Role of Rules of Thumb in Forecasting Foreign Tourist Arrival: A Case Study of India

Bhattacharya, Kaushik

Indian Institute of Management Lucknow

31 January 2011

Role of Rules of Thumb in Forecasting Foreign Tourist Arrival: A Case Study of India

By

Kaushik Bhattacharya*

Indian Institute of Management Lucknow

January 2011

Abstract

The paper examines forecast performances of some popular rules of thumb vis-à-vis more sophisticated time series models in the specific context of foreign tourist arrival in India. Among all forecasting approaches attempted in the study, exponential smoothing (ES) and ARIMA provided the best short-term forecasts, closely followed by autoregressive distributed lag (ADL) models. These results are largely in agreement with cross-country findings on tourism forecast. Foreign tourist arrival data in India, however, displayed a regularity that did not change substantially even in the face of major global or local events. Given the regularity, our study suggests that rules of thumb can play an important practical part in short-term forecasts of tourist arrival in India. Our study, however, reveals that forecasts from such thumb rules could be improved substantially through simple residual corrections and incorporation of other information available in the public domain.

Keywords: Tourism, Tourist Arrival, Forecasting, Rules of Thumb, Exponential Smoothing, ARIMA, ADL

Journal of Economic Literature Classification: C22, L83

---

* The views expressed in this paper are personal and not that of the institution the author belongs to. An earlier version of the paper was presented in a conference entitled "Tourism in Global Village" in Indian Institute of Management Lucknow (April 8–10, 2009). The author is grateful to conference participants whose comments and suggestions have substantially improved the earlier exposition. The responsibility for any error, however, lies with the author. For the data and the computer programmes used in this study, the author could be contacted at his email address: kbhattacharya@iiml.ac.in.

Role of Rules of Thumb in Forecasting Foreign Tourist Arrival: A Case Study of India

1 Introduction

Academic efforts over the years have led to substantial improvement in forecasting tools. The standard toolkit of today's forecaster not only includes smoothing techniques, but also ADL, ARIMA type models and a host of other complex nonlinear single and multiple-equation tools. Concomitant development of software has made many of these tools popular and routine in business applications.

While the advanced forecasting tools have increased options, from a practical perspective correct uses of them are not without costs. This is because more complex the tool, more tends to be the cost in information collection. Further, correct use of complex tools requires an expertise that may not be readily available within an organization. Because of these, any practical effort in forecasting cannot ignore the so called "rules of thumb". It is easy to jump to the conclusion that the widespread use of rules of thumb is evidence in favor of "sloppy workmanship" on the part of business management. However, as demonstrated by Baumol and Quandt (1964) long back in the context of pricing, rules of thumb could often be "more efficient pieces of equipment of optimal decision making".

In the context of forecasting, rules of thumb are often used in a limited way in terms of the so called "naïve forecasts". These "naïve forecasts" provide benchmarks against which the efficacy of an econometric model is assessed. In the context of exchange rate forecasting, a seminal study of Meese and Rogoff (1983) demonstrated that even such naïve forecasts could outperform forecasts from more sophisticated time series models. However, rules of thumb may not necessarily be naïve forecasts alone, they could easily be generalized to encompass moving average and standard smoothing techniques.

The purpose of this paper is to assess the relative strengths and weaknesses of rules of thumb vis-à-vis more sophisticated forecasting techniques in the specific context of tourism. At its core, tourism is primarily a human activity that involves cultural, economic and social aspects. In terms of gross output and value added, it is an important sector in any part of the world. Forecasting is particularly crucial in the tourism industry due the perishable nature of the tourism product. As in other business, long term projections based on structural features help the tourism industry in improving the existing infrastructure. As considerable amount of planning and risk is involved, such an exercise should ideally be done as rigorously as possible. Internationally, patterns in annual tourist arrival data have been examined by many researchers for different countries.[1] The general finding from cross-country experience appears to be that no single method or model provides uniformly superior forecasts across countries or over time. Further, the benchmark survey of Witt and Witt (1995b) highlights that in some cases forecasts based on naïve models outperform all other models, questioning the very role of rigorous modeling in practical tourism related projections.

Like long term forecasts, short term forecasts also play an important part in tourism. These forecasts help all tourism related business to prepare for the logistical requirements. Empirical studies based on monthly time series data highlight some success in short-term forecasts in recent years, especially in explaining and projecting monthly tourist arrival (Goh and Law, 2002). However, a limitation in these studies is that due to the frequency and the short-time horizon involved and due to the perishable nature of the tourism product, relevant information collection is often not practicable. Such studies are, therefore, often based on movements of a single variable (e.g., aggregate or country-wise data on tourist arrival), data on which can be collected without cost from an official source. Given the cost in data collection and the limited time available, good rules of thumb could be of immense help in such situations.

This paper compares forecast performances of some popularly used rules of thumb with those from a few more sophisticated time series models. The rules of thumb that we specify are standard and non-technical ones that can be understood by any one non-initiated in forecasting. Forecasts based on these rules of thumbs are compared to those based on smoothing techniques, ADL and ARIMA type models. In particular, a major purpose of our study is to examine whether residual corrections coupled with incorporation of other available information in the public domain, would lead to improvements in the standard rules of thumbs.

The empirical application in this paper is restricted to monthly tourist arrival data for India. Among all the countries in the world, India would be an interesting case study not only because of the large size of its domestic economy, but also because of three other factors. First, a notable feature of international tourism growth has been the gradual shift in the preference for destinations from Europe and North America to South Asia. India, with its diversity of natural attractions and rich tapestry of cultural heritage is an important South-Asian tourist destination. Despite this, rigorous efforts in projecting tourist arrival have been limited in case of India. This is not surprising because at least till the early 1990s, the Indian economy was a relatively closed economy driven by the forces of command and control. Also, with the opening up of the economy and dismantling of the earlier license-permit regime, the business potential for tourism increased substantially. The National Policy on Tourism was framed and announced in India in 2002. Concurrently, the Government of India also started a multi-pronged

[1] For example, Witt and Witt (1995a, 1995b) is an early survey that summarizes the findings till early 1990s. Frechtling (2001) provides a comparatively recent survey.

approach for promotion of tourism. For example, road shows in key source markets of Europe and the 'Incredible India' campaign on prominent TV channels and in magazines across the world were a few important steps taken to advertise Indian tourism. It is likely that consistent policy efforts have led to a structural shift in tourist arrival in case of India, posing a challenge to modelling. Third, India is one of the major countries in the world affected by terrorism. Indian case study can reveal to what extent terrorism can affect tourism business at a macro-level in an emerging democracy in the short-run.

The plan of the paper is as follows: Section 2 describes the data and carries out a descriptive analysis. Section 3 specifies the thumb rules and then describes in brief the other more technical methods used. Section 4 compares the quality and accuracy of forecasts from these models. Section 5 examines the impact of specific events like terrorist attacks on tourist arrivals. Finally Section 6 concludes the paper.

2 Monthly Data on Tourist Arrival in India: A Descriptive Analysis

Data on tourist arrival is compiled by the Ministry of Tourism, Government of India. Such data are disseminated regularly in the Monthly Statistical Abstract published by the Central Statistical Organisation (CSO). Private organizations like Indiastat maintains a time series database on tourism based on these publications. The data used in this study were collected from Indiastat. In this study we use the monthly time series data on foreign tourist arrival in India from January 1992 to December 2007, e.g., 192 observations. By tourist arrival, all arrivals by land, sea or air have been aggregated for the entire India. Due to long time lag in data finalization, post 2007 provisional data have not been considered in this study.

Figure 1: Number of Tourist Arrivals in India

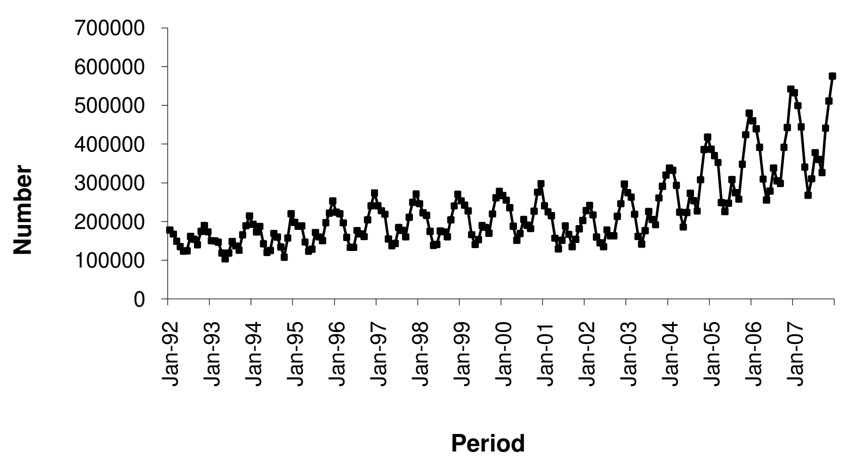


Figure 1 plots the graph of monthly tourist arrival in India. As the number of arrivals is a “level” data, it reveals the presence of an upward trend. Interestingly, Figure 1 suggests the presence of a structural break circa 2002. Prior to 2002, the trend growth appears to be slow. However, post-2002, the data show strong evidence in favor of increased growth in tourist arrival. We discuss the possible reasons of the structural shift in further detail in Section 5.

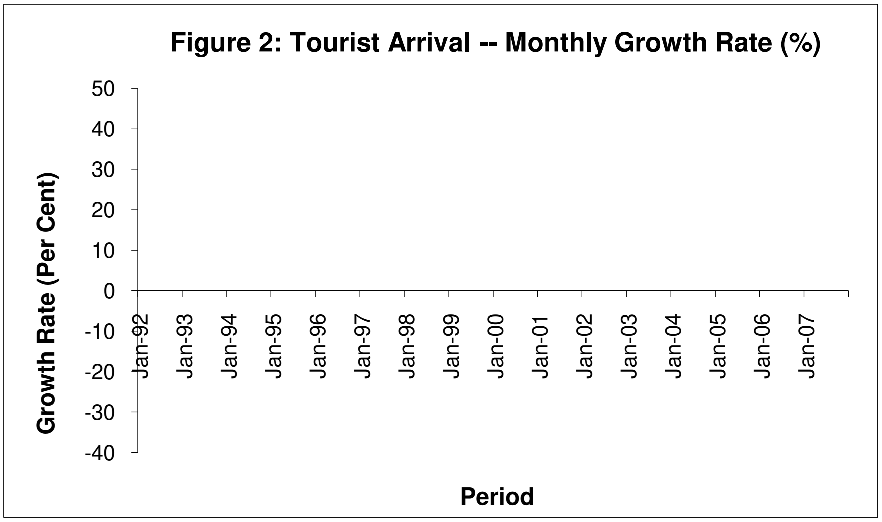


Figure 1 also presents strong evidence of seasonality in monthly tourist arrival. To examine the nature and the extent of seasonality, Figure 2 plots the monthly growth rates of tourist arrival. Figure 2 reveals the presence of two peaks within a year. It may be noted that the occurrence of two peaks in tourism data is fairly common across countries. The months in which the peaks fall could be different across countries because these peaks depend upon the vacation periods and festival seasons in the source countries and the country of destination (Goh and Law, 2002). In case of India, the first peak occurs during the month of July, the second one during October. *Prima facie*, the first peak in case of July in India is due to the fact that it is contemporaneous with the period of summer vacation in the Western economies. The reason behind the second peak appears to be local. It is well known that months of October and November fall in the festival season in large parts of India. Festivals like Durga Puja, Dussera and Diwali tend to fall during this period. Interestingly, during the festival season in December in the Western economies, monthly growth in tourist arrival in India appears to be limited.

To examine the stability in the seasonal pattern, we plot year-over-year charts of monthly tourist arrival for the entire period in Figure 3. To have clarity, the plots are presented in three separate graphs, viz., Figures 3a, 3b and 3c. The year-over year charts reveal a surprisingly stable pattern in seasonality. Occasionally, data for a year show a large dip or increase for one or two months, as in case of

1994 (October and November), 1996 (December) and 2001 (post-September 11). In general, such deviations tend to be aberrations rather than the norm. The practical implication of the regularity is that in the Indian case, simpler models are likely to be as effective and useful as more complex and sophisticated models. Another practical implication is that, *prima facie*, terrorist attacks did not seem to have caused irregular patterns in tourist arrival on sustained basis.

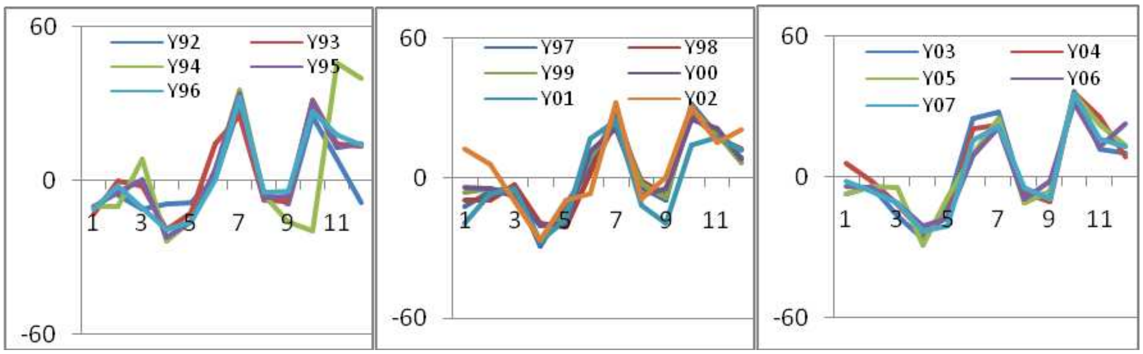


Figure 3A Figure 3B Figure 3C

Figure 3: Year-Over-Year Chart of the Monthly Growth Rates of Tourist Arrival

3 Choice of Rules of Thumbs and Models

In this paper, we compare performances of six forecasting techniques. First, we consider two popularly used thumb rules. In case of the first thumb rule, projection for a month is obtained by multiplying the tourist arrival during the previous month with one plus the monthly growth rate observed during the corresponding month last year (Model 1). In case of the second thumb rule, for a particular month, a similar method based on the last known annual month-to-month growth rate for that month is used.

For example, suppose we want to project X_{t+1} based on the data available till month t. In case, of Model 1, the forecast is:

(1) $F_{t+1} = X_t\,[1 + \{(X_{t-11} - X_{t-12})/X_{t-12}\}\,]$

In case of the second model, the forecast is:

(2) $F_{t+1} = X_{t-11}\,[1+ \{(X_{t-11} - X_{t-23})/X_{t-23}\}\,]$

It may be noted that the above two thumb rules also play the part of naïve forecasts (shortened as Naïve 1 and Naïve 2 respectively) that are used as benchmarks for comparing forecast performance.

The third model that we use is the relatively unsophisticated method of exponential smoothing (Model 3, shortened as ES). In this case, the type of smoothing technique that takes care of both trend and seasonality has been adopted. It is well known that forecasts based on ES could be different for different values of level, trend and seasonality parameters. In our case, these parameters have been chosen based on their in-sample forecast performance.

In Model 4, the monthly growth rates have been regressed on relevant seasonal dummies. A few remarks are, however, necessary here. First, at this stage, we do not test the existence of unit roots. Our purpose at this stage is not to identify the correct or the best econometric model, but just to observe forecast performance. Second, Model 4 is basically a model of seasonal indicators pertaining to monthly growth rates of tourist arrival. If no intercept is used in the model, the seasonal indicators will be identical to the averages of the monthly growth rates pertaining to a specific month. As the number of years in our sample is fairly large, such averaging is likely to estimate the seasonal indicators with a fair degree of accuracy, especially as the descriptive analysis carried out in Section 2 reveals that they are reasonably stable.

Model 5 is an autoregressive distributed lag (ADL) type regression model. In Model 5, we consider not only the month-dummies, but also lags for the previous 36 months. Obviously, inclusion of so many variables would affect parsimony. To choose the effective lags, general to specific type of algorithm – as suggested by Hendry – is used. The approach involves specifying an initial regression equation with as many variables as one can conceive of. A backward stepwise algorithm that drops variables as per some set criteria (e.g., based on level of significance of a variable) gradually reduces the number of variables.

Finally, Model 6 is an ARIMA model. Initially, within the ARIMA framework, a few competing models are specified after obtaining the autocorrelations and partial autocorrelations of the series. The best model among them is chosen based on the diagnostic tests and AIC and BIC criteria. Appendix A presents the technical details about the choices and the final specifications for all models.

In each case, data for 2006 and 2007 have been used for testing out of sample forecast performance of all models. As our focus is on short-term forecast, the forecast horizon has been consciously kept as one month only. Thus, for each model, we compute twenty-four one-step ahead forecasts recursively using a fixed window.

To evaluate forecast performance, we consider five standard criteria: (i) mean absolute error (MAE), (ii) root-mean square error (RMSE), (iii) mean absolute percentage error (MAPE), (iv) root-mean-square percentage error (RMSPE) and (v) Theil-U statistic.[2] Among these five, the last one is specifically meant for directional accuracy.

[2] The details of the formulas etc. are available in any standard time series textbook.

4 Empirical Results

Results of forecasts from the six different approaches have been presented in Table 1. Table 1 reveals that among these, performance of ES (Model 3) and ARIMA (Model 6) turn out to be the best. The relative ranking among these two depends on the criterion used. For all criteria, however, the differences in accuracy of Models 3 and 6 are not substantially greater than that of Model 5 (ADL). Interestingly, both Model 1 and Model 2 perform reasonably well in terms of forecast performance. MAPEs for both the models are less than 5.0 per cent. Among the two naïve models, however, performance of Model 1 turns out to be better.

Table 1: One-Month Ahead Forecast - Performance of Different Models

Criteria	Model 1 (Naïve 1)	Model 2 (Naïve 2)	Model 3 (ES)	Model 4 (SI)	Model 5 (ADL)	Model 6 (ARIMA)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
MAE	16893.66	18549.04	11454.70	15882.20	12270.62	11455.20
RMSE	21262.81	22713.02	15000.70	19115.44	16425.93	14360.15
MAPE	4.37	4.81	2.96	4.06	3.09	3.14
RMSPE	5.38	5.87	3.71	4.73	3.91	3.84
Theil U	0.35	0.37	0.25	0.31	0.27	0.24

It may be noted that the results reported in Table 1 appear to be consistent with cross-country experience. Internationally, for tourist arrival series, ES and ADL type models tend to outperform other models (Witt and Witt, 1995b). In case of India, the performances of ARIMA and ES models are better than other models, with the ADL model being marginally inferior.

Table 2 presents the results on the relationship of the residuals with its first lag.[3] The reason for restricting our attention on the first lag and not higher ones is to examine whether simple residual corrections for bias etc. can improve forecasts based on the thumb rules. Tests reveal that other than Model 1 and Model 3, the residuals, in general, are uncorrelated with their first lags.

[3] It may be noted that for Model 5 and Model 6, we carried out detailed diagnostic checks. For both the models, the Ljung-Box statistics were not found to be statistically significant at 5.0 per cent level up to 36 lags.

Table 2: The Relationship between Residuals with Their First Lags

Model	Estimated Equation	Goodness of Fit (R^2)
(1)	(2)	(3)
Model 1 (Naïve 1)	e_t = -2018.07 – 0.45 e_{t-1} (-0.47) (-2.02)#	0.16
Model2 (Naïve 2)	e_t = -4377.95 -0.31 e_{t-1} (-0.93) (-1.45)	0.09
Model 3 (ES)	e_t = -1403.03 + 0.38 e_{t-1} (-0.41) (1.89)@	0.14
Model 4 (SI)	e_t = 83.01 - 0.19 e_{t-1} (0.02) (-0.88)	0.04
Model 5 (ADL)	e_t = 1498.64 – 0.12 e_{t-1} (0.41) (-0.53)	0.01
Model 6 (ARIMA)	e_t = 198.18 – 0.27 e_{t-1} (0.06) (-1.27)	0.07

Note: The bracketed figures are t-statistics. Here, # and @ denote significance at 5% and 10% level respectively.

5 Impact of Specific Events on Tourist Arrival in India

In this section, we examine the impact of some specific non-seasonal events on foreign tourist arrival in India. Depending upon the source and the nature of the events, the impact on tourist arrival could be positive (e.g., major international sports or cultural events) or negative (e.g., riots or terrorist attacks). It may be noted that these events may not necessarily be constrained to the domestic economy. For example, terror attacks like September 11 clearly affected the tourism sector globally, at least in the short-run.

In this study, we have identified major terrorist attacks both in India and the world from the Global Terrorism Database. Timing of other events (e.g., riots, World Cup Cricket etc.) is well known and detailed documentation on them is available in the public domain.

Besides these events, another major factor that can initiate decisive shift in time series pattern is Government policy. Depending upon its nature, such policies could contribute in both positive and negative manner. A major difference in the impact of policy variables is that it is likely to lead to a structural break, rather than a transitory shock. Accordingly, the econometric treatment in case of policy shocks is also different from that in case of transitory shocks. For the first type of shocks, the intervention variable is a dummy which takes the value unity only during the period of the event and both its contemporaneous and lagged impact of the variable is examined. In contrast, for the second type of shocks, the intervention dummy is unity for all the observations happening after that event.

Table 3 presents some of the major events in India chronologically during the reference period, the term in first bracket reflecting the name of the dummy variable in the study and the term in third bracket reflecting the number of lags included in this exercise to study the lagged impact. In case, the date of a particular event is on the later half of a month, the next month is considered as the starting point for that event. Thus, for Gujarat Riot, March 2002 has been considered as the starting point.

Table 3: A Chronology of Major Global and Local Events that Could Potentially Affect Tourist Arrival in India

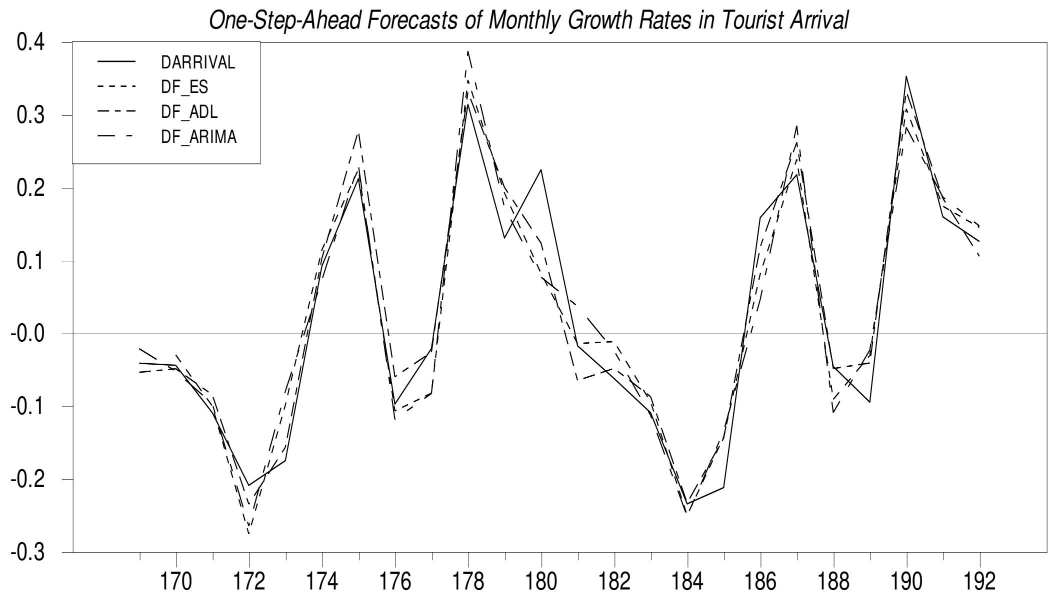
Year	Month / Date	Event Name (Dummy Variable Name) [Lags]
(1)	(2)	(3)
1992-93	Dec 92- Jan 93	Bombay Riot (BomRiot)
1993	Mar-12	Bombay Serial Blasts (BomSerialBlast1)
1996	Feb-Mar	World Cup Cricket (WC_Cricket) [0]
1998	May-11	Atom Bomb Blasts by India (AtomBomb) [3]
2001	Sep-11	Attack on Twin Towers (TwinTour) [3]
2001	Dec-13	Attack on Indian Parliament (IndianParliament) [3]
2002	Feb 27, Mar-Jun	Gujarat Riots (GujaratRiot) [3]
2002	Oct-12	Bali Bombings (BaliBombing) [3]
2002	----	New Tourism Policy of Government of India (NewPolicy)
2004	Mar-11	Madrid Train Bombings (MadridBombing) [3]
2005	Jul-07	London Bombings (LondonBombing) [3]
2006	Jul-11	Bombay Serial Blasts (BomSerialBlast2) [3]

It may be noted that it would not be possible to assess the impact of Bombay Riot and the first set of Bombay Serial Blasts within the ADL and ARIMA framework specified by us. This is because our models test the impact of lags up to a length of 36. A negative side of long-lag specification is that in this process, any extraneous information during the first 36 month period is lost. Similarly, the second set of Bombay Serial Blasts falls in the forecast period only and its impact cannot be estimated from data.

For each event whose impact could be assessed rigorously, we estimate their impact through dummy variables. In the regression set up, such dummies are specified not only for that particular period, but also for a few subsequent lags to assess their lagged impact. In case of ARIMA models, they are integrated to the main model in the form of intervention dummies. In general, we have considered three lags of all relevant dummies pertaining to riots or terrorism related events. In case of World Cup Cricket, only contemporaneous dummies have been specified. In case of the dummy pertaining to New Tourism Policy, all observations since January 2003 have been specified as unity.

Figure 4 presents the graph of actual and the forecasted monthly growth in tourist arrival in case of ES, ADL and ARIMA models after the incorporation of all corrections. In Figure 4,, the forecasted series are named DF_ES, DF_ADL and DF_ARIMA respectively. For all three models, Figure 4 indicates a good performance in terms of both magnitude and directions.

Figure 4: Performances of ES, ADL and ARIMA Models



The summary statistics on forecast performance after these corrections are summarized in Table 4. It may be noted that in case of Model 2 and Model 4, no events turned out to be statistically significant. Hence, these dummies corresponding to events were dropped from these models and as a result, there are no changes in their forecast performances reported in Table 1.

Table 4: Forecast Performance after Residual Corrections and Inclusion of Specific Events during the Reference Period

Criteria	Model 1 (Naïve 1)	Model 2 (Naïve 2)	Model 3 (ES)	Model 4 (SI)	Model 5 (ADL)	Model 6 (ARIMA)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
MAE	15818.36	18549.04	11176.38	15882.20	11997.51	11252.59
RMSE	19384.76	22713.02	14865.94	19115.44	15661.65	14237.54
MAPE	4.00	4.81	2.89	4.06	3.05	3.08
RMSPE	4.66	5.87	3.67	4.73	3.86	3.78
Theil U	0.32	0.37	0.24	0.31	0.26	0.23

Table 4 reveals that residual and other information corrections lead to some improvements in Model 1 and Model 3. For example, after these corrections, Model 1 now outperforms Model 4 in terms of MAE. Similarly, inclusion of the impact of specific events has led to modest improvements in forecasts for both the ADL and the ARIMA model. After all corrections, ES tends to outperform ARIMA as per MAE, MAPE and RMSPE but the opposite happens in case of RMSE and Theil U. In fact, the measures of forecast performances in case of the two approaches are so close that they are almost at par – whatever be the criteria. As in the earlier case, the forecast performance of ADL model also is close to ES and ARIMA.

