Stochastic variation, New Zealand visitor arrivals, and the effects of 11 September 2001

John Haywood[*]   John Randal[†]

Abstract

We analyse daily and monthly short-term visitor arrival time series for New Zealand, to assess the effect of the 11 September 2001 terrorist attacks on the number of visitors to New Zealand. Somewhat misleading reports from the media, based on these data, are highlighted. In particular, we show that the stochastic nature of this series precludes attribution of short-term fluctuations in visitors to any one event, such as the terrorist attacks of September 11. We recommend that measures of forecast uncertainty, such as prediction intervals, should be routinely provided. Users of forecasts could then make more informed commentary and would have better guidance for decision-making.

Key words: causation; forecast uncertainty; prediction intervals; stochastic seasonality; terrorism effects.

[*]School of Mathematical and Computing Sciences, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand. email: John.Haywood@mcs.vuw.ac.nz.

[†]School of Economics and Finance, Victoria University of Wellington.

1 Introduction

It is widely appreciated that many observed time series exhibit seasonal vari-
ation. In temperate climates, for example, any activity that is related to
periods of growth obviously varies through the year. Another seasonal class
of time series are retail trade figures, which are heavily influenced by pre-
Christmas spending and post-Christmas sales. Two clearly seasonal New
Zealand time series are displayed in Figure 1: (a) daily hours of daylight,
recorded every 10 days (available from `http://www.rasnz.org.nz/`) and
(b) daily short term visitor arrivals, shown smoothed using a six-month (182
day) centred moving average. Both series share seasonal highs in summer and
lows in winter; within-year variations seem quite similar, although slightly
out of phase. These two series are fundamentally different, however, in one
crucial respect: while the variation in hours of daylight is entirely determin-
istic and hence predictable to within a couple of minutes, the seasonality in
visitor arrivals is stochastic.

As we shall illustrate, variations in the total number of short term arrivals
to New Zealand can have many possible sources, including such rare and un-
predictable events as the terrorist attacks of 11 September 2001. Further, we
note that explanations which attribute movements in seasonal patterns to
any single cause, such as the September 11 incidents, are commonly made by
various analysts. Such explanations are almost always too simplistic and of-
ten leave much that is unexplained on closer scrutiny. We show that sensible
statistical modelling should lead to a more realistic appreciation of the fun-
damental characteristics of any data set and consequently to more informed
commentary and decision-making regarding such data.

In Section 2 we highlight somewhat misleading reports from the media
concerning these arrivals data, which motivate the need for appropriate mod-
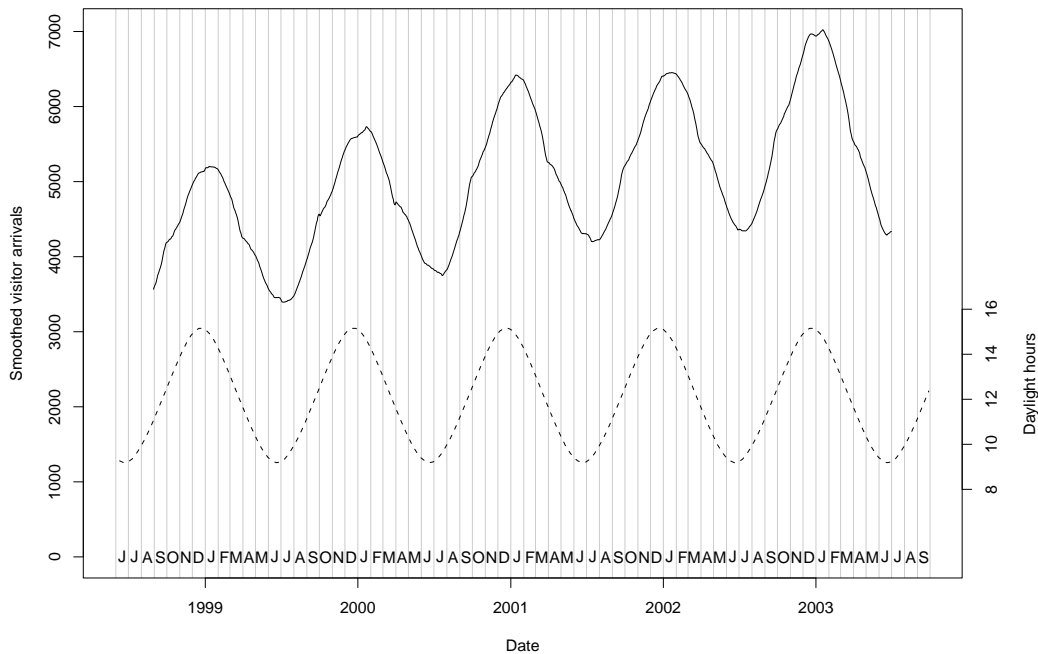
Figure 1. (a) Daily hours of daylight in Wellington, New Zealand, recorded every 10 days (dashed line and RHS axis); (b) A six month, centred moving average of the total daily short-term visitor arrivals to New Zealand (solid line and LHS axis). Both series are recorded for the period 1 June 1998 to 30 September 2003. Vertical grey lines are shown on the first of each month, and the respective months shown by their first letter along the bottom axis.

elling of stochastic variation in this and more general situations. Section 3 illustrates some of the sources of variability for New Zealand visitor arrivals but also notes that to identify all such sources is impractical. Two simple stochastic models for daily and monthly arrivals are presented in Section 4, while in Section 5 we give some concluding comments.

2 Motivation: the reported effects of 9/11 on short-term visitor arrivals to New Zealand

As an illustration of the perceived effects of the September 11 terrorist incident on New Zealand short-term visitor arrivals, we reproduce here several short passages from The Dominion Post (2002) newspaper, one of New Zealand's widely read daily papers. All the quotes appeared in the period August to December 2002, thus occurring around the one year anniversary of 11 September 2001. Note that much of the reported analysis was conducted by specialists within the tourism industry or the government and not by Dominion Post reporters. Article headlines are shown in bold.

> **Tourist firm's profit plunges 98 per cent**
> Tourism Holdings' profit nose-dived 98 per cent to $255,000 in the year to June 30; a result of last year's airline failures and *the September 11 terrorist attacks.*
> > (The Dominion Post, 29 August 2002, our emphasis)

Actually that firm's sales only fell 9 per cent, as reported later in the same article. The firm also reported, "mounting costs . . . including higher depreciation, property, lease and marketing costs." Thus the explanation of a very large change in profits by the September 11 incident seems at best partial, yet it was given prominence in the article.

