Trending seasonal data with multiple structural breaks. NZ visitor arrivals and the minimal effects of 9/11

John Haywood[*] John Randal[†]

5 November 2008

Abstract

We demonstrate the poor performance, with seasonal data, of existing methods for endogenously dating multiple structural breaks. Motivated by iterative nonparametric techniques, we present a new approach for estimating parametric structural break models that performs well, and which amalgamates the Macaulay cycle with modern structural break estimation. We suggest that iterative estimation methods are a simple but important feature of this approach when modelling seasonal data. The methodology is illustrated by simulation and then used for an analysis of monthly short term visitor arrival time series to New Zealand, to assess the effect of the 9/11 terrorist attacks. While some historical events had a marked structural effect on those arrivals, we show that 9/11 did not.

Key words: Break dates; Endogenous dating of structural changes; Iterative fitting; Seasonal adjustment; Trend extraction.

[*]School of Mathematics, Statistics and Operations Research, Victoria University of Wellington.

[†]Corresponding author. School of Economics and Finance, Victoria University of Wellington, PO Box 600, Wellington, New Zealand. email: john.randal@vuw.ac.nz.

1 Introduction

The economic importance of tourism to New Zealand is high and has increased considerably in recent years. As Pearce (2001) noted in his review article, international visitor arrivals increased by 65% over the period 1990 to 1999 and foreign exchange earnings increased by 120% (in current terms). More recently, for the year ended March 2004 tourism expenditure was $17.2 billion Statistics New Zealand (2005). In that year, the tourism industry made a value added contribution to GDP of 9.4%, while 5.9% of the total employed workforce had work directly engaged in tourism. Further, tourism's 18.5% contribution to exports was greater than that of all other industries including dairy products, which in turn was greater than the contributions from meat and meat products, wood and wood products, and seafood.

The time series of monthly short term visitor arrivals to New Zealand is one direct and easily recorded measurement of the international tourist contribution to the New Zealand economy. A useful precursor to development of tourism policy or business strategy is an understanding of the dynamic behaviour of these seasonal data. A classical time series decomposition includes unobserved components representing an evolving trend, a seasonal encapsulating regular deviation from the trend on a within-year basis, and an irregular, which is the residual or unexplained variation in the data. There are various ways to estimate these components, using both parametric and nonparametric approaches; see for example Harvey (1989), Hamilton (1994), Findley *et al.* (1998), Franses (1998) and Makridakis, Wheelwright & Hyndman (1998). Such a decomposition then allows an interpretation of the dynamic behaviour of visitor arrivals in terms of the estimated components.

There seems little doubt that the terrorist attacks of 11 September 2001 have had a pronounced influence on world events since that time. For example, see US Department of State (2004), for a summary of 100 editorial opinions from media in 57 countries around the world, commenting on the three years following September 2001. Those terrorist events and their subsequent effects have been used to explain apparent movements in many time series, and in this paper we concentrate on a particular example: the number of short term visitor arrivals to New Zealand.

Our focus is to detect any longer term, or structural, changes in trend or seasonal components of the arrivals as a result of the 9/11 events. We also wish to compare the magnitude of any 9/11 effects with those due to other causes. Consequently we did not wish to specify the dates of any structural changes, but rather estimate the number and position of these endogenously. To achieve this we investigated the use of Bai & Perron's (1998, 2003) procedures for estimating multiple structural changes in a linear model. Their approach permits periods of stable dynamic behaviour between relatively infrequent but significant changes to the parameters of the model. However, there is clearly no empirical requirement that changes in the trend and seasonal components occur simultaneously. As we demonstrate, for the visitor arrivals data changes typically occur more frequently in the trend. In contrast, direct application of Bai & Perron's (1998, 2003) methodology fits components simultaneously and yields a relatively poor decomposition, as we show via a simulation study and analysis of the visitor arrivals. We propose a new iterative fitting procedure for seasonal data, based on Bai & Perron (1998, 2003) and using existing R packages (R Development Core Team, 2008), which gives much improved performance in terms of flexibility of fitted trends (via more appropriate placement of breaks) and lack of residual serial correlation.

Throughout the paper the term 'trend' (or trend component) is used to describe the evolving, underlying behaviour of a time series. That underlying behaviour reflects both long term movements and medium term cyclical fluctuations, where long term and medium term are in relation to the (shorter) period of the evolving seasonal component that we also consider explicitly. This notion of trend agrees with that used by many national statistical offices; e.g., see section 2.1, Australian Bureau of Statistics (2003). Certainly we agree with Busetti & Harvey (2008), that the strong but common assumption in the econometrics literature of a constant underlying slope when testing for a trend is often implausible. Since our focus is on breaks in the structure of the trend and seasonal components, we choose to model those components piecewise, with endogenously estimated changes in the linear trend slope and/or seasonal pattern corresponding to identified structural changes.

We find there is actually little to suggest that the September 11 incidents had much effect on New Zealand visitor arrivals, when viewed in the context of 'normal' historically observed movements. In contrast, we suggest some other historical events which do appear to have affected visitor arrivals to New Zealand quite markedly. We make no attempt to forecast the arrivals data using structural break models; we suggest other approaches, such as ARIMA modelling (Box & Jenkins, 1976), would be more suitable if prediction was the aim. In fact Haywood & Randal (2004) used that approach to demonstrate that the 9/11 events did not significantly affect New Zealand visitor arrivals, by showing that the observations post-9/11 were contained within out of sample prediction intervals computed using a seasonal ARIMA ('airline') model, fitted to arrivals data up to 9/11. In this paper though, the focus is explicitly on identifying structural changes if they exist in the arrivals data, in trend and/or seasonal components.

In Section 2 we present an exploratory data analysis (EDA) of New Zealand visitor arrivals and a discussion of some apparent sources of variability in the data. Section 3 motivates and presents the iterative estimation of a parametric model that allows separate structural changes in the trend and seasonal components. Simulated data is used to illustrate the good performance of the new methodology. In Section 4 we use our iterative approach to model the arrivals data and in Section 5 we give some concluding comments.

2 EDA of short term visitor arrivals to New Zealand

We consider 25 complete years of monthly short term visitor arrival series from January 1980 to December 2004. The arrivals are from the seven most important countries of origin, ranked by current proportion of the total: Australia, UK, USA, Japan, Korea, China, Germany, as well as a residual series from 'Other' origins. We analyse these series individually along with their aggregate, denoted 'Total' (Figure 1).

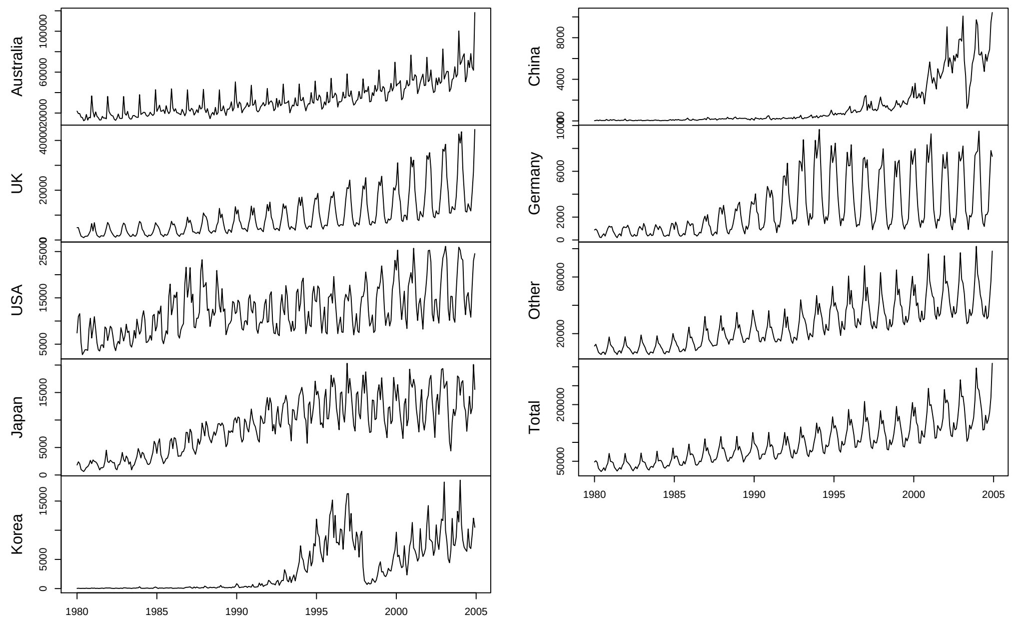

Figure 1. Monthly short term visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The vertical scales are not equal.

