ISSN 1440-771X

Department of Econometrics and Business Statistics

http://www.buseco.monash.edu.au/depts/ebs/pubs/wpapers/

Hierarchical forecasts for Australian domestic tourism

George Athanasopoulos, Roman A Ahmed and Rob J Hyndman

August 2007

Working Paper 12/07

Hierarchical forecasts for Australian domestic tourism

George Athanasopoulos
Department of Econometrics and Business Statistics,
Monash University, VIC 3800
Australia.
Email: George.Athanasopoulos@buseco.monash.edu.au

Roman A Ahmed
Department of Econometrics and Business Statistics,
Monash University, VIC 3800
Australia.
Email: Roman.Ahmed@buseco.monash.edu.au

Rob J Hyndman
Department of Econometrics and Business Statistics,
Monash University, VIC 3800
Australia.
Email: Rob.Hyndman@buseco.monash.edu.au

23 August 2007

JEL classification: C13,C22,C53

Hierarchical forecasts for Australian domestic tourism

Abstract: In this paper we explore the hierarchical nature of tourism demand time series and produce short-term forecasts for Australian domestic tourism. The data and forecasts are organized in a hierarchy based on disaggregating the data for different geographical regions and for different purposes of travel. We consider five approaches to hierarchical forecasting: two variations of the top-down approach, the bottom-up method, a newly proposed top-down approach where top-level forecasts are disaggregated according to forecasted proportions of lower level series, and a recently proposed optimal combination approach. Our forecast performance evaluation shows that the top-down approach based on forecast proportions and the optimal combination method perform best for the tourism hierarchies we consider. By applying these methods, we produce detailed forecasts for the Australian domestic tourism market.

Keywords: Australia, exponential smoothing, hierarchical forecasting, innovations state space models, optimal combination forecasts, top-down method, tourism demand.

1 Introduction

Quarterly tourism demand is measured by the number of "visitor nights", the total nights spent away from home. The data is disaggregated by geographical region and by purpose of travel, thus forming a natural hierarchy of quarterly time series. In this paper we take advantage of this hierarchical structure, using hierarchical forecasting methods to produce forecasts for several levels of disaggregation for the Australian domestic tourism market.

Australia can be divided into six states: New South Wales (NSW), Victoria (VIC), Queensland (QLD), South Australia (SA), Western Australia (WA) and Tasmania (TAS), and the Northern Territory (NT). (For the purposes of this analysis, we treat the Australian Capital Territory as part of NSW and refer to the Northern Territory as a "state".) Business planners require forecasts for the whole of Australia, for each state, and for smaller regions.

In Section 2 we present two hierarchical time series structures for Australian domestic tourism data. In the first hierarchy, we initially disaggregate the data by purpose of travel and then by geographical region. In the second hierarchy, we disaggregate the data on geographical region alone.

The most common approaches to forecasting hierarchical time series are the top-down and bottom-up approaches. The majority of the literature on hierarchical forecasting has focused on comparing the performance of these two methods with some favouring the top-down approaches (see for example Grunfeld and Griliches, 1960; Fogarty et al., 1990; Narasimhan et al., 1994; Fliedner, 1999) others the bottom-up approaches (see for example Orcutt et al., 1968; Edwards and Orcutt, 1969; Kinney, 1971; Dangerfield and Morris, 1992; Zellner and Tobias, 2000) and some finding either method to be uniformly superior (see for example Weatherby, 1984; Fliedner and Mabert, 1992; Shing, 1993). In Section 3 we introduce some notation which neatly generalises hierarchical forecasting approaches. We then present two new hierarchical forecasting methods. First we propose a new top-down approach which is based on disaggregating the top-level forecasts according to forecasted rather than the conventional historical (and therefore static) proportions. Second, we present the newly proposed optimal combination approach of Hyndman et al. (2007). The optimal combination approach is based on forecasting all series at all levels and then using a regression model to optimally combine these forecasts. The resulting revised forecasts display some desirable properties not found in forecasts from other approaches.

We present our modelling procedure in Section 4. For each series, at each level of the hierarchies, we obtain forecasts using an innovations state space model. Considering regional tourism demand and tourism demand by purpose of travel, allows for specific characteristics and dynamics in the data to surface at different levels of the hierarchy. We believe that the greatest advantage of the two new approaches we consider compared to the conventional methods, is that with these approaches we are able to capture the various characteristics through the individual modelling of all the series.

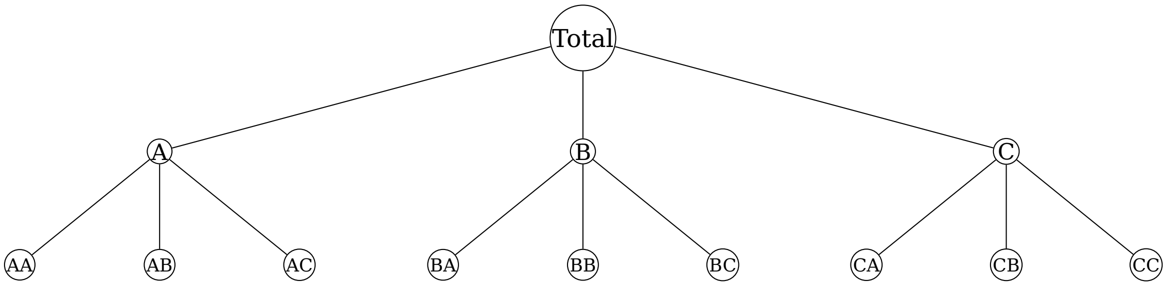
In order to evaluate the performance of the alternative hierarchical approaches, we perform an out-of-sample forecast evaluation in Section 5. We conclude that the best performing hierarchical approach for this application is our newly proposed top-down method followed by the optimal combination approach.

We apply the two new approaches in Section 6, where we forecast tourism demand for Australia and the states from both hierarchies. Our forecasts show a decline in the aggregate domestic tourism demand for Australia over the next two years. This decline is mainly driven by a decline in tourism demand in the states of New South Wales and Victoria. Continuing with the top-down approach based on forecasted proportions, we produce forecasts for all levels of the hierarchies and draw some useful conclusions for policy makers. We present a summary of our findings and concluding remarks in Section 7.

2 Hierarchical time series

Consider the hierarchical structure of Figure 1. We denote the completely aggregated "Total" series as level 0, the first level of disaggregation as level 1, and so on down to the bottom level K, which comprises the most disaggregated series. Hence, the hierarchy depicted in Figure 1 is a $K = 2$ level hierarchy. Let $Y_{X,t}$ be the tth observation ($t = 1, \dots, n$), of series Y_X which corresponds

Figure 1: *A two level hierarchical tree diagram.*



to node X on the hierarchical tree. We use a sequence of letters to denote the individual nodes as depicted in Figure 1. For example, $Y_{A,t}$ denotes the tth observation of the series corresponding to node A at level 1, $Y_{AB,t}$ denotes the tth observation of the series corresponding to node AB at level 2, and so on. Notice that the actual letter sequence depicts the individual node and the length of the letter sequence denotes the level. For the total aggregate level, the tth observation is denoted by Y_t. We let m_i denote the total number of series for level i and $m = m_0 + m_1 + \ldots + m_K$ denotes the total number of series in the hierarchy.

We study two hierarchical time series structures for the Australian domestic tourism market. The structures of the hierarchies are presented in Tables 1 and 2. Each hierarchy allows for different dynamics to surface which we attempt to capture in our modeling process.

For each domestic tourism demand time series, we have quarterly observations on the number of visitor nights which we use as an indicator of tourism activity. The sample begins with the first quarter of 1998 and ends with the final quarter of 2006. The series are obtained from the National Visitor Survey which is managed by Tourism Research Australia. The data were collected using computer-assisted telephone interviews from approximately 120,000 Australians aged 15 years and over on an annual basis (Tourism Research Australia 2005).

Table 1: *Hierarchy 1.*

Level	Number of series	Total series per level
Australia	1	1
Purpose of Travel	4	4
States and Territories	7	28
Capital city versus other	2	56

For more details on this structure refer to Appendix A.1.

Table 2: *Hierarchy 2.*

Level	Total series per level
Australia	1
States and Territories	7
Zones	27
Regions	82

For more details on this structure refer to Appendix A.2.

3 Alternative approaches to hierarchical forecasting

In order to generalise the notation for the various approaches to hierarchical forecasting, we let vector $\boldsymbol{Y}_{i,t}$ contain all the observations in level i at time t. We stack all the observations of all series at time t in a column vector defined as $\boldsymbol{Y}_t = [Y_t, \boldsymbol{Y}'_{1,t}, \dots, \boldsymbol{Y}'_{K,t}]'$. Now we can write

$$\boldsymbol{Y}_t = \boldsymbol{S}\boldsymbol{Y}_{K,t} \tag{1}$$

where $\boldsymbol{S}$ is a "summing" matrix of order $m \times m_K$ that aggregates the bottom level series all the way up the hierarchy. For example, for the hierarchy of Figure 1 we have

$$\underset{(13\times1)}{\begin{bmatrix} Y_t \\ Y_{\mathrm{A},t} \\ Y_{\mathrm{B},t} \\ Y_{\mathrm{C},t} \\ Y_{\mathrm{AA},t} \\ Y_{\mathrm{AB},t} \\ \vdots \\ Y_{\mathrm{CB},t} \\ Y_{\mathrm{CC},t} \end{bmatrix}} = \underset{(13\times9)}{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ & & & & & & & & \\ & & & & \boldsymbol{I}_9 & & & & \\ & & & & & & & & \end{bmatrix}} \underset{(9\times1)}{\begin{bmatrix} Y_{\mathrm{AA},t} \\ Y_{\mathrm{AB},t} \\ Y_{\mathrm{AC},t} \\ Y_{\mathrm{BA},t} \\ Y_{\mathrm{BB},t} \\ Y_{\mathrm{BC},t} \\ Y_{\mathrm{CA},t} \\ Y_{\mathrm{CB},t} \\ Y_{\mathrm{CC},t} \end{bmatrix}}$$

where $\boldsymbol{I}_k$ denotes an identity matrix of order $k \times k$.

In hierarchical forecasting we are interested in working with the forecasts rather than the actual observations of each series. Suppose we have generated h-step-ahead forecasts for each individual series Y_{X}, denoted by $\hat{Y}_{\mathrm{X},n}(h)$. We should clarify that these forecasts are based on a sample of $t = 1, \dots, n$, hence they are the forecasts for time $n + h$. Therefore, $\hat{Y}_{\mathrm{AB},n}(h)$ denotes the h-step-ahead base forecast of series Y_{AB} using the sample $Y_{\mathrm{AB},1}, \dots Y_{\mathrm{AB},n}$. For level i, all h-step-ahead base forecasts will be represented by $\hat{\boldsymbol{Y}}_{i,n}(h)$, and the h-step-ahead base forecasts for the whole hierarchy are represented by the vector $\hat{\boldsymbol{Y}}_n(h)$ which contains all the base forecasts stacked in the same order as $\boldsymbol{Y}_t$.

Using this notation, all existing hierarchical methods can be represented by the general form

$$\tilde{\boldsymbol{Y}}_n(h) = \boldsymbol{S}\boldsymbol{P}\hat{\boldsymbol{Y}}_n(h) \tag{2}$$

where $\boldsymbol{S}$ is the $m \times m_K$ summing matrix as in equation (1) and $\boldsymbol{P}$ is a matrix of order $m_K \times m$. The role of $\boldsymbol{P}$ is different depending on the hierarchical approach. This will become clear in the sections that follow. This general representation shows that the final revised forecasts $\tilde{Y}_n(h)$ produced by any hierarchical forecasting approach are the result of linearly combining the independent base forecasts, $\hat{Y}_n(h)$.

3.1 The bottom-up approach

Arguably the most commonly applied method to hierarchical forecasting is the bottom-up approach (see for example Theil, 1954; Orcutt et al., 1968; Shlifer and Wolff, 1979; Dunn et al., 1976; Dangerfield and Morris, 1992; Zellner and Tobias, 2000). To represent this approach by the general form of equation (2) we denote

$$\boldsymbol{P} = [\mathbf{0}_{m_K \times (m-m_K)} \mid \boldsymbol{I}_{m_K}] \tag{3}$$

where $\mathbf{0}_{i \times j}$ is the $i \times j$ null matrix. The role of $\boldsymbol{P}$ here is to extract the bottom level forecasts, which are then aggregated by the summation matrix $\boldsymbol{S}$ to produce the revised forecasts for the whole hierarchy. The greatest advantage of this approach is that by modelling the data at the most disaggregated bottom level we do not lose any information due to aggregation. Hence we can better capture the dynamics of the individual series. However, bottom level data can be quite noisy and therefore more challenging to model.

3.2 Top-down approaches based on historical proportions

The other commonly applied method in hierarchical forecasting is the top-down approach (see for example Grunfeld and Griliches, 1960; McLeavey and Narasimhan, 1985; Lütkephol, 1984; Fliedner, 1999). The most common form of the top-down approach is to disaggregate the forecasts of the "Total" series and distribute these down the hierarchy based on the historical proportions of the data. In terms of the general form of equation (2), we write

$$\boldsymbol{P} = [\boldsymbol{p} \mid \mathbf{0}_{m_K \times (m-1)}] \tag{4}$$

where $\boldsymbol{p} = [p_1, p_2, \ldots, p_{m_K}]'$ are a set of proportions for the bottom level series. So the role of $\boldsymbol{P}$ here is to distribute the top level forecasts to forecasts for the bottom level series.

In this paper we consider two versions of this approach which performed quite well in Gross and Sohl (1990). For the first one

$$p_j = \sum_{t=1}^{n} \frac{Y_{j,t}}{Y_t} \Big/ n \qquad (5)$$

for $j = 1, \ldots, m_K$. We label this TDHP1 (top-down historical proportions 1) in the tables that follow. Each proportion p_j reflects the average of the historical proportions of the bottom level series $\{Y_{j,t}\}$ over the period $t = 1, \ldots, n$ relative to the total aggregate $\{Y_t\}$; i.e., vector $\boldsymbol{p}$ reflects the *average historical proportions*.

In the second version we consider

$$p_j = \sum_{t=1}^{n} \frac{Y_{j,t}}{n} \Big/ \sum_{t=1}^{n} \frac{Y_t}{n} \qquad (6)$$

for $j = 1, \ldots, m_K$. We label this TDHP2 in the tables that follow. Each p_j proportion here captures the average historical value of the bottom level series $\{Y_{j,t}\}$ relative to the average value of the total aggregate $\{Y_t\}$; i.e., vector $\boldsymbol{p}$ reflects the *proportions of the historical averages*.

The simplicity of the application of these top-down approaches is their greatest attribute. One only needs to model and produce forecasts for the most aggregated top level series. These approaches seem to produce quite reliable forecasts for the aggregate levels and they are very useful with low count data. On the other hand, their greatest disadvantage is the loss of information due to aggregation. With these top-down approaches, we are unable to capture and take advantage of individual series characteristics such as time dynamics, special events, etc. Finally, with these methods we base the disaggregation of the "Total" series forecasts on historical and static proportions, and these proportions will miss any trends in the data.

3.3 Top-down approach based on forecasted proportions

To improve on the above historical and static nature of the proportions used to disaggregate the top level forecasts, we introduce a top-down method for which the proportions for disaggregating the top level forecasts are based on forecasted proportions of lower level series. As the results that follow will show, this method has worked well with the tourism hierarchies we consider in this paper. The greatest disadvantage of this method, which in fact is a disadvantage of any top-down approach, is that the top-down approaches do not produce unbiased revised forecasts even if the base forecasts are unbiased (refer to the discussion of equation (5) in Hyndman et al., 2007).

As with the previous two top-down approaches,

$$\boldsymbol{P} = [\boldsymbol{p} \mid \boldsymbol{0}_{m_K \times (m-1)}] \tag{7}$$

where $\boldsymbol{p} = [p_1, p_2, \dots, p_{m_K}]'$ are a set of proportions for the bottom level series. In order to present a general form for the bottom level proportions we need to introduce some new notation. Let $\hat{Y}_{j,n}^{(i)}(h)$ be the h-step-ahead forecast of the series that corresponds to the node which is i levels above j. Also let $\Sigma(\hat{Y}_{i,n}(h))$ be the sum of the h-step-ahead forecasts below node i which are directly connected to node i. For example in Figure 1, $\Sigma(\hat{Y}_{i,n}^{(2)}(h)) = \Sigma(\hat{Y}_{\text{Total},n}(h)) = \hat{Y}_{\text{A},n}(h) + \hat{Y}_{\text{B},n}(h) + \hat{Y}_{\text{C},n}(h)$. Then

$$p_j = \prod_{i=0}^{K-1} \frac{\hat{Y}_{j,n}^{(i)}(h)}{\Sigma(\hat{Y}_{j,n}^{(i+1)}(h))} \tag{8}$$

for $j = 1, 2, \dots, m_K$.

If we generate $\hat{Y}_{\text{Total},n}(h)$ for the top level series of the hierarchy in Figure 1, the revised final forecasts moving down the farthest left branch of the hierarchy will be,

$$\tilde{Y}_{\text{A},n}(h) = \left(\frac{\hat{Y}_{\text{A},n}(h)}{\hat{Y}_{\text{A},n}(h) + \hat{Y}_{\text{B},n}(h) + \hat{Y}_{\text{C},n}(h)}\right) \times \hat{Y}_{\text{Total},n}(h)$$

and

$$\tilde{Y}_{\text{AA},n}(h) = \left(\frac{\hat{Y}_{\text{AA},n}(h)}{\hat{Y}_{\text{AA},n}(h) + \hat{Y}_{\text{AB},n}(h) + \hat{Y}_{\text{AC},n}(h)}\right) \times \tilde{Y}_{\text{A},n}(h).$$

Therefore

$$\tilde{Y}_{\text{AA},n}(h) = \left(\frac{\hat{Y}_{\text{AA},n}(h)}{\hat{Y}_{\text{AA},n}(h) + \hat{Y}_{\text{AB},n}(h) + \hat{Y}_{\text{AC},n}(h)}\right)\left(\frac{\hat{Y}_{\text{A},n}(h)}{\hat{Y}_{\text{A},n}(h) + \hat{Y}_{\text{B},n}(h) + \hat{Y}_{\text{C},n}(h)}\right) \times \hat{Y}_{\text{Total},n}(h)$$

and so

$$p_1 = \left(\frac{\hat{Y}_{\text{AA},n}(h)}{\hat{Y}_{\text{AA},n}(h) + \hat{Y}_{\text{AB},n}(h) + \hat{Y}_{\text{AC},n}(h)}\right)\left(\frac{\hat{Y}_{\text{A},n}(h)}{\hat{Y}_{\text{A},n}(h) + \hat{Y}_{\text{B},n}(h) + \hat{Y}_{\text{C},n}(h)}\right).$$

Other proportions are similarly obtained.

3.4 The optimal combination approach

The final approach to hierarchical forecasting we consider is the "optimal combination approach" introduced in Hyndman, Ahmed and Athanasopoulos (2007). This approach optimally combines the base forecasts to produce the set of revised forecasts. Unlike any other existing method, this approach uses all the information available within a hierarchy, it allows for correlations and interactions between series at each level of the hierarchy, it accounts for ad hoc adjustments of forecasts at any level and it produces unbiased forecasts which are consistent across the levels of the hierarchy. Furthermore, this approach can also produce estimates of forecast uncertainty that are consistent across levels of the hierarchy.

The general idea is derived from the representation of the h-step-ahead base forecasts of a hierarchy by the linear regression model

$$\hat{\boldsymbol{Y}}_n(h) = \boldsymbol{S}\boldsymbol{\beta}_h + \boldsymbol{\varepsilon}_h \tag{9}$$

where $\boldsymbol{\beta}_h = \mathrm{E}[\hat{\boldsymbol{Y}}_{K,n}(h) \mid \boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n]$ is the unknown mean of the base forecasts of the bottom level K, and $\boldsymbol{\varepsilon}_h$ has zero mean and covariance matrix $V[\boldsymbol{\varepsilon}_h] = \Sigma_h$. If we knew Σ_h then we could use generalised least squares estimation to obtain the minimum variance unbiased estimate of $\boldsymbol{\beta}_h$. In general, this is not known but Hyndman et al. (2007) show that under the reasonable assumption that $\boldsymbol{\varepsilon}_h \approx \boldsymbol{S}\boldsymbol{\varepsilon}_{K,h}$ where $\boldsymbol{\varepsilon}_{K,h}$ contains the forecast errors in the bottom level, the best linear unbiased estimator for $\boldsymbol{\beta}_h$ is $\hat{\boldsymbol{\beta}}_h = (\boldsymbol{S}'\boldsymbol{S})^{-1}\boldsymbol{S}'\hat{\boldsymbol{Y}}_n(h)$. This leads to the revised forecasts given by $\tilde{\boldsymbol{Y}}_n(h) = \boldsymbol{S}\hat{\boldsymbol{\beta}}_h$ and hence in the general form of equation (2),

$$\boldsymbol{P} = (\boldsymbol{S}'\boldsymbol{S})^{-1}\boldsymbol{S}'. \tag{10}$$

In some circumstances, simpler forecasting equations can be obtained. Note that hierarchy 1 is balanced which means that at each node within a level the same degree of disaggregation takes place; i.e., the number of series at each node varies across levels but not within a level. Therefore the simple ANOVA method presented in equations (12) and (13) in Hyndman et al. (2007) can be applied for producing the revised forecasts for the optimal combination approach.

3.5 Prediction intervals

Hyndman et al. (2007) also show that, for all of the methods that can be represented by (2), the variance of the forecasts is given by

$$\mathrm{Var}[\tilde{\boldsymbol{Y}}_n(h)] = \boldsymbol{S}\boldsymbol{P}\Sigma_h\boldsymbol{P}'\boldsymbol{S}' \tag{11}$$

where $\boldsymbol{\Sigma}_h$ is the variance of the base forecasts, $\hat{\boldsymbol{Y}}_n(h)$. Thus, prediction intervals on the revised forecasts can be obtained provided $\boldsymbol{\Sigma}_h$ can be reliably estimated. However, estimation of $\boldsymbol{\Sigma}_h$ is a difficult problem and we leave this to a later paper. Consequently, we do not provide prediction intervals for the forecasts presented here.

4 Forecasting individual series

The classification of the exponential smoothing methods in Table 3 originated with the Pegels (1969) taxonomy, which was further advanced by Gardner (1985), Hyndman et al. (2002) and Taylor (2003). Each of the fifteen methods listed has a trend and a seasonal component. Hence, cell (N,N) describes the simple exponential smoothing method, cell (A,N) Holt's linear method, and so on. We model and forecast all series in the hierarchy individually at all levels for each hierarchical structure using exponential smoothing based on innovations state space models. Hyndman et al. (2002) developed a statistical framework for most of the exponential smoothing methods presented in Table 3. The statistical framework incorporates stochastic models, likelihood calculations, prediction intervals and procedures for model selection. We extend their framework here to include Taylor's (2003) multiplicative damped method.

For each method, there are two possible state space models: one corresponds to a model with additive errors and the other to a model with multiplicative errors. Table 4 presents the fifteen models with additive errors and their forecast functions. The multiplicative error models can be obtained by replacing ε_t by $\mu_t \varepsilon_t$. Empirically, we have found that the purely additive models (models with additive error, trend and seasonality) give better forecast accuracy. Consequently, we selected models by minimising the AIC amongst all additive models. In a few cases, the forecasts from the additive models did not have face validity (e.g., the forecasts were negative), and the models for these series were then replaced by models with multiplicative components. The models selected for each series are given in Appendix A.

Table 3: *Classification of exponential smoothing methods.*

		Seasonal Component		
		N	A	M
Trend Component		(None)	(Additive)	(Multiplicative)
N	(None)	N,N	N,A	N,M
A	(Additive)	A,N	A,A	A,M
$\mathrm{A_d}$	(Additive damped)	$\mathrm{A_d}$,N	$\mathrm{A_d}$,A	$\mathrm{A_d}$,M
M	(Multiplicative)	M,N	M,A	M,M
$\mathrm{M_d}$	(Multiplicative damped)	$\mathrm{M_d}$,N	$\mathrm{M_d}$,A	$\mathrm{M_d}$,M

Table 4: *State space equations for each additive error model in the classification. Multiplicative error models are obtained by replacing ε_t by $\mu_t\varepsilon_t$. In each case, ℓ_t denotes the level of the series at time t, b_t denotes the slope at time t, s_t denotes the seasonal component of the series at time t, and m denotes the number of seasons in a year; α, β, γ and ϕ are constants with $0 < \alpha, \gamma, \phi < 1$ and $0 < \beta < \alpha$; $\hat{Y}_{t+h|t}$ denotes the h-step-ahead forecast based on all the data up to time t; $\phi_h = \phi + \phi^2 + \ldots + \phi^h$; $\hat{Y}_{t+h|t}$ denotes a forecast of Y_{t+h} based on all the data up to time t, and $h_m^+ = \big[(h-1) \text{ mod } m\big] + 1$.*

Trend component	Seasonal component **N** (none)	 **A** (additive)	 **M** (multiplicative)
N (none)	$\mu_t = \ell_{t-1}$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t$	$\mu_t = \ell_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t + s_{t-m+h_m^+}$	$\mu_t = \ell_{t-1}s_{t-m}$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/\ell_{t-1}$ $\hat{Y}_{t+h\|t} = \ell_t s_{t-m+h_m^+}$
A (additive)	$\mu_t = \ell_{t-1} + b_{t-1}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t + hb_t$	$\mu_t = \ell_{t-1} + b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t + hb_t + s_{t-m+h_m^+}$	$\mu_t = (\ell_{t-1} + b_{t-1})s_{t-m}$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + b_{t-1})$ $\hat{Y}_{t+h\|t} = (\ell_t + hb_t)s_{t-m+h_m^+}$
A_d (additive damped)	$\mu_t = \ell_{t-1} + \phi b_{t-1}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t + \phi_h b_t$	$\mu_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t + \phi_h b_t + s_{t-m+h_m^+}$	$\mu_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t/s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1} + \phi b_{t-1})$ $\hat{Y}_{t+h\|t} = (\ell_t + \phi_h b_t)s_{t-m+h_m^+}$
M (multiplicative)	$\mu_t = \ell_{t-1}b_{t-1}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t/\ell_{t-1}$ $\hat{Y}_{t+h\|t} = \ell_t b_t^h$	$\mu_t = \ell_{t-1}b_{t-1} + s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t/\ell_{t-1}$ $s_t = s_{t-m} + \gamma\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t b_t^h + s_{t-m+h_m^+}$	$\mu_t = \ell_{t-1}b_{t-1}s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t/(s_{t-m}\ell_{t-1})$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1}b_{t-1})$ $\hat{Y}_{t+h\|t} = \ell_t b_t^h s_{t-m+h_m^+}$
M_d (multiplicative damped)	$\mu_t = \ell_{t-1}b_{t-1}^{\phi}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha\varepsilon_t$ $b_t = b_{t-1}^{\phi} + \beta\varepsilon_t/\ell_{t-1}$ $\hat{Y}_{t+h\|t} = \ell_t b_t^{\phi_h}$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi} + s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha\varepsilon_t$ $b_t = b_{t-1}^{\phi} + \beta\varepsilon_t/\ell_{t-1}$ $s_t = s_{t-m} + \gamma\varepsilon_t$ $\hat{Y}_{t+h\|t} = \ell_t b_t^{\phi_h} + s_{t-m+h_m^+}$	$\mu_t = \ell_{t-1}b_{t-1}^{\phi}s_{t-m}$ $\ell_t = \ell_{t-1}b_{t-1}^{\phi} + \alpha\varepsilon_t/s_{t-m}$ $b_t = b_{t-1}^{\phi} + \beta\varepsilon_t/(s_{t-m}\ell_{t-1})$ $s_t = s_{t-m} + \gamma\varepsilon_t/(\ell_{t-1}b_{t-1})$ $\hat{Y}_{t+h\|t} = \ell_t b_t^{\phi_h} s_{t-m+h_m^+}$

5 Forecast performance evaluation

In order to evaluate the forecasting performance of each of the hierarchical approaches presented in Section 3, we perform an out-of-sample forecast evaluation for both of the Australian domestic tourism hierarchies considered in this paper. We initially select models (as in Section 4) using the whole sample. We then re-estimate the models based on the first 12 observations (1998:Q1–2001:Q4) and produce 1 to 8-step-ahead forecasts. We increase the sample size by one observation and re-estimate the models and again produce 1 to 8-step-ahead forecasts. This process is iterated until 2005:Q3 and it produces 24 1-step-ahead forecasts, 23 2-steps-ahead forecasts, 22 3-step-ahead forecasts, up to 17 8-step-ahead forecasts. We use these forecasts to evaluate the out-of-sample forecast performance of each of the hierarchical methods considered. We calculate the mean absolute percentage error (MAPE) for each forecast horizon and for each of the alternative hierarchical approaches. The results for hierarchy 1 are presented in Table 5 and the results for hierarchy 2 are presented in Table 6. The first four panels in each table are self explanatory. In these we present the MAPEs for the alternative hierarchical approaches for each of the four levels in the hierarchy. In the final panel labeled "Total" we present the aggregate MAPEs across the whole of the hierarchy. Finally, the final column on each table labeled "Average" shows the average MAPE across all the forecast horizons for each approach.

For both hierarchies it can be seen that the two top-down approaches based on static historical proportions are only useful for forecasting the very top level of the hierarchies. This is not surprising. With the top-down strategies, no disaggregation takes place at the top level. All we do here is to model the time series at the top level independently of the hierarchical structure. However, as we move down the hierarchy the performance of the top-down approaches is shown to deteriorate. These two methods are easily identified as the overall worst performing methods and are not recommended.

From the three alternative approaches remaining, it seems that the overall best performing method for both hierarchies is the top-down approach based on the forecasting proportions. This approach is clearly the best performing for hierarchy 2. For hierarchy 1 the optimal combination approach seems to also be performing well. The surprising feature of this analysis is the better-than-expected performance of the bottom-up approach. We believe that the good performance of this approach can be attributed to the nature of the data. Even at the very bottom level the data is well behaved with a prominent seasonal component for most series.

Furthermore, this method is also advantaged by the short-term forecasts we are producing in this forecast evaluation exercise. If the forecast horizon was longer the performance of this method would deteriorate as it misses the trends in the series. For example, for hierarchy 2, none of the selected bottom level models include a trend so that the bottom-up approach produces flat forecasts for all series at all levels. However, at level 1 there is a strong downward trend for the New South Wales series which comprises 33% of the total tourism demand for Australia. This trend is captured by both the top-down method based on forecasted proportions and the optimal combination approach.

Table 5: *Out-of-sample forecasting performance: Hierarchy 1.*

	Forecast Horizon (h)								
MAPE	1	2	3	4	5	6	7	8	Average
				Top Level: Australia					
Bottom-up	**3.48**	**3.30**	3.81	4.04	3.90	4.56	4.53	4.58	4.03
Top-down HP1	3.89	3.71	**3.41**	**3.90**	3.91	**4.12**	**4.27**	**4.27**	**3.93**
Top-down HP2	3.89	3.71	**3.41**	**3.90**	3.91	**4.12**	**4.27**	**4.27**	**3.93**
Top-down FP	3.89	3.71	**3.41**	**3.90**	3.91	**4.12**	**4.27**	**4.27**	**3.93**
Optimal	3.80	3.64	3.48	3.94	**3.85**	4.22	4.34	4.35	3.95
				Level 1: Purpose of travel					
Bottom-up	6.15	6.22	6.49	6.99	7.80	8.15	8.21	7.88	7.24
Top-down HP1	9.83	9.34	9.34	9.67	9.81	9.52	9.88	9.81	9.65
Top-down HP2	10.01	9.56	9.55	9.84	9.98	9.71	10.06	9.97	9.84
Top-down FP	5.73	5.78	**5.58**	6.15	**6.80**	**7.28**	**7.56**	7.68	**6.57**
Optimal	**5.63**	**5.71**	5.74	**6.14**	6.91	7.35	7.57	**7.64**	6.59
				Level 2: States					
Bottom-up	**21.34**	**21.75**	**21.81**	**22.39**	**23.76**	23.26	23.01	23.31	**22.58**
Top-down HP1	32.63	30.98	31.49	31.91	32.23	30.11	30.51	30.91	31.35
Top-down HP2	32.92	31.23	31.72	32.13	32.47	30.32	30.67	31.01	31.56
Top-down FP	22.15	21.96	21.94	22.52	23.79	23.18	22.96	**23.07**	22.70
Optimal	22.17	21.80	22.33	23.53	24.26	**23.15**	**22.76**	23.90	22.99
				Bottom Level: Capital city versus other					
Bottom-up	**31.97**	31.65	31.39	32.19	33.93	33.70	32.67	33.47	32.62
Top-down HP1	42.47	40.19	40.57	41.12	41.71	39.67	39.87	40.68	40.79
Top-down HP2	43.04	40.54	40.87	41.44	42.06	39.99	40.21	40.99	41.14
Top-down FP	32.16	31.30	31.24	**32.18**	34.00	**33.25**	**32.42**	**33.22**	**32.47**
Optimal	32.31	**30.92**	**30.87**	32.41	**33.92**	33.35	32.47	34.13	32.55
				Total					
Bottom-up	**62.94**	62.91	63.50	65.59	69.39	69.66	68.43	69.25	66.46
Top-down HP1	88.82	84.23	84.82	86.59	87.66	83.42	84.53	85.67	85.72
Top-down HP2	89.85	85.04	85.55	87.31	88.42	84.14	85.21	86.24	86.47
Top-down FP	63.93	62.76	**62.16**	**64.75**	**68.49**	**67.82**	67.22	**68.24**	**65.67**
Optimal	63.92	**62.08**	62.42	66.02	68.94	68.06	**67.14**	70.03	66.07

Table 6: *Out-of-sample forecasting performance: Hierarchy 2.*

MAPE	*Forecast Horizon (h)* 1	2	3	4	5	6	7	8	Average
				Top Level: Australia					
Bottom-up	**3.79**	**3.58**	3.92	4.01	4.12	4.55	4.30	**4.24**	4.06
Top-down HP1	3.89	3.71	3.41	3.90	3.91	4.12	4.27	4.27	**3.93**
Top-down HP2	3.89	3.71	**3.41**	3.90	**3.91**	**4.12**	**4.27**	4.27	**3.93**
Top-down FP	3.89	3.71	**3.41**	3.90	**3.91**	**4.12**	**4.27**	4.27	**3.93**
Optimal	3.83	3.66	3.46	**3.88**	3.92	4.19	4.30	4.25	3.94
				Level 1: States					
Bottom-up	10.70	10.52	10.68	10.85	11.37	11.46	11.43	11.27	11.03
Top-down HP1	19.30	18.63	19.00	18.94	19.48	18.68	19.27	19.85	19.14
Top-down HP2	19.17	18.57	18.90	18.78	19.34	18.60	19.13	19.67	19.02
Top-down FP	**10.58**	**10.29**	**10.20**	**10.54**	**10.94**	**10.90**	**11.08**	**11.18**	**10.71**
Optimal	11.07	10.58	10.67	11.13	11.60	11.62	11.89	12.21	11.35
				Level 2: Zones					
Bottom-up	14.99	14.97	14.88	14.98	15.73	15.69	15.63	15.65	15.32
Top-down HP1	24.14	23.55	23.84	23.94	24.46	23.66	24.28	24.33	24.03
Top-down HP2	24.32	23.77	24.07	24.11	24.60	23.88	24.53	24.51	24.22
Top-down FP	**14.82**	**14.83**	**14.58**	**14.78**	**15.44**	**15.36**	**15.51**	**15.54**	**15.11**
Optimal	15.16	15.06	14.78	15.27	15.85	15.74	15.87	16.15	15.48
				Bottom Level: Regions					
Bottom-up	33.12	32.54	31.86	32.26	33.97	33.74	34.01	33.96	33.18
Top-down HP1	41.95	40.36	40.87	41.09	41.77	40.51	41.43	41.76	41.22
Top-down HP2	42.50	40.96	41.45	41.61	42.28	41.03	41.97	42.23	41.75
Top-down FP	**31.82**	**31.50**	**30.80**	**31.53**	**32.58**	**32.50**	**33.16**	**33.29**	**32.15**
Optimal	35.89	33.86	33.04	34.26	36.22	36.06	36.64	37.49	35.43
				Total					
Bottom-up	62.59	61.61	61.34	62.11	65.18	65.44	65.37	65.13	63.60
Top-down HP1	89.28	86.25	87.12	87.86	89.62	86.96	89.26	90.21	88.32
Top-down HP2	89.88	87.02	87.83	88.40	90.12	87.62	89.90	90.68	88.93
Top-down FP	**61.11**	**60.33**	**58.98**	**60.74**	**62.87**	**62.87**	**64.03**	**64.28**	**61.90**
Optimal	65.96	63.16	61.94	64.54	67.59	67.61	68.70	70.10	66.20

6 Forecasts

The approach that performed best in the forecast evaluation exercise was our proposed top-down approach based on forecasted proportions. The next-best performing method was the optimal combination approach. In this section we use these two methods to forecast tourism demand for Australia and the Australian states from both hierarchical structures. The large number of series in each hierarchy prevents us from presenting the raw data forecasts. In order to summarise our forecast results in a useful manner, we present the average forecasted rate of growth/decline per annum, as calculated over the next two years for each series.

6.1 Forecasts for Australia and the states

In Table 7 we present the forecast average rate of growth/decline for tourism demand for Australia and the Australian states (refer to Figure 2 for a map of Australia). The forecasted rates from the top-down approach based on forecasted proportions are labeled "Top-down FP" and the optimal combination forecasted rates are labeled "Optimal". The forecasted rates from all sources seem to be consistent in terms of direction. There are only two exceptions: the case of Victoria where the top-down forecasted rate from hierarchy 2 is positive (although very small) in contrast to the decline shown by all the other sources, and the case of Western Australia where the forecasted rates from hierarchy 1 for both top-down and optimal approaches are negative in contrast to the forecasts from hierarchy 2 which are positive for both approaches.

Table 7: *Forecast average rate of growth/decline per annum over 2007 and 2008.*

	Australia	NSW	VIC	QLD	SA	WA	TAS	NT
Top-down FP hierarchy 1	−0.29	−2.03	−0.61	1.05	2.22	−0.04	−2.21	5.32
Optimal hierarchy 1	−0.24	−1.84	−0.80	0.78	2.76	−0.10	−1.22	6.46
Top-down FP hierarchy 2	−0.29	−2.29	0.06	1.07	2.70	0.00	−1.15	0.15
Optimal hierarchy 2	−0.35	−2.22	−0.19	0.59	3.03	0.11	−0.28	1.16
Proportion		*0.33*	*0.19*	*0.26*	*0.10*	*0.07*	*0.03*	*0.02*

The *Proportion* entry denotes the historical proportion of tourism in the corresponding area to total Australian tourism.

The consensus from the methods is that there will be a decline in domestic tourism demand for Australia over the next two years. The most conservative rate of decline of 0.24% p.a. is produced by the optimal combination approach from hierarchy 1. The least conservative rate of decline of 0.35% p.a. is given by the optimal combination approach from hierarchy 2. This rate of decline seems to be driven mainly by the decline in the states of New South Wales and Victoria

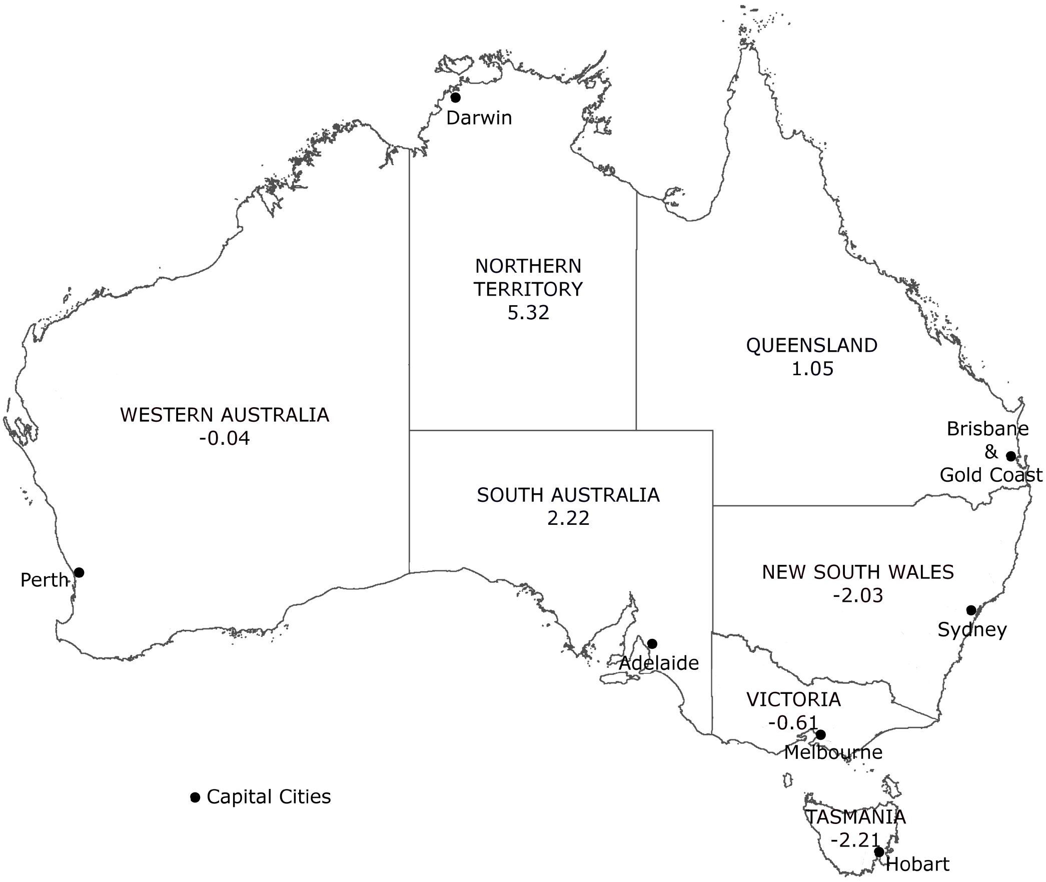


Figure 2: *Average forecast rate of growth/decline per annum over 2007 and 2008 for the states of Australia. The rates are produced by the top-down approach based on forecasted proportions from Hierarchy 1.*

which make up approximately 52% of Australia's aggregate domestic tourism demand. The areas showing some signs of growth are the states of Queensland, South Australia and the Northern Territory.

In the left panel of Figure 3, we plot the quarterly data for the aggregate Australian domestic tourism demand. The plot reveals the nature of the data (a prominent seasonal component) and the nature of the forecasts produced by the innovations state space models. We plot only the top-down forecasts here as the forecasts from the optimal combination approach are not very different. In the right panel of Figure 3, we plot the annual data and the forecasts from the three alternative approaches. As the methods give such similar forecasts, we will only present forecasts from the top-down forecasted proportions approach for lower levels of the hierarchies.

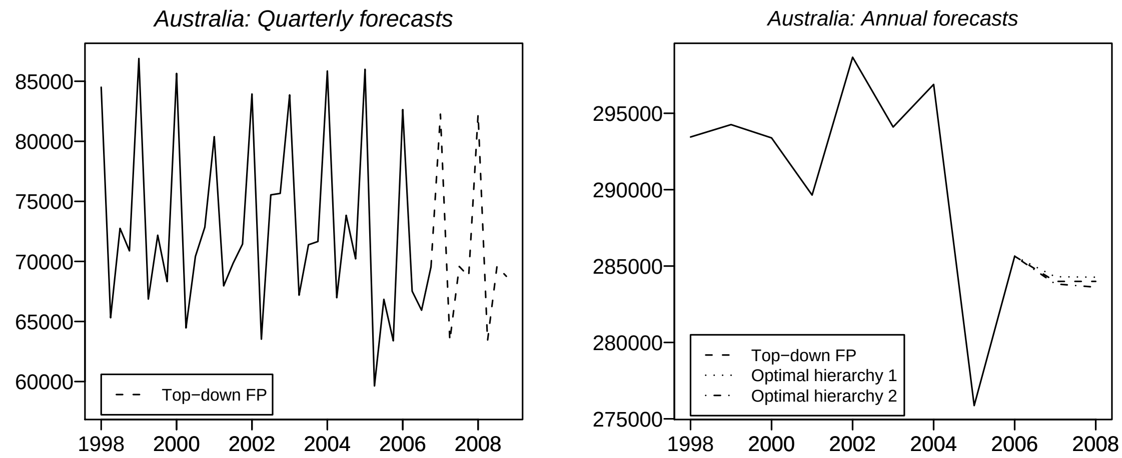


Figure 3: *Quarterly and annual forecasts for Australian tourism demand for the period 2007 and 2008.*

6.2 Further forecasts from hierarchy 1

The structure of hierarchy 1 allows us to model tourism demand based on the four purposes of travel. The forecasted growth rates per purpose of travel presented in the "Total" row in Table 8 show a decline in three of the four components. The only component showing an increase is "Other" which is a relatively small component and has little impact on aggregate domestic tourism demand. The two main components which make up 78% of domestic tourism are "Holiday" and VFR travel. We forecast an increase in "Holiday" travel for the states of Queensland and South Australia and an increase in VFR for the states of South Australia and Western Australia. For all other areas at this level, our forecasts show a decline in these two main components. Note that the only state for which we forecast growth over the next two years for all four components is South Australia.

Table 8: *Average forecast rate of growth/decline per annum over 2007 and 2008 by purpose of travel.*

	Holiday	VFR	Business	Other	Total	*Proportion*
NSW	−4.39	−1.00	0.32	5.38	−2.03	*0.33*
VIC	−0.46	−1.40	1.32	−1.36	−0.61	*0.19*
QLD	1.82	−0.17	−3.80	14.87	1.05	*0.26*
SA	0.74	1.19	4.94	10.45	2.22	*0.10*
WA	−0.84	1.08	−4.89	14.27	−0.04	*0.07*
TAS	−4.38	−0.92	8.98	−5.53	−2.21	*0.03*
NT	−2.60	−4.32	22.40	56.35	5.32	*0.02*
Total	−1.20	−0.61	−0.20	9.25		
Proportion	*0.47*	*0.31*	*0.15*	*0.07*		

The *Proportion* entry denotes the aggregate historical proportion of the corresponding purpose of travel to total Australian tourism.

Table 9: *Average percentage growth/decline per annum over 2007 and 2008 for capital city versus the rest of the state.*

	Holiday		VFR		Business		Other		Total	
	Cap City	Other	Cap City	Other	Cap City	Other	Cap City	Other	Cap City	Other
NSW	−2.74	−4.68	−0.50	−1.21	−3.09	2.90	−1.60	10.72	−1.78	−2.12
VIC	−2.59	0.54	−1.93	−0.93	5.39	−4.16	−3.23	0.59	−0.96	−0.36
QLD	2.63	1.25	2.36	−2.09	−6.31	−2.01	13.17	16.23	1.81	0.49
SA	0.06	−1.24	2.33	−0.51	−0.28	−7.09	8.88	20.25	1.70	−1.18
WA	0.74	0.74	2.81	−0.63	4.53	5.41	17.78	2.23	4.01	1.02
TAS	−7.81	−2.46	2.35	−3.20	8.20	9.77	6.98	−12.00	−2.07	−2.29
NT	2.36	−7.08	4.21	−13.97	22.48	22.33	70.62	47.25	9.87	1.17

Brisbane and Gold Coast are considered the combined capital city for the state of Queensland

The largest growth shown in Table 9 in terms of "Holiday" travel comes from the state of Queensland and in particular its capital city. It should be noted that we consider both Brisbane and the Gold Coast as the capital city of this state. (Please refer to Figure 2 for a map of Australia showing the capital city for each state.) The Gold Coast (and Queensland in general) is arguably the most developed and most promoted Australian holiday destination, which explains the forecasted growth over the next two years.

For the other two main states (New South Wales and Victoria), our forecasts show a decline in "Holiday" travel for their respective capital cities, Sydney and Melbourne. For the rest of New South Wales we forecast a significant decline, and we forecast a moderate growth for the rest of Victoria. For the remaining areas the results are mixed in terms of "Holiday" travel. For South Australia we forecast a slight increase for Adelaide and a decrease for the rest of the state. We forecast "Holiday" travel to increase uniformly across the state of Western Australia and decrease uniformly across the state of Tasmania, with quite a large decline for Hobart. For the Northern Territory, we forecast an increase in Darwin and a decline for the rest of the territory.

Travel for VFR is forecast to decline for all the states outside their respective capital cities. A decline is forecast for Sydney and Melbourne, but we forecast a moderate growth for the rest of the capital cities.

The forecast growth rates for "Business" travel show an increase for both New South Wales and Victoria. Our forecasts show that this increase will come from different sources. "Business" travel will grow in New South Wales due to an increase outside the city of Sydney. In contrast, for Victoria, the growth in "Business" travel will come from an increase within the Melbourne area. For the rest of Australia, "Business" travel will either uniformly decline across the states (in Queensland and South Australia), or will uniformly grow, (in Western Australia, Tasmania and the Northern Territory).

6.3 Further forecasts from hierarchy 2

Hierarchy 2 allows us to analyse our top level Australian forecasts in more depth in terms of location. In Table 10 and Figure 4, we present the forecast growth rates for the tourism zones of Australia as they are classified in Appendix A.2.

Table 10: *Average forecast rate of growth/decline per annum over 2007 and 2008 for Australian tourism zones as classified in Appendix A.2.*

	NSW		π		VIC		π		QLD		π
AA	Metro	−4.61	*0.32*	BA	Metro	−0.05	*0.47*	CA	Metro	2.97	*0.57*
AB	Nth Coast	−2.45	*0.27*	BB	West Coast	2.05	*0.09*	CB	Central Coast	0.39	*0.16*
AC	Sth Coast	−1.02	*0.10*	BC	East Coast	−0.88	*0.13*	CC	Nth Coast	−3.17	*0.18*
AD	Sth	0.45	*0.11*	BD	Nth East	0.32	*0.16*	CD	Inland	0.02	*0.09*
AE	Nth	−0.37	*0.14*	BE	Nth West	−0.20	*0.15*				
AF	ACT	−0.13	*0.06*								
	SA		π		WA		π		TAS		π
DA	Metro	2.69	*0.46*	EA	West Coast	0.06	*0.75*	FA	Sth	−2.41	*0.47*
DB	Sth Coast	−1.57	*0.18*	EB	Nth	−3.23	*0.14*	FB	Nth East	−0.51	*0.33*
DC	Inland	5.56	*0.20*	EC	Sth	4.83	*0.11*	FC	Nth West	1.03	*0.21*
DD	West Coast	4.35	*0.16*								
	NT		π								
GA	Nth Coast	−0.79	*0.59*								
GB	Central	1.59	*0.41*								

The π entries denote the aggregate historical proportion of the corresponding zone to the aggregate historical tourism of the state or the territory the zone belongs to.

The decline forecasted for the state of New South Wales is mainly driven by a decline in the costal zones. In particular, the Metro, North Coast and South Coast zones, which comprise approximately 70% of the tourism demand in New South Wales, all show a significant decline. The only zone which shows some growth is the inland South zone. For the state of Victoria, the major contributors to the forecasted decline are the Metro, the East Coast and the North West inland zones which make up approximately 75% of tourism demand in Victoria.

The state of Queensland is the second largest contributor to domestic tourism. The overall moderate forecasted growth is driven by the growth in the Metro and the Central Coast zones. These zones comprise approximately 73% the state's domestic tourism. As we have previously mentioned, these areas are arguably the most well-known and well-developed tourist destinations in Australia. In the North Coast region of the state, a significant decline is forecasted. The other

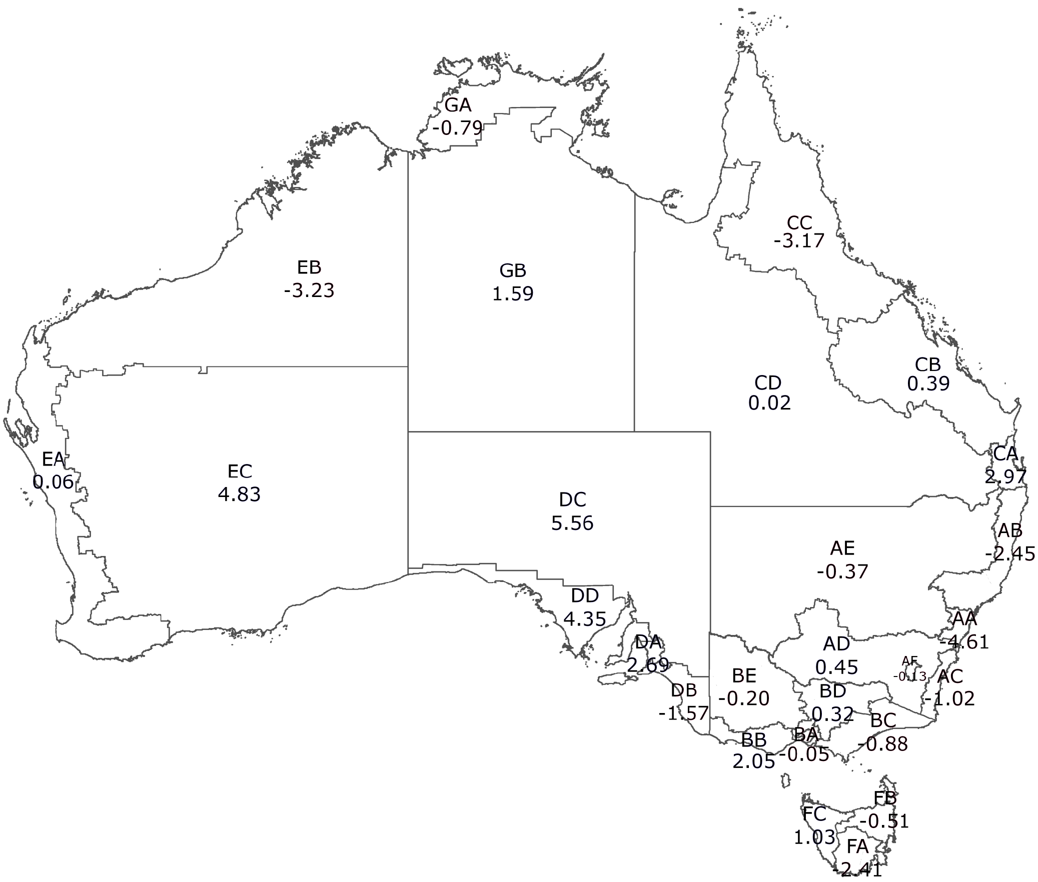


Figure 4: *Average forecast rate of growth/decline per annum over 2007 and 2008 for the states of Australia.*

state for which we also forecast growth is South Australia. Moderate to high growth is shown for most of the state with the exception of the South Coast for which we forecast a decline.

Consider the costal areas of the three main states: New South Wales, Victoria and Queensland. The zones that comprise these areas are (starting from the south of Victoria—see Figure 4): West Coast (BB), Metro (BA) and East Coast (BC) Victoria; South Coast (AC), Metro (AA) and North Coast (AB) New South Wales; Metro (CA), Central (CB) and North Coast (CC) Queensland. Based on the historical data, tourism demand in these areas comprises approximately 60% of the aggregate Australian tourism. For these areas combined, we forecast a decline of 0.74% per annum over the next two years. If we exclude the growth forecasted for Metro and Central Coast Queensland, the decline drops to 1.93% per annum for the next two years. For the rest of Australia (i.e., excluding the east cost zones), we forecast a growth of 0.39% per annum for the next two years. These results show the importance of the east cost areas to Australian domestic tourism and hence the tourism authorities should pay significant attention to this.

Table 11: *Average forecast rate of growth/decline per annum over 2007 and 2008 for Australian tourism regions as classified in Appendix A.2.*

NSW		π	VIC		π	QLD		π	SA		π
AAA	−3.58	*0.02*	BAA	−0.62	*0.36*	CAA	2.39	*0.21*	DAA	3.00	*0.42*
AAB	−4.51	*0.26*	BAB	4.66	*0.06*	CAB	1.87	*0.22*	DAB	−1.74	*0.03*
AAC	−5.71	*0.04*	BAC	−0.82	*0.05*	CAC	5.81	*0.14*	DAC	2.76	*0.01*
ABA	−3.29	*0.06*	BBA	2.05	*0.09*	CBA	1.20	*0.05*	DBA	1.76	*0.08*
ABB	−0.78	*0.14*	BCA	0.12	*0.04*	CBB	2.86	*0.03*	DBB	−4.02	*0.08*
ABC	−4.40	*0.08*	BCB	−1.02	*0.05*	CBC	−0.10	*0.05*	DBC	−2.89	*0.02*
ACA	−1.02	*0.10*	BCC	−1.55	*0.04*	CBD	−2.16	*0.03*	DCA	10.94	*0.04*
ADA	1.38	*0.03*	BDA	0.93	*0.04*	CCA	−2.59	*0.03*	DCB	9.80	*0.04*
ADB	0.34	*0.02*	BDB	−4.51	*0.02*	CCB	−7.49	*0.05*	DCC	−4.72	*0.02*
ADC	−3.13	*0.03*	BDC	−0.24	*0.06*	CCC	−1.00	*0.10*	DCD	10.04	*0.06*
ADD	0.67	*0.03*	BDD	1.29	*0.01*	CDA	−0.06	*0.04*	DCE	−2.28	*0.04*
AEA	2.67	*0.06*	BDE	10.34	*0.01*	CDB	0.12	*0.04*	DDA	4.89	*0.08*
AEB	−2.74	*0.04*	BDF	0.95	*0.02*				DDB	3.75	*0.07*
AEC	6.52	*0.02*	BEA	−2.52	*0.05*						
AED	−3.75	*0.02*	BEB	−3.05	*0.01*						
AFA	−0.13	*0.06*	BEC	3.17	*0.04*						
			BED	6.46	*0.01*						
			BEE	−5.38	*0.01*						
			BEF	−3.49	*0.02*						
			BEG	10.46	*0.02*						
WA		π	**TAS**		π	**NT**		π			
EAA	−5.54	*0.12*	FAA	−1.51	*0.38*	GAA	0.56	*0.47*			
EAB	0.50	*0.40*	FAB	−5.56	*0.08*	GAB	13.73	*0.06*			
EAC	2.79	*0.23*	FBA	0.14	*0.11*	GAC	−20.33	*0.04*			
EBA	−3.23	*0.14*	FBB	14.95	*0.05*	GAD	7.53	*0.02*			
ECA	4.83	*0.11*	FBC	−3.48	*0.17*	GBA	9.50	*0.12*			
			FCA	2.19	*0.16*	GBB	13.55	*0.04*			
			FCB	−3.44	*0.05*	GBC	12.30	*0.06*			
						GBD	−6.78	*0.17*			
						GBE	17.35	*0.03*			

The π entries denote the aggregate historical proportion of the corresponding tourism region to the aggregate historical tourism of the state or the territory the region belongs to.

In Table 11, we present the average forecasted rates of growth/decline per annum over the period 2007–2008 for the series in the bottom level of hierarchy 2, which are the tourism regions as classified in Appendix A.2. In Figure 5 we colour code the forecasted rates and plot these on the corresponding regions on the map of Australia. The darkest shade of grey is given to the regions for which our forecasted rate shows a severe decline, i.e., an average decline of more than 3% per annum over the next two years. As the forecasted rates improve, i.e., less of a decline and moving into growth, the shaded grey gets lighter. The lightest shaded regions, i.e., the regions coloured white show a significant average growth of more than 3% per annum over the next two years.

These results present a plethora of information for local governments and tourism authorities. We will only highlight some general observations. Figure 5 allows us to recognise the spatial correlations that exist around Australia. For example, consider the state of New South Wales. There is a clear clustering of inland regions (ADA, ADB, ADD, AEA and AEC) with an average positive forecasted growth rate over the next two years. On the other hand there is a clustering of costal regions (AAA, AAB, AAC, AED and ABA) with a very pessimistic average forecasted rate of decline of more than 3% per annum over the next two years. The same conclusions can be drawn for the state of Queensland. A clustering of regions with a positive forecasted growth rate can be found around Brisbane and the Gold Coast (CAA, CAB, CAC, CBA and CBB). A clustering of regions with an average forecasted decline over the next two years can be found further north (CBC, CBD, CCA, CCB and CCC).

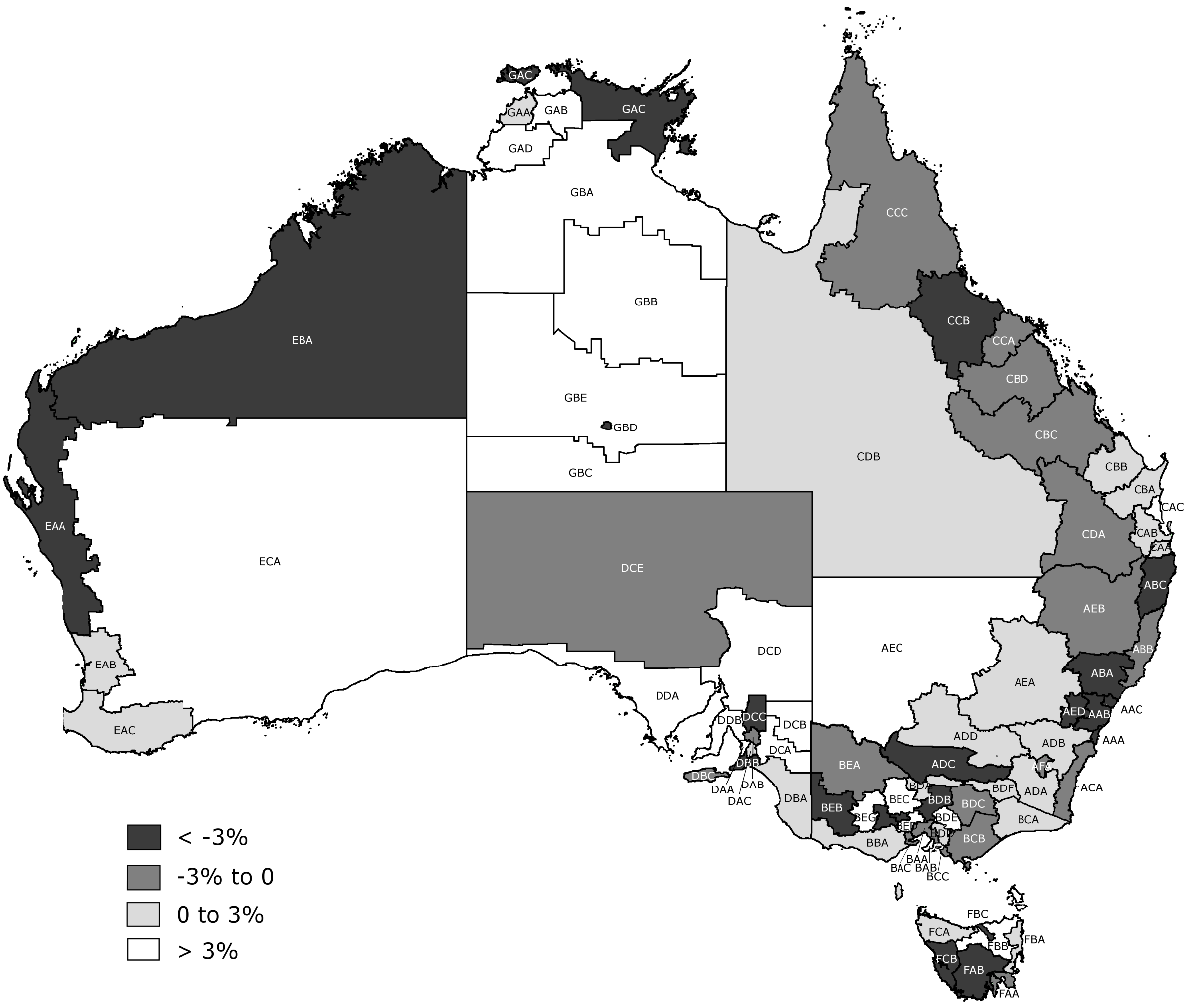


Figure 5: *Average forecast rate of growth/decline per annum over 2007 and 2008 for Australian tourism regions as classified in Appendix A.2.*

7 Summary and conclusions

In this paper we have applied hierarchical forecasting to the domestic tourism market for Australia. We have considered five methods of hierarchical forecasting. The first two are variations of the conventional top-down approach: in the first one the top-level forecasts are distributed to lower levels according to average historical proportions; and in the second approach the top-level forecasts are distributed to lower levels according to the proportions of historical averages. The third approach considered is the conventional bottom-up approach. We then consider two new approaches. Our new top-down approach improves on the conventional top-down methods by distributing the top-level forecasts to lower levels according to forecasted proportions of the lower levels and not the historical static proportions of the conventional methods. Finally, we consider the optimal combination approach recently introduced by Hyndman et al. (2007). Our evaluation of the forecast performance of all five approaches shows that the best performing method for the two tourism time series hierarchies we consider are the top-down method based on forecasted proportions and the optimal combination approach.

Our forecasts show a decline for aggregate Australian domestic tourism over the next two years. This is consistent with Athanasopoulos and Hyndman (2007) who produce longer-term forecasts for Australian domestic tourism demand. Applying the hierarchical approach has allowed us identify sources that will considerably contribute to this decline. Disaggregating the data by purpose of travel we forecast a decline in the three main purposes of travel: "Holiday", "VFR" and "Business". Geographically, the aggregate decline is mainly driven by a decline for the states of New South Wales and Victoria. In both hierarchies, Queensland is identified as a state with the highest forecast growth. Further geographical disaggregation of the data has allowed us to identify the costal areas in the east coast of Australia, with the exception of the metro and central coast of Queensland, as major contributors to the aggregate decline of Australian domestic tourism.

References

Athanasopoulos, G. and R. J. Hyndman (2007) Modelling and forecasting Australian domestic tourism, *Tourism Management*, **forthcoming**.

Dangerfield, B. J. and J. S. Morris (1992) Top-down or bottom-up: Aggregate versus disaggregate extrapolations, *International Journal of Forecasting*, **8**, 233–241.

Dunn, D. M., W. H. Williams and T. L. DeChaine (1976) Aggregate versus subaggregate models in local area forecasting, *Journal of the American Statistical Association*, **71**(353), 68–71.

Edwards, J. B. and G. H. Orcutt (1969) Should aggregation prior to estimation be the rule?, *The Review of Economics and Statistics*, **51**(4), 409–420.

Fliedner, E. B. and V. A. Mabert (1992) Constrained forecasting: some implementation guidelines, *Decision Sciences*, **23**, 1143–1161.

Fliedner, G. (1999) An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation, *Computers and Operations Research*, **26**, 1133–1149.

Fogarty, D. W., J. H. Blackstone and T. R. Hoffman (1990) *Production and Inventory Management*, South-Western Publication Co., Cincinnati, 2nd ed.

Gardner, Jr, E. S. (1985) Exponential smoothing: the state of the art, *Journal of Forecasting*, **4**, 1–28.

Gross, C. and J. Sohl (1990) Dissagregation methods to expedite product line forecasting, *Journal of Forecasting*, **9**, 233–254.

Grunfeld, Y. and Z. Griliches (1960) Is aggregation necessarily bad?, *The Review of Economics and Statistics*, **42**(1), 1–13.

Hyndman, R. J., R. A. Ahmed and G. Athanasopoulos (2007) Optimal combination forecasts for hierarchical time series, Working paper 09/07, Department of Econometrics and Business Statistics, Monash University. **URL:** *http://www.robhyndman.info/papers/hierarchical.htm*

Hyndman, R. J., A. B. Koehler, R. D. Snyder and S. Grose (2002) A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, **18**, 439–454.

Kinney, Jr., W. R. (1971) Predicting earnings: entity versus subentity data, *Journal of Accounting Research*, **9**(1), 127–136.

Lütkephol, H. (1984) Forecasting contemporaneously aggregated vector arma processes, *Journal of Business and Economic Statistics*, **2**, 201–214.

McLeavey, D. W. and S. Narasimhan (1985) *Production Planning and Inventory Control*, Allyn and Bacon, Inc.

Narasimhan, S. L., D. W. McLeavey and P. Billington (1994) *Production Planning and Inventory Control*, Allyn & Bacon, 2nd ed.

Orcutt, G. H., W. H. Watt and J. B. Edwards (1968) Data aggregation and infromation loss, *The American Economic Review*, **58**(4), 773–787.

Pegels, C. C. (1969) Exponential smoothing: some new variations, *Management Science*, **12**, 311–315.

Shing, N. K. (1993) *A study of bottom-up and top-down forecasting methods*, Master's thesis, Royal Melbourne Institute of Technology.

Shlifer, E. and R. W. Wolff (1979) Aggregation and proration in forecasting, *Management Science*, **25**(6), 594–603.

Taylor, J. (2003) Exponential smoothing with a damped multiplicative trend, *International Journal of Forecasting*, **19**, 715–725.

Theil, H. (1954) *Linear Aggregation of Economic Relations*, North-Holland Publishing Company.

Tourism Research Australia (2005) *Travel by Australians, September Quarter 2005*, Tourism Australia, Canberra.

Weatherby, G. (1984) *Aggregation, disaggregation and combination of forecasts*, Ph.D. thesis, Georgia Institute of Technology.

Zellner, A. and J. Tobias (2000) A note on aggregation, disaggregation and forecasting performance, *Journal of Forecasting*, **19**, 457–469.

A Hierarchies and Models

A.1 Hierarchy 1

Table 12: *Refer to Figure 2 on page 17 for a geographical division of Australia to the states including the capital cities.*

		Top level	*Model*
1	Total	Australia	ANA
		Level 1: Purpose of travel	
2	A	Holiday	ANA
3	B	VFR	ANA
4	C	Business	ANA
5	D	Other	ANA
		Level 2: States	
6	AA	*Hol*- NSW	AA_dA
7	AB	*Hol*- VIC	$A\ A_dA$
8	AC	*Hol*- QLD	ANA
9	AD	*Hol*- SA	ANA
10	AE	*Hol*- WA	ANA
11	AF	*Hol*- TAS	ANA
12	AG	*Hol*- NT	ANA
13	BA	*VFR*- NSW	ANA
14	BB	*VFR*- VIC	ANA
15	BC	*VFR*- QLD	ANA
16	BD	*VFR*- SA	ANA
17	BE	*VFR*- WA	ANA
18	BF	*VFR*- TAS	ANA
19	BG	*VFR*- NT	ANA
20	CA	*Bus*- NSW	ANA
21	CB	*Bus*- VIC	ANN
22	CC	*Bus*- QLD	ANA
23	CD	*Bus*- SA	ANN
24	CE	*Bus*- WA	ANA
25	CF	*Bus*- TAS	ANN
26	CG	*Bus*- NT	ANN
27	DA	*Oth*- NSW	ANA
28	DB	*Oth*- VIC	ANN
29	DC	*Oth*- QLD	ANA
30	DD	*Oth*- SA	ANN
31	DE	*Oth*- WA	ANA
32	DF	*Oth*- TAS	ANN
33	DG	*Oth*- NT	MNM
		Bottom level: Capital city versus other	
34	AAA	*Hol*- NSW - Sydney	ANA
35	AAB	*Hol*- NSW - Other	AA_dA
36	ABA	*Hol*- VIC - Melbourne	ANA
37	ABB	*Hol*- VIC - Other	AAA
38	ACA	*Hol*- QLD - Bris + GC	ANA
39	ACB	*Hol*- QLD - Other	ANA
40	ADA	*Hol*- SA - Adelaide	ANA
41	ADB	*Hol*- SA - Other	ANA
42	AEA	*Hol*- WA - Perth	ANA
43	AEB	*Hol*- WA - Other	ANA
44	AFA	*Hol*- TAS - Hobart	ANA
		Bottom level (continued)	
45	AFB	*Hol*- TAS - Other	ANA
46	AGA	*Hol*- NT - Darwin	ANA
47	AGB	*Hol*- NT - Other	ANA
48	BAA	*VFR*- NSW - Sydney	ANA
49	BAB	*VFR*- NSW - Other	ANA
50	BBA	*VFR*- VIC - Melbourne	ANA
51	BBB	*VFR*- VIC - Other	ANA
52	BCA	*VFR*- QLD - Bris + GC	ANA
53	BCB	*VFR*- QLD - Other	ANA
54	BDA	*VFR*- SA - Adelaide	ANA
55	BDB	*VFR*- SA - Other	ANA
56	BEA	*VFR*- WA - Perth	ANA
57	BEB	*VFR*- WA - Other	ANN
58	BFA	*VFR*- TAS - Hobart	ANA
59	BFB	*VFR*- TAS - Other	ANA
60	BGA	*VFR*- NT - Darwin	ANA
61	BGB	*VFR*- NT - Other	ANN
62	CAA	*Bus*- NSW - Sydney	ANA
63	CAB	*Bus*- NSW - Other	ANN
64	CBA	*Bus*- VIC - Melbourne	ANN
65	CBB	*Bus*- VIC - Other	ANN
66	CCA	*Bus*- QLD - Bris + GC	ANN
67	CCB	*Bus*- QLD - Other	ANA
68	CDA	*Bus*- SA - Adelaide	ANN
69	CDB	*Bus*- SA - Other	ANN
70	CEA	*Bus*- WA - Perth	ANA
71	CEB	*Bus*- WA - Other	ANA
72	CFA	*Bus*- TAS - Hobart	ANN
73	CFB	*Bus*- TAS - Other	ANN
74	CGA	*Bus*- NT - Darwin	ANN
75	CGB	*Bus*- NT - Other	ANN
76	DAA	*Oth*- NSW - Sydney	ANN
77	DAB	*Oth*- NSW - Other	ANA
78	DBA	*Oth*- VIC - Melbourne	ANN
79	DBB	*Oth*- VIC - Other	ANN
80	DCA	*Oth*- QLD - Bris + GC	ANN
81	DCB	*Oth*- QLD - Other	ANA
82	DDA	*Oth*- SA - Adelaide	ANN
83	DDB	*Oth*- SA - Other	ANA
84	DEA	*Oth*- WA - Perth	ANN
85	DEB	*Oth*- WA - Other	ANA
86	DFA	*Oth*- TAS - Hobart	ANN
87	DFB	*Oth*- TAS - Other	ANN
88	DGA	*Oth*- NT - Darwin	ANA
89	DGB	*Oth*- NT - Other	MNM

Note: Brisbane and the Gold Coast are considered the combined capital city for Queensland.

A.2 Hierarchy 2

Table 13: *Refer to Figure 4 on page 21 for a geographical division of Australia to the Zones and Figure 5 on page 23 for a geographical division of Australia down to the tourism Regions as shown in this table.*

		Top level	*Model*
1	Total	Australia	ANA
		Level 1: States	
2	A	NSW	AAA
3	B	VIC	ANA
4	C	QLD	ANA
5	D	SA	ANA
6	E	WA	ANA
7	F	TAS	ANA
8	G	NT	ANA
		Level 2: Zones	
9	AA	Metro NSW	ANA
10	AB	Nth Coast NSW	ANA
11	AC	Sth Coast NSW	ANA
12	AD	Sth NSW	ANA
13	AE	Nth NSW	ANN
14	AF	ACT	ANN
15	BA	Metro VIC	ANA
16	BB	West Coast VIC	ANA
17	BC	East Coast VIC	ANA
18	BD	Nth East VIC	ANA
19	BE	Nth West VIC	ANA
20	CA	Metro QLD	ANA
21	CB	Central Coast QLD	ANA
22	CC	Nth Coast QLD	ANA
23	CD	Inland QLD	ANA
24	DA	Metro SA	ANA
25	DB	Sth Coast SA	ANA
26	DC	Inland SA	ANN
27	DD	West Coast SA	ANA
28	EA	West Coast WA	ANA
29	EB	Nth WA	ANA
30	EC	Sth WA	ANN
31	FA	Sth TAS	ANA
32	FB	Nth East TAS	ANA
33	FC	Nth West TAS	ANA
34	GA	Nth Coast NT	ANA
35	GB	Central NT	ANA
		Bottom level: Regions	
36	AAA	102 Illawarra	ANA
37	AAB	104 Sydney	ANA
38	AAC	118 Central Coast	ANA
39	ABA	110 Hunter	ANA
40	ABB	112 Nth Coast NSW + 120 Lrd Howe Isl	ANA
41	ABC	113 Northern Rivers Tropical NSW	ANA
42	ACA	101 South Coast	ANA
43	ADA	105 Snowy Mountains	ANA
44	ADB	106 Capital Country	ANA
45	ADC	107 The Murray	ANA
46	ADD	108 Riverina	ANN
47	AEA	109 Explorer Country	ANN
48	AEB	114 New England North West	ANA
49	AEC	115 Outback NSW	ANN
50	AED	119 Blue Mountains	ANN
51	AFA	117 Canberra	ANN
52	BAA	201 Melbourne	ANA
53	BAB	207 Peninsula	ANA
54	BAC	214 Geelong	ANA
55	BBA	204 Western	ANA
56	BCA	211 Lakes	ANA
57	BCB	212 Gippsland	ANA

		Bottom level (continued)	
58	BCC	221 Phillip Island	ANA
59	BDA	208 Central Murray	ANA
60	BDB	209 Goulburn	ANN
61	BDC	210 High Country	ANA
62	BDD	213 Melbourne East	ANN
63	BDE	219 Upper Yarra	ANA
64	BDF	220 Murray East	ANN
65	BEA	202 Wimmera + 203 Mallee	ANN
66	BEB	205 Western Grampians	ANN
67	BEC	206 Bendigo Loddon	ANA
68	BED	215 Macedon	ANN
69	BEE	216 Spa Country	ANN
70	BEF	217 Ballarat	ANN
71	BEG	218 Central Highlands	ANA
72	CAA	301 Gold Coast	ANA
73	CAB	302 Brisbane	ANA
74	CAC	303 Sunshine Coast	ANA
75	CBA	304 Hervey Bay/Maryborough	ANA
76	CBB	307 Bundaberg	ANA
77	CBC	308 Fitzroy	ANA
78	CBD	309 Mackay	ANA
79	CCA	310 Whitsundays	ANN
80	CCB	311 Northern	ANA
81	CCC	312 Tropical North Queensland	ANA
82	CDA	306 Darling Downs	ANN
83	CDB	314 Outback	ANA
84	DAA	404 Adelaide	ANA
85	DAB	405 Barossa	ANN
86	DAC	408 Adelaide Hills	ANN
87	DBA	401 Limestone Coast	ANA
88	DBB	403 Fleurieu Peninsula	ANA
89	DBC	413 Kangaroo Island	ANA
90	DCA	402 Murraylands	ANA
91	DCB	406 Riverland	ANN
92	DCC	407 Clare Valley	ANN
93	DCD	409 Flinders Ranges	ANA
94	DCE	410 Outback SA	ANA
95	DDA	411 Eyre Peninsula	ANA
96	DDB	412 Yorke Peninsula	ANA
97	EAA	550 Australia's Coral Coast	ANA
98	EAB	553 Experience Perth	ANA
99	EAC	552 Australia's South West	ANA
100	EBA	551 Australia's North West	ANA
101	ECA	554 Australia's Golden Outback	ANN
102	FAA	601 Greater Hobart	ANA
103	FAB	602 Southern	ANA
104	FBA	603 East Coast	ANA
105	FBB	604 Northern	ANA
106	FBC	605 Greater Launceston	ANA
107	FCA	606 North West	ANA
108	FCB	607 West Coast	ANA
109	GAA	801 Darwin	ANA
110	GAB	802 Kakadu	ANN
111	GAC	803 Arnhem	ANN
112	GAD	809 Daly	ANA
113	GBA	804 Katherine	ANA
114	GBB	805 Tablelands	ANA
115	GBC	806 Petermann	ANA
116	GBD	807 Alice Springs	ANA
117	GBE	808 MacDonnell	ANA