> Tourism Ministry figures show there were 26,000, or 1.4 per cent, fewer tourists in total last year *than forecast*, and much of the shortfall *can be attributed to September 11.*
> > (The Dominion Post, 7 September 2002, our emphasis)

In fact, the number of visitor arrivals from 1 September 2001 to 31 August 2002 was 1,959,886, an increase of 42,102 (or 2.2%) on the previous year. The wish to explain precisely why a forecast of a stochastic series was wrong is a failure to acknowledge the naturally occurring variability in that series: any forecast will be wrong, but a sensible quantification of the likely size of the error is a realistic goal.

The following was on the front page:

Sept 11 costs Kiwis $1 billion

The main costs incurred *as a direct result of the attacks* include: $64 million in lost income from tourists.

(The Dominion Post, 7 September 2002, our emphasis)

September 11 hits tourist numbers

Uncertainty about travelling on September 11 contributed to a 3 per cent drop in visitor arrivals *in August compared with the same month last year.* Visitors from New Zealand's biggest market, Australia, dropped 12 per cent while Japanese tourist numbers fell 6 per cent. Tourism Holdings general manager Shaun Murray said, "[w]e are *very sure* that speculation that people were wary of travelling around the time of 11 September was in fact quite correct." Tourism New Zealand chief executive George Hickton agreed.

(The Dominion Post, 21 September 2002, our emphasis)

On a more positive note however, we also have the following:

Tourists spending more

Tourism Minister Mark Burton said the International Visitors

Survey showed *tourist spending* was up 14 per cent on last year.

<div align="right">(The Dominion Post, 21 November 2002, our emphasis)</div>

Tourist boom expected this summer

Tourist operators are expecting a boom tourist season this summer. "We are going to top last summer," Mr Hickton [Tourism New Zealand chief executive] said. "October was better than last year, and we think that will continue." The tourist industry in December welcomed two million visitors in a year *for the first time.*

<div align="right">(The Dominion Post, 30 December 2002, our emphasis)</div>

So in contrast to articles published three months earlier, by the end of 2002, the reported performance within the tourist industry was of record breaking numbers over the previous year and an expected additional boom in future tourist arrivals.

Explanations for variations in seasonality or trend that attribute movements to any one cause, such as the September 11 incidents noted in the quotes above, are almost always too simplistic. Comparisons are often made with the 'equivalent' period from the preceding year; however this comparison may not always correctly reflect observed features of the data. For example, what if the previous year was 'high', rather than 'normal'? Should a decrease then be viewed pessimistically? That would assume no variability was the norm, other than 'bigger is better'.

For monthly data, the 'equivalent' period from the preceding year often means 'same month'. However, that ignores known calendar effects like Easter – see, for example, Zhang, McLaren & Leung (2001), who discuss an approach suitable for an Australian Easter holiday effect. An example of a
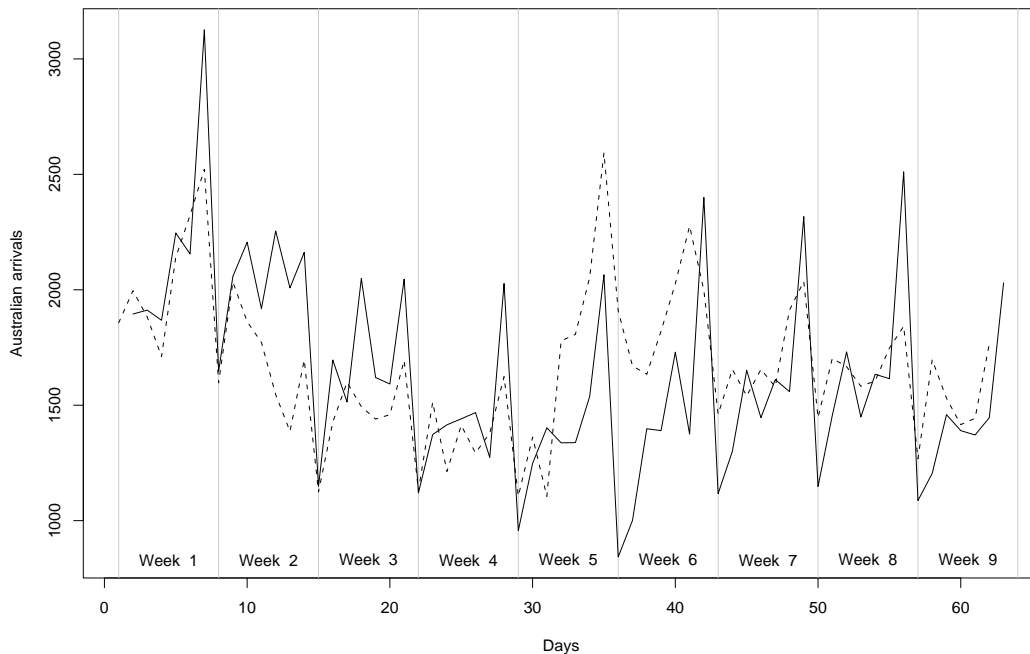
Figure 2. Daily visitor arrivals from Australia to New Zealand for July and August 2001 and 2002. The data are aligned by weekday. Consequently the 2001 series (dashed line) starts at observation 1 (a Sunday) and the 2002 series (solid line) starts at observation 2 (a Monday). Vertical grey lines are shown at each Sunday.

more obscure 'calendar effect' in these data is shown in Figure 2, which plots July and August visitor arrivals from Australia for 2001 and 2002. These series both have a seven day 'seasonal' pattern, a feature which appears in the total arrivals as well as this subseries. For a typical week the smallest number of arrivals is on Sunday, shown by the vertical grey lines in Figure 2, while the largest number of visitors arrive on Saturday.

On Saturday 13 July 2002, New Zealand's All Blacks hosted the Australian Wallabies for a rugby test match. The apparent disruption to the visitor arrivals from Australia is seen in the solid line of Figure 2, in Week 2 of the plot. Saturday's peak survives, but not as the weekly high, and abnormally large numbers are seen earlier in the week. In contrast, in 2001

the All Blacks' home game was in August. The two teams met on Saturday 11 August 2001, and this is evident in the dashed line in Week 6. Here we see relatively low arrivals on Saturday, with the peak arrivals for the week on Friday.

From the perspective of year to year changes within the months of July and August, the net result of this trans-Tasman rugby 'calendar effect' is to observe an increase in July from 2001 to 2002 but a corresponding decrease in August. Attribution of the observed August movements to a September 11 effect, as reported elsewhere and noted above, is clearly not the whole story. Our purpose here is not to provide reasons for all observed changes in arrival numbers. Rather, we hope to illustrate that sensible modelling of stochastic variation is both necessary and sufficient to identify factors that do have unusually large effects on arrivals.

3 Variability in short-term visitor arrivals to New Zealand

The economic importance of tourism to New Zealand has increased considerably in the relatively recent past. As Pearce (2001) notes in his review article on tourism in New Zealand in the 1990s, international visitor arrivals increased by 65% over the period 1990 to 1999 (from 976,010 to 1,607,241 per year) while foreign exchange earnings (in current terms) increased by 120%. Pearce also noted, however, that volatility in visitor arrivals had increased from the level observed over the 1980s. That increase in volatility, or in unpredictable variation, is salient for our main argument in this paper: ignoring that volatility may lead to forecast errors that are larger than expected and to misleading explanations for observed behaviour. More recently, Statistics

New Zealand (2003) note that for the year ended March 2002, tourism expenditure was \$14.6 billion. In that year, the tourism industry made a value added contribution to GDP of 9%, split evenly between direct and indirect contributions, and 89,711 people (FTE) had work that was directly engaged in tourism: 5.5% of the total employed workforce. In the year ended March 2002, tourism's 14.3% contribution to exports was slightly less than that of dairy products (16.9%) but greater than the contribution of meat products (10.3%), which in turn was greater than the contributions from wood and wood products, and from seafood. So tourism now ranks among the most important industries in New Zealand.

The short-term New Zealand visitor arrival series is one direct and easily recorded measurement of the international tourist contribution to the New Zealand economy; see Statistics New Zealand (2003) for a breakdown of recent expenditure by domestic and international tourists. On a monthly basis, Statistics New Zealand releases official monthly totals, and these are commonly reported in the media, often with comparison made to the same month one year ago. The daily arrivals series exhibits additional structure (like the seven day 'seasonal' pattern evident in the Australian arrivals plotted in Figure 2); however it also has inherent stochastic variation, which must be accounted for in any analysis.

We analyse daily and monthly data from 1 June 1998 to 30 September 2003, looking at the total arrivals and also the most important countries of origin, ranked by proportion of the total: Australia, the United Kingdom, the United States of America, Japan, Korea, China, and Germany. For some analysis, we omit the first four months' data and use five complete years from 1 October 1998 to 30 September 2003, to give data spanning an integer multiple of the (monthly) seasonal period.

	Minimum	LQ	Median	Mean	UQ	Maximum
Australia	15.9	27.6	33.2	33.6	38.8	64.1
UK	3.5	7.6	9.8	10.5	12.8	24.5
USA	3.2	8.3	9.9	10.4	11.9	28.6
Japan	0.9	5.6	8.0	8.7	11.2	25.5
Korea	0.0	2.2	3.7	4.0	5.5	14.1
China	0.0	1.3	2.1	2.4	3.2	11.0
Germany	0.0	1.2	2.0	2.4	3.2	11.0
Other	15.6	25.1	27.8	28.0	30.7	45.8

Table 1. Summary statistics for the daily proportion of visitors to New Zealand by origin.

As already noted, Figure 1 shows the within-year seasonal variation exhibited in total visitor arrivals, with numbers reaching a local maximum in the summer months December to February, and a local minimum in the winter months June and July.

3.1   Australian arrivals

Visitors to New Zealand from Australia made up 32.7% of the total visitors to New Zealand over the data period (3,152,070 out of 9,631,676). The maximum proportion on a day was 64.1% on 16 September 2000 (3,474 out of 5,417) and the minimum was 15.9% on 28 January 2001 (921 out of 5,794). The maximum occurred during the September school holidays in New South Wales, Victoria and Queensland for that year. The minimum was a Sunday shortly before the start of the first school term throughout Australia.

Table 1 shows that Australia is by far the single biggest source of visitors to New Zealand, with a contribution similar to that of the 'Others'. This
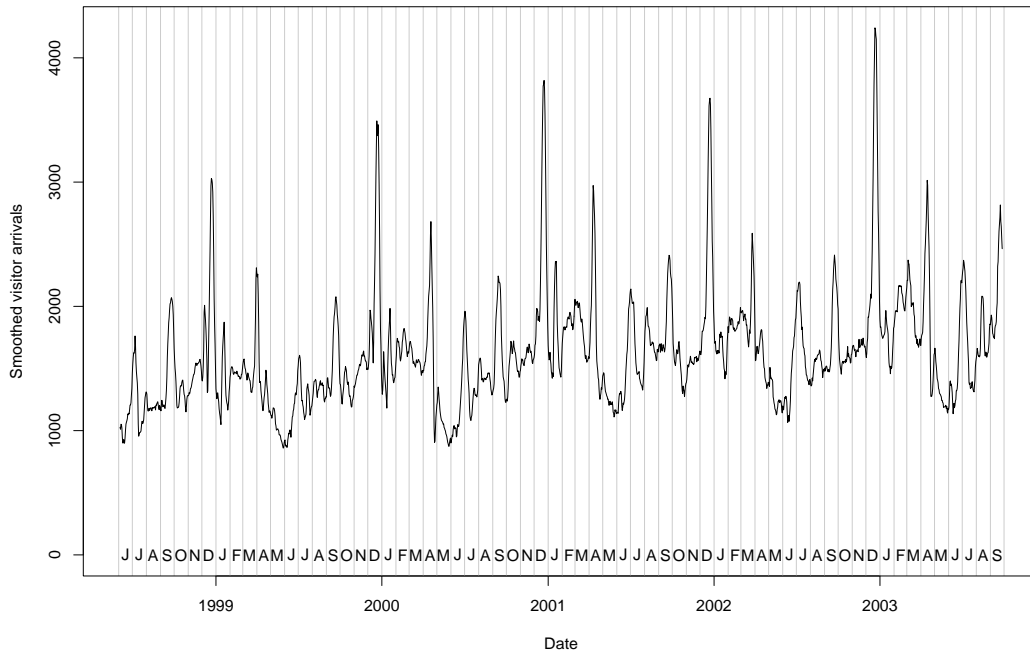
Figure 3. A seven point, centred moving average of the daily short-term visitor arrivals to New Zealand from Australia.

is significant to the total arrivals series, because as the nearest neighbour to an already geographically isolated country, the Australian arrivals exhibit variation not seen in the remaining data. A seven point moving average of the daily visitors from Australia is shown in Figure 3. This series shows that the numbers from Australia peak at four times during the year, coinciding with school holidays: Easter (March or April), mid-winter (June or July), spring (September) and at Christmas. While the Christmas arrivals regularly peak in the two weeks leading to Christmas day, the Easter arrivals reflect the movements in the Easter festival. Easter Sunday fell on 4 April 1999, 23 April 2000, 15 April 2001, 31 March 2002 and 20 April 2003 and this is reflected in the arrival data from Australia, which peak on 31 March to 3 April 1999, 19-21 April 2000, 10-13 April 2001, 27-29 March 2002 and 17-18 April 2003.

11

Year to year comparisons for the Australian series are certainly affected by these Easter movements, and because Australia is such a large contributor to total arrivals, this also feeds into year-to-year comparisons of those totals. In particular, we note that the last three Easters have alternated between April and March. March on March proportionate increases have been 8% and -2% for Australia and 14% and -4% for the Total for 2001-2002 and 2002-2003 respectively. In contrast the April proportionate increases were -15% and 28% for Australia and -6% and 5% overall.

One reason for the high numbers of Australian arrivals during school holidays is that any New Zealanders who live in Australia are classified as 'Australian' if they return to New Zealand for a visit: current place of residence determines country of origin classification for short-term visitors (Richard Penny, Statistics New Zealand, personal communication).

3.2   Non-Australian arrivals

Figure 4 shows that an Easter increase is not present in the non-Australian arrivals. In this plot we have used a smoothing window of 21 days, and we see the Easter peaks clearly in the (dashed) Australian estimates, and also their influence on the total (solid) smoothed visitor arrivals. The third (dotted) series, which is all visitors of non-Australian origin, shows no regularly repeating increase at Easter. This plot hence illustrates the unique effect, due to weight of numbers, that dynamic changes in the Australian arrivals can have on total arrivals, in contrast to the effects of arrivals from other individual countries.

Monthly subseries for the countries having the largest numbers of arrivals to New Zealand are shown in Figure 5, along with the total arrivals, for the reduced data set October 1998 – September 2003. This plot shows the
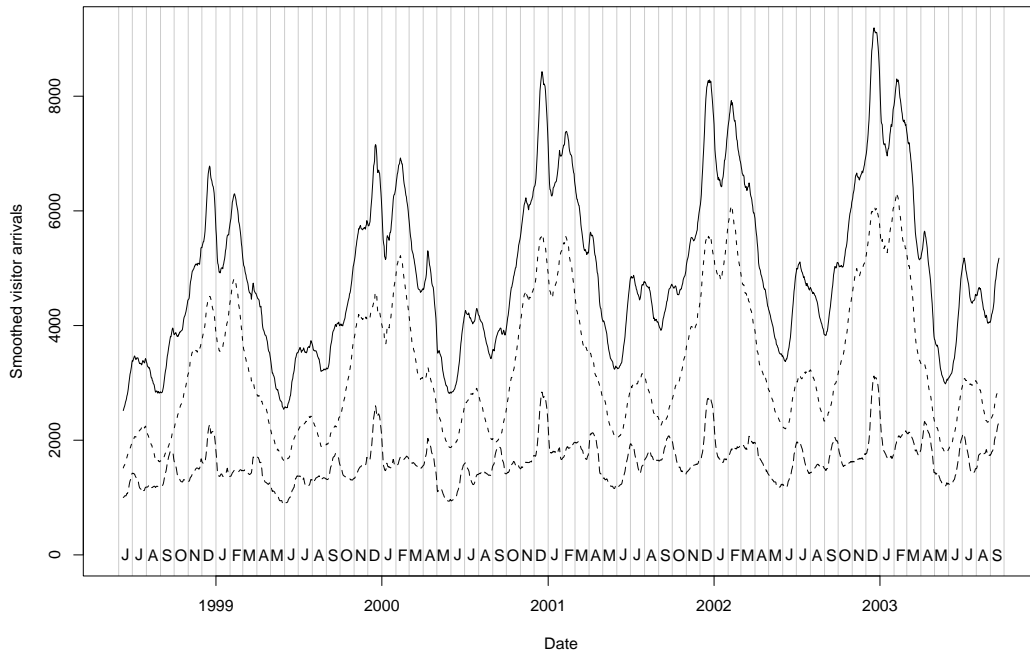
12

Figure 4. A 21-point, centred moving average of the total number of daily visitors to New Zealand from any origin (solid line), from any origin except Australia (dotted line) and from Australia (dashed line).

monthly subseries (of five observations) and allows us to see not only the trend in these subseries around the horizontal medians for each month, but also the within-year seasonality, visible as the pattern in the medians. We see quite different seasonal patterns and trends across the eight subgroups, and the total arrivals series. Visitors from the UK and Germany exhibit a fairly traditional annual seasonal pattern, linked to New Zealand's summer weather. Arrivals from the USA, Other and Total have a similar pattern but have a local maximum during winter, perhaps related to the New Zealand ski season. Korea and China both exhibit very strong growth up to 2003. The decline evident from April 2003 onwards is (partially) explained by SARS; see also Section 4. Growth is particularly strong in the number of Chinese visitors in February; this is probably related to the beginning of the academic year
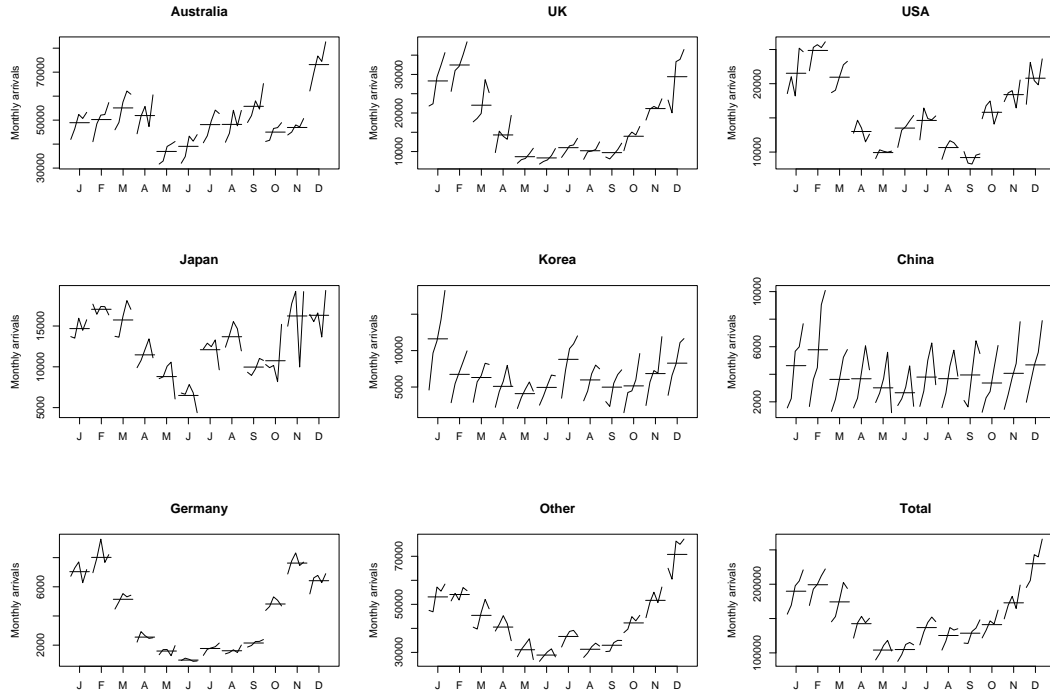
13

Figure 5. Monthly subseries of the arrivals to New Zealand; countries of origin are indicated above each plot. Each month has five observations, from October 1998 to September 2003. The vertical scales are not equal.

at New Zealand universities. Note that Chinese students, who would have student visas, are not classed as short-term visitors, but any accompanying family members or friends would be.

The exploratory data analysis presented in this section makes clear that variation in the total arrivals series comes from many sources, some of which affect only certain countries of origin. To hope to identify all such sources of variation would be wildly optimistic; to correctly anticipate the direction and magnitude of those various effects in advance would be even more difficult, and impossible in the case of completely unexpected events like those of September 11. The use of a stochastic model, however, allows explicitly for

14

'naturally occurring' variability from whatever source and hence makes it possible to infer when a particular event, such as the September 11 attacks, had an effect that was particularly noteworthy. Such models also quantify the range of outcomes that can reasonably be expected, and hence provide 'insurance' against 'typical' unforeseen events, provided that uncertainty is taken into account when forming plans. In the next section we present two simple stochastic models, which we believe well illustrate these points.

4 Modelling the short-term visitor arrivals

Although the previous sections highlight the complexities of the constituent arrival series, we proceed to model the total arrivals with a view to determining whether disruption to travel into New Zealand following 11 September 2001 was bigger than usual, given historically observed stochastic variation. For both daily and monthly data we take natural logs to help stabilise variance. Furthermore, the commonly reported comparisons with an 'equivalent' period one year earlier, noted in Section 2, are usually expressed as proportionate changes. A difference in logarithms is approximately equal to the corresponding proportionate difference in the original data, provided that difference is quite small. Hence the use of logged data could also be motivated from a tourism industry perspective.

We then use seasonal ARIMA models, popularised by Box & Jenkins (1976). For data y_t, the general form of these multiplicative models is

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)e_t, \qquad (1)$$

where B is the back shift operator, s is the seasonal period, d and D are the degrees of non-seasonal and seasonal differencing required to induce stationarity, $\phi_p(B) = (1 - \phi_1 B - \cdots - \phi_p B^p)$ is the non-seasonal autoregressive

15

component, $\Phi_P(B^s) = (1 - \Phi_1 B^s - \cdots - \Phi_p B^{Ps})$ is the seasonal autoregressive component, $\theta_q(B) = (1 - \theta_1 B - \cdots - \theta_q B^q)$ is the non-seasonal moving average component, $\Theta_Q(B^s) = (1 - \Theta_1 B^s - \cdots - \Theta_Q B^{Qs})$ is the seasonal moving average component, and e_t is Gaussian white noise with variance σ^2. The stochastic process generated by model (1) is said to have order $(p, d, q) \times (P, D, Q)_s$.

Multiplicative ARIMA models are conceptually straightforward, but do capture the main features of the arrivals data and also quantify the stochastic variation. Extensions to these models are certainly possible; for example, we do not include regression terms to model the Easter calendar effect, as detailed in Zhang et al. (2001).

For the daily data, our approach regarding Easter effects is to limit our estimation period to the year immediately preceding 11 September 2001, thereby including only one Easter period. While this may seem to ignore much of the data, in practice it is well motivated. With these daily data, we are interested in out of sample behaviour following the terrorist attacks; hence our estimation period must end by 10 September 2001. Further, in this daily context, the predictable within-year patterns obvious in Figure 1 are modelled as a local trend and the relevant seasonal pattern is the within-week one, illustrated in Figure 2, with a period of seven days. Hence we have 52 'replicates' of the seasonal pattern in our 365 observations, thus allowing a reasonable sample size for ARIMA estimation, but also restricting attention to the more recently observed seasonality.

Our final choice of model for the (logged) daily arrivals is the 'airline' ARIMA $(0, 1, 1) \times (0, 1, 1)_7$ model, used by Box & Jenkins (1976) with $s = 12$, not 7, to model a monthly time series (Jan 1949 – Dec 1960) of international airline passengers (hence the name) and, more importantly, to illustrate in

depth their approach to multiplicative seasonal ARIMA modelling. The airline model has been widely used for seasonal data, with considerable success. Much of that success may be explained by the parsimonious form of the model, requiring only two estimated parameters, which yet allows the (constrained) modelling of autocorrelations in the differenced and seasonally differenced data at lags 1, $s-1$, s and $s+1$. The form of the model's forecast function is intuitively appealing, consisting of a linear trend and seasonal dummies, continually updated to take account of the most recently observed behaviour; see Newbold (1988) for discussion of such predictors.

Motivation for the form of the airline model is also intuitive, when considered from the point of view of past data. The model takes an exponentially weighted moving average (EWMA) as the prediction for each season; e.g. this Monday is predicted as a linear combination of the preceding Sunday, Saturday, Friday, ..., with exponentially decreasing weights. To allow for season-specific effects, or days of the week in this case, the errors from that EWMA made in previous weeks are themselves tracked by a seasonal EWMA; i.e. this Monday's forecast error depends on the error from the preceding Monday's forecast, on the error from the Monday before that, and so on. That the forecast for a future day depends on recent changes in the level of the series and on behaviour typical for that day of the week is certainly sensible.

As shown in Figure 6, the in-sample correlogram of the differenced data and the unstructured residuals after model fitting, respectively suggest and support use of the airline model. This had estimated coefficients (P-values): $\theta_1 = 0.2485$ (< 0.001) and $\Theta_1 = 0.7396$ (< 0.001). One standard time series model check is a portmanteau test for lack of residual correlation, such as the Ljung-Box test (Ljung & Box 1978). When calculated from
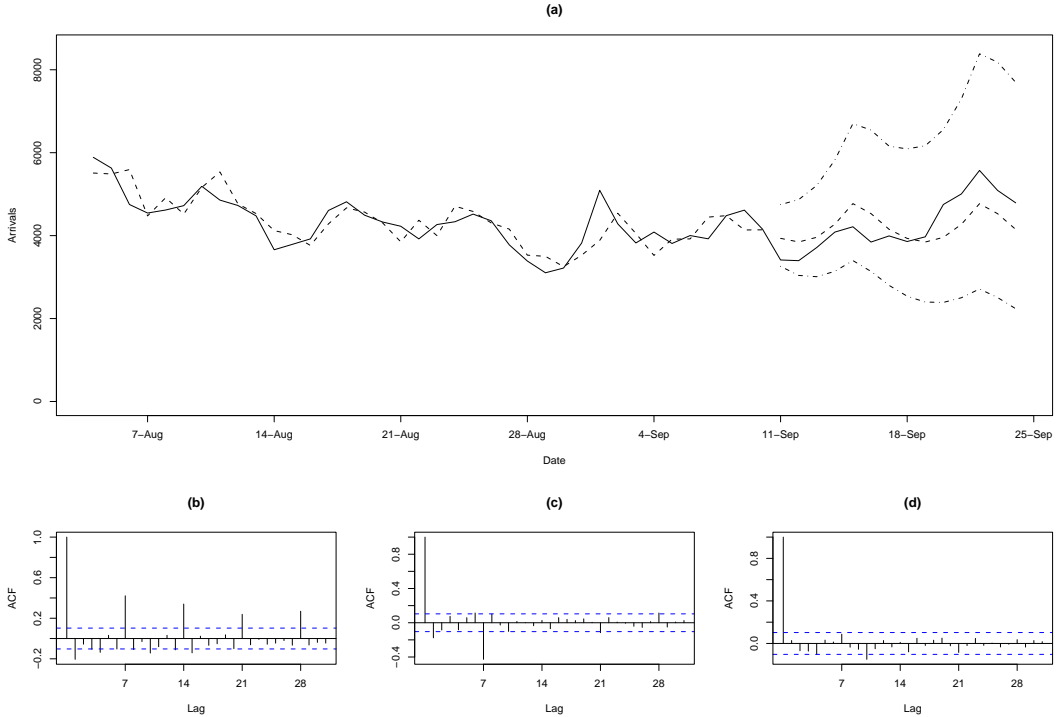
17

Figure 6. (a) Data (solid line), forecasts (dashed line) and out of sample prediction intervals (dot-dashed line) for daily arrivals data, 4 August to 24 September 2001. The forecasts are from the 'airline' model, estimated up to 10 September 2001. (b) The correlogram for the differenced, logged arrivals data. (c) The correlogram for the differenced and seasonally differenced logged arrivals data. (d) The correlogram for the residuals from the 'airline' model. All correlograms have approximate 95% confidence intervals.

residual sample autocorrelations up to lag 25, the Ljung-Box (chi-squared) statistic is $Q_{LB}(25) = 31.866$, which is not significant at a 10% level on 23 degrees of freedom. The choice of the airline model is also backed up by the use of information criteria, with the lowest recorded SBC value of all models considered after differencing and seasonally differencing the data (Schwarz 1978).

Even with a reasonable Ljung-Box statistic, the significance of individual

18

residual autocorrelations at lags 4 and 10, visible in Figure 6, could arguably be of some concern, although neither lag is related to the seasonal (weekly) pattern. One alternative to the airline model, which results in no significant residual autocorrelations, is the ARIMA $(1,1,1) \times (1,0,1)_7$ model. Instead of seasonal differencing, a first order seasonal autoregressive parameter is included, along with an additional first order non-seasonal autoregressive parameter. After allowing for the different number of available observations, due to the lack of seasonal differencing, this latter model is preferable on the basis of SBC, and has $Q_{LB}(25) = 19.511$; clearly not significant on 21 degrees of freedom.

Ultimately however, we prefer the airline model for two reasons: parsimony, with two estimated parameters rather than four, and for protection against structural changes, as advocated by Hendry in a series of papers with several coauthors: for example, Clements & Hendry (1996), Clements & Hendry (2001) and Hendry & Mizon (2001). In this instance, one relevant type of structural change that we protect against with seasonal (weekly) differencing is changes to the potential arrivals. Such changes could be caused by new entrants to the airline industry, airline closures or restructuring. A good example of that type of change occurred on 14 September 2001, when Ansett Australia and Ansett International suspended all flights, including those into New Zealand, following a move to voluntary administration on 12 September. Ansett was a major airline in the Australasian region, and its financial collapse and termination of flights clearly affected flight schedules and capacities within New Zealand and on routes from Australia. Of course, use of the airline model in this case does not protect against that particular structural change, since it occurred outside the estimation period and no modelling approach can guard against such out of sample occurrences.

19

Also shown in Figure 6 are forecasts and out of sample prediction intervals for 11-24 September 2001, the period immediately following the terrorist attacks in New York (that is, forecasts made from the airline model estimated only up to 10 September 2001). Note that the prediction intervals widen quickly due to the integration in the airline model, and also to the inherent variability in the arrivals data. On the basis of the forecasts shown, there is no evidence that anything 'out of the ordinary' occurred in this period, and the out of sample data are fully consistent with the stochastic nature of the observed seasonality. Similar conclusions are reached if the ARIMA $(1, 1, 1) \times (1, 0, 1)_7$ model is used: all out of sample observations fall within the prediction intervals. Hence our final choice of model does not affect our central argument: for the daily data, observed departures from expectation were no greater after September 11 than they were beforehand.

Although we have only displayed two weeks of out of sample forecasts, we make no suggestion that the effects of the 11 September attacks would all be observed within that period. What is clear, however, is that given the rate of divergence of the prediction intervals as the forecast horizon increases, any subsequent observations would certainly fall within those intervals. Such wide prediction intervals reflect the appropriately acknowledged uncertainty about the future behaviour of the modelled series.

Daily arrivals are naturally much more volatile than arrivals over longer periods, so we next consider monthly arrivals data, aggregated from the daily arrivals. Analysis of the sample autocorrelation and partial autocorrelation functions suggests use of an $ARIMA(1, 0, 0) \times (0, 1, 1)_{12}$ model for the log data, with a constant included in the model. This had estimated coefficients (P-values): constant $= 0.0692$ (< 0.001), $\phi_1 = 0.4468$ (< 0.001) and $\Theta_1 = 0.8923$ (0.173) with $Q_{LB}(12) = 5.839$, which is clearly not significant
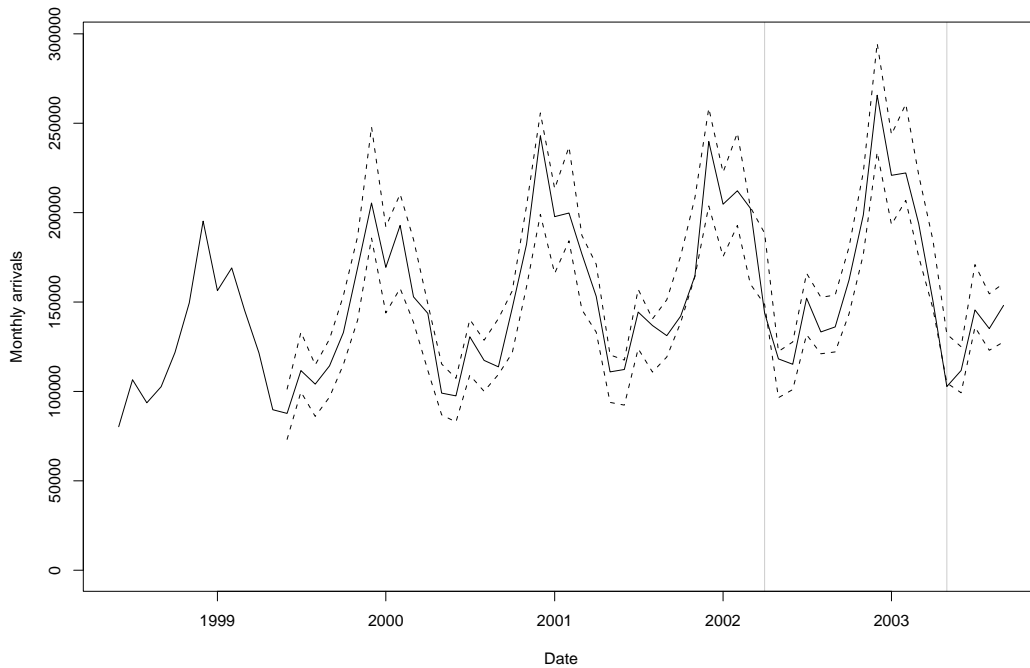
20

Figure 7. Data (solid line) and prediction intervals for one step forecasts (dashed line) for monthly arrivals data, June 1998 to September 2003. Vertical grey lines at April 2002 and May 2003 indicate data that fell outside the intervals (see text).

on 10 degrees of freedom. Even though the first order seasonal moving average term is not significant at conventional levels, its inclusion is strongly supported by the use of information criteria. Also, without it the residual sample autocorrelation at the relevant seasonal lag, now lag 12, is significant at a 5% level. A constant is necessary to model the (local) upward trend that is evident in the series as a positive drift term in the seasonal random walk.

Since we have only five years of data, giving just five complete seasonal cycles, we do not consider out of sample forecasts viable with the monthly data. Instead we focus on in sample forecasts; Figure 7 shows the full monthly series, as well as 95% prediction intervals for the one-step in sample forecasts from the chosen model.

21

The observed data lie inside the prediction intervals in all but two instances, April 2002 and May 2003, neither of which seem related to events in September 2001. The first instance is directly attributable to the Easter calendar effect. Easter fell in April in 2001, and so April 2001 arrivals were higher than forecast; see Section 3.1 for relevant dates of peak arrivals. Due to the seasonal differencing in the model and the first order seasonal moving average term, the forecast for April 2002 is influenced by the previous April's error and consequently is forecast as high. In 2002 however, Easter fell in March; as a consequence the April arrivals were lower than forecast, and in this case outside the prediction intervals of the model.

The second observation outside the prediction intervals in Figure 7 occurs in May 2003. We suggest this is probably related to SARS, but realise that such direct attribution of cause and effect is exactly what we warn against more generally here! Referring back to Figure 5, note that while the number of Australian, British, American and German visitors increased in May 2003 relative to May 2002, the visitor arrivals from the Asian countries (Japan, Korea and China) and those from 'Other' origins, were dramatically down on 2002, supporting the suggested linkage to SARS.

Regarding the events of September 2001, certainly the data in Figure 7 do lie just above the lower prediction interval in October and November 2001. We have already noted that in addition to the terrorist attacks, one of the major airlines in the Australasian region ceased flights within and into New Zealand on 14 September. Hence while arrival numbers were lower than forecast following September 2001, they were not significantly so when modelled appropriately. Also there is more than one plausible causal mechanism for the observed movements, in contrast to the majority of reports cited in Section 2.

Our chosen model for the monthly data provides forecasts that are consistent with an earlier analysis of a longer data set. Thomson (1986) modelled logged monthly arrivals to New Zealand using a semi-parametric approach, taking as forecast functions an estimated straight line for each month, hence allowing both intercept and slope differences for each season. Thomson's modelling was thus far less parsimonious than ours, but his focus was on exposition of possible approaches to such data, including simple but informative exploratory data analysis. By comparison the airline model, which also fits this monthly data well, would allow intercept differences but impose a common slope, as does the model we chose here with the slope included via the drift term. Central to all such forecasts is the continually updated, stochastic, repeating monthly seasonal pattern and (local) trending behaviour. One further difference between Thomson's approach and ours is that parametric models, such as the multiplicative ARIMA class (1), produce prediction intervals routinely; the use of these intervals is central to our argument concerning 'usual' and 'larger than expected' deviations from a point forecast. In contrast, Thomson's semi-parametric approach would not provide appropriate prediction intervals, for increasing lead times, without considerable further effort.

5 Discussion

The growth in the number of visitor arrivals to New Zealand was certainly lower than expected in the latter part of 2001, as reported elsewhere and noted in Section 2. However we do not feel that there is conclusive evidence to attribute this fall solely to the terrorist events of September 2001. As mentioned above, the termination of flights by Ansett Australia and Ansett
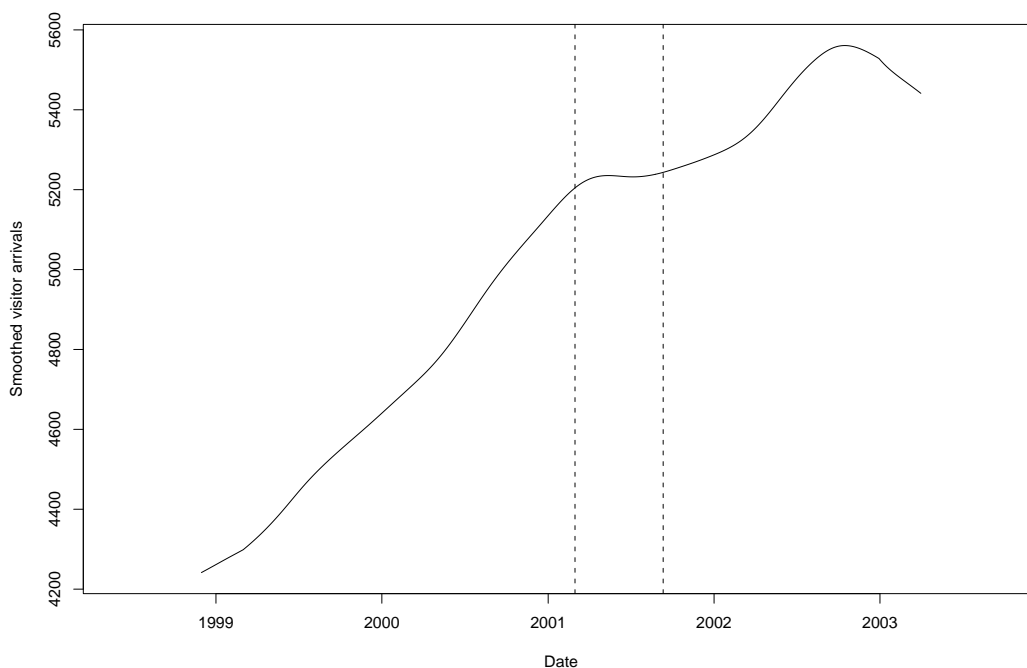
Figure 8. A trend estimate of the daily visitors to New Zealand from any origin. The vertical lines are at 1 March 2001 and 11 September 2001.

International on 14 September 2001 certainly affected capacity and timing of arrivals to New Zealand. Distinguishing the Ansett effect from those of the terrorist attacks is problematic, given the three day separation between them. A further plausible cause for the lower than forecast visitors to New Zealand is the US recession dated March 2001 (Hall, Feldstein, Bernanke, Frankel, Gordon & Zarnowitz 2001), along with the world wide flow-on effects from a slow down in the US economy. The recession predates September 2001 by six months but it is consistent with observed features of the NZ visitor arrivals data. A trend estimate for the daily arrivals produced using STL (Cleveland, Cleveland, McRae & Terpenning 1990) and displayed in Figure 8 shows a levelling off just after March 2001. Arrivals increase from about mid-2001, passing through September and a further acceleration occurs early in 2002.

The apparent large reduction in visitors from late 2002 onwards is not due

24

to SARS but rather is primarily an 'end effect', caused by the data stopping soon after the winter months when arrivals are low. The estimation of trend by loess-based filters, such as STL, is known to be problematic at the ends of series, in particular because of large revisions to new observations; see Gray & Thomson (1990). While SARS has probably reduced the trend slightly from what would otherwise have occurred, the observed effect will be far less than that suggested by Figure 8. In contrast, trend movements visible in 2001 are sufficiently far back in time to be safe from revisions due to updating.

We have highlighted several events that have affected the numbers and the timing of visitors to New Zealand, some of which will continue in the future, like the trans-Tasman rugby 'calendar effect'. Part of our purpose was to illustrate that there are a vast number of such events, affecting visitors from different countries of origin. Rather than need to try and identify all such sources of variation, a sensibly chosen (simple) stochastic model explicitly quantifies such variability and uses it when providing prediction intervals for optimal forecasts of future visitor numbers. As shown in Section 4, for either daily or monthly arrivals data, the observations after 11 September 2001 lie within 95% prediction intervals obtained from such stochastic models. Hence while several events in 2001 almost certainly did affect arrivals to New Zealand, none of those events, alone or combined, resulted in variation larger than might reasonably have been expected, given historically observed variability.

Forecasts of future visitor arrivals to New Zealand are commonly made and reported. For example, Tourism Research Council New Zealand (2003) provide annual forecasts for 2003 to 2009, broken down by country of origin. They also give monthly forecasts for July 2003 to June 2005 for a subset of those countries. Perhaps of most concern is the lack of any assessment

of forecast uncertainty, other than to once note that, "...the assumptions underlying the forecasts ...are subject to uncertainty. The forecasts should therefore be used as guidelines rather than exact forecasts." Tourism Research Council New Zealand (2003), p. 19. Our suggestion is that those guidelines would be much more useful if they included a plausible range of values for the forecasts that reflected observed uncertainty, and would help to prevent the misleading attribution of subsequent variation solely to individual 'newsworthy' events, such as the terrorist attacks of September 2001.

A common complaint is that prediction intervals quickly become 'ridiculously wide' even at short forecast horizons, and therefore are of no use. Such wide intervals appropriately reflect the uncertainty present in forecasts; to suggest they are of no use reveals a desire to view observed time series as more predictable than is warranted. We hope such misguided views will be reported less often in future.

References

Box, G. E. P. & Jenkins, G. M. (1976), *Time Series Analysis, Forecasting and Control*, second edn, Oakland: Holden-Day.

Clements, M. P. & Hendry, D. F. (1996), 'Intercept corrections and structural change', *Journal of Applied Econometrics* **11**, 475–494.

Clements, M. P. & Hendry, D. F. (2001), 'An historical perspective on forecast errors', *National Institute Economic Review* **177**, 100–112.

Cleveland, R. B., Cleveland, W. S., McRae, J. & Terpenning, I. (1990), 'STL: A seasonal-trend decomposition procedure based on loess', *Journal of Official Statistics* **6**, 3–73.

Gray, A. G. & Thomson, P. J. (1990), 'Invited commentary on 'STL: A seasonal-trend decomposition procedure based on loess' by R. B. Cleveland et al.', *Journal of Official Statistics* **6**, 47–54.

Hall, R., Feldstein, M., Bernanke, B., Frankel, J., Gordon, R. & Zarnowitz, V. (2001), The business-cycle peak of March 2001, Technical report, Business Cycle Dating Committee, National Bureau of Economic Research, USA. Available at `http://www.nber.org/cycles/november2001/`.

Hendry, D. F. & Mizon, G. E. (2001), Forecasting in the presence of structural breaks and policy regime shifts, Technical report, Available at `http://www.nuff.ox.ac.uk/Users/Hendry/Papers/papersprogress.html`.

Ljung, G. M. & Box, G. E. P. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**, 297–303.

Newbold, P. (1988), 'Predictors projecting linear trend plus seasonal dummies', *The Statistician* **37**, 111–127.

Pearce, D. (2001), 'Tourism', *Asia Pacific Viewpoint* **42**, 75–84.

Schwarz, G. (1978), 'Estimating the dimension of a model', *The Annals of Statistics* **6**, 461–464.

Statistics New Zealand (2003), *Provisional Tourism Satellite Account 2000-2002*, Wellington: Statistics New Zealand.

The Dominion Post (2002), Wellington: The Dominion Post.

Thomson, P. J. (1986), Forecasting visitor arrivals to New Zealand: An elementary analysis, *in* R. J. Brook, G. C. Arnold, T. H. Hassard & R. M. Pringle, eds, 'The Fascination of Statistics', New York: Marcel Dekker, pp. 351–373.

Tourism Research Council New Zealand (2003), *International Visitor Arrivals Forecasts 2003-2009*, Wellington: Ministry of Tourism.

Zhang, X., McLaren, C. H. & Leung, C. C. S. (2001), 'An Easter proximity effect: Modelling and adjustment', *Australian & New Zealand Journal of Statistics* **43**, 269–280.