Forecasting in a Mixed Up World: Nowcasting Hawaii Tourism

Ashley Hirashima, James Jones, Carl S. Bonham, Peter Fuleky*

University of Hawaii Economic Research Organization and Department of Economics, University of Hawaii, 540 Saunders Hall, 2424 Maile Way, Honolulu, HI 96822, United States

1. Introduction

World wide tourism growth has led to an extensive literature seeking to predict visitor volumes. In many destinations, predictions of tourism activity are used in planning and business operations. Because tourism services are perishable, firms rely on forecasts of tourism demand to efficiently manage supply and prices. Tourism agencies use forecasts to design promotional strategies and set performance targets; other governmental organizations incorporate tourism forecasts into larger macroeconomic models. When data are published with a long release lag, agencies may be interested in predicting current or real time demand, referred to as nowcasting, or even predicting the recent past, referred to as backcasting. Both concepts fall under the general concept of forecasting.

Tourism activity is often forecast using a wide variety of indicators published at different frequencies and with different release lags. In practice, the data is usually transformed to a single frequency. But the aggregation process eliminates valuable intra-period information that could be used to update the forecast. Nowcasting performance, in particular, may be influenced by use of the most recent and highest frequency data. Such practical issues lead to two questions: (1) how can data released with different lags and frequencies be combined in the generation of multi-period forecasts, and (2) what benefits can be derived from such combinations.

There exists a wide range of methods used to forecast tourism demand, but the method that offers the best performance varies by application. In two meta-studies, Li, Song, and Witt (2005) and Song and Li (2008) review 22 and 55 published arti-
cles, respectively, that compare alternative forecasting methods in tourism applications. Both studies conclude that no single method dominates all others. In a more recent meta-study, Peng, Song, and Crouch (2014) review 262 articles published between 1980 and 2011 and reach essentially the same conclusions noting that forecast performance is data- and context-dependent.

The most popular methods used to forecast tourism demand can be classified into two groups: univariate time series models and multivariate regressions. Univariate methods, like Exponential Smoothing (ES) and Autoregressive Integrated Moving Average (ARIMA) models, use only the history of the variable of interest for prediction. In contrast econometric models, like Autoregressive Distributed Lag (ARDL) and Error Correction (EC) models, incorporate information from a set of explanatory variables. We provide a brief outline of several of these models and their use in the tourism literature in Section 2. Although in recent years several non-regression based approaches have appeared in the tourism forecasting literature, many using Artificial Intelligence models, these are beyond the scope of our study (for more information on these methods see Kon & Turner (2005)).

A common feature of conventional models is that they operate at a single frequency, with the variable of interest and any explanatory variables aggregated to the same frequency before estimation. Consider predicting quarterly tourist arrivals using the monthly number of inbound airline passengers as an explanatory variable. If passenger counts are highly correlated with tourist arrivals, then the monthly passenger counts may contain substantial predictive power. For example, if the forecaster observes that passenger counts were much higher than expected after the first month of the quarter, she may want to incorporate this information to update the tourist arrivals forecast for the quarter. However with a single frequency model the passenger count series must be aggregated to the quarterly frequency, and the intra-quarter information cannot be used for prediction. Some forecasters make ad hoc adjustments to their short term forecasts to incorporate information from incomplete periods at the end of sample. But this process can be cumbersome and, by definition, subjective. Using a model-based approach to incorporate high frequency information streamlines the forecasting process and may improve forecasting accuracy.

Several approaches have been recently developed to directly use high frequency regressors to predict a low frequency variable of interest. This is a rapidly growing area of research with over 50 studies in the last decade (see for example Camacho, Perez-Quiros, & Poncela, 2013). The simple example above could be formulated as an Unrestricted Mixed Data Sampling (U-MIDAS) model developed by Foroni, Marcellino, and Schumacher (2015). U-MIDAS uses high frequency regressors to predict a low frequency variable, so that monthly passenger counts can be directly used to predict quarterly tourist arrivals without any aggregation. Consequently, all of the monthly information can be used as soon as it is available. However the U-MIDAS model results in parameter proliferation when there is a large frequency mismatch between the variable of interest and the high frequency regressors. The Mixed Data Sampling (MIDAS) method of Ghysels, Santa-Clara, and Valkanov (2004) and Ghysels, Sinko, and Valkanov (2007) addresses this problem by imposing non-linear restrictions on the model, but it sacrifices the simplicity of ordinary least squares parameter estimation. Foroni et al. (2015) showed that when the frequency mismatch is small, as in the case of using monthly data to predict a quarterly series, U-MIDAS models tend to outperform MIDAS models.

Mixed frequency forecasting methods have been used extensively in macroeconomic applications. The survey by Camacho et al. (2013), investigating a variety of short-term forecasting methods, and the study by Jansen, Jin, and de Winter (2012), evaluating eleven different models to forecast real GDP for several European countries, found that incorporating high frequency information improves predictive accuracy for the current period (nowcasting), but gains in forecasting one or two periods ahead appear to be muted. Baumeister, Guérin, and Kilian (2015) noted that forecast precision depends on whether the high frequency data provides a useful signal or simply introduces additional noise.

Despite their appeal, mixed-frequency forecasting methods have been used only by few tourism studies. In fact, we are aware of only one such study: Bangwayo-Skeete and Skeete (2015) used weekly Google search data in a MIDAS model to predict monthly tourist arrivals to several Caribbean destinations. They found that the mixed frequency model outperformed several univariate methods. A related study by Jackman and Naitram (2015) used weekly Google Trends data and air passenger arrivals to produce short term weekly forecasts (nowcasts) of tourist arrivals in Barbados. The study suggested that high frequency Google Trends data can significantly improve forecasting performance in some markets, but did not use mixed-frequency methods.

Li, Pan, Law, and Huang (2017) forecast tourism demand with a composite search index that they constructed using a generalized dynamic factor model. To deal with data sampled at different frequencies, the authors converted monthly tourist volumes into a weekly series and used it alongside weekly tourism search engine data. They found that this method performed better than a traditional time series model or a model using a principal component analysis based index. Mixed frequency factor models based on the Kalman filter have also been used for prediction (see for example Fuleky & Bonham, 2015), but they are beyond the scope of our study. While Bai, Ghysels, and Wright (2013) found that factor models and the MIDAS models analyzed in our paper have similar forecasting performances, the implementation of the former is more complex.

Our study contributes to the tourism forecasting literature by providing a thorough explanation of forecasting methods that systematically incorporate high frequency information via covariates. We evaluate the performance of these methods in two tourism specific applications and present practical implementation guidelines to facilitate their adoption. Our results indicate that the incorporation of timely intra-period data in the forecasting process results in significant gains in predictive accuracy.
دریافت فوری
متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات