Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 25 (2009) 146–166
international journal of forecasting
www.elsevier.com/locate/ijforecast

Hierarchical forecasts for Australian domestic tourism

George Athanasopoulos*, Roman A. Ahmed, Rob J. Hyndman

Department of Econometrics and Business Statistics, Monash University, VIC 3800, Australia

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 according to geographical regions and 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 the 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 of 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, and use hierarchical forecasting methods to produce forecasts of the Australian domestic tourism market for several levels of disaggregation.

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

* Corresponding author.
E-mail addresses: George.Athanasopoulos@buseco.monash.edu.au (G. Athanasopoulos), Roman.Ahmed@buseco.monash.edu.au (R.A. Ahmed), Rob.Hyndman@buseco.monash.edu.au (R.J. Hyndman).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2008.07.004

purpose of travel and then by geographical region. The purposes of travel we consider are: holiday, visiting friends and relatives (VFR), business, and other. In the second hierarchy, we disaggregate the data by 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 performances of these two methods, with some favouring the top-down approaches (see for example Fliedner (1999), Fogarty, Blackstone, and Hoffman (1990), Grunfeld and Griliches (1960) and Narasimhan, McLeavey, and Billington (1994)), while others favour the bottom-up approaches (see for example Dangerfield and Morris (1992), Edwards and Orcutt (1969), Kinney (1971), Orcutt, Watt, and Edwards (1968) and Zellner and Tobias (2000)), and still others find neither method to be uniformly superior (see for example Fliedner and Mabert (1992), Shing (1993) and Weatherby (1984)). In Section 3 we introduce a 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 proportions, rather than the conventional historical (and therefore static) proportions. Second, we present the newly proposed "optimal combination" approach of Hyndman, Ahmed, and Athanasopoulos (2007). The optimal combination approach is based on forecasting all series at all levels, and then using a regression model to obtain the minimum variance unbiased combination of the 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, and at each level of the hierarchies, we obtain forecasts using a single source of error (or innovations) state space model (see Aoki (1987)). These models have been very successful when applied to data from forecasting competitions (e.g., Hyndman, Koehler, Snyder, and Grose (2002) and Makridakis and Hibon (2000)). Considering regional tourism demand and tourism demand by purpose of travel allows 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 of the series. Both the innovations state space models we use to forecast the individual series, and the hierarchical forecasting approaches we apply to combine these forecasts, are novel to the tourism literature. Neither appeared in the comprehensive reviews of Li, Song, and Witt (2005) and Song and Li (2008), which jointly cover tourism forecasting studies from 1990 to 2006. The only exception is Athanasopoulos and Hyndman (2008), who also use innovations state space models as a time series alternative to econometric modelling approaches, and forecast Australian domestic tourism demand at an aggregate level, and also by purpose of travel.

In order to evaluate the performance of the alternative hierarchical approaches, we perform an out-of-sample forecast evaluation in Section 5. We also test for significant differences in predictive accuracy between the approaches by applying the Diebold–Mariano test. 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 Fig. 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 Fig. 1 is a $K = 2$ level hierarchy. Let $Y_{X,t}$ be the tth observation ($t = 1, \ldots, n$) of series Y_X, which corresponds to node X on the hierarchical tree.