As seen in Figure 1 a 'U'-shaped seasonal pattern is common, with visitor numbers reaching a local maximum in the summer months December to February, and a local minimum in the winter months June and July. Further, it is apparent that the amplitude of the seasonal variation tends to increase with the level of the series, indicating a multiplicative relationship between trend and seasonal components. Australian and UK arrivals appear to be growing at a relatively steady rate. In contrast, a large downturn in arrivals from the USA is evident in the late 1980s, a period which closely followed the stock market crash of October 1987. The trend in Japanese arrivals levels off over the last 15 years. The effect of the Asian financial crisis of 1997 is evident especially in the Korean data, with visitor numbers dramatically reduced just after this event. Arrivals from China contain perhaps the most visible short term effect in these series, which is due to the SARS epidemic that virtually eliminated international travel by Chinese nationals during May and June 2003. German arrivals show a clear change from exponential growth prior to the early 1990s to a more stable pattern in recent times. The Other arrivals show a SARS effect much less prominent than that seen in the Chinese arrivals, as do some further series including Total arrivals. One of the more obvious shifts in the aggregate Total series appears to be linked to the Korean downturn, which can be attributed to the Asian financial crisis.

The Asian financial crisis of 1997-1998 markedly affected stock markets and exchange rates in nine East Asian countries: Hong Kong, Indonesia, Japan, Korea, Malaysia, Philippines, Singapore, Taiwan and Thailand. See Kaminsky & Schmukler (1999) for a chronology of the crisis, from the official onset marked by the devaluation of the Thai baht on 2 July 1997 up to the resignation of Indonesian President Suharto in May 1998. Kaminsky & Schmukler (1999) suggest the presence of important contagion effects in those markets, based on an analysis of identified market jitters. More recent analysis by Dungey *et al.* (2004) suggests, however, that increased exchange rate volatility observed in Australia and New Zealand around that time was not due to contagion from Asian countries, or unanticipated factors, but rather to common (anticipated) world factors such as trade linkages. This is one context in which changes in short term visitor arrivals to New Zealand from Asian countries around 1997-1998 can be viewed, since tourism has become such an important sector of the New Zealand economy, as noted above. In particular, Korea is one of the five source countries with the largest recent (2000-2004) proportion of visitors to New Zealand (Table 1).

Table 1. Summary statistics for the monthly proportion of visitors to New Zealand, by origin. The final three columns give proportions of the Total for the entire 25 year sample period, and the five-year periods 1980-1984 and 2000-2004, respectively.

	Min	LQ	Median	UQ	Max	80-04	80-84	00-04
Australia	21.8	30.0	35.9	41.6	58.8	33.8	44.9	33.3
UK	3.4	6.3	8.0	10.6	18.3	9.8	7.6	11.8
USA	6.3	10.3	13.0	16.3	29.4	12.4	16.7	10.0
Japan	2.8	7.1	9.1	11.0	17.8	9.2	5.9	7.8
Korea	0.0	0.2	1.0	4.3	10.5	3.4	0.2	4.8
China	0.0	0.2	0.4	1.2	4.7	1.4	0.1	3.1
Germany	0.8	1.5	2.2	3.4	7.5	2.9	1.8	2.6
Other	17.9	23.4	26.1	28.6	34.3	27.0	22.8	26.7

Table 1 shows that Australia is by far the biggest single source of visitors to New Zealand, accounting for almost exactly one-third of visitors in the 2000-2004 five year period and slightly more over the entire data period. The maximum proportion in a month from Australia was 58.8% in June 1985, and the minimum was 21.8% in February 1997. An Australian influence is notable in the Total arrivals, because as the nearest neighbour to a geographically isolated country, arrivals from Australia exhibit variation not seen in the remaining data. As seen in Figure 1, the Australian data has a regular seasonal pattern which is quite different from that of any other country. A closer examination indicates three peaks per year before 1987 and four thereafter; we discuss this further in Section 4.

One way of estimating unobserved trend and seasonal components is to use a robust, non-parametric technique such as STL (Cleveland *et al.*, 1990); here we use STL as implemented in R (R Development Core Team, 2008). This procedure consists of an iterated cycle in which the data is detrended, then the seasonal is updated from the resulting detrended seasonal subseries, after which the trend estimate is updated. At each iteration, robustness weights are formed based on the estimated irregular component and these are used to down-weight outlying observations in subsequent calculations. A typical STL decomposition is shown in Figure 2 for the natural logarithm of the Total arrivals. The log transformation is commonly used to stabilise a seasonal pattern which increases with the level of the series, and effectively transforms a multiplicative decomposition into an additive one.

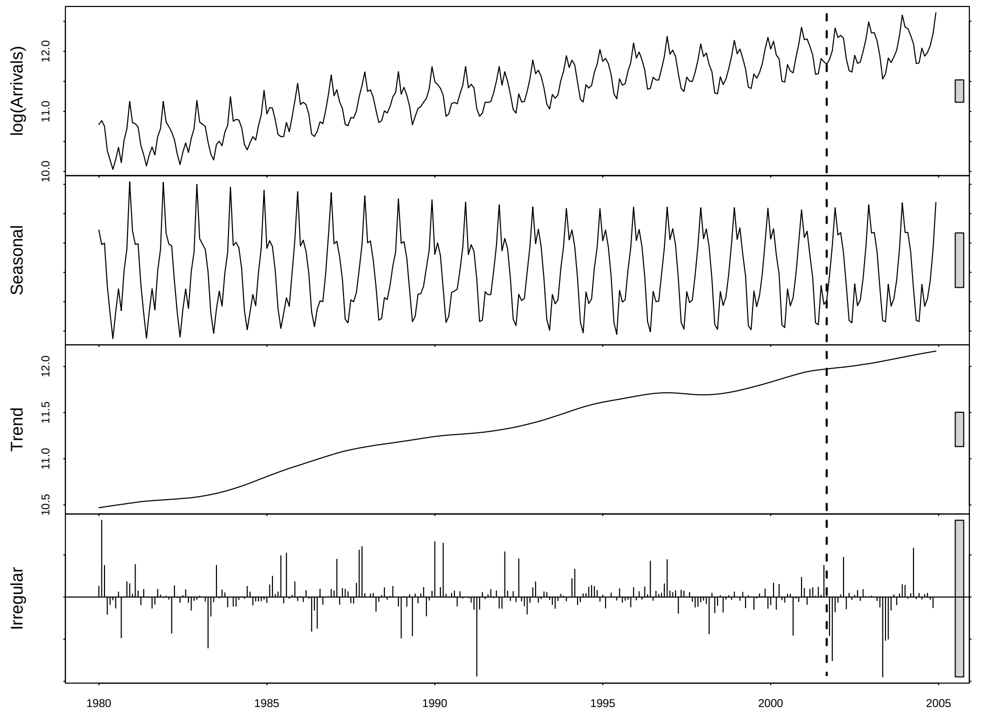

Figure 2. The STL decomposition of the log aggregate monthly visitor arrivals to New Zealand from January 1980 to December 2004. The vertical dashed line is at September 2001, and the solid bars on the right hand side of the plot are all the same height, to aid comparisons.

Figure 2 shows an evolving seasonal pattern, an upward trend with several changes in slope, and a relatively small irregular component. A vertical line is added to indicate September 2001. There is no obvious (structural) change in the trend at or about this month, although there is a reduction in the slope of the trend nearer the start of 2001, which we discuss further in

Section 5 below. More prominent is a cluster of negative irregulars immediately following 9/11, the largest of which is the third largest negative irregular in the sample period. Jointly though, these irregulars are smaller and less persistent than those occurring at the time of the SARS outbreak in 2003. Our exploratory analysis with STL thus suggests that while the events of 9/11 may have had a moderate short term (irregular) effect, there is nothing to suggest that a longer term (structural) effect occurred. We investigate this hypothesis more formally in Section 4.

3 Iterative break estimation for seasonal data

Bai & Perron (1998, 2003) present a methodology for fitting a linear model with structural breaks, in which the break points, i.e. the times at which the parameters change, are determined optimally. The optimal positions of m break points are determined by minimising the residual sum of squares, for each positive integer $m \leq m_{\max}$. The optimal number of break points ($0 \leq m^* \leq m_{\max}$) may then be determined by, for example, minimising an information criterion such as BIC (Schwarz, 1978). Given a sample of T observations, the selected break points are estimated consistently, with rate T convergence of the estimated break fractions (that is, the proportions of the data between consecutive breaks).

The maximum number of break points, $m_{\max}$, is determined by the number of observations relative to the number of parameters in the model. In general, for a model with m breaks and q parameters, at least q observations are needed between each pair of break points, requiring at least $T \geq (m+1)q$ observations in total. Clearly if the model has many parameters, fewer break points can be estimated from a given sequence of observations.

We consider implementing this approach for a time series of the form

$$Y_t = T_t + S_t + I_t \qquad t = 1, \ldots, T$$

where Y_t are the observed data (transformed if necessary), T_t is an unobserved trend component, S_t is an unobserved seasonal component with seasonal period s, and I_t is an unobserved irregular component. Many observed time series do not follow an additive decomposition, including the NZ visitor arrivals as noted in Section 2; however, we assume a suitable stabilizing transformation can be applied (e.g., see Section 4). In the following model, evolution of trend and seasonal components is explicitly modelled as structural changes occurring at endogenously identified break points. Short term, random changes may also occur, but these are modelled by the irregular component I_t. We assume that between two break points t^*_{j-1} and t^*_j ($j = 1, \ldots, m+1$), the trend T_t is linear,

$$T_t = \alpha_j + \beta_j t \qquad t = t^*_{j-1} + 1, \ldots, t^*_j$$

and, again between break points, the seasonal component is fixed,

$$S_t = \sum_{i=1}^{s-1} \delta_{i,j} D_{i,t} \qquad t = t_{j-1}^* + 1, \ldots, t_j^*$$

where $D_{i,t}$ are seasonal dummies, equal to one if time t is "season" i and zero otherwise. We use the convention that $t_0^* = 0$ and $t_{m+1}^* = T$ (Bai & Perron, 1998). Under these assumptions, we note that for daily or monthly data (with $s = 7$ and $s = 12$ respectively), and for quarterly data (with $s = 4$) to a lesser extent, the trend component will be parsimonious relative to the seasonal component.

Bai & Perron's (1998, 2003) methodology offers two alternatives for estimating the unknown break points, t_j^* ($j = 1, \ldots, m$), in such a model. The first is that the coefficients of one component are fixed over the entire sample period (a partial structural change model); the second is that parameters in both components should have the same break points (a pure structural change model). We demonstrate below that neither of these options is satisfactory for the type of data examined in this paper, i.e. seasonal time series with evolving trends, and large s (in this case, $s = 12$).

When considering trend extraction and assuming that structural breaks will be required, in general we wish to allow break points in the seasonal component, which is inconsistent with a partial structural change model. Conversely, we would not necessarily wish to constrain any seasonal break points to occur at the same places as the trend break points, as required in a pure structural change model. On the face of it, this requirement is not necessarily restrictive, since the parameter estimates of one component are not forced to change from one segment of the data to the next. However, when selecting the optimal number of break points using a penalised likelihood criterion, e.g. BIC, this compromises the method's ability to select break points in the data, i.e. the estimated number of breaks may be too low. One example of where these issues may be important is in arrivals from Australia. As noted in Section 2, the Australian arrivals seem to have a seasonal break point in 1987 (changing from three peaks to four), with no apparent change in trend.

To address this concern we estimate the trend and seasonal components separately, using a new iterative approach motivated by the Macaulay cycle seasonal decomposition method (Macaulay, 1931), e.g. as implemented in STL using modern weighted averages. This allows more flexible structural break estimation than fitting both components simultaneously. As above, we assume that the time series can be decomposed into a piecewise linear time trend and a piecewise constant seasonal pattern. Each component is then estimated using the methodology of Bai & Perron (1998, 2003), implemented in R (R Development Core Team, 2008) using the strucchange package of Zeileis *et al.* (2002, 2003). We employ the default method of selecting the number of breaks, which uses BIC.

The trend of the data Y_t is estimated using a piecewise linear model for the seasonally adjusted time series $V_t = Y_t - \hat{S}_t$, i.e.,

$$V_t = \alpha_j + \beta_j t + \epsilon_t \qquad t = t^*_{j-1} + 1, \ldots, t^*_j$$

for $j = 1, \ldots, m+1$, where ϵ_t is a zero-mean disturbance and t^*_j, $j = 1, \ldots, m$, are the unknown trend break points. For the first iteration, we set $\hat{S}_t = 0$ for all t.

Once the trend has been estimated, we estimate the seasonal component of Y_t using a piecewise seasonal dummy model for the detrended data $W_t = Y_t - \hat{T}_t$, i.e.,

$$W_t = \delta_{0,j} + \sum_{i=1}^{s-1} \delta_{i,j} D_{i,t} + \nu_t \qquad t = t'_{j-1} + 1, \ldots, t'_j$$

for $j = 1, \ldots, m'+1$, where $D_{i,t}$ are the seasonal dummies, ν_t is a zero-mean disturbance and t'_j, $j = 1, \ldots, m'$, are the unknown seasonal break points. As before, we take $t'_0 = 0$ and $t'_{m'+1} = T$. The estimates $\hat{\delta}_{i,j}$ are adjusted at the end of each iteration so that they add to zero within each full seasonal cycle (between seasonal breaks), to prevent any change in trend appearing as a result of a seasonal break happening 'mid-year'. That is,

$$\sum_{i=0}^{s-1} \hat{\delta}_{i,j} = 0 \qquad \text{for all } j.$$

This estimation process is then iterated to convergence of the estimated break points.

We are thus able to estimate a trend which, due to its parsimonious representation, is able to react to obvious shifts in the general movement of the data. If required, we are able to identify important changes in the seasonal pattern separately. Since the trend and seasonal break points, t^*_j and t'_j respectively, are estimated independently, they are not constrained to coincide. Of course, this does not preclude (some) trend and seasonal break points coinciding if appropriate. In all data analysis and simulations we have followed the recommendations of Bai & Perron (2003) and Zeileis *et al.* (2003), concerning the fraction of data needed between breaks. For monthly seasonal data, we used three full years (36 observations) as a minimum, corresponding to 12% of a 25 year data span. However, a further consequence of the iterative estimation of trend and seasonal breaks is that while any two breaks of the same type must have a minimum separation (three years here), the distance between a trend break and a seasonal break has no constraints. This feature is of practical importance, e.g. as shown for the arrivals data in Section 4, and is another desirable feature of the new iterative approach.

Table 2. Basic structure of each of the data generating processes used in the simulation study.

Model	Trend component	Seasonal component	# of breaks
1	No breaks	No breaks	0
2	Break at $\frac{T}{2}$	No breaks	1
3	Break at $\frac{T}{2}$	Break at $\frac{T}{2}$	1
4	Break at $\frac{T}{3}$	Break at $\frac{2T}{3}$	2
5	Breaks at $\frac{T}{3}$ and $\frac{2T}{3}$	Break at $\frac{2T}{3}$	2
6	Breaks at $\frac{T}{4}$ and $\frac{3T}{4}$	Break at $\frac{T}{2}$	3
7	Breaks at $\frac{T}{4}$, $\frac{T}{2}$ and $\frac{3T}{4}$	Break at $\frac{T}{2}$	3
8	Breaks at $\frac{T}{4}$, $\frac{T}{2}$ and $\frac{3T}{4}$	Break at $\frac{3T}{8}$	4
9	Breaks at $\frac{T}{4}$, $\frac{T}{2}$ and $\frac{3T}{4}$	Break at $\frac{T}{2}$	3
10	Breaks at $\frac{T}{4}$, $\frac{T}{2}$ and $\frac{3T}{4}$	Break at $\frac{3T}{8}$	4

The importance of this method is now illustrated using simulated data. We considered 20 different data generating processes, as described in Table 2. For each of two slope coefficients (either 0.05, or 0.1), the trend components are ramp-like, and alternate between the positive slope and a zero slope. They are continuous, except models 9 and 10 where the trend in the second half of the data is an exact replication of the trend in the first half, requiring a discontinuity at $\frac{T}{2}$. The seasonal component is piecewise constant. The two seasonal cycles used in models 3 to 10 are shown in Figure 3, and are identical except for the ordering of the Jan/Feb, Mar/Apr, Jul/Aug and Sep/Oct values. In models 1 and 2, only the first cycle is used.

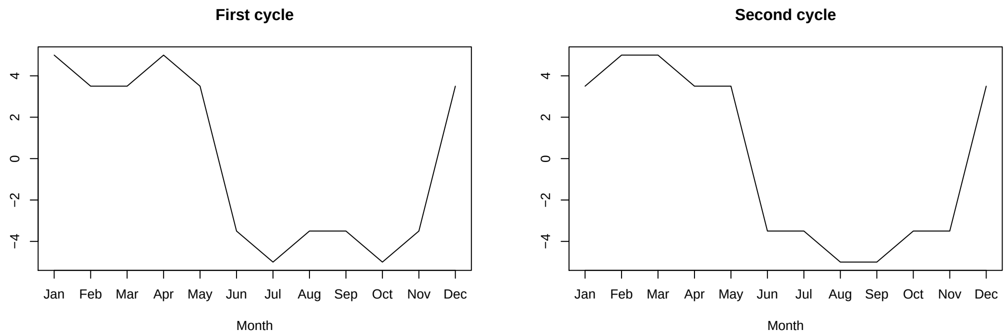

Figure 3. The two seasonal cycles used for the seasonal component in the illustrative simulations.

The data are given by

$$Y_t = T_t + S_t + I_t \qquad t = 1, \ldots, T$$

where $I_t \sim$ i.i.d. $N(0, 1)$. The number of break points in the data for each data generating process is given in the final column of Table 2. In the case of models 1, 2, 4, 6, 8 and 10, the total number of break points is equal to $m + m'$, i.e. the number of trend break points plus the number of seasonal break points. In models 3, 5, 7 and 9, the seasonal break point coincides with a trend break point, so the total number of break points is $m + m' - 1$.

For each of models 1 to 10, we simulated $N = 10000$ series of length $T = 288$ for the two different trend slopes. For each of these series, we estimated the trend and seasonal components simultaneously using the Bai & Perron (1998, 2003) approach – the "complete approach". In particular, we restricted the parameters in both the trend and seasonal components to change simultaneously; i.e., we fitted a pure structural change model. Between breaks, the constant term in the estimated trend was corrected so that the seasonal component added to zero. We also applied the new iterated methodology to the same series, fitting the trend and seasonal components separately.

In a decomposition exercise, we wish to estimate the components well, and the break points are the means to this end. While we will comment on the distribution of the estimated break points (number, and position), overall decomposition quality will be established via mean squared errors for the components and observations. Note that in a forecasting exercise, the focus will typically be on out-of-sample performance, and this will be based on the position of the final break point, and the quality of the trend and seasonal components beyond this break point.

Table 3 presents summary information on the numbers of estimated break points, i.e., the sample distribution, by model and estimation technique. The root mean squared errors reported are based on the average squared deviations (across series) of the estimated number of break points from the actual number (M and m for the complete and iterated approaches respectively). In the case of the complete approach and the arguments presented above, we expected that the estimated number of break points would be too small, and this is exactly what the top panel in the table confirms. While the complete approach performs very well in cases where the trend and seasonal break points coincide (or do not exist), and performance improves when the trend slope increases, it performs very poorly when the seasonal component does not have a coinciding trend break point, namely, models 4, 6, 8 and 10. For no series is the estimated number of break points higher than the true number, so there is a downward bias in the estimated number of break points when fitting the two components simultaneously.

In contrast, the iterated approach does a much better job of estimating the number of break points. While it never estimates the number correctly for all 10000 series within a model (unlike the complete approach for models 1, 2 and 3 for both slopes, and models 5 and 9 for the higher slope), the root mean squared errors are typically very close to zero, and usually much smaller than the corresponding RMSEs for the complete approach. With the exception of models 7 and

Table 3. Numbers of estimated break points using the complete approach (in the top panel) and for the trend component using the iterated approach (in the bottom panel). M is the total number of break points in the data, and m is the number of trend break points. On the left are the results when the trend slope is 0.05, on the right, when the slope is 0.1. The number of series where the number of break points is correctly estimated is shown in bold. The RMSE figures are for the number of break points.

		0	1	2	3	4	RMSE	0	1	2	3	4	RMSE
Model	M	Complete, slope 0.05						Complete, slope 0.1					
1	0	**10000**					0	**10000**					0
2	1		**10000**				0		**10000**				0
3	1		**10000**				0		**10000**				0
4	2		5324	**4676**			0.73		3831	**6169**			0.62
5	2	1	2034	**7965**			0.45			**10000**			0
6	3	2	8974	978	**46**		1.92			7267	**2733**		0.85
7	3	1482	7839	620	**59**		2.13	1		362	**9637**		0.19
8	4	370	9145	485	0	**0**	3.00		15	6420	3555	**10**	1.71
9	3		9490	381	**129**		1.96				**10000**		0
10	4		6500	3008	492	**0**	2.67				9932	**68**	1.00
	m	Iterated, slope 0.05						Iterated, slope 0.1					
1	0	**9889**	107	4			0.11	**9889**	107	4			0.11
2	1		**9832**	164	4		0.13		**9824**	173	3		0.14
3	1		**9779**	214	7		0.16		**9769**	226	5		0.16
4	1		**9801**	193	6		0.15		**9799**	196	5		0.15
5	2		31	**9876**	101	1	0.12			**9805**	194	1	0.14
6	2			**9888**	112		0.11			**9799**	199	2	0.14
7	3	3	24	5869	**4100**	4	0.77			2	**9839**	105	0.10
8	3	5	27	5717	**4244**	7	0.77			1	**9868**	131	0.11
9	3		2	51	**9918**	29	0.09				**9911**	89	0.09
10	3		2	72	**9905**	21	0.10				**9917**	83	0.09

8, the iterated approach is much less sensitive to the slope in the trend component. We conducted an additional simulation with no seasonal component, and are able to attribute the poor performance to the "difficult" specification of the trend in this particular instance. The complete approach also performs poorly here; in fact much worse than the iterated approach. The two exceptions aside, sample distributions and RMSEs for the iterated approach are generally very similar across trend slopes.

Of particular note are the similarities (for a fixed trend slope) when comparing model 7 to model 8, and when comparing model 9 to model 10. In both of these cases, the position of the seasonal break shifts from a point where it coincides with a trend break, to one where it does not. This harms the complete approach estimates dramatically, but as we might expect, has little or no effect on trend break estimates from the iterated approach.

The estimated seasonal break points from the iterated approach are shown in Table 4 (recall the seasonal break points from the complete approach coincide with the trend breaks in Table 3). For the iterated approach there were no "false positives" in Models 1 and 2, with zero break

Table 4. Numbers of estimated seasonal break points using the iterated approach. On the left are the results when the trend slope is 0.05, on the right, when the slope is 0.1. The number of series where the number of break points is correctly estimated is shown in bold (always more than 95%). The RMSE figures are for the number of break points.

	Slope 0.05			Slope 0.1		
Model	0	1	RMSE	0	1	RMSE
1	**10000**		0	**10000**		0
2	**10000**		0	**10000**		0
3	153	**9847**	0.12	154	**9846**	0.12
4	449	**9551**	0.21	467	**9533**	0.22
5	438	**9562**	0.21	455	**9545**	0.21
6	128	**9872**	0.11	127	**9873**	0.11
7	180	**9820**	0.13	124	**9876**	0.11
8	333	**9667**	0.18	235	**9765**	0.15
9	99	**9901**	0.10	116	**9884**	0.11
10	217	**9783**	0.15	242	**9758**	0.16

points being estimated in every case. In models 3 to 10, only small numbers of series had no seasonal break point estimated, with the correct number identified in over 95% of cases. The quality of estimation appears to be independent of the slope in the trend components, but somewhat related to the exact specification of the trend. When identified, the position of the seasonal break point was very precisely estimated.

Summarising the position of the estimated break points is difficult for the complete approach, and particularly when the estimated number of break points is too low. When the estimated number of break points is correct, we could calculate bias, and RMSE for the estimated dates. However, it is difficult to provide a fair comparison on this basis without incorporating a penalty for the number of series correctly estimated. In addition, while it is clear exactly which breaks (trend and/or seasonal) should be estimated using the iterated approach, it is less clear for the complete approach.

By way of illustration, we present the results for model 9 with a trend slope of 0.05 in Figure 4. The complete approach favours a single estimated break point coinciding very precisely with the combined seasonal and trend break point at $\frac{T}{2}$. This accounts for 94.9% of all cases, and underestimates the number of break points by 2. When two break points are estimated (in 3.81% of the cases), it is not clear which of the true break points are estimated, but given the proximity constraint, it is likely that the two will comprise a very precise estimate of the middle break point (at least 75% of the cases) and an imprecise estimate of one or other of the trend break points. In a vast minority of cases (1.29% of the series), the correct number of break points are estimated in the complete approach, and these appear to be unbiased for the true break points, with varying degrees of precision. In contrast, the iterated approach favours the

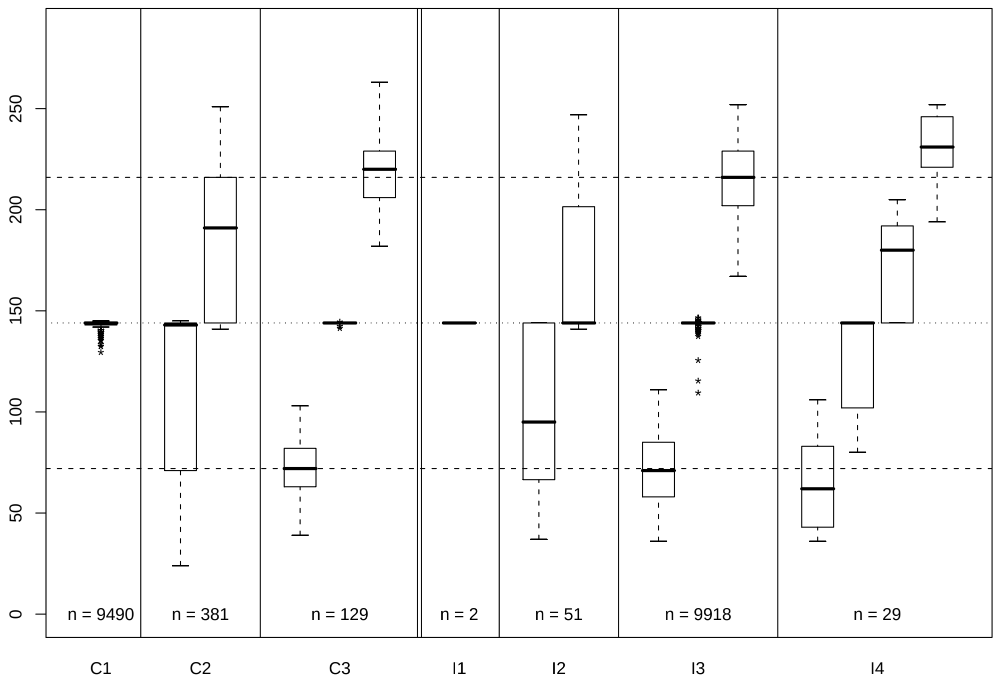

Figure 4. The estimated break points for model 9, with a trend slope of 0.05. The seasonal break point and contemporaneous discontinuous trend break point are shown by the dotted line, and the two (continuous) trend break points are shown by the horizontal dashed lines. To the left of the double vertical lines are the estimated break points from the complete approach. The panels are labeled C for complete approach or I for iterated approach, followed by the number of breaks estimated. Each boxplot corresponds to the sample distribution of a single estimated break point. The number of instances in each case (also shown in Table 3) is listed along the bottom of the plot.

correct number of estimated break points in 99.18% of the series. These appear to be unbiased for the true break points, with the trend plus seasonal break point being estimated much more precisely than the trend-only break points.

Table 5 gives a measure of the quality of the estimated components, and the predictions of the data, for the simulated series. Namely, we present "RMSEs" via the square root of the median (over the 10000 series) of the following statistics: for the trend component, $\text{MSE}_T = \frac{1}{T}\sum_t (T_t - \hat{T}_t)^2$, the seasonal component, $\text{MSE}_S = \frac{1}{T}\sum_t (S_t - \hat{S}_t)^2$, the trend plus seasonal, $\text{MSE}_{T+S} = \frac{1}{T}\sum_t (T_t + S_t - \hat{T}_t - \hat{S}_t)^2$, and the data $\text{MSE}_Y = \frac{1}{T}\sum_t (Y_t - \hat{T}_t - \hat{S}_t)^2$.

Estimation of the trend worsens in both approaches as the complexity of the data generating process increases. In the case of the complete approach, a marked shift occurs from model 5 to model 6, where there are three break points, but only one component changes at each, in contrast to two break points with a common shift in trend and seasonal. In contrast, MSE_T for the iterated approach appears insensitive to the placement of the seasonal break, i.e. the figures are almost identical for models 5 and 6, models 7 and 8, and models 9 and 10. Increasing the trend slope tends to improve the MSEs for the complete approach, the exception being model 8. A similar, though typically less dramatic improvement, is also a feature of the iterated

Table 5. Square root of the median MSE for estimated components. On the left are the results for the lower trend slope, on the right, the results for the higher slope. In the top panel are the results for the complete approach, in the bottom panel the results for the iterated approach. The MSEs are averaged over the 288 observations in the series, and the median is calculated over the 10000 series.

	T_t	S_t	$T_t + S_t$	Y_t	T_t	S_t	$T_t + S_t$	Y_t
Model	Complete, slope 0.05				Complete, slope 0.1			
1	1.18	3.22	3.52	16.57	1.18	3.22	3.52	16.57
2	2.46	4.85	5.54	16.05	2.38	4.77	5.42	16.08
3	1.89	4.92	5.32	16.12	1.92	4.90	5.30	16.12
4	3.26	7.48	9.32	16.67	2.80	6.89	7.59	16.01
5	3.28	6.33	7.23	15.68	2.92	6.10	6.81	15.60
6	9.01	5.48	10.40	18.08	3.82	8.81	9.66	16.54
7	8.89	6.24	10.67	18.26	3.79	7.23	8.16	15.04
8	7.70	5.22	9.27	17.66	9.58	7.16	11.26	17.04
9	9.07	4.71	10.25	18.31	3.66	6.96	7.90	15.09
10	8.97	7.67	11.86	18.77	4.19	8.22	9.34	15.77
	Iterated, slope 0.05				Iterated, slope 0.1			
1	1.19	3.22	3.53	16.57	1.19	3.22	3.53	16.57
2	2.53	3.22	4.20	16.41	2.43	3.22	4.14	16.43
3	2.54	4.95	5.64	16.03	2.45	4.96	5.61	16.04
4	2.56	4.98	5.69	16.04	2.44	4.98	5.64	16.05
5	3.67	4.99	6.28	15.87	3.42	4.98	6.12	15.90
6	3.62	4.95	6.21	15.86	3.40	4.95	6.09	15.89
7	5.27	4.99	7.19	15.89	4.26	4.95	6.60	15.73
8	5.25	5.01	7.18	15.87	4.27	4.99	6.63	15.72
9	4.06	4.95	6.45	15.79	3.72	4.95	6.26	15.83
10	4.05	4.98	6.48	15.78	3.73	4.98	6.28	15.83

approach.

Estimation of the seasonal component using the complete approach seems very dependent on the exact specification of the data generating process. The RMSEs themselves are quite variable, and do not seem to have systematic patterns with the changing models, nor with the increase in the trend slope. In contrast, the iterated approach yields seasonal patterns which are estimated in a way that is insensitive to the trend component. In the case where the seasonal component does have a break point (models 3 to 10), the RMSEs are virtually identical across both model and slope. In all but model 1, where the RMSEs are identical, the seasonal component is better estimated using the iterated approach than the complete approach.

Similar variation in the quality of estimation of the "signal" $T_t + S_t$ is seen for the complete approach as in the individual components. There do not appear to be systematic changes across models and trend slope. With the exception of models 1 and 3, the iterated approach always does better than the complete approach. The iterated figures appear to get larger from models 1 to 8, with model 9 and 10 providing an easier challenge than the others. As the slope increases, the trend component is better estimated, and this results in better estimation of the complete signal.

Finally, we look at the prediction errors for the actual observations. Here, we see that the complete approach actually leads to comparable performance against the iterated approach, even when the individual components, or the signal, are not well estimated. As trend slope increases, the complete approach tends to improve, whereas it makes little difference to the iterated approach. Indeed, the iterated approach appears to have a virtually constant RMSE across models, whereas the complete approach has some variation.

Table 6 perhaps gives the strongest indication that the complete approach is failing to fit the data well. While the prediction errors are similar in size to those from the iterated approach (as seen by the summary of MSE_Y in Table 5), there remains significant structure in those errors. Ljung-Box tests were conducted at lag 20, and the p-values of the tests are collated in Table 6. If the residuals were indeed uncorrelated, only 5% of these p-values should be below 5%. Reported are the actual percentages across the 10000 simulated series, by model and estimation method.

Table 6. Percentage of the 10000 series with residual autocorrelation significant at the 5% level, based on a Ljung-Box test up to lag 20.

Model	Complete, 0.05	Complete, 0.1	Iterated, 0.05	Iterated, 0.1
1	9.1	9.1	9.0	9.0
2	16.6	16.5	10.3	10.1
3	16.4	16.3	17.7	17.7
4	52.6	49.5	19.3	19.4
5	43.3	30.9	22.3	22.2
6	94.0	55.3	20.0	19.9
7	97.0	54.1	19.1	22.9
8	95.6	78.8	21.1	23.7
9	98.9	52.2	23.4	22.7
10	95.4	45.5	24.0	23.5

The results show that for the models with larger numbers of breaks and the lower trend slope coefficient, the complete approach is leaving significant autocorrelation in the residuals in almost all cases. This reduces somewhat for the larger slope, but the proportions are still very large. The iterated approach also has more rejections than would be expected under the null; however, these are typically much less frequent than in the complete approach. In addition, this particular aspect of quality seems largely independent of the trend slope when the two components are estimated iteratively.

Overall, this simulation shows the undesirable consequences of fitting two components simultaneously when a parameter-rich seasonal component breaks at times other than those of a relatively parsimonious trend component. Our new iterated approach to fitting such components largely addresses this concern, and generally provides much better estimation of individual components. Next we apply this iterated technique to the arrivals data.

4 Modelling the arrivals using an iterated approach

The seasonal variation of the arrivals typically increases with the level of the series (Figure 1). Applying the new iterated approach directly to the untransformed data would certainly require seasonal breaks to account for the changes in amplitude of the seasonal component. This is clearly undesirable because such changes usually evolve smoothly, so should not be modelled as abrupt changes. Consequently a stabilising transformation is needed.

A log transformation is one obvious possibility, but this does not yield an optimal stabilising transformation for all these series and instead we estimated a power transformation, identified using the robust spread-vs-level plots described in Hoaglin *et al.* (1983). For each individual series we calculated the median and interquartile range (IQR) of the monthly arrivals for each of the 25 calendar years, then regressed log IQR on log median. The appropriate stabilising transformation is $x^{1-\text{slope}}$, and the transformed series are shown in Figure 5, with the estimated powers.

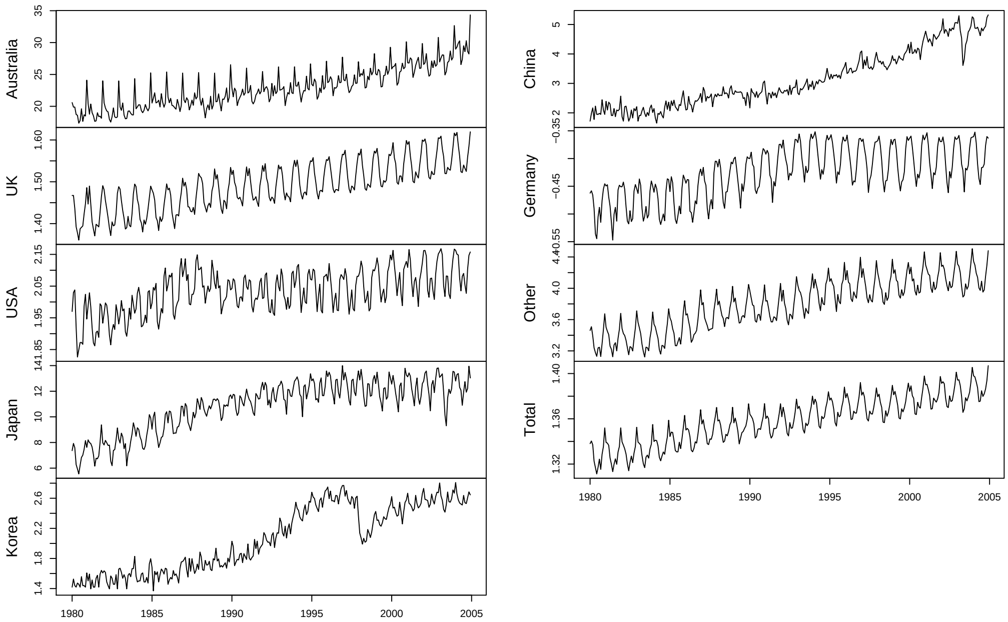

Figure 5. Power transformed monthly short term visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The power transformations are: Australia 0.3, UK 0.05, USA 0.08, Japan 0.27, Korea 0.11, China 0.18, Germany -0.11, Other 0.13, and Total 0.03.

Confidence intervals for the slopes in these spread-vs-level regressions support the use of logs only in the case of the UK, USA and Total arrivals (i.e. a power of zero, or a slope of one). In the case of Germany the estimated power is negative, so $-x^{1-\text{slope}}$ is used to preserve order in the transformed arrivals. All further analysis is conducted on the transformed data.

In the case of the transformed arrivals data, each linear time trend requires two parameters, and each dummy seasonal an additional $s - 1 = 11$. Figure 5 indicates that for most series a linear time trend would need breaks. Further, while the seasonal patterns generally have constant variation over the length of the series due to the power transformations, we do not wish to preclude seasonal changes during the data period. As the simulation study demonstrated, the parameter-rich trend-plus-seasonal (complete) model would severely limit our ability to appropriately fit the data, since the large number of seasonal dummies would reduce the possible number of breaks, especially when selected by BIC.

As with the simulated data, we used a minimum period between breaks of 36 observations for estimation of both trend and seasonal components. In fact, when estimating the trend and seasonal components iteratively, there is scope to reduce that minimum period for estimation of the trend component, since it only requires two parameters between breaks. This possibility further increases the flexibility of trends estimated using our new iterated approach. However, to simplify comparisons we have not pursued this option here. For the iterative approach, three iterations were sufficient to ensure convergence of the estimated break points in all cases but Other and Total, which each required four.

The estimated trend break points are shown in Table 7 along with estimated 95% confidence intervals. The confidence intervals were formed with heteroscedasticity and autocorrelation consistent (HAC) estimates of the covariance matrix (Andrews, 1991). These confidence intervals are computed and displayed as standard output using the R package sandwich; they make use of a quadratic spectral kernel with vector autoregressive prewhitening, as recommended by Andrews & Monahan (1992). Details of the R implementation are given in Zeileis (2004, 2006). Figure 6 displays the estimated parametric trends and break points (with confidence intervals), along with nonparametric trends estimated by STL. September 2001 is included in only two confidence intervals, indicating the possibility that the terrorist events of 9/11 may be linked to a structural break in the trend of arrivals for those two origins: Australia and Other. Other is difficult to interpret given its composite nature, although it is plausible that the 9/11 events did have an effect on tourist behaviour in some of these countries. An alternative (or perhaps complementary) explanation is discussed in Section 5.

In the case of Australia, a break is estimated in the month following 9/11, which results in an increased trend slope but a decreased intercept. A relevant confounding effect is the collapse of the airline Ansett Australia, which occurred just three days after the terrorist attacks of 9/11; hence it is impossible to separate these two effects with monthly data. The termination of flights by Ansett Australia and Ansett International on 14 September 2001 certainly affected capacity

Table 7. Estimated trend break points for the transformed monthly visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The middle column gives the estimated break points, while the first and third columns give the lower and upper 95% confidence limits respectively, estimated using a HAC estimate of the covariance matrix.

	Australia	
1984(5)	1985(1)	1985(2)
1989(3)	1989(4)	1990(5)
1997(10)	1997(12)	1998(1)
2001(1)	2001(10)	2001(11)
	UK	
1985(10)	1986(1)	1986(4)
1990(7)	1990(8)	1996(5)
	USA	
1982(12)	1983(3)	1986(12)
1988(9)	1988(10)	1990(1)
1998(6)	1998(8)	2001(6)
	Japan	
1987(3)	1987(6)	1987(8)
1996(2)	1996(8)	1996(10)
	Korea	
1982(8)	1983(12)	1984(4)
1990(9)	1990(10)	1990(11)
1994(9)	1994(11)	1994(12)
1997(10)	1997(11)	1997(12)
2000(10)	2000(11)	2001(1)

	China	
1984(7)	1984(8)	1985(1)
1988(10)	1989(7)	1989(9)
1997(1)	2000(11)	2000(12)
	Germany	
1986(6)	1986(7)	1986(11)
1994(5)	1994(6)	1994(7)
1999(6)	1999(8)	2000(11)
	Other	
1983(1)	1983(3)	1983(4)
1985(6)	1986(8)	1986(9)
1990(7)	1990(10)	1990(12)
1992(11)	1994(1)	1994(3)
1997(3)	1997(6)	1997(8)
2001(4)	2001(7)	2001(9)
	Total	
1982(12)	1983(1)	1983(3)
1987(10)	1987(12)	1988(4)
1989(8)	1990(12)	1991(4)
1997(1)	1997(3)	1997(6)

and timing of arrivals to New Zealand. In addition, in the following week, strike action targeted at Air New Zealand occurred at Melbourne and Perth airports (Air New Zealand had acquired control of Ansett Australia during the year preceding its collapse). Those strikes required the cancellation of all Air New Zealand trans-Tasman flights operating from Melbourne and Perth. These physical constraints on passenger numbers are a plausible explanation for the observed decrease in intercept, while the subsequent increase in the rate of arrivals from New Zealand's nearest neighbour is unlikely to have any causal links from the terrorist events of September 2001.

Focusing on Figure 6 more generally, we note that it is often difficult to distinguish between the two alternative trend estimates; i.e. those from our iterated approach and STL. In particular, the iterative parametric method achieves similar flexibility in its trend estimate to the nonparametric technique, with the latter essentially fitting linear time trends at each point in the series using only a local window of observations to estimate parameters. The break point technology allows instantaneous changes in the trend however, unlike the STL technique. In effect, STL is requiring an 'innovational outlier' approach to any structural changes in the data, while our parametric procedure models the changes directly and permits an 'additive outlier'

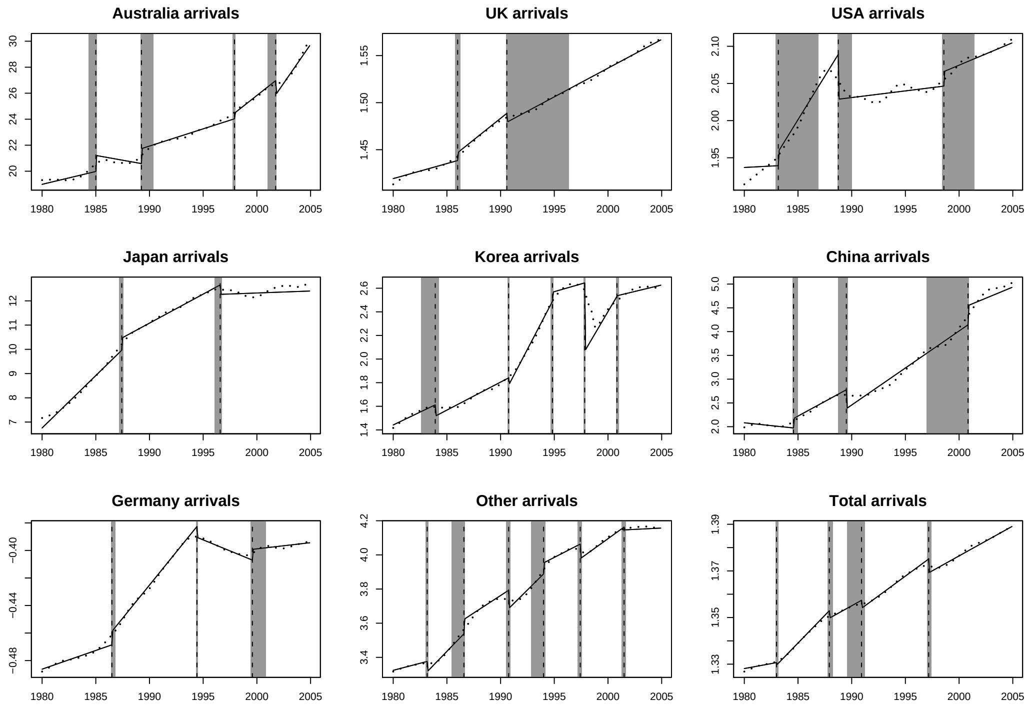

Figure 6. Estimated trends and trend break points for the transformed monthly visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The solid line is the piecewise linear time trend, while the dotted line is the estimated STL trend. The vertical dashed lines and grey regions respectively indicate the fitted break points and their 95% confidence intervals, estimated using a HAC estimate of the covariance matrix.

approach. (In a series of papers, Perron and coauthors popularised the use of these 'outlier' terms, to describe an approach which is attributed to the intervention analysis work of Box & Tiao (1975).) An obvious contrast between the two approaches is seen in the Korean data at the time of the Asian financial crisis. The parametric break point is dated at November 1997 (with a narrow 95% confidence interval of October to December), which corresponds exactly to the month that the financial crisis first affected Korea (Kaminsky & Schmukler, 1999). However STL spreads the downward impact of the crisis over a number of months, in contrast to the observed behaviour.

Table 8 gives the estimated seasonal break points for the transformed arrivals, with the estimated seasonal components shown in Figure 7. Korea, China, Germany and Other have no estimated seasonal break points. As the power transformations have effectively stabilised the seasonal variation, any changes in the seasonal patterns more likely reflect behavioural changes in the time of year when visitors arrive. For example, in Australia's seasonal pattern the 'middle' peak has moved and one extra peak has been added, reflecting a shift from a three-term school year to a four-term year in New South Wales in 1987 (NSW Department of Education, 1985). The placement of the seasonal break point coincides exactly with the final month under the old three term system, with the first holiday in the new sequence occurring in July 1987. The UK data show a shift in arrivals from the second half of the year to the first and a shift in the

Table 8. Estimated seasonal break points for the transformed monthly visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The middle column gives the estimated break points, while the first and third columns give the lower and upper 95% confidence limits respectively, estimated using a HAC estimate of the covariance matrix. Korea, China, Germany and Other have no estimated seasonal break points.

Origin	Point estimate and 95% CI		
Australia	1987(1)	1987(6)	1987(9)
UK	1985(11)	1986(6)	1987(9)
USA	1995(1)	1995(4)	1995(12)
Japan	1987(10)	1988(6)	1988(12)
Total	1987(3)	1987(7)	1988(1)

peak arrivals from December to February. The USA and Japanese arrivals have had relatively complex changes, while the Total series has seen most change in the winter months. This shift reflects the Australian change in structure, and occurs at the same time. Note here the practical relevance of allowing breaks in the seasonal component of any given series to be independent of those in the trend, with no minimum separation between them: four of these five seasonal breaks (all except USA) are less than three years away from at least one corresponding trend break (see Tables 7 and 8). However, none of the dates for trend and seasonal breaks coincide in any given series, which reinforces the need to allow the components to break separately for additional flexibility in the fitted model.

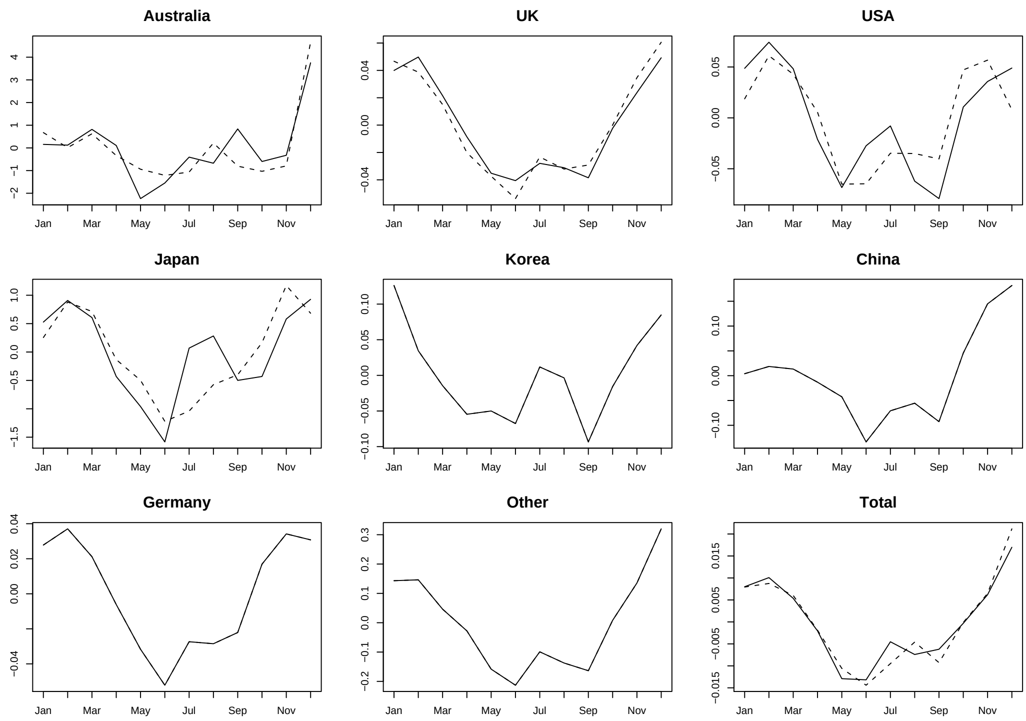

Figure 7. The estimated seasonal components for visitors to New Zealand by origin. The solid line is the final estimated seasonal component; it is the only estimate in four of the nine cases, where no seasonal breaks were detected. The five dashed lines are the seasonal components prior to the seasonal break points listed in Table 8.

To conclude this section, we compare the trend estimates obtained from our new iterated approach to the trends obtained fitting a complete structural break model (with 13 parameters between breaks), and using STL. In Figure 8 we present trends for the Korean arrivals and those from Other origins. We also show sample autocorrelation functions for the three sets of residuals from each series. The trends are all similar, but the agreement is closest for the iterated approach and STL. Some differences are evident particularly at the end of the series though, which would be important for prediction. For Other arrivals, the number of parameters required for the complete model clearly restricts the estimated number of breaks, leading to greater departures from the STL trend than achieved by iteration. The irregular components also favour the iterated approach over STL and the complete model, as the residuals for the latter are highly autocorrelated, especially at low lags (see Figure 8). In contrast, the residuals of the iterated method exhibit far less autocorrelation, indicating a better overall decomposition; this is as expected, following the simulation evidence presented in Table 6.