Appendix A presents two typical estimated equations, one pertaining to ADL and the other pertaining to ARIMA with intervention dummies. Among specific events, September 11 seemed to have a significant impact on tourist arrival in India during September and October of 2001. Interestingly, domestic riots or terror related events whose impacts could be studied within our framework did not seem to have much effect on tourist arrival. So far as Government policies are concerned, the NewPolicy dummy turns out to be significant in the ADL model. However, in ARIMA model with intervention terms, the coefficient pertaining to NewPolicy was not found to be statistically significant.

6 Conclusion

The paper examined forecast performances of some simple and popular rules of thumb vis-à-vis more sophisticated time series models in the specific context of tourism. Our case study on foreign tourist arrival data in India suggested that such thumb rules had a key role to play in practical decision making. Given the perishable nature of tourism product, their effectiveness in providing a starting point in a typical forecasting exercise cannot be ignored. In the Indian context, in particular, monthly growth in tourist arrival displayed a regularity that did not change substantially even in the face of major global or local events. Such events, at best, had only a limited impact in explaining short-term growth in tourism. Given the regularity, our study suggested that rules of thumb should play an important part in short-term forecasts of foreign tourist arrival in India for years to come, unless extraneous events shift the seasonality. Our analysis, however, revealed that forecasts from simple thumb rules could be improved significantly after simple residual corrections and incorporation of some other information data on which are available in the public domain.

Among all the approaches that have been attempted in this study, ES and ARIMA models provided the best short-term forecasts, closely followed by ADL models. Earlier literature on tourism forecast for different countries had highlighted the importance of simple approaches like ES. Our results in the context of India are largely in agreement with cross-country experience in tourism forecasts.

In the context of short-term forecasts, the scope of this study is limited because it has not used real-time data. In the short-run, decisions are based on provisional rather than the final data. In case there is a systematic and explainable difference between these two sets of data, the results presented in this study might be different. Also, this study ignored the impact of some other short-term factors like air fare. A systematic research agenda in the context of tourism in India would be to improve the quality and quantity of the information content that serve as ingredients in any model building exercise.

References

Chan YM, 1993: 'Forecasting tourism: a sine wave time series regression approach', *Journal of Travel Research*, **32**, 58–60.

Chu FL, 1998: 'Forecasting tourism: a combined approach', *Tourism Management*, **19**, 515–520.

Frechtling DC, 2001: *Forecasting Tourism Demand: Methods and Strategies,* Butterworth-Heinemann.

Goh C and R Law, 2002: 'Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention', *Tourism Management*, **23**, 499–510.

Lim C, 2001: 'Monthly seasonal variations: Asian tourism to Australia', *Annals of Tourism Research*, **28**, 68–82.

Meese R and K Rogoff, 1983: 'Empirical exchange rate models of the 1970s: do they fit out of sample?' *Journal of International Economics,* **14**, 3 – 24.

Qu H and HQ Zhang, 1996: 'Projecting international tourist arrivals in East Asia and the Pacific to the year 2005', *Journal of Travel Research*, **35**, 27–34.

Witt SF and CA Witt, 1995a: *Modeling and Forecasting Demand in Tourism*, San Diego: Academic Press.

Witt SF and CA Witt, 1995b: 'Forecasting tourism demand: a review of empirical research', *International Journal of Forecasting*, **11**, 447–475.

Wong KKF, 1997: 'The relevance of business cycles in forecasting international tourist arrivals', *Tourism Management*, **18**, 581–586.

Appendix A: Forecasting Models -Technical Details

1. Exponential Smoothing (ES) Model: (Estimation Period: Jan-1992 – Dec-2005)

Figure 1 in Section 2 of this study suggested that the best form of trend would be exponential, with multiplicative form of seasonality. For such a combination, the ES model was fitted for different values of the level (α), trend (γ) and seasonality (δ) parameters. Best parameter combination for the sample period was obtained through grid search.

The chosen parameters for the ES model were: α=0.4, γ=0.2, δ=0.5

2. Autoregressive Distributed Lag (ADL) Model: (Estimation Period: Jan-1992 to Dec-2005)

Dependent Variable: DARRIVAL (e.g., monthly growth in tourist arrival)
Total Observations: 156; R^2 = 0.93; D-W Statistic = 2.11

Variable	Coefficient	Standard Error	T-Statistics	Significance
1. Constant	0.1165	0.0109	10.66	0.00
2. DARRIVAL{1}	-0.1707	0.0718	-2.38	0.02
3. DARRIVAL{2}	-0.2591	0.06339	-4.09	0.00
4. DARRIVAL{3}	-0.2301	0.049291670	-4.67	0.00
5. DARRIVAL{4}	-0.1827	0.051926916	-3.52	0.00
6. DARRIVAL{5}	-0.1523	0.056531481	-2.69	0.01
7. DARRIVAL{6}	-0.1445	0.049628794	-2.91	0.00
8. DARRIVAL{17}	0.1422	0.054252065	2.62	0.01
9. DARRIVAL{26}	-0.1535	0.051001321	-3.01	0.00
10. DARRIVAL{28}	-0.1035	0.044198888	-2.34	0.02
11. MAR	-0.1528	0.034648489	-4.41	0.00
12. APR	-0.3297	0.033266451	-9.91	0.00
13. MAY	-0.3460	0.035159238	-9.84	0.00
14. JUN	-0.1646	0.029985565	-5.49	0.00
15. AUG	-0.2231	0.026535154	-8.41	0.00
16. SEP	-0.1234	0.034244140	-3.60	0.00
17. OCT	0.1991	0.027013030	7.37	0.00
18. NOV	0.1113	0.032494122	3.42	0.00
19. DEC	0.1273	0.036681144	3.47	0.00
20. TWINTOWER	-0.1334	0.049700145	-2.68	0.01
21. TWINTOWER{1}	-0.1804	0.050193470	-3.59	0.00
22. NEWPOLICY	0.0182	0.008011492	2.27	0.02

Note: The term in second bracket indicates lag of a variable.

3. ARIMA Model: (Estimation Period: Jan-1992 – Dec-2005)

<u>LARRIVAL (e.g., log of monthly tourist arrival)</u>

Autocorrelations

1:	0.90	0.73	0.63	0.55	0.45	0.39	0.44	0.53	0.60	0.69	0.84	0.92
13:	0.82	0.67	0.57	0.48	0.39	0.32	0.36	0.44	0.51	0.60	0.74	0.82

Partial Autocorrelations

1:	0.90	-0.35	0.32	-0.17	-0.05	0.25	0.42	-0.02	0.28	0.42	0.48	0.04
13:	-0.49	-0.09	-0.08	-0.07	-0.04	-0.23	0.03	-0.12	0.10	0.08	0.20	-0.01

<u>DARRIVAL (e.g. first difference of log of monthly tourist arrival)</u>

Autocorrelations

1:	0.26	-0.28	-0.10	0.09	-0.18	-0.52	-0.19	0.09	-0.11	-0.24	0.27	0.82
13:	0.27	-0.23	-0.07	0.09	-0.13	-0.51	-0.20	0.05	-0.13	-0.24	0.28	0.82

Partial Autocorrelations

1:	0.26	-0.37	0.11	-0.02	-0.27	-0.44	-0.04	-0.26	-0.38	-0.52	-0.13	0.48
13:	0.06	0.08	0.08	0.01	0.22	-0.01	0.10	-0.11	-0.07	-0.22	-0.03	0.28

D12ARRIVAL (e.g., annual point-to-point growth rate in tourist arrival)

Autocorrelations

1:	0.72	0.50	0.39	0.31	0.25	0.26	0.28	0.28	0.22	0.12	-0.01	-0.13
13:	0.00	0.08	0.10	0.15	0.20	0.14	0.10	0.08	0.12	0.10	0.05	0.03

Partial Autocorrelations

1:	0.72	-0.04	0.09	0.01	0.02	0.13	0.07	0.05	-0.09	-0.10	-0.17	-0.15
13:	0.39	-0.03	0.04	0.07	0.08	-0.05	0.05	0.02	0.02	-0.18	-0.18	-0.08

Box-Jenkins - Estimation of the Model with Dependent Variable LARIVAL by Gauss-Newton

Dependent Variable: LARRIVAL

Usable Observations: 143; R^2 = 0.98; D-W Statistic = 2.21; Convergence in 20 Iterations.

Variable	Coefficients	Stanard Error	T-Statistics	Significance
1. AR{6}	0.0131	0.0102	1.28	0.20
2. AR{12}	1.0176	0.0111	91.38	0.00
3. MA{1}	-0.1198	0.0584	-2.05	0.04
4. MA{12}	-0.7732	0.0588	-13.15	0.00
5. N_TWINTOWER{0}	-0.1488	0.0503	-2.96	0.00
6. N_TWINTOWER{1}	-0.1381	0.0503	-2.74	0.01

Note: The term in second bracket denotes lag of a variable.