Taylor & Francis
Taylor & Francis Group

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://www.tandfonline.com/loi/uasa20

Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing

Alysha M. De Livera, Rob J. Hyndman & Ralph D. Snyder

To cite this article: Alysha M. De Livera, Rob J. Hyndman & Ralph D. Snyder (2011) Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing, Journal of the American Statistical Association, 106:496, 1513-1527, DOI: 10.1198/jasa.2011.tm09771

To link to this article: http://dx.doi.org/10.1198/jasa.2011.tm09771

Published online: 24 Jan 2012.

Submit your article to this journal

Article views: 5261

View related articles

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=uasa20

Download by: [University of Washington Libraries] **Date:** 30 March 2016, At: 13:41

Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing

Alysha M. DE LIVERA, Rob J. HYNDMAN, and Ralph D. SNYDER

An innovations state space modeling framework is introduced for forecasting complex seasonal time series such as those with multiple seasonal periods, high-frequency seasonality, non-integer seasonality, and dual-calendar effects. The new framework incorporates Box–Cox transformations, Fourier representations with time varying coefficients, and ARMA error correction. Likelihood evaluation and analytical expressions for point forecasts and interval predictions under the assumption of Gaussian errors are derived, leading to a simple, comprehensive approach to forecasting complex seasonal time series. A key feature of the framework is that it relies on a new method that greatly reduces the computational burden in the maximum likelihood estimation. The modeling framework is useful for a broad range of applications, its versatility being illustrated in three empirical studies. In addition, the proposed trigonometric formulation is presented as a means of decomposing complex seasonal time series, and it is shown that this decomposition leads to the identification and extraction of seasonal components which are otherwise not apparent in the time series plot itself.

KEY WORDS: Fourier series; Prediction intervals; Seasonality; State space models; Time series decomposition.

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

1. INTRODUCTION

Many time series exhibit complex seasonal patterns. Some, most commonly weekly series, have patterns with a non-integer period. Weekly U.S. finished motor gasoline products in thousands of barrels per day, as shown in Figure 1(a), has an annual seasonal pattern with period $365.25/7 \approx 52.179$.

Other series have high-frequency multiple seasonal patterns. The number of retail banking call arrivals per 5-minute interval between 7:00 a.m. and 9:05 p.m. each weekday, as depicted in Figure 1(b), has a daily seasonal pattern with period 169 and a weekly seasonal pattern with period $169 \times 5 = 845$. A longer version of this series might also exhibit an annual seasonal pattern. Further examples where such multiple seasonal patterns can occur include daily hospital admissions, requests for cash at ATMs, electricity and water usage, and access to computer web sites.

Yet other series may have dual-calendar seasonal effects. Daily electricity demand in Turkey over nine years from 1 January 2000 to 31 December 2008, shown in Figure 1(c), has a weekly seasonal pattern and two annual seasonal patterns: one for the Hijri calendar with a period of 354.37; and the other for the Gregorian calendar with a period of 365.25. The Islamic Hijri calendar is based on lunar cycles and is used for religious activities and related holidays. It is approximately 11 days shorter than the Gregorian calendar. The Jewish, Hindu, and Chinese calendars create similar effects that can be observed in time series affected by cultural and social events (e.g., electricity demand, water usage, and other related consumption data), and need to be accounted for in forecasting studies (Lin and Liu 2002; Riazuddin and Khan 2005). Unlike the multiple periods seen with hourly and daily data, these dual calendar effects involve non-nested seasonal periods.

Most existing time series models are designed to accommodate simple seasonal patterns with a small integer-valued period (such as 12 for monthly data or 4 for quarterly data). Important exceptions (Harvey and Koopman 1993; Harvey, Koopman, and Riani 1997; Taylor 2003, 2010b; Pedregal and Young 2006; Gould et al. 2008; Taylor and Snyder 2009) handle some but not all of the above complexities. Harvey, Koopman, and Riani (1997), for example, used a trigonometric approach for single seasonal time series within a traditional multiple source of error state space framework. The single source of error approach adopted in this article is similar in some respects, but admits a larger effective parameter space with the possibility of better forecasts (see Hyndman et al. 2008, chap. 13), allows for multiple nested and non-nested seasonal patterns, and handles potential nonlinearities. The articles by Pedregal and Young (2006) and Harvey and Koopman (1993) have models for double seasonal time series, but they have not been sufficiently developed for time series with more than two seasonal patterns, and are not capable of accommodating the nonlinearity found in many time series in practice. Similarly, in modeling complex seasonality, the existing exponential smoothing models (e.g., Taylor 2003, 2010b; Gould et al. 2008; Taylor and Snyder 2009) suffer from various weaknesses such as overparameterization, and the inability to accommodate both non-integer period and dual-calendar effects. In contrast, we introduce a new innovations state space modeling framework based on a trigonometric formulation which is capable of tackling all of these seasonal complexities. Using the time series in Figure 1, we demonstrate the versatility of the proposed approach for forecasting and decomposition.

In Section 2.1 we review the existing seasonal innovations state space models including an examination of their weaknesses, particularly in relation to complex seasonal patterns. We then introduce in Sections 2.2 and 2.3 two generalizations designed to overcome some or all of these problems, one relying

Alysha M. De Livera is Research Fellow, Faculty of Science, The University of Melbourne, Victoria 3010, Australia (E-mail: *alyshad@unimelb.edu.au*). Rob J. Hyndman is Professor, Department of Econometrics and Business Statistics, Monash University, Victoria 3800, Australia (E-mail: *rob.hyndman@monash.edu*). Ralph D. Snyder is Associate Professor, Department of Econometrics and Business Statistics, Monash University, Victoria 3800, Australia (E-mail: *ralph.snyder@monash.edu*). The first author acknowledges the support provided by the Commonwealth Scientific and Industrial Research Organisation, Australia. The authors thank Dr. Peter Toscas from the Commonwealth Scientific and Industrial Research Organisation, Australia, the editor, the associate editor, and two referees for comments that improved the clarity and quality of the article.

© 2011 American Statistical Association
Journal of the American Statistical Association
December 2011, Vol. 106, No. 496, Theory and Methods
DOI: 10.1198/jasa.2011.tm09771

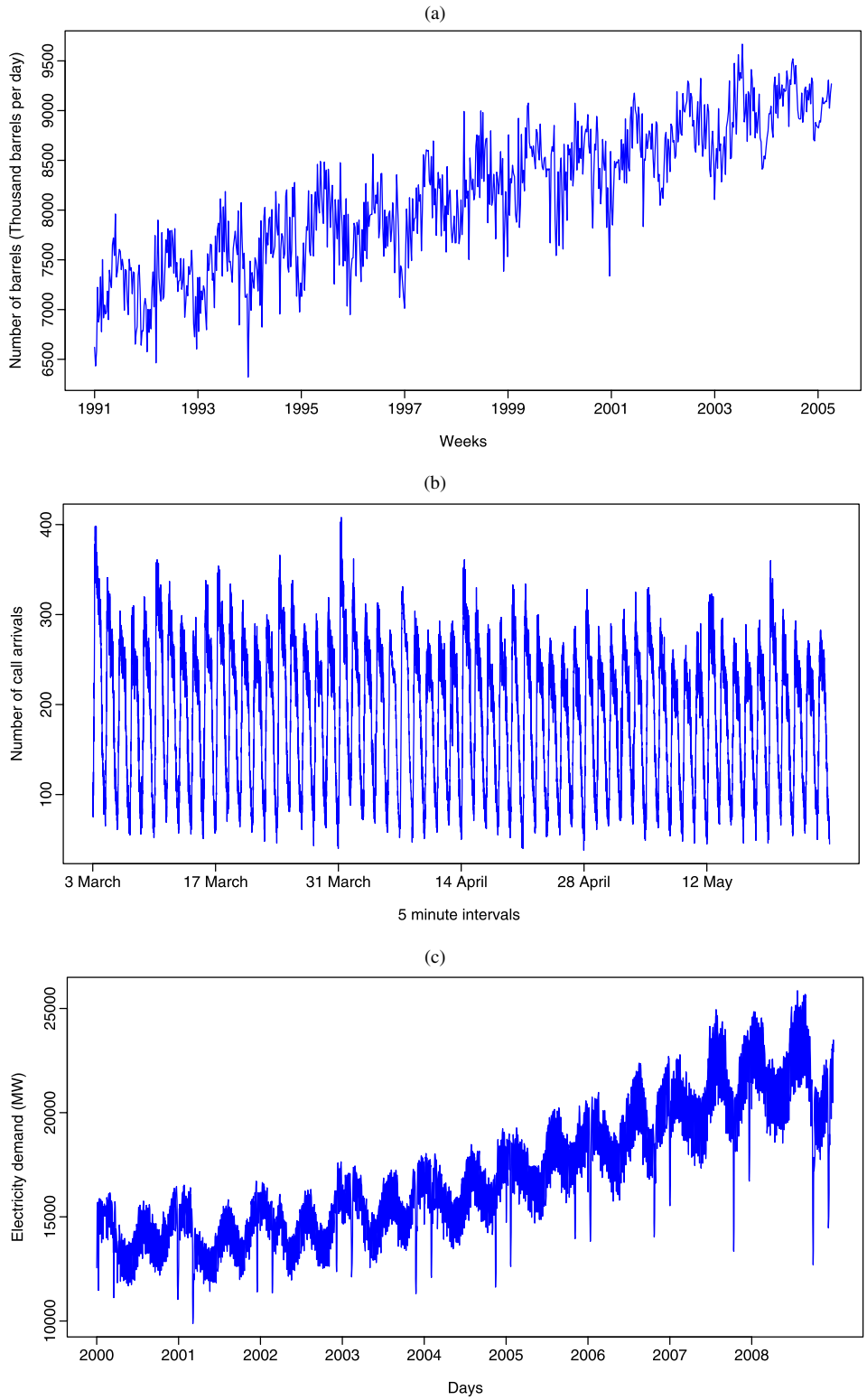

Figure 1. Examples of complex seasonality showing (a) non-integer seasonal periods, (b) multiple nested seasonal periods, and (c) multiple non-nested and non-integer seasonal periods. (a) U.S. finished motor gasoline products supplied (thousands of barrels per day), from February 1991 to July 2005. (b) Number of call arrivals handled on weekdays between 7 a.m. and 9:05 p.m. from March 3, 2003, to May 23, 2003 in a large North American commercial bank. (c) Turkish electricity demand data from January 1, 2000, to December 31, 2008. The online version of this figure is in color.

on trigonometric representations for handling complex as well as the usual single seasonal patterns in a straightforward manner with fewer parameters. Section 3 contains a new method for the calculation of maximum likelihood estimators, formulas for point and interval predictions, and the description of the model selection methodology. It will be seen that the proposed

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

estimation procedure is sufficiently general to be applied to *any* innovations state space model while possessing some important advantages over an existing approach. The proposed models are then applied in Section 4 to the time series from Figure 1 where it will be seen that the trigonometric formulation leads to better forecasts and may be used for decomposition. Some conclusions are drawn in Section 5.

2. EXPONENTIAL SMOOTHING MODELS FOR SEASONAL DATA

2.1 Traditional Approaches

Single seasonal exponential smoothing methods, which are among the most widely used forecasting procedures in practice (Makridakis et al. 1982; Makridakis and Hibon 2000; Snyder, Koehler, and Ord 2002), have been shown to be optimal for a class of innovations state space models (Ord, Koehler, and Snyder 1997; Hyndman et al. 2002). They are therefore best studied in terms of this framework because it then admits the possibility of likelihood calculation, the derivation of consistent prediction intervals, and model selection based on information criteria. The single source of error (innovations) state space model is an alternative to its common multiple source of error analogue (Harvey 1989) but it is simpler, more robust, and has several other advantages (Hyndman et al. 2008).

The most commonly employed seasonal models in the innovations state space framework include those underlying the well-known Holt–Winters additive and multiplicative methods. Taylor (2003) extended the linear version of the Holt–Winters method to incorporate a second seasonal component as follows:

$$y_t = \ell_{t-1} + b_{t-1} + s_t^{(1)} + s_t^{(2)} + d_t, \tag{1a}$$

$$\ell_t = \ell_{t-1} + b_{t-1} + \alpha d_t, \tag{1b}$$

$$b_t = b_{t-1} + \beta d_t, \tag{1c}$$

$$s_t^{(1)} = s_{t-m_1}^{(1)} + \gamma_1 d_t, \tag{1d}$$

$$s_t^{(2)} = s_{t-m_2}^{(2)} + \gamma_2 d_t, \tag{1e}$$

where m_1 and m_2 are the periods of the seasonal cycles and d_t is a white-noise random variable representing the prediction error (or disturbance). The components ℓ_t and b_t represent the level and trend components of the series at time t, respectively, and $s_t^{(i)}$ represents the ith seasonal component at time t. The coefficients α, β, γ_1, and γ_2 are the so-called smoothing parameters, and ℓ_0, b_0, $\{s_{1-m_1}^{(1)}, \ldots, s_0^{(1)}\}$, and $\{s_{1-m_2}^{(2)}, \ldots, s_0^{(2)}\}$ are the initial state variables (or "seeds").

The parameters and seeds must be estimated, but this can be difficult when the number of seasonal components is large. This problem is partly addressed by noting that there is a redundancy when m_2 is an integer multiple of m_1, something that seems to have previously gone unnoticed. Consider a time series $\{r_t\}$ consisting of repeated sequences of the constants $c_1, \ldots, c_{m_1}$, one for each season in the smaller cycle. Then the seasonal equations can be written as

$$s_t^{(1)} + r_t = \left(s_{t-m_1}^{(1)} + r_t\right) + \gamma_1 d_t, \tag{2a}$$

$$s_t^{(2)} - r_t = \left(s_{t-m_2}^{(2)} - r_t\right) + \gamma_2 d_t. \tag{2b}$$

When these are summed, the effect of r_t disappears. This suggests that the m_1 seed seasonal effects for the smaller seasonal cycle can be set to zero without constraining the problem in any way. Alternatively, each sub-season repeats itself m_2/m_1 times within the longer seasonal pattern. We can impose the constraint that the seed seasonal effects associated with each sub-season must sum to zero. For example, the period 10:00–11:00 a.m. repeats itself seven times in a week. We can insist that the seven seasonal effects associated with this particular hour sum to zero and that this is repeated for each of the 24 hour periods in a day. Analogues of these restrictions can be developed when there are three or more seasonal patterns.

Despite this correction, a large number of initial seasonal values remain to be estimated when some of the seasonal patterns have large periods, and such a model is likely to be overparameterized. For double seasonal time series Gould et al. (2008) attempted to reduce this problem by dividing the longer seasonal length into sub-seasonal cycles that have similar patterns. However, their adaptation is relatively complex and can only be used for double seasonal patterns where one seasonal length is a multiple of the other. To avoid the potentially large optimization problem, the initial states are usually approximated with various heuristics (Taylor 2003, 2010b; Gould et al. 2008), a practice that does not lead to optimized seed states. We will propose an alternative estimation method, one that relies on the principle of least squares to obtain optimized seed states—see Section 3.

A further problem is that none of the approaches based on (1) can be used to handle complex seasonal patterns such as non-integer seasonality and calendar effects, or time series with non-nested seasonal patterns. One of our proposed models will allow for all these features.

The nonlinear versions of the state space models underpinning exponential smoothing, although widely used, suffer from some important weaknesses. Akram, Hyndman, and Ord (2009) showed that most nonlinear seasonal versions can be unstable, having infinite forecast variances beyond a certain forecasting horizon. For some of the multiplicative error models which do not have this flaw, Akram, Hyndman, and Ord (2009) proved that sample paths will converge almost surely to zero even when the error distribution is non-Gaussian. Furthermore, for nonlinear models, analytical results for the prediction distributions are not available.

The models used for exponential smoothing assume that the error process $\{d_t\}$ is serially uncorrelated. However, this may not always be the case. In an empirical study, using the Holt–Winters method for multiplicative seasonality, Chatfield (1978) showed that the error process there is correlated and can be described by an AR(1) process. Taylor (2003), in a study of electricity demand forecasting using a double-seasonal Holt–Winters multiplicative method, found a similar problem. Others such as Gardner (1985), Reid (1975), and Gilchrist (1976) have also mentioned this issue of correlated errors, and the possibility of improving forecast accuracy by explicitly modeling it. The source of this autocorrelation may be due to features of the series not explicitly allowed for in the specification of the states. Annual seasonal effects may impact on the call center data, for example, but the limited sample size means that it cannot be explicitly modeled.

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

2.2 Modified Models

We now consider various modifications of the state space models for exponential smoothing to handle a wider variety of seasonal patterns, and to also deal with the problems raised above.

To avoid the problems with nonlinear models that are noted above, we restrict attention to linear homoscedastic models but allow some types of nonlinearity using Box–Cox transformations (Box and Cox 1964). This limits our approach to only positive time series, but most series of interest in practice are positive. The notation $y_t^{(\omega)}$ is used to represent Box–Cox transformed observations with the parameter ω, where y_t is the observation at time t.

We can extend model (1) to include a Box–Cox transformation, ARMA errors, and T seasonal patterns as follows:

$$y_t^{(\omega)} = \begin{cases} \dfrac{y_t^{\omega} - 1}{\omega}, & \omega \neq 0, \\ \log y_t, & \omega = 0, \end{cases} \tag{3a}$$

$$y_t^{(\omega)} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t, \tag{3b}$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha d_t, \tag{3c}$$

$$b_t = (1-\phi) b + \phi b_{t-1} + \beta d_t, \tag{3d}$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t, \tag{3e}$$

$$d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t, \tag{3f}$$

where $m_1, \ldots, m_T$ denote the seasonal periods, ℓ_t is the local level in period t, b is the long-run trend, b_t is the short-run trend in period t, $s_t^{(i)}$ represents the ith seasonal component at time t, d_t denotes an ARMA(p, q) process, and ε_t is a Gaussian white-noise process with zero mean and constant variance σ^2. The smoothing parameters are given by α, β, and γ_i for $i = 1, \ldots, T$. We adopt the Gardner and McKenzie (1985) damped trend with damping parameter ϕ, but follow the suggestion in the article by Snyder (2006) to supplement it with a long-run trend b. This change ensures that predictions of future values of the short-run trend b_t converge to the long-run trend b instead of zero. The damping factor is included in the level and measurement equations as well as the trend equation for consistency with the work of Gardner and McKenzie (1985), but identical predictions are obtained (see Snyder 2006) if it is excluded from the level and measurement equations.

The identifier BATS is an acronym for key features of the model shown by (3): Box–Cox transform, ARMA errors, Trend, and Seasonal components. It is supplemented with arguments $(\omega, \phi, p, q, m_1, m_2, \ldots, m_T)$ to indicate the Box–Cox parameter, damping parameter, ARMA parameters (p and q), and the seasonal periods $(m_1, \ldots, m_T)$. For example, BATS$(1, 1, 0, 0, m_1)$ represents the underlying model for the well-known Holt–Winters additive single seasonal method. The double seasonal Holt–Winters additive seasonal model described by Taylor (2003) is given by BATS$(1, 1, 0, 0, m_1, m_2)$, and that with the residual AR(1) adjustment in the model of Taylor (2003, 2008) is given by BATS$(1, 1, 1, 0, m_1, m_2)$. The Holt–Winters additive triple seasonal model with AR(1) adjustment in the article by Taylor (2010b) is given by BATS$(1, 1, 1, 0, m_1, m_2, m_3)$.

The BATS model is the most obvious generalization of the traditional seasonal innovations models to allow for multiple seasonal periods. However, it cannot accommodate non-integer seasonality, and it can have a very large number of states; the initial seasonal component alone contains m_T nonzero states. This becomes a huge number of values for seasonal patterns with high periods.

2.3 Trigonometric Seasonal Models

In the quest for a more flexible parsimonious approach, we introduce the following trigonometric representation of seasonal components based on Fourier series (Harvey 1989; West and Harrison 1997):

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)}, \tag{4a}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t, \tag{4b}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t, \tag{4c}$$

where $\gamma_1^{(i)}$ and $\gamma_2^{(i)}$ are smoothing parameters and $\lambda_j^{(i)} = 2\pi j/m_i$. We describe the stochastic *level* of the ith seasonal component by $s_{j,t}^{(i)}$, and the stochastic *growth* in the *level* of the ith seasonal component that is needed to describe the change in the seasonal component over time by $s_{j,t}^{*(i)}$. The number of harmonics required for the ith seasonal component is denoted by k_i. The approach is equivalent to index seasonal approaches when $k_i = m_i/2$ for even values of m_i, and when $k_i = (m_i - 1)/2$ for odd values of m_i. It is anticipated that most seasonal components will require fewer harmonics, thus reducing the number of parameters to be estimated. A deterministic representation of the seasonal components (e.g., Abraham and Box 1978) can be obtained by setting the smoothing parameters equal to zero.

A new class of innovations state space models is obtained by replacing the seasonal component $s_t^{(i)}$ in Equation (3) by the trigonometric seasonal formulation, and the measurement equation by $y_t^{(\omega)} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + d_t$. This class is designated by TBATS, the initial T connoting "trigonometric." To provide more details about their structure, this identifier is supplemented with relevant arguments to give the designation TBATS$(\omega, \phi, p, q, \{m_1, k_1\}, \{m_2, k_2\}, \ldots, \{m_T, k_T\})$.

A TBATS model requires the estimation of $2(k_1 + k_2 + \cdots + k_T)$ initial seasonal values, a number which is likely to be much smaller than the number of seasonal seed parameters in a BATS models. Because it relies on trigonometric functions, it can be used to model non-integer seasonal frequencies. A TBATS model should be distinguished from two other related (Proietti 2000) multiple source of error seasonal formulations presented by Hannan, Terrel, and Tuckwell (1970) and Harvey (1989). Some of the key advantages of the TBATS modeling framework are: (i) Being an innovations state space model, it admits a larger effective parameter space with the possibility of better forecasts (Hyndman et al. 2008, chap. 13); (ii) it allows for the accommodation of nested and non-nested multiple seasonal components; (iii) it handles typical nonlinear features that are

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

often seen in real time series; (iv) it allows for any autocorrelation in the residuals to be taken into account; and (v) it involves a much simpler, yet efficient estimation procedure (see Section 3).

2.4 Innovations State Space Formulations

The above models are special cases of the linear innovations state space model (Anderson and Moore 1979) adapted here to incorporate the Box–Cox transformation to handle nonlinearities. It then has the form

$$y_t^{(\omega)} = \mathbf{w}'\mathbf{x}_{t-1} + \varepsilon_t, \tag{5a}$$

$$\mathbf{x}_t = \mathbf{F}\mathbf{x}_{t-1} + \mathbf{g}\varepsilon_t, \tag{5b}$$

where $\mathbf{w}'$ is a row vector, $\mathbf{g}$ is a column vector, $\mathbf{F}$ is a matrix, and $\mathbf{x}_t$ is the unobserved state vector at time t.

2.4.1 TBATS Model. The state vector for the TBATS model with a nonstationary growth term can be defined as $\mathbf{x}_t = (\ell_t, b_t, \mathbf{s}_t^{(1)}, \ldots, \mathbf{s}_t^{(T)}, d_t, d_{t-1}, \ldots, d_{t-p+1}, \varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q+1})'$ where $\mathbf{s}_t^{(i)}$ is the row vector $(s_{1,t}^{(i)}, s_{2,t}^{(i)}, \ldots, s_{k_i,t}^{(i)}, s_{1,t}^{*(i)}, s_{2,t}^{*(i)}, \ldots, s_{k_i,t}^{*(i)})$. Let $\mathbf{1}_r = (1, 1, \ldots, 1)$ and $\mathbf{0}_r = (0, 0, \ldots, 0)$ be row vectors of length r; let $\boldsymbol{\gamma}_1^{(i)} = \gamma_1^{(i)}\mathbf{1}_{k_i}$, $\boldsymbol{\gamma}_2^{(i)} = \gamma_2^{(i)}\mathbf{1}_{k_i}$, $\boldsymbol{\gamma}^{(i)} = (\boldsymbol{\gamma}_1^{(i)}, \boldsymbol{\gamma}_2^{(i)})$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}^{(1)}, \ldots, \boldsymbol{\gamma}^{(T)})$, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_p)$, and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)$; let $\mathbf{O}_{u,v}$ be a $u \times v$ matrix of zeros, let $\mathbf{I}_{u,v}$ be a $u \times v$ rectangular diagonal matrix with element 1 on the diagonal, and let $\mathbf{a}^{(i)} = (\mathbf{1}_{k_i}, \mathbf{0}_{k_i})$ and $\mathbf{a} = (\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(T)})$. We shall also need the matrices $\mathbf{B} = \boldsymbol{\gamma}'\boldsymbol{\varphi}$, $\mathbf{C} = \boldsymbol{\gamma}'\boldsymbol{\theta}$,

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{C}^{(i)} & \mathbf{S}^{(i)} \\ -\mathbf{S}^{(i)} & \mathbf{C}^{(i)} \end{bmatrix}, \qquad \tilde{\mathbf{A}}_i = \begin{bmatrix} \mathbf{0}_{m_i-1} & 1 \\ \mathbf{I}_{m_i-1} & \mathbf{0}'_{m_i-1} \end{bmatrix},$$

and $\mathbf{A} = \bigoplus_{i=1}^T \mathbf{A}_i$, where $\mathbf{C}^{(i)}$ and $\mathbf{S}^{(i)}$ are $k_i \times k_i$ diagonal matrices with elements $\cos(\lambda_j^{(i)})$ and $\sin(\lambda_j^{(i)})$, respectively, for $j = 1, 2, \ldots, k_i$ and $i = 1, \ldots, T$, and where $\bigoplus$ denotes the direct sum of the matrices. Let $\tau = 2\sum_{i=1}^T k_i$.

Then the matrices for the TBATS model can be written as $\mathbf{w} = (1, \phi, \mathbf{a}, \boldsymbol{\varphi}, \boldsymbol{\theta})'$, $\mathbf{g} = (\alpha, \beta, \boldsymbol{\gamma}, 1, \mathbf{0}_{p-1}, 1, \mathbf{0}_{q-1})'$, and

$$\mathbf{F} = \begin{bmatrix} 1 & \phi & \mathbf{0}_\tau & \alpha\boldsymbol{\varphi} & \alpha\boldsymbol{\theta} \\ 0 & \phi & \mathbf{0}_\tau & \beta\boldsymbol{\varphi} & \beta\boldsymbol{\theta} \\ \mathbf{0}'_\tau & \mathbf{0}'_\tau & \mathbf{A} & \mathbf{B} & \mathbf{C} \\ 0 & 0 & \mathbf{0}_\tau & \boldsymbol{\varphi} & \boldsymbol{\theta} \\ \mathbf{0}'_{p-1} & \mathbf{0}'_{p-1} & \mathbf{O}_{p-1,\tau} & \mathbf{I}_{p-1,p} & \mathbf{O}_{p-1,q} \\ 0 & 0 & \mathbf{0}_\tau & \mathbf{0}_p & \mathbf{0}_q \\ \mathbf{0}'_{q-1} & \mathbf{0}'_{q-1} & \mathbf{O}_{q-1,\tau} & \mathbf{O}_{q-1,p} & \mathbf{I}_{q-1,q} \end{bmatrix}.$$

These matrices apply when all of the components are present in the model. When a component is omitted, the corresponding terms in the matrices must be omitted.

2.4.2 BATS Model. The state space form of the BATS model can be obtained by letting $\mathbf{s}_t^{(i)} = (s_t^{(i)}, s_{t-1}^{(i)}, \ldots, s_{t-(m_i-1)}^{(i)})$, $\mathbf{a}^{(i)} = (\mathbf{0}_{m_i-1}, 1)$, $\boldsymbol{\gamma}^{(i)} = (\gamma_i, \mathbf{0}_{m_i-1})$, $\mathbf{A} = \bigoplus_{i=1}^T \tilde{\mathbf{A}}_i$, and by replacing $2k_i$ with m_i in the matrices presented above for the TBATS models.

2.4.3 Reduced Forms. It is well known that linear forecasting systems have equivalent ARIMA (Box and Jenkins 1970) reduced forms, and it has been shown that the forecasts from some exponential smoothing models are identical to the forecasts from particular ARIMA models (McKenzie 1984; Chatfield and Yar 1991). The reduced forms of BATS and TBATS models can be obtained by

$$\varphi_p(L)\eta(L)y_t^{(\omega)} = \theta_q(L)\delta(L)\varepsilon_t, \tag{6}$$

where L is the lag operator, $\eta(L) = \det(\mathbf{I} - \mathbf{F}^*L)$, $\delta(L) = \mathbf{w}^* \operatorname{adj}(\mathbf{I} - \mathbf{F}^*L)\mathbf{g}^*L + \det(\mathbf{I} - \mathbf{F}^*L)$, $\varphi_p(L)$ and $\theta_q(L)$ are polynomials of length p and q, $\mathbf{w}^* = (1, \phi, \mathbf{a})$, $\mathbf{g}^* = (\alpha, \beta, \boldsymbol{\gamma})'$, and

$$\mathbf{F}^* = \begin{bmatrix} 1 & \phi & \mathbf{0} \\ 0 & \phi & \mathbf{0} \\ \mathbf{0}' & \mathbf{0}' & \mathbf{A} \end{bmatrix},$$

with the corresponding parameters defined as above. (Refer to De Livera 2010b, chap. 4 for the proofs.) For BATS models with a nonstationary growth, the reduced form is then given by (6), with

$$\eta(L) = (1 - \phi L)(1 - L)\prod_{j=1}^T (L^{m_j-1} + L^{m_j-2} + \cdots + L + 1),$$

$$\begin{aligned} \delta(L) = {} & \prod_{j=1}^T (L^{m_j-1} + L^{m_j-2} + \cdots + L + 1) \\ & \times [L^2(\phi - \phi\alpha) + L(\alpha + \phi\beta - \phi - 1) + 1] \\ & + (1 - \phi L) \\ & \times \sum_{i=1}^T \prod_{j=1, i\neq j}^T (L^{m_j-1} + L^{m_j-2} + \cdots + L + 1)\gamma_i L^{m_i}. \end{aligned}$$

