals in the Canary Islands. Tourism Economics, 14, 123-129.

Diebold, F.X., & Rudebusch, G.D. (1989). Long memory and persistence in the aggregate output. Journal of Monetary Economics, 24, 189-209.

Diebold, F.X., & Rudebusch, G.D. (1991). Is consumption too smooth? Long memory and the Deaton paradox. The Review of Economics and Statistics, 73, 1-9.

Dritsakis, N. (2004). Cointegration analysis of German and British tourism demand for Greece. Tourism Management, 25, 111-119.

Garcia-Ferrer, A., & Queralt, R.A. (1997). A note on forecasting international tourism demand in Spain. International Journal of Forecasting, 13, 539-549.

Gil-Alana, L.A. (2000). Fractional integration in the purchasing power parity. Economics Letters, 69, 285-288.

Gil-Alana, L.A. (2004). The use of the Bloomfield model as an approximation to ARMA processes in the context of fractional integration. Mathematical and Computer Modelling, 39, 429-436.

Gil-Alana, L.A. (2005). Modelling International Monthly Arrivals using Seasonal Univariate Long-Memory Processes. Tourism Management, 26, 867-878.

Gil-Alana, L.A., Cuñado, J., &. Perez de Gracia, F. (2004). Seasonal fractional integration in the Spanish monthly quarterly time series. Journal of Travel Research, 42, 408-414.

Gil-Alana, L.A., & Robinson, P.M. (1997). Testing of unit roots and other nonstationary hypotheses in macroeconomic time series. Journal of Econometrics, 80, 241-268.

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

Granger, C.W.J. (1966). The typical spectral shape of an economic variable. Econometrica, 37, 150-161.

Granger, C.W.J. (1980). Long memory relationships and the aggregation of dynamic models. Journal of Econometrics, 14, 227-238.

Gustavsson, P., & Nordström, J. (2001). The impact of seasonal unit roots and vector ARMA modelling on forecasting monthly tourism flows. Tourism Economics, 7, 117-133.

Hoti, S., León, C., & McAleer, M. (2006). International tourism demand and volatility models for the Canary Islands, Mimeo.

Hoti, S., McAleer, M,, & Shareef, R. (2007). Modelling international tourism and country risk spillovers for Cyprus and Malta. Tourism Management, 18, 1472-1484.

Kim, J.H. (1999). Forecasting monthly tourist departures from Australia. Tourism Economics, 5, 277-291.

Kim, S., & Song, H. (1998). Analysis of inbound tourism demand in South Korea. A cointegration and error correction approach. Tourism Analysis, 3, 25-41.

Kim, S., & Wong, K. (2006). Effects of news shock on inbound tourist demand volatility in Korea. Journal of Travel Research, 44, 457-466.

Kon, S.C. and W.L. Turner (2005), Neural network forecasting of tourism demand, Tourism Economics 11, 301-328.

Kulendran, N. (1996). Modelling quarterly tourism flows to Australia. Tourism Economics, 2, 203–222.

Kulendran, N., & Witt, S.F. (2001). Cointegration versus least square regressions. Annals of Tourist Research, 28, 2, 291-311.

Kulendran, N., & King, M.L. (1997). Forecasting international quarterly tourist flows using error-correction and time-series models. International Journal of Forecasting, 13, 319-327.

Lim, C. (2004). The major determinants of Korean outbound travel to Australia. Mathematics and Computers in Simulation, 64, 477-485.

Lim, C., & McAleer, M. (2001). Cointegration analysis of quarterly tourism demand by Hong Kong and Singapore for Australia. Applied Economics, 33, 1599-1619.

Lim, C., & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. Tourism Management, 23, 389-396.

Mandelbrot, B.B. (1977). Fractals, form, chance and dimension, Freeman, San Francisco.

Narayan, P.K. (2005). The structure of tourist expenditure in Fiji: evidence for unit root structural break tests. Applied Economics, 37, 1157-1161.

Naude, W., & Saayman, A. (2005). The determinants of tourist arrivals in Africa. A panel data regression analysis. Tourism Economics, 11, 365-392.

Nelson, C.R., & Plosser, C.I. (1982). Trends and Random Walks in Macroeconomic Time Series. Journal of Monetary Economics, 10, 139-162.

Robinson, P.M. (1978). Statistical inference for a random coefficient autoregressive model. Scandinavian Journal of Statistics, 5, 163-168.

Robinson, P.M. (1994). Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association, 89, 1420-1437.

Saayman, M. (2006). Tourism Marketing: Back to Basics, Leisure Consultants and Publications, Potchefstroom.

Saayman, A., & Saayman, M. (2008). Determinants of inbound tourism to South Africa. Tourism Economics, 14, 1, 81-96.

Seddighi, H., & Shearing D.F. (1997). The demand for tourism in North East England with special reference to Northumbria an empirical analysis. Tourism Management, 18, 499–511.

Sowell, F. (1992). Modelling long run behaviour with the fractional ARIMA model. Journal of Monetary Economics, 29, 277-302.

Syriopoulos, T. (1995). A dynamic model of demand for Mediterranean tourism. International Review of Applied Economics, 9, 318–336.

Vogt, M.G., & Wittayakorn, C. (1998). Determinants of the demand for Thailand's exports of tourism. Applied Economics, 30, 711–715.

Wang, C.H. (2004). Predicting tourism demand using fuzzy time series and hybrid grey theory, Tourism Management 25, 367-374.

Table 1: Time Series analyzed in the article

Acronym	Description	Number
SUA	Stay Units Available	A64100
SUNS	Stay Units Nights Sold	B64100
OR	Occupancy Rate	C64100
IPSUNS	Income Per Stay Units Nights Sold	D64100
IFA	Income From Accommodation	E64100
IFRBS	Income From Restaurant and Bar	F64100
OI	Other Income	G64100
TI	Total Income	H64100

Data source: Statistics South Africa (http://www.statssa.gov.za).

Table 2: Estimates of d under the assumption of white noise errors

Time Series	No regressors	An intercept	A linear time trend
SUA	0.935 (0.778, 1.165)	0.758 (0.636, 0.953)	0.763 (0.644, 0.954)
SUNS	0.869 (0.681, 1.132)	0.475 (0.343, 0.711)	0.440 (0.255, 0.710)
OR	0.865 (0.679, 1.124)	0.493 (0.334, 0.771)	0.464 (0.243, 0.770)
IPSUNS	0.909 (0.706, 1.180)	1.013 (0.800, 1.330)	1.012 (0.756, 1.332)
IFA	0.847 (0.560, 1.168)	0.800 (0.635, 1.103)	0.789 (0.561, 1.103)
IFRBS	0.741 (0.449, 1.039)	0.512 (0.405, 0.714)	0.409 (0.189, 0.704)
OI	0.765 (0.481, 1.172)	0.782 (0.556, 1.163)	0.776 (0.501, 1.163)
TI	0.852 (0.568, 1.188)	0.812 (0.640, 1.123)	0.805 (0.582, 1.123)

Table 3: Parameter estimates in the selected models across Table 2

Time Series	Fract. Diff. Par.	Intercept	Linear time trend
SUA	0.758 (0.636, 0.953)	107.388 (135.856)	----------
SUNS	0.475 (0.343, 0.711)	1520.029 (22.938)	----------
OR	0.493 (0.334, 0.771)	46.089 (22.126)	----------
IPSUNS	1.013 (0.800, 1.330)	409.433 (19.156)	----------
IFA	0.800 (0.635, 1.103)	608.060 (8.630)	----------
IFRBS	0.512 (0.405, 0.714)	210.219 (10.911)	1.696 (3.025)
OI	0.782 (0.556, 1.163)	152.970 (2.238)	----------
TI	0.812 (0.640, 1.123)	970.876 (7.401)	----------

Table 4: Estimates of d under the assumption of Bloomfield-type errors

Time Series	No regressors	An intercept	A linear time trend
SUA	0.826 (0.507, 1.216)	0.726 (0.512, 1.015)	0.738 (0.537, 1.015)
SUNS	0.588 (0.019, 1.083)	0.333 (0.142, 0.594)	0.051 (-0.785, 0.546)
OR	0.604 (0.008, 1.087)	0.260 (0.057, 0.555)	-0.173 (-0.769, 0.503)
IPSUNS	0.539 (0.190, 1.150)	0.678 (0.524, 1.164)	0.104 (-0.380, 1.167)
IFA	0.241 (0.169, 0.917)	0.578 (0.424, 0.878)	0.086 (-0.633, 0.875)
IFRBS	0.143 (0.085, 0.966)	0.428 (0.258, 0.643)	-0.317 (-0.740, 0.490)
OI	0.274 (0.139, 0.630)	0.415 (0.232, 0.735)	0.126 (-0.319, 0.706)
TI	0.266 (0.181, 0.884)	0.576 (0.417, 0.884)	0.259 (-0.381, 0.872)

Table 5: Parameter estimates in the selected models across Table 4

Time Series	Fract. Diff. Par.	Intercept	Linear time trend
SUA	0.726 (0.512, 1.015)	107.450 (139.27)	----------
SUNS	0.051 (-0.785, 0.546)	1465.291 (46.648)	3.852 (4.367)
OR	-0.173 (-0.769, 0.503)	44.062 (85.676)	0.115 (7.360)
IPSUNS	0.104 (-0.380, 1.167)	387.135 (52.768)	4.185 (20.592)
IFA	0.086 (-0.633, 0.875)	561.863 (23.757)	8.496 (12.913)
IFRBS	-0.317 (-0.740, 0.490)	197.973 (65.136)	2.144 (21.933)
OI	0.126 (-0.319, 0.706)	122.405 (4.935)	4.178 (6.107)
TI	0.259 (-0.381, 0.872)	913.898 (13.950)	13.654 (7.612)

Table 6: Estimates of d under the assumption of seasonal AR errors

Time Series	No regressors	An intercept	A linear time trend
SUA	0.939 (0.793, 1.158)	0.775 (0.654, 0.962)	0.779 (0.660, 0.962)
SUNS	0.924 (0.760, 1.157)	0.668 (0.559, 0.825)	0.664 (0.545, 0.826)
OR	0.926 (0.765, 1.152)	0.666 (0.554, 0.827)	0.660 (0.536, 0.826)
IPSUNS	0.879 (0.697, 1.121)	0.857 (0.737, 1.126)	0.792 (0.562, 1.128)
IFA	0.806 (0.609, 1.032)	0.796 (0.698, 0.953)	0.766 (0.633, 0.945)
IFRBS	0.781 (0.596, 1.000)	0.624 (0.536, 0.763)	0.555 (0.408, 0.744)
OI	0.769 (0.481, 1.178)	0.784 (0.557, 1.170)	0.778 (0.501, 1.170)
TI	0.849 (0.631, 1.116)	0.853 (0.687, 1.131)	0.847 (0.651, 1.130)

Table 7: Parameter estimates in the selected models across Table 6

Time Series	Fract. Diff. Par.	Intercept	Linear time trend	AR
SUA	0.775 (0.654, 0.962)	107.360 (133.639)	----------	-0.153
SUNS	0.668 (0.559, 0.825)	1471.330 (15.522)	----------	0.792
OR	0.666 (0.554, 0.827)	45.225 (15.943)	----------	0.819
IPSUNS	0.792 (0.562, 1.128)	408.940 (19.181)	3.336 (2.487)	0.563
IFA	0.796 (0.698, 0.953)	608.862 (8.541)	----------	0.735
IFRBS	0.555 (0.408, 0.744)	211.674 (8.839)	1.613 (1.933)	0.762
OI	0.784 (0.557, 1.170)	152.793 (2.225)	----------	0.020
TI	0.853 (0.687, 1.131)	959.258 (7.137)	----------	0.485