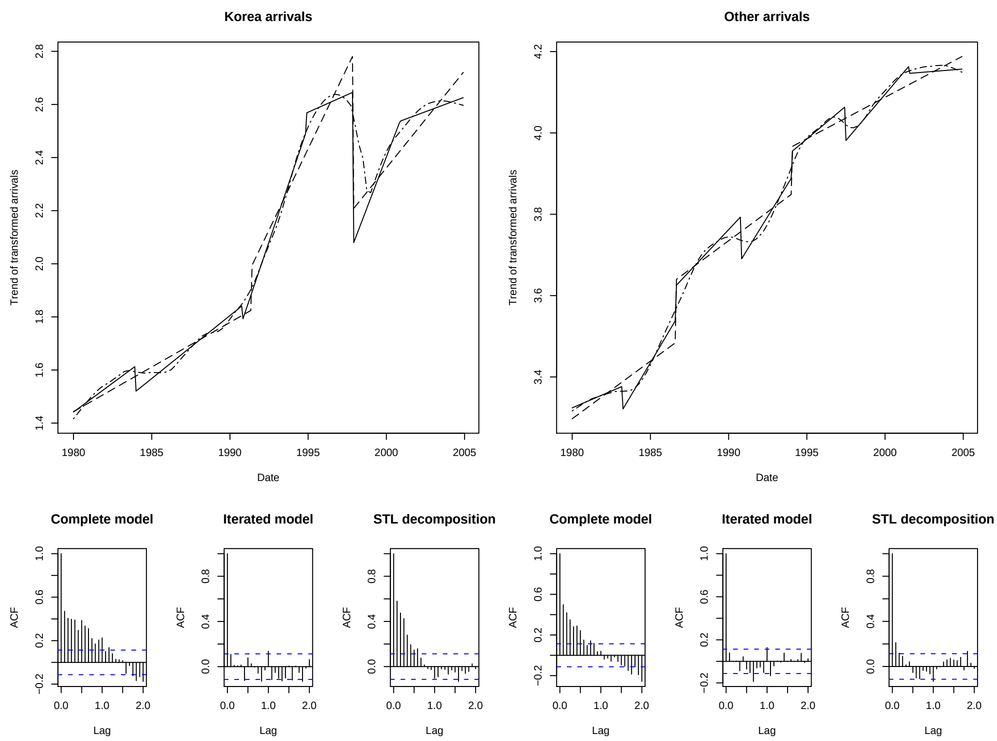

Figure 8. Trend estimates for the transformed Korean arrivals and those from Other origins. The trend estimates are based on the complete model (dashed), the new iterated approach (solid) and STL (dot-dashed). Also shown are sample autocorrelation functions for the residuals from the three methods.

5 Discussion

The growth in the number of visitor arrivals to New Zealand was lower than expected in late 2001 (e.g., by the New Zealand Ministry of Tourism, as noted in Haywood and Randal, 2004), yet there is no conclusive evidence to attribute this forecast error solely to the terrorist events of 9/11. The termination of flights by Ansett Australia on 14 September 2001 certainly affected capacity and timing of arrivals from Australia to New Zealand, and that would have affected Total arrivals in September 2001 somewhat as well. Indeed Australia is the only (individual) country of origin with a structural change in trend identified close to 9/11. The subsequent rate of Australian arrivals to New Zealand in fact shows an *increase*, following an initial drop which is plausibly explained by the Ansett effect; see Table 7 and Figure 6.

A further plausible cause for the lower than forecast number of visitors is the US recession dated March 2001 (Hall *et al.*, 2001), along with the world wide flow-on effects from a slow down in the US economy. The recession predates 9/11 by six months but that is consistent with observed features of the data. In particular, March 2001 corresponds exactly to the minimum in the second difference of an STL trend of Total monthly (log) arrivals, indicating a maximum decrease in the slope at that time. It is possible that the slow down seen in the Other (composite) arrivals series, dated July 2001, may be due in part to the flow-on effects from this US recession.

It seems quite clear that the events of 9/11 did not have much influence on the longer term numbers of visitors to New Zealand, and especially not a negative influence. In contrast our analysis suggests other events which have had marked structural effects on these data, especially from certain countries of origin (refer to Tables 7 and 8, and Figures 6 and 7). In particular and as already discussed in Section 4, the Asian financial crisis of 1997-1998 precipitated a massive drop in arrivals from Korea, and the estimated intercept and slope of Total arrivals both decreased too, in 1997. Again as noted above, the estimated change in the Australian seasonal pattern is explained perfectly (in both timing and effect) by the 1987 switch from a three-term to a four-term school year in New South Wales. The stock market crash of October 1987 preceded a dramatic decline in arrivals from the USA, followed by a sustained period of only moderate growth. Further, both the intercept and slope of Total arrivals decreased late in 1987. In contrast, the SARS epidemic affected arrivals from China in a different way, with a very short-lived but large reduction, which we class as temporary and not structural. The overall effects of 9/11 might also be seen as temporary and negative, but of a smaller magnitude than those associated with SARS.

Estimation of structural breaks was facilitated by a new implementation of Bai & Perron's (1998, 2003) work that is recommended for seasonal data. Specifically, use of an iterative approach to estimate the trend and seasonal components separately enabled us to locate structural breaks in the data, and to attribute these to either changes in the trend or the seasonal pattern. Estimating these components simultaneously did not achieve the same flexibility in the estimated components, nor in the location of the break points. The agreement between the estimated parametric trends from the iterated approach and the nonparametric STL trends is especially pleasing, as is the lack of residual structure around those parametric trends when compared to other trend estimates.

Acknowledgments

Statistics New Zealand kindly supplied the data. We gratefully acknowledge the assistance of Ray Brownrigg with implementing the simulation on MSOR's computation grid. We thank those who commented on presentations at Statistics New Zealand, the Reserve Bank of New Zealand, Victoria Management School, the ASC/NZSA 2006 Conference, the TSEFAR 2006 Conference, and the WCDM07 Conference. We also thank Peter Thomson for some helpful suggestions that improved the paper.

References

Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59,** 817–858.

Andrews, D.W.K. & Monahan, J.C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, **60**, 953–966.

Australian Bureau of Statistics (2003). Information Paper: A Guide to Interpreting Time Series – Monitoring Trends, ABS Cat. No. 1349.0. Australian Bureau of Statistics: Canberra.

Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* **66,** 47–78.

Bai, J. & Perron, P. (2003). Computation and analysis of multiple structural change models. *J. Appl. Econometrics* **18,** 1–22.

Box, G.E.P. & Jenkins, G.M. (1976) *Time Series Analysis, Forecasting and Control*, Holden-Day: Oakland, 2nd edition.

Box, G.E.P. & Tiao, G.C. (1975). Intervention analysis with applications to economic and environmental problems. *J. Amer. Statist. Assoc.* **70,** 70–79.

Busetti, F. & Harvey, A. (2008). Testing for trend. *Econometric Theory* **24,** 72–87.

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

Dungey, M., Fry, R. & Martin, V.L. (2004). Currency market contagion in the Asia-Pacific region. *Aust. Econ. Papers* **43,** 379–395.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. & Chen, B.C. (1998) New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business and Economic Statistics*, **16**, 127–177.

Franses, P.H. (1998). *Time Series Models for Business and Economic Forecasting.* Cambridge: Cambridge University Press.

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/

Hamilton, J.D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Haywood, J. & Randal, J. (2004). Stochastic seasonality, New Zealand visitor arrivals, and the effects of 11 September 2001. Research Report 04-1, School of Mathematics, Statistics and Computer Science, Victoria University of Wellington, NZ.

Hoaglin, D.C., Mosteller, F. & Tukey, J.W. (Eds.) (1983). *Understanding Robust and Exploratory Data Analysis.* New York: Wiley.

Kaminsky, G.L. & Schmukler, S.L. (1999). What triggers market jitters? A chronicle of the Asian crisis. *J. Int. Money & Finance* **18,** 537–560.

Macaulay, F.R. (1931). *The Smoothing of Time Series.* New York: National Bureau of Economic Research.

Makridakis, S., Wheelwright, S.C., & Hyndman, R.J. (1998). *Forecasting: Methods and Applications*, 3rd edn. New York: Wiley.

NSW Department of Education (1985). 1987: four term future. *Perspectives: Looking at Education* **8(4)** (April).

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

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

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

Statistics New Zealand (2005). *Tourism Satellite Account 2004*. Wellington: Statistics New Zealand.

US Department of State (2004). Three years after 9/11: mixed reviews for war on terror. Available at
http://www.globalsecurity.org/security/library/news/2004/09/wwwh40915.htm

Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software*, **11**(10), 1–17.

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software*, **16**(9), 1–16.

Zeileis, A., Kleiber, C., Krämer, W. & Hornik, K. (2003). Testing and dating of structural changes in practice. *Comput. Statist. & Data Anal.* **44,** 109–123.

Zeileis, A., Leisch, F., Hornik, K. & Kleiber, C. (2002). strucchange: an R package for testing for structural change in linear regression models. *Journal of Statistical Software*, **7**(2), 1–38.