For TBATS models with a nonstationary growth, the reduced form is then given by (6), with

$$\eta(L) = (1 - L)(1 - \phi L)\prod_{i=1}^T \prod_{j=1}^{k_i} \left(1 - 2\cos\lambda_j^{(i)} L + L^2\right),$$

$$\begin{aligned} \delta(L) = {} & [L^2\phi(1 - \alpha) + L(\alpha + \phi\beta - \phi - 1) + 1] \\ & \times \prod_{i=1}^T \prod_{j=1}^{k_i} \left(1 - 2\cos\lambda_j^{(i)} L + L^2\right) \\ & + (1 - L)(1 - \phi L) \\ & \times \sum_{i=1}^T \sum_{j=1}^{k_i} \prod_{\tilde{i}=1, \tilde{i}\neq i}^T \prod_{\tilde{j}=1, \tilde{j}\neq j}^{k_{\tilde{i}}} \left(1 - 2\cos\lambda_{\tilde{j}}^{(\tilde{i})} L + L^2\right) \\ & \times \left[\left(\cos\lambda_j^{(i)}\gamma_{1i} + \sin\lambda_j^{(i)}\gamma_{2i}\right)L^2 - \gamma_{1i}L^3\right] \\ & + (1 - L)(1 - \phi L)L \\ & \times \prod_{i=1}^T \prod_{j=1}^{k_i} \left(1 - 2\cos\lambda_j^{(i)} L + L^2\right) \sum_{i=1}^T k_i\gamma_{1i}. \end{aligned}$$

One of the benefits of using this ARIMA reduced form compared to the existing innovations state space methodology is that it allows for the derivation of exact likelihood estimates without treating initial state values as additional parameters (Gardner, Harvey, and Phillips 1980; Melard 1984;

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Kohn and Ansley 1985). It also allows for the derivation of forecast error variance, computation of forecast intervals, and residual analysis. However, the reduced form of the TBATS model has a relatively complex ARIMA structure which is dependent on the number of terms k_i chosen for the ith seasonal component, and encompasses trigonometric coefficients that rely on the frequency of each seasonal pattern. Several other advantages of state space modeling over ARIMA modeling have been described by Durbin (2001, chap. 3, sect. 3.5).

3. ESTIMATION, PREDICTION, AND MODEL SELECTION

3.1 Estimation

The typical approach with linear state space models is to estimate unknown parameters like the smoothing parameters and the damping parameter using the sum of squared errors or the Gaussian likelihood (see Hyndman et al. 2008, chap. 3). In our context it is necessary to also estimate the unknown Box–Cox transformation parameter ω, and the ARMA coefficients.

The seed states of state space models are usually treated as random vectors. Given trial values of the unknown parameters, the joint steady-state distributions of stationary states are derived, and then assigned to associated seed states. Thus, for given values of ϕ and σ^2, the seed short-run growth rate would be assigned an $\mathrm{N}(0, \sigma^2/(1-\phi^2))$ distribution. Most states, however, are nonstationary, and they are presumed to have Gaussian distributions with arbitrarily large variances (Ansley and Kohn 1985). The Kalman filter is typically used to obtain one-step-ahead prediction errors and associated variances needed for evaluating fitting criteria for given trial values of the parameters. The Kalman filter in the work of Snyder (1985b) would be appropriate for innovations state space models in particular. However, it would need to be augmented with additional equations (De Jong 1991) to handle the nonstationary states.

A simpler alternative is available in the context of innovations state space models. By conditioning on all the seed states and treating them as unknown fixed parameters, exponential smoothing can be used instead of an augmented Kalman filter to generate the one-step-ahead prediction errors needed for likelihood evaluation. In this case both the parameters and seed states are selected to maximize the resulting *conditional* likelihood function. If not for the different treatment of stationary states, the exponential smoothing and augmented Kalman filter approaches yield the same conditional distribution of the final state vector and so yield identical prediction distributions of future series values (Hyndman et al. 2008).

The conditional likelihood of the observed data $\mathbf{y} = (y_1, \ldots, y_n)$ is derived on the assumption that $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$. This implies that the density of the transformed series is $y_t^{(\omega)} \sim \mathrm{N}(\mathbf{w}'\mathbf{x}_{t-1}, \sigma^2)$ so that the density of the transformed data is

$$p\big(\mathbf{y}^{(\omega)} \mid \mathbf{x}_0, \boldsymbol{\vartheta}, \sigma^2\big) = \prod_{t=1}^{n} p\big(y_t^{(\omega)} \mid \mathbf{x}_{t-1}, \boldsymbol{\vartheta}, \sigma^2\big) = \prod_{t=1}^{n} p(\varepsilon_t) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(\frac{-1}{2\sigma^2}\sum_{t=1}^{n}\varepsilon_t^2\right),$$

where $\boldsymbol{\vartheta}$ is a vector containing the Box–Cox parameter, smoothing parameters, and ARMA coefficients. Therefore, the density of the original series, using the Jacobian of the Box–Cox transformation, is

$$\begin{aligned} p(y_t \mid \mathbf{x}_0, \boldsymbol{\vartheta}, \boldsymbol{\sigma}^2) &= p\big(y_t^{(\omega)} \mid \mathbf{x}_0, \boldsymbol{\vartheta}, \sigma^2\big)\left|\det\left(\frac{\partial y_t^{(\omega)}}{\partial y}\right)\right| \\ &= p\big(y_t^{(\omega)} \mid \mathbf{x}_0, \boldsymbol{\vartheta}, \sigma^2\big)\prod_{t=1}^{n} y_t^{\omega-1} \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(\frac{-1}{2\sigma^2}\sum_{t=1}^{n}\varepsilon_t^2\right)\prod_{t=1}^{n} y_t^{\omega-1}. \end{aligned}$$

On concentrating out the variance σ^2 with its maximum likelihood estimate

$$\hat{\sigma}^2 = n^{-1}\sum_{t=1}^{n}\varepsilon_t^2, \tag{7}$$

we obtain the log-likelihood given by

$$\mathcal{L}(\mathbf{x}_0, \boldsymbol{\vartheta}, \sigma^2) = \frac{-n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\varepsilon_t^2 + (\omega-1)\sum_{t=1}^{n}\log y_t. \tag{8}$$

Substituting (7) into (8), multiplying by -2, and omitting constant terms, we get

$$\mathcal{L}^*(\mathbf{x}_0, \boldsymbol{\vartheta}) = n\log\left(\sum_{t=1}^{n}\varepsilon_t^2\right) - 2(\omega-1)\sum_{t=1}^{n}\log y_t. \tag{9}$$

The quest is to minimize the quantity (9) to obtain maximum likelihood estimates, but the dimension of the seed states vector $\mathbf{x}_0$ makes this computationally challenging. Our approach to this problem is based on the observation that ε_t is a linear function of the seed vector $\mathbf{x}_0$. Thus, we show that it is possible to concentrate the seed states out of the likelihood, and so substantially reduce the dimension of the numerical optimization problem. This concentration process is the exponential smoothing analogue of de Jong's method for augmenting Kalman filter to handle seed states with infinite variances.

The innovation ε_t can be eliminated from the transition equation in (5) to give $\mathbf{x}_t = \mathbf{D}\mathbf{x}_{t-1} + \mathbf{g}y_t$ where $\mathbf{D} = \mathbf{F} - \mathbf{g}\mathbf{w}'$. The equation for the state, obtained by back-solving this recurrence equation to period 0, can be used in conjunction with the measurement equation to obtain

$$\begin{aligned} \varepsilon_t &= y_t^{(\omega)} - \mathbf{w}'\sum_{j=1}^{t-1}\mathbf{D}^{j-1}\mathbf{g}y_{t-j}^{(\omega)} - \mathbf{w}'\mathbf{D}^{t-1}\mathbf{x}_0 \\ &= y_t^{(\omega)} - \mathbf{w}'\tilde{\mathbf{x}}_{t-1} - \mathbf{w}_{t-1}'\mathbf{x}_0 \\ &= \tilde{y}_t - \mathbf{w}_{t-1}'\mathbf{x}_0, \end{aligned} \tag{10}$$

where $\tilde{y}_t = y_t^{(\omega)} - \mathbf{w}'\tilde{\mathbf{x}}_{t-1}$, $\tilde{\mathbf{x}}_t = \mathbf{D}\tilde{\mathbf{x}}_{t-1} + \mathbf{g}y_t$, $\mathbf{w}_t' = \mathbf{D}\mathbf{w}_{t-1}'$, $\tilde{\mathbf{x}}_0 = 0$, and $\mathbf{w}_0' = \mathbf{w}'$ (see Snyder 1985a, for the derivation). Thus, the relationship between each error and the initial state vector $\mathbf{x}_0$ is linear. It can also be seen from (10) that the seed vector $\mathbf{x}_0$ corresponds to a regression coefficients vector, and so it may be estimated using conventional linear least squares methods.

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Thus, the problem reduces to minimizing the following with respect to $\boldsymbol{\vartheta}$:

$$\mathcal{L}^*(\boldsymbol{\vartheta}) = n\log(\mathrm{SSE}^*) - 2(\omega - 1)\sum_{t=1}^{n}\log y_t, \tag{11}$$

where SSE^* is the optimized value of the sum of squared errors for given parameter values.

In contrast to the existing estimation procedure (Hyndman et al. 2008, chap. 14) where a heuristic scheme is employed to find the values of the initial states, our approach concentrates out the initial state values from the likelihood function, leaving only the much smaller parameter vector for optimization, a tactic in our experience that leads to better forecasts. It may also be effective in reducing computational times instead of invoking the numerical optimizer to directly estimate the seed state vector.

We can constrain the estimation to the *forecastability region* (Hyndman, Akram, and Archibald 2007) so that the characteristic roots of $\mathbf{D}$ lie within the unit circle, a concept that is equivalent to the invertibility condition for equivalent ARIMA models. The coefficients $\mathbf{w}'\mathbf{D}^{j-1}\mathbf{g}$ are the matrix analogues of the weights in an exponentially weighted average, and this constraint ensures that their effect is to reduce the importance placed on older data. When some roots lie on the unit circle, the discounting effect is lost (although this possibility admits some important special cases). For integer period seasonality, the seasonal values can also be constrained when optimizing, so that each seasonal component sums to zero.

3.2 Prediction

The prediction distribution in the transformed space for future period $n + h$, given the final state vector $\mathbf{x}_n$ and given the parameters $\boldsymbol{\vartheta}, \sigma^2$, is Gaussian. The associated random variable is designated by $y^{(\omega)}_{n+h|n}$. Its mean $\mathrm{E}(y^{(\omega)}_{n+h|n})$ and variance $\mathrm{V}(y^{(\omega)}_{n+h|n})$ are given, after allowing for the Box–Cox transformation, by the equations (Hyndman et al. 2005):

$$\mathrm{E}\big(y^{(\omega)}_{n+h|n}\big) = \mathbf{w}'\mathbf{F}^{h-1}\mathbf{x}_n, \tag{12a}$$

$$\mathrm{V}\big(y^{(\omega)}_{n+h|n}\big) = \begin{cases} \sigma^2, & \text{if } h = 1, \\ \sigma^2\left[1 + \sum_{j=1}^{h-1} c_j^2\right], & \text{if } h \geq 2, \end{cases} \tag{12b}$$

where $c_j = \mathbf{w}'\mathbf{F}^{j-1}\mathbf{g}$. The prediction distribution of $y_{n+h|n}$ is not normal. Point forecasts and forecast intervals, however, may be obtained using the inverse Box–Cox transformation of appropriate quantiles of the distribution of $y^{(\omega)}_{n+h|n}$. The point forecast obtained this way is the median, a minimum mean absolute error predictor (Pankratz and Dudley 1987; Proietti and Riani 2009). The prediction intervals retain the required probability coverage under back-transformation because the Box–Cox transformation is monotonically increasing. To simplify matters we use the common *plug-in* approach to forecasting. The pertinent parameters and final state are replaced by their estimates in the above formulas. This ignores the impact of estimation error, but the latter is a second-order effect in most practical contexts.

3.3 Model Selection

3.3.1 The Use of an Information Criterion. In this article, the $\mathrm{AIC} = \mathcal{L}^*(\hat{\boldsymbol{\vartheta}}, \hat{\mathbf{x}}_0) + 2K$ is used for choosing between the models, where K is the total number of parameters in $\boldsymbol{\vartheta}$ plus the number of free states in $\mathbf{x}_0$, and $\hat{\boldsymbol{\vartheta}}$ and $\hat{\mathbf{x}}_0$ denote the estimates of $\boldsymbol{\vartheta}$ and $\mathbf{x}_0$. When one of the smoothing parameters takes the boundary value 0, the value of K is reduced by 1 as the model simplifies to a special case. For example, if $\beta = 0$, then $b_t = b_0$ for all t. Similarly, when either $\phi = 1$ or $\omega = 1$, the value of K is reduced by 1 in each instance to account for the resulting simplified model (e.g., when $\omega = 1$, the model simplifies to a linear model without a Box–Cox transformation, and when $\phi = 1$, the model reduces to a model without a damping effect in the trend component). In the applications of our article, the Nelder–Mead algorithm (Nelder and Mead 1965), which is being successfully employed in recent R packages (e.g., *forecast* package for R Hyndman 2010), was used for optimization, with a tolerance level of $1e$–08 for such boundary values. In an empirical study, Billah, Hyndman, and Koehler (2005) indicated that information criterion approaches, such as the AIC, provide the best basis for automated model selection, relative to other methods such as prediction-validation. Alternative information criteria such as the AICc (Burnham and Anderson 2002) may also be used.

3.3.2 Selecting the Number of Harmonics k_j in the Trigonometric Models. The forecasts from the TBATS model depend on the number of harmonics k_i used for the seasonal component i. It is impractical to consider all possible combinations in the quest for the best combination. After much experimentation we found that the following approach leads to good models and that further improvement can rarely be achieved (see De Livera 2010b, chap. 3).

De-trend the first few seasons of the transformed data using an appropriate de-trending method. In this article, we employed the method described by Hyndman et al. (2008, chap. 2). Approximate the resulting de-trended data using the linear regression $\sum_{i=1}^{T}\sum_{j=1}^{k_i} a_j^{(i)}\cos(\lambda_j^{(i)}t) + b_j^{(i)}\sin(\lambda_j^{(i)}t)$. Starting with a single harmonic, gradually add harmonics, testing the significance of each one using F-tests. Let k_i^* be the number of significant harmonics (with $p < 0.001$) for the ith seasonal component. Then fit the required model to the data with $k_i = k_i^*$ and compute the AIC. Considering one seasonal component at a time, repeatedly fit the model to the estimation sample, gradually increasing k_i but holding all other harmonics constant for each i, until the minimum AIC is achieved. This approach, based on multiple linear regression, is preferred over letting $k_i^* = 1$ for each component, as the latter was found to be unnecessarily time-consuming.

3.3.3 Selecting the ARMA Orders p and q for the Models. In selecting a model, suitable values for the ARMA orders p and q must also be found. We do this using a two-step procedure. First, a suitable model with no ARMA component is selected. Then the automatic ARIMA algorithm of Hyndman and Khandakar (2008) is applied to the residuals from this model in order to determine the appropriate orders p and q (we assume the residuals are stationary). The selected model is then fitted again but with an ARMA(p, q) error component, where

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

the ARMA coefficients are estimated jointly with the rest of the parameters. The ARMA component is only retained if the resulting model has lower AIC than the model with no ARMA component. Our subsequent work on the proposed models on a large number of real time series (De Livera 2010a) has indicated that fitting ARMA in a two-step approach yielded the best out-of-sample predictions, compared to several alternative approaches.

4. APPLICATIONS OF THE PROPOSED MODELS

The results obtained from the application of BATS and TBATS to the three complex time series in Figure 1 are reported in this section.

In addition, it is shown that the TBATS models can be used as a means of decomposing complex seasonal time series into trend, seasonal, and irregular components. In decomposing time series, the trigonometric approach has several important advantages over the traditional seasonal formulation. First, seasonal components obtained from the BATS model are not normalized (cf. the seasonal model of Harrison and Stevens 1976). Although normalized components may not be necessary if one is only interested in the forecasts and the prediction intervals, when the seasonal component is to be analyzed separately or used for seasonal adjustment, normalized seasonal components are required (Archibald and Koehler 2003; Hyndman et al. 2008). Thus, BATS models have to be modified, so that the seasonal components are normalized for each time period, before using them for time series decomposition (see De Livera 2010b, chap. 5 for a normalized version of the BATS model). In contrast, the trigonometric terms in TBATS models do not require normalization, and so are more appropriate for decomposition. Second, in estimating the seasonal components using BATS, a large number of parameters are required, which often leads to noisy seasonal components. In contrast, a smoother seasonal decomposition is expected from TBATS where the smoothness of the seasonal component is controlled by the number of harmonics used. Furthermore, a BATS model cannot be used to decompose time series with non-integer seasonality and dual-calendar effects. Using TBATS models for complex seasonal time series, the overall seasonal component can be decomposed into several individual seasonal components with different frequencies. These individual seasonal components are given by $s_t^{(i)}$ $(i = 1, \ldots, T)$ and the trend component is obtained by ℓ_t. Extracting the trend and seasonal components then leaves behind a covariance stationary irregular component, denoted by d_t. In particular, this approach leads to the identification and extraction of one or more seasonal components, which may not be apparent in the time series plots themselves.

4.1 Application to Weekly U.S. Gasoline Data

Figure 1(b) shows the number of barrels of motor gasoline product supplied in the United States, in thousands of barrels per day, from February 1991 to July 2005 (see *www.forecastingprinciples.com/files/T_competition_new.pdf* for details). The data are observed weekly and show a strong annual seasonal pattern. The length of seasonality of the time series is $m_1 = 365.25/7 \approx 52.179$. The time series exhibits an upward additive trend and an additive seasonal pattern, that is, a pattern for which the variation does not change with the level of the time series.

The series, which consists of 745 observations, was split into two segments: an estimation sample period (484 observations) and a test sample (261 observations). The estimation sample was used to obtain the maximum likelihood estimates of the initial states and the smoothing parameters, and to select the appropriate number of harmonics and ARMA orders. Following the procedure for finding the number of harmonics to start with, it was found that only one harmonic was highly significant. The model was then fitted to the whole estimation sample of 484 values by minimizing the criterion equation (11). The values of the AIC decreased until $k_1 = 7$, and then started to increase.

Out-of-sample performance was measured by the root mean squared error (RMSE), defined as

$$\mathrm{RMSE}_h = \sqrt{\frac{1}{p-h+1}\sum_{t=n}^{n+p-h}(y_{t+h} - \hat{y}_{t+h|t})^2}, \qquad (13)$$

where $p = 261$ is the length of the test sample, $n = 484$ is the length of the estimation sample, and h is the length of the forecast horizon. Further analysis showed that changing the value of k_1 from 7 generated worse out-of-sample results, indicating that the use of the AIC as the criterion for this model selection procedure is a reasonable choice. In this way, the TBATS(0.9922, 1, 0, 0, {365.25/7, 7}) model was obtained. As a second step, ARMA models were fitted to the residuals with (p, q) combinations up to $p = q = 5$, and it was discovered that the TBATS(0.9922, 1, 0, 1, {365.25/7, 7}) model minimizes the AIC.

The BATS model was considered next with $m_1 = 52$, and following the above procedure, it was discovered that the BATS(0.9875, 1, 0, 1, 52) model minimized the AIC. Figure 2 shows the out-of-sample RMSEs obtained for the two models, and it can be seen that the trigonometric model performs better for all lead times.

The BATS model cannot handle the non-integer periods, and so has to be rounded off to the nearest integer. It may also be overparameterized, as 52 initial seasonal values have to be estimated. Both these problems are overcome in the trigonometric formulation.

Tables 1 and 2 show the estimated parameters obtained for the TBATS and BATS models, respectively. The estimated values of 0 for β and 1 for ϕ imply a purely deterministic growth rate with no damping effect. The models also imply that the irregular component of the series is correlated and can be described by an ARMA(0, 1) process, and that a strong transformation is not necessary to handle nonlinearities in the series.

The decomposition of the Gasoline time series, obtained from the fitted TBATS model, is shown in Figure 3. The vertical bars at the right side of each plot represent equal heights plotted on different scales, thus providing a comparison of the size of each component. The trigonometric formulation in TBATS allows for the removal of more randomness from the seasonal component without destroying the influential bumps.