Table 8: First 12 impulse responses

	SUA	SUNS	OR	IPSUNS	IFA	IFRBS	OI	TI
0	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
1	**0.775**	**0.668**	**0.666**	**0.792**	**0.796**	**0.555**	**0.784**	**0.853**
2	**0.687**	**0.557**	**0.554**	**0.709**	**0.714**	0.431	**0.699**	**0.790**
3	**0.636**	0.495	0.493	**0.660**	**0.666**	0.367	**0.648**	**0.751**
4	**0.600**	0.454	0451	**0.626**	**0.632**	0.326	**0.613**	**0.723**
5	**0.573**	0.424	0.421	**0.600**	**0.606**	0.297	**0.587**	**0.702**
6	**0.552**	0.400	0.398	**0.579**	**0.585**	0.275	**0.566**	**0.685**
7	**0.534**	0.381	0.379	**0.562**	**0.568**	0.257	**0.548**	**0.671**
8	**0.519**	0.365	0.363	**0.547**	**0.554**	0.243	**0.533**	**0.658**
9	**0.506**	0.352	0.349	**0.534**	**0.541**	0.231	**0.521**	**0.647**
10	0.494	0.340	0.338	**0.523**	**0.530**	0.222	**0.509**	**0.638**
11	0.484	0.330	0.327	**0.513**	**0.520**	0.212	0.499	**0.629**

In bold the responses which are above 0.500.

Table 9: First 10 seasonal impulse responses

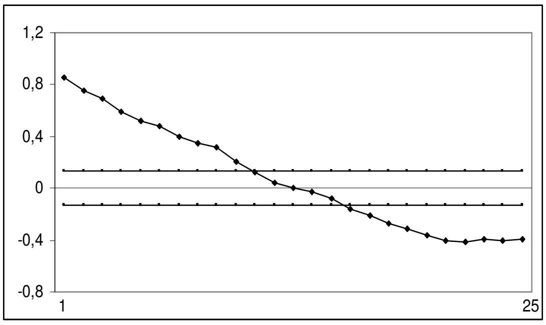
	SUA	SUNS	OR	IPSUNS	IFA	IFRBS	OI	TI
0	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
12	0.322	**1.113**	**1.137**	**1.067**	**1.247**	0.966	0.510	**1.107**
24	0.359	**1.138**	**1.186**	**1.039**	**1.362**	0.887	0.434	**1.100**
36	0.318	**1.125**	**1.193**	0.988	**1.412**	0.802	0.397	**1.065**
48	0.301	**1.095**	**1.179**	0.937	**1.425**	0.722	0.372	**1.026**
60	0.286	**1.057**	**1.153**	0.890	**1.418**	0.651	0.356	0.990
72	0.275	**1.016**	**1.121**	0.851	**1.399**	0.589	0.342	0.960
84	0.266	0.974	**1.086**	0.818	**1.375**	0.535	0.331	0.935
96	0.259	0.934	**1.050**	0.790	**1.347**	0.490	0.321	0.914
108	0.252	0.895	**1.014**	0.766	**1.319**	0.451	0.313	0.895
120	0.246	0.860	0.979	0.746	**1.291**	0.417	0.306	0.880

In bold the responses which are above 1.000.

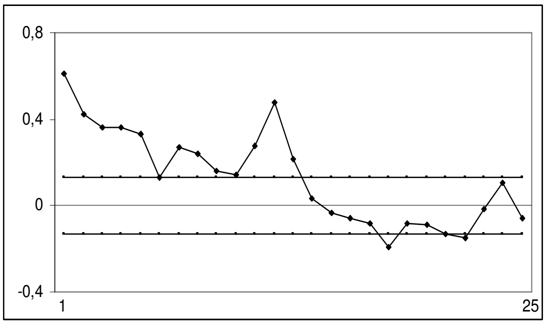
Figure 1: Time series plots

Figure 2: Sample autocorrelation functions

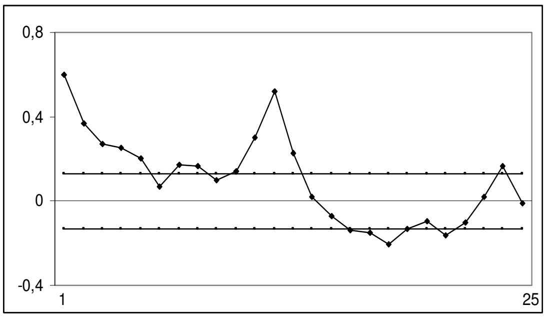
SUA



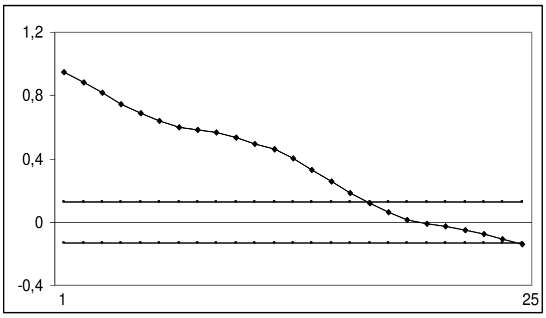
SUNS



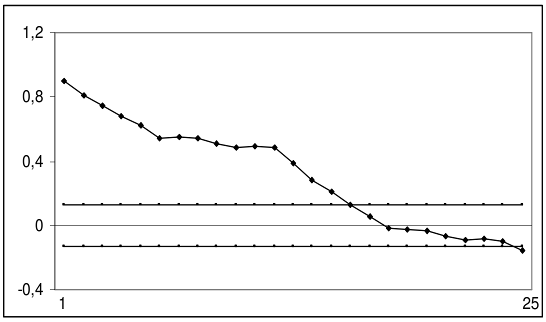
OR



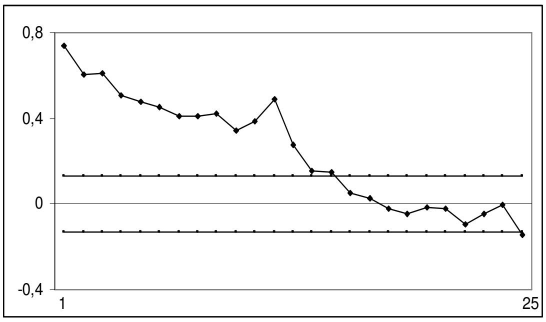
IPSUNS



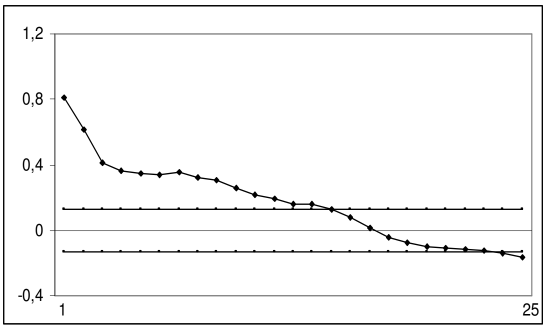
IFA



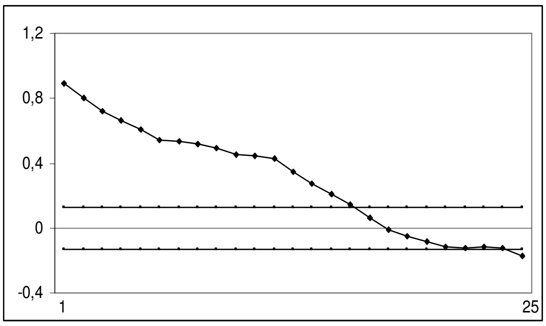
IFRBS



OI



TI



ERROR: undefined
OFFENDING COMMAND: r

STACK:

4
1811
4625
666