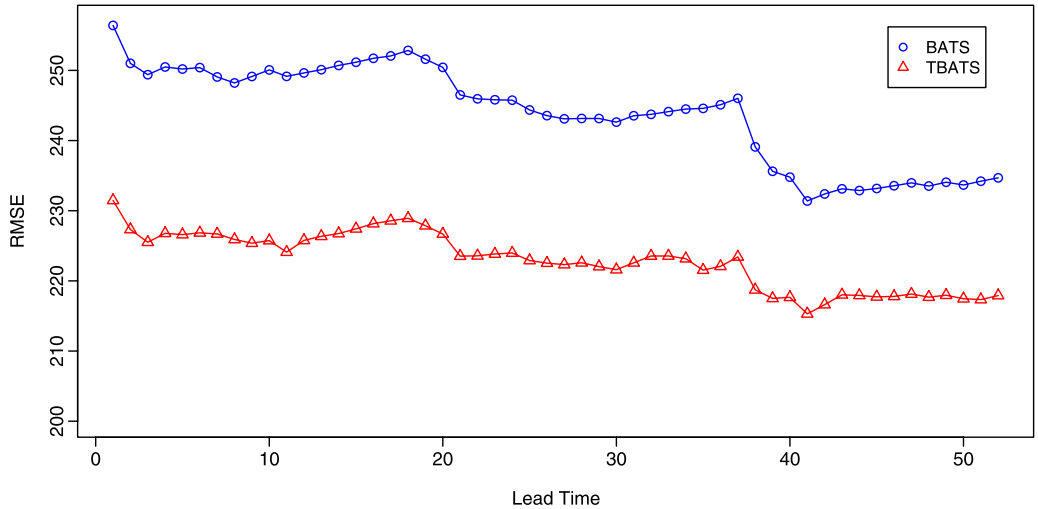

Figure 2. Out-of-sample results for the U.S. gasoline data using BATS(0.9875, 1, 0, 1, 52) and TBATS(0.9922, 1, 0, 1, {365.25/7, 7}). The online version of this figure is in color.

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

4.2 Application to Call Center Data

The call center series in Figure 1(a) consists of 10,140 observations, that is, 12 weeks of data starting from 3 March 2003 (Weinberg, Brown, and Stroud 2007). It contains a daily seasonal pattern with period 169 and a weekly seasonal pattern with period $169 * 5 = 845$. The fitting sample consists of 7605 observations (9 weeks). As the trend appears to be close to zero, the growth rate b_t was omitted from the models.

The selection procedure led to the models TBATS(1, NA, 3, 1, {169, 29}, {845, 15}) and BATS(0.4306, NA, 3, 0, 169, 845). Other BATS models with $\omega = 1$ were also tried, but their forecasting performance was worse.

The post-sample forecasting accuracies of the selected BATS and TBATS models are compared in Figure 4. Again TBATS, which requires fewer parameters to be estimated, is more accurate than BATS.

The estimated parameters for the TBATS model shown in Table 1 imply that no Box–Cox transformation is necessary for this time series, and that the weekly seasonal component is more variable than the daily seasonal component. The irregular component is modeled by an ARMA(3, 1) process.

The decomposition obtained from TBATS, as shown in Figure 5, clearly exhibits strong daily and weekly seasonal components. The weekly seasonal pattern evolves considerably over time but the daily seasonal pattern is relatively stable. As is seen from the time series plot itself, the trend component is very small in magnitude compared to the seasonal components.

4.3 Application to the Turkey Electricity Demand Data

The Turkey electricity demand series shown in Figure 1(c) has a number of important features that should be reflected in the model structure. Three seasonal components with frequencies $m_1 = 7$, $m_2 = 354.37$, and $m_3 = 365.25$ exist in the series. The sharp drops seen in the seasonal component with period 354.37 are due to the Seker (also known as *Eid ul-Fitr*) and Kurban (also known as *Eid al-Adha*) religious holidays, which follow the Hijri calendar, while those seen in the seasonal component with frequency 365.25 are due to national holidays which follow the Gregorian calendar. Table 3 gives the dates of the holidays from the Hijri and Gregorian calendars. Seker is a three-day festival when sweets are eaten to celebrate the end of the fast of Ramadan. Kurban is a four-day festival when sacrificial sheep are slaughtered and their meat distributed to the poor. In addition, there are national holidays which follow the Gregorian calendar as shown in the table.

In this study, the series, which covers a period of 9 years, is split into two parts: a fitting sample of $n = 2191$ observations (6 years) and a post-sample period of $p = 1096$ observations (3 years). The model selection procedure was followed to give the TBATS(0.1393, 1, 3, 2, {7, 3}, {354.37, 23}, {365.25, 3}) and BATS(0.0013, 1, 0, 0, 7, 354, 365) models.

Figure 6 shows that a better post-sample forecasting performance is again obtained from the TBATS model. The poor performance of the BATS model may be explained by its inability to capture the dual seasonal calendar effects and the large num-

Table 1. Parameters chosen for each application of the TBATS model

	Parameters														
Data	ω	ϕ	α	β	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	θ_1	θ_2	φ_1	φ_2	φ_3
Gasoline	0.9922	1	0.0478	0	0.0036	−0.0005					−0.2124				
Call center	1		0.0921		0.0006	−0.0002	0.0022	0.0020			−0.1353		0.1776	−0.0144	−0.0200
Electricity	0.1393	1	0.8019	0	0.0034	−0.0037	0.0001	0.0001	0.1171	0.251	−1.2610	0.3128	1.0570	−0.2991	−0.1208

Table 2. Parameters chosen for each application of the BATS model

	Parameters											
Data	ω	ϕ	α	β	γ_1	γ_2	γ_3	θ_1	θ_2	φ_1	φ_2	φ_3
Gasoline	0.9875	1	0.0457	0	0.2246			−0.2186				
Call center	0.4306		0.0368		0.0001	0.0001				0.0552	0.1573	0.1411
Electricity	0.0013	1	0.2216	0	0	0	0					

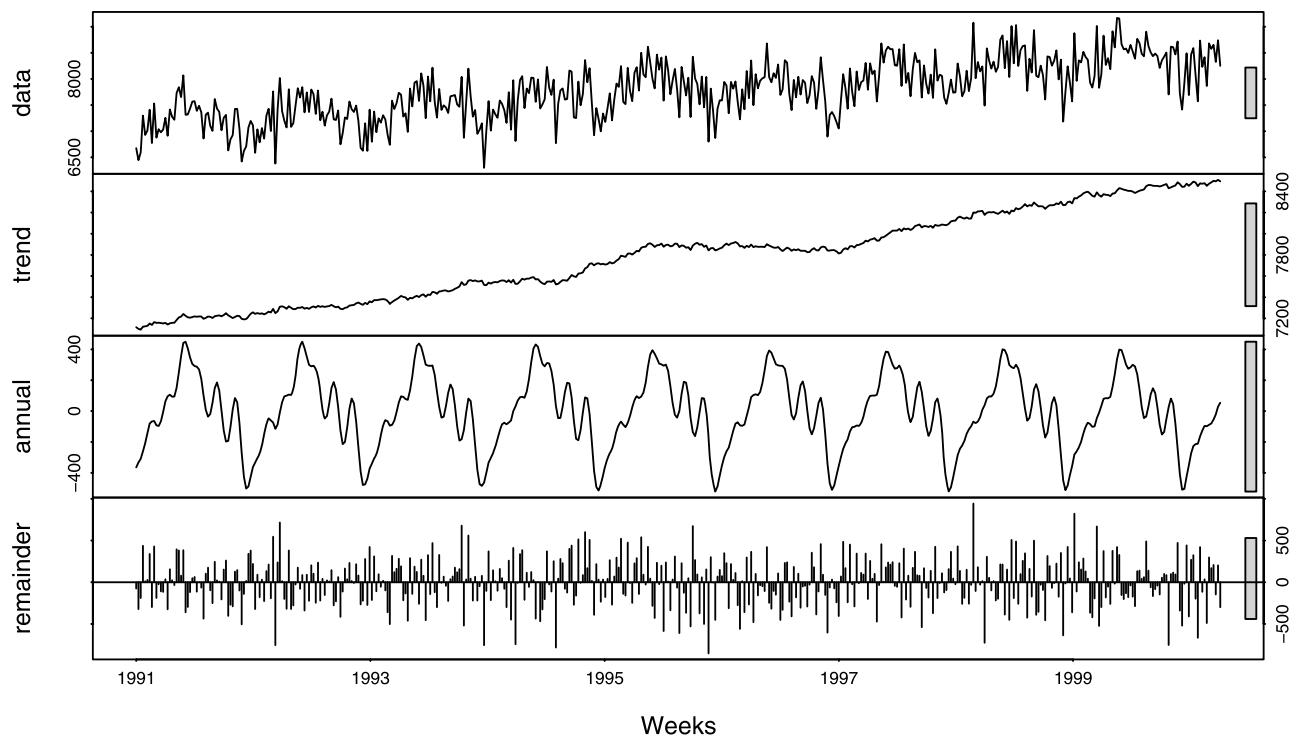

Figure 3. Trigonometric decomposition of the U.S. gasoline data. The within-sample RMSE was 279.9.

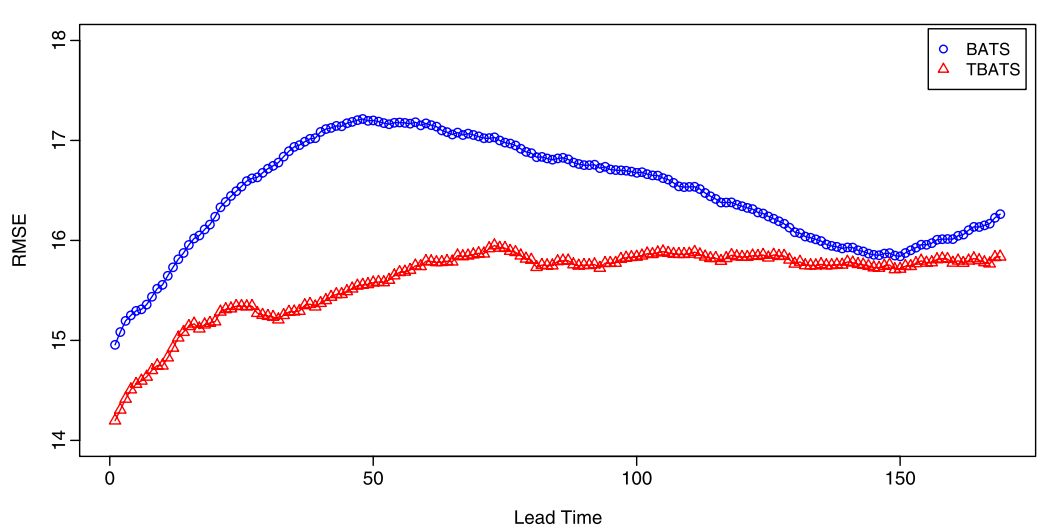

Figure 4. Out-of-sample results for the call center data using BATS(0.4306, NA, 3, 0, 169, 845) and TBATS(1, NA, 3, 1, {169, 29}, {845, 15}). The online version of this figure is in color.

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

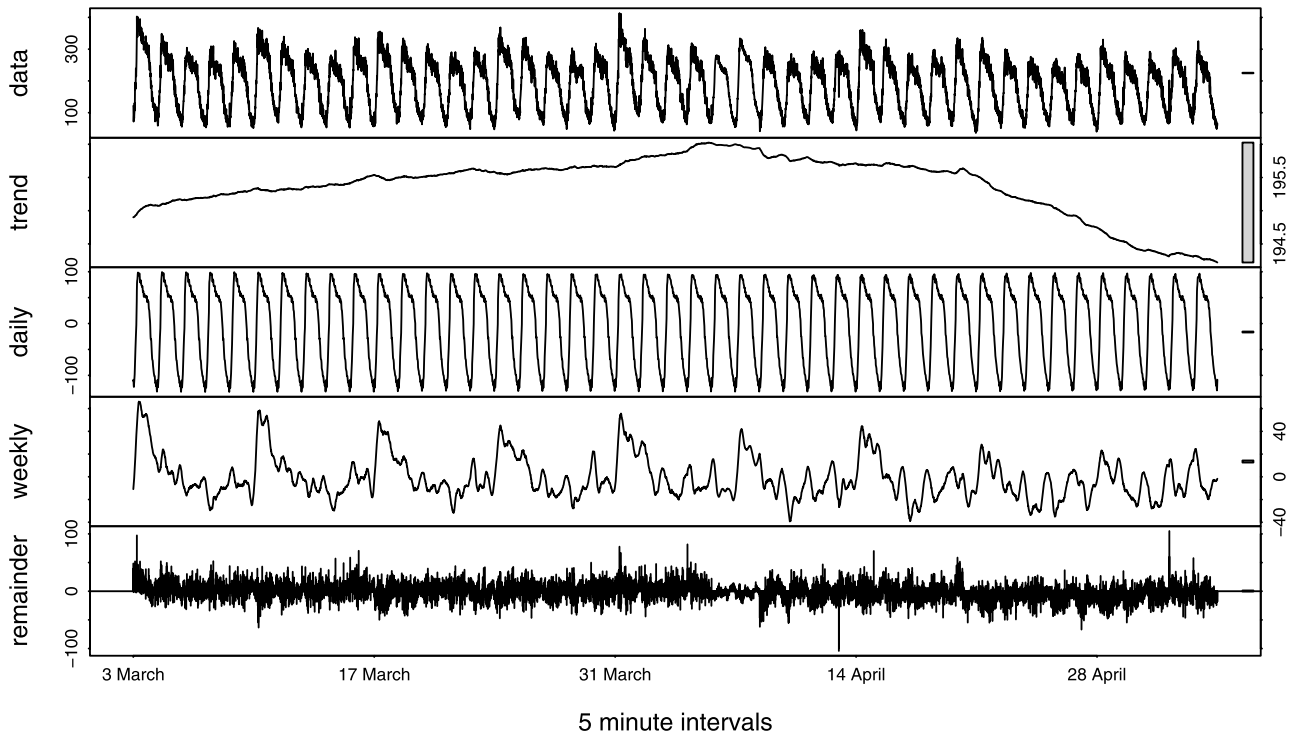

Figure 5. Trigonometric decomposition of the call center data. The within-sample RMSE was 14.7.

ber of values that are required to be estimated. The estimated zero values for the smoothing parameters shown in Table 2 for the BATS solution suggest stable seasonal components. The Hijri seasonal component based on the TBATS solution displays a similar level of stability. However, moderate change is implied by the TBATS model in the weekly and Gregorian seasonal components. Both models required strong Box–Cox transformations in order to handle the obvious nonlinearity in the time series plot.

The decomposition of the series obtained by using the chosen TBATS model is shown in Figure 7. The first panel shows the transformed observations and the second shows the trend component. The third panel shows the weekly seasonal component with period 7, and the fifth and the sixth panels show the seasonal component based on the Hijri calendar with period 354.37 and the seasonal component based on the Gregorian calendar with period 365.25, respectively. The seasonal components shown in the fifth and sixth panels may initially appear to be mirror images. However, their combined effect, shown in the fourth panel, indicates that this is not the case. Interpreting their combined effect as *the annual seasonal component* would be misleading as there is no unique annual calendar in this situation: both constituent calendars are of different lengths.

The rather wiggly components of the decomposition are probably due to the use of a large number of harmonics in each seasonal component. This is necessary to capture the sharp drops seen in the time series plot. If we were to augment the stochastic seasonal component by deterministic holiday effects (given in Table 3) represented by dummy variables, the number of harmonics required might be reduced. Using a trend component, a seasonal component, and holiday dummy variables, regression was performed on the transformed $y_t^{(\omega)}$ values. The term $\sum_{i=1}^{3}\sum_{j=1}^{k_i} a_j^{(i)}\cos(\lambda_j^{(i)}t) + b_j^{(i)}\sin(\lambda_j^{(i)}t)$ was used to cap-

Table 3. The dates of Turkish holidays between 1 January 2000 and 31 December 2006

	Religious holidays		
Year	Seker holiday	Kurban holiday	National holidays
2000	08 Jan–10 Jan 27 Dec–29 Dec	16 Mar–19 Mar	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct
2001	16 Dec–18 Dec	05 Mar–08 Mar	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct
2002	05 Dec–07 Dec	22 Feb–25 Feb	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct
2003	25 Nov–27 Nov	11 Feb–14 Feb	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct
2004	14 Nov–16 Nov	01 Feb–04 Feb	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct
2005	03 Nov–05 Nov	20 Jan–23 Jan	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct
2006	23 Oct–25 Oct	10 Jan–13 Jan 31 Dec	01 Jan, 23 Apr, 19 May, 30 Aug, 29 Oct

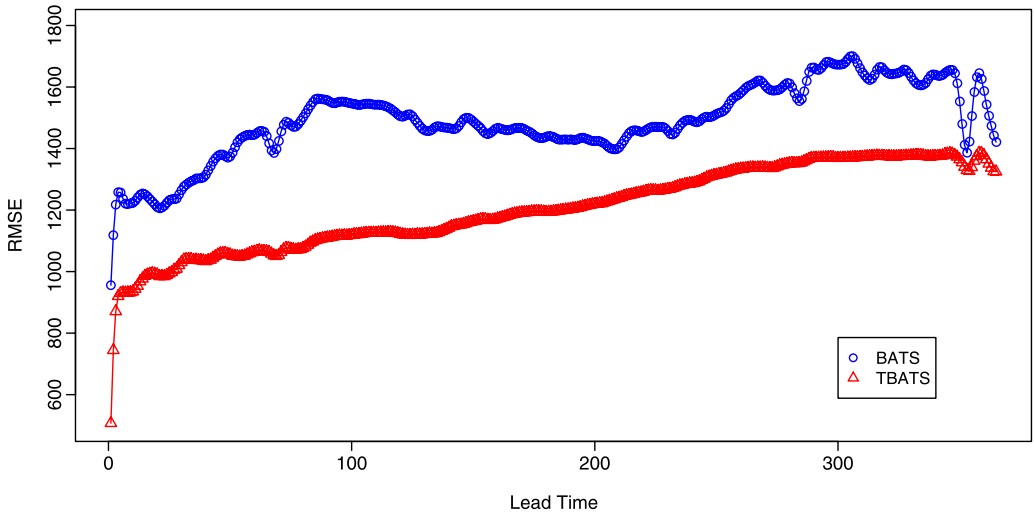

Figure 6. Out-of-sample results for the Turkey electricity demand data using BATS(0.0013, 1, 0, 0, 7, 354, 365) and TBATS(0.1393, 1, 3, 2, {7, 3}, {354.37, 23}, {365.25, 3}). The online version of this figure is in color.

ture the multiple seasonality with $k_1 = 3$ and $k_2 = k_3 = 1$. The estimated holiday effect was then removed from the series and the remainder was decomposed using TBATS to achieve the result shown in Figure 8, which provides a much smoother seasonal decomposition. Again, the sum of the Hijri seasonal component and the Gregorian seasonal component shown in the fourth panel illustrates that the Hijri and Gregorian seasonal components do not offset each other.

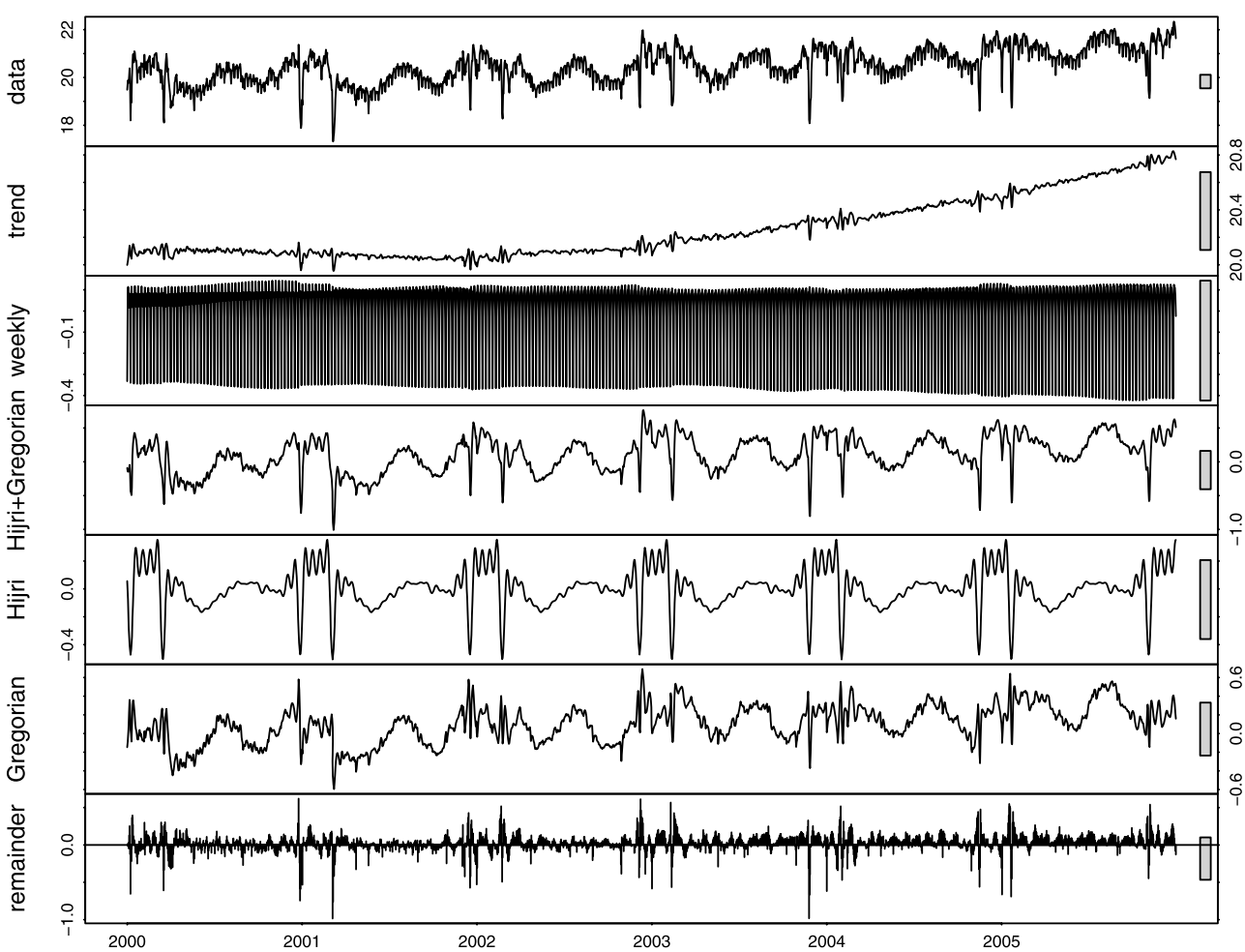

Figure 7. Trigonometric decomposition of the Turkey electricity demand data. The within-sample RMSE was 0.1346.

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

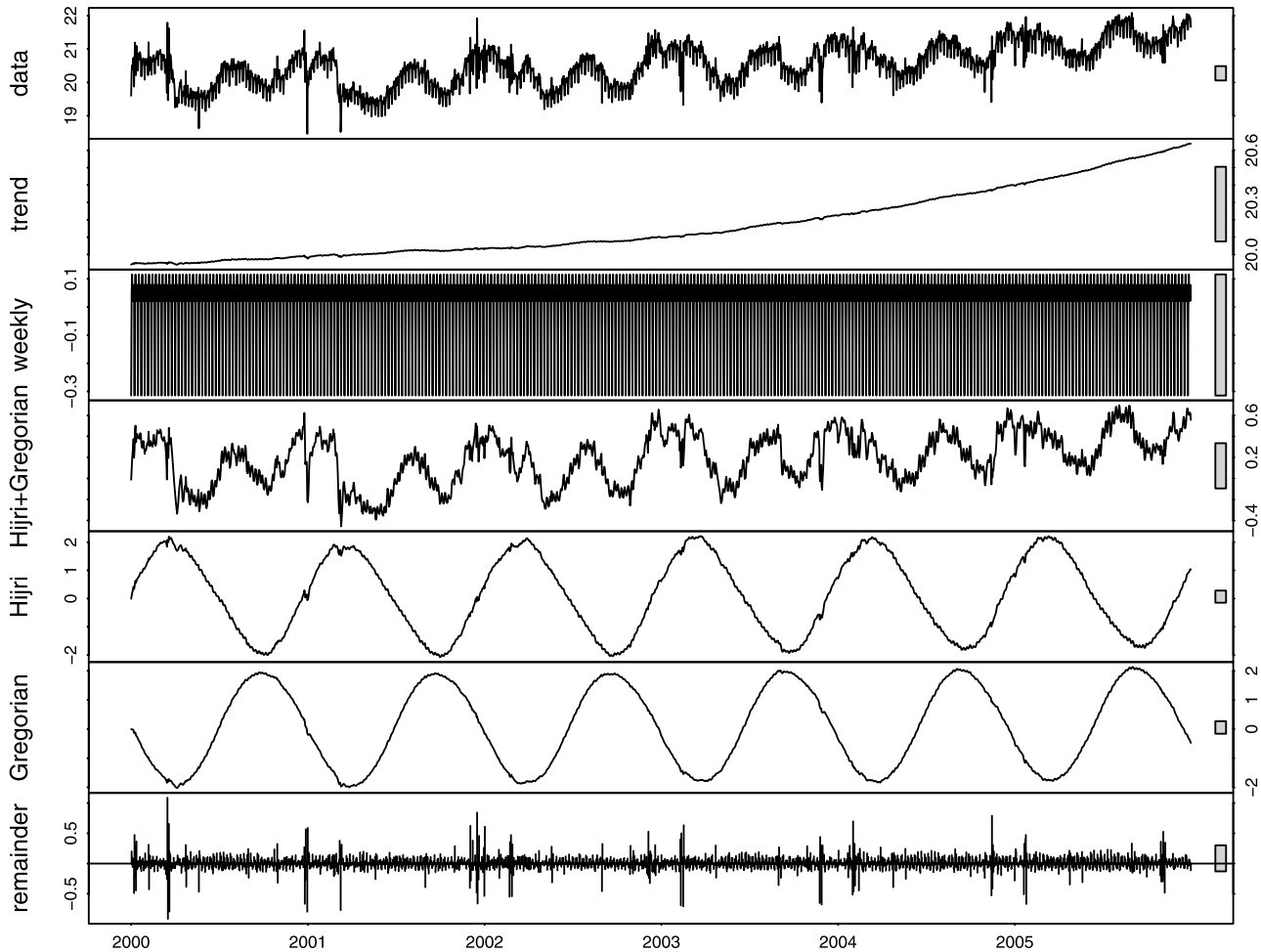

Figure 8. Trigonometric decomposition of the regressed Turkey electricity demand data. The within-sample RMSE was 0.1296.

This analysis demonstrates the capability of our trigonometric decomposition in extracting those seasonal components which are otherwise not apparent in graphical displays.

In forecasting complex seasonal time series with such deterministic effects, both BATS and TBATS models may be extended to accommodate regressor variables, allowing additional information to be included in the models (see De Livera 2010a, chap. 7 for a detailed discussion of the BATS and TBATS models with regressor variables).

5. CONCLUDING REMARKS

A new state space modeling framework, based on the innovations approach, was developed for forecasting time series with complex seasonal patterns. The new approaches offer alternatives to traditional counterparts, providing several advantages and additional options. A key feature of the proposed trigonometric framework is its ability to model both linear and nonlinear time series with single seasonality, multiple seasonality, high period seasonality, non-integer seasonality, and dual-calendar effects. We are not aware of another modeling procedure that is able to forecast and decompose all these complex seasonal time series features within a single framework.

In addition, the framework consists of a new estimation procedure which is sufficiently general to be applied to *any* innovations state space model. By relying on maximum likelihood estimation, it avoids the ad hoc startup choices with unknown statistical properties commonly used with exponential smoothing. By incorporating the least squares criterion, it streamlines the process of obtaining the maximum likelihood estimates.

The applications of the proposed modeling framework to three complex seasonal time series demonstrated that the trigonometric models led to a better out-of-sample performance with substantially fewer values to be estimated than traditional seasonal exponential smoothing approaches (see Table 4). The trigonometric approach was also illustrated as a means of decomposing complex seasonal time series. The ability to handle such complex seasonality is a key advantage of the trigonometric approach over most traditional decomposition methods (e.g., Cleveland et al. 1990; Findley et al. 1998; Gómez and Maravall 1998) which are mainly designed to handle monthly or quarterly data.

A further advantage of the proposed framework is its adaptability. It can be altered to encompass various deterministic effects that are often seen in real-life time series. For instance, the moving holidays such as Easter and irregular holidays can be handled by incorporating dummy variables in the models (see De Livera 2010a, chap. 7), and the varying length of months can be managed by adjusting the data for trading days before modeling (see Makridakis, Wheelwright, and Hyndman 1998).

Table 4. Number of estimated parameters for each model in each application

Data	Model	No. parameters
Gasoline	BATS(0.9875, 1, 0, 1, 52)	60
	TBATS(0.9922, 1, 0, 1, {365.25/7, 7})	23
Call center	BATS(0.4306, NA, 3, 0, 169, 845)	1026
	TBATS(1, NA, 3, 1, {169, 29}, {845, 15})	102
Electricity	BATS(0.0013, 1, 0, 7, 354, 365)	735
	TBATS(0.1393, 1, 3, 2, {7, 3}, {354.37, 23}, {365.25, 3})	79

In handling moving holiday effects further forecasting performance may be obtained by considering their impact on neighboring days. For a detailed discussion on moving holiday effects refer to the work of Findley and Soukup (2000).

The framework can also be adapted to handle data with zero and negative values. The use of a Box–Cox transformation limits our approach to positive time series, as is often encountered in complex seasonal time series. However, the inverse hyperbolic sine transformation (Johnson 1949) can be used in its place should the need arise.

The derivation of likelihood estimates of the proposed approach relies on the assumption of a Gaussian distribution for the errors, something that is often a reasonable approximation when the level of the process is sufficiently far from the origin (Hyndman et al. 2008). In cases where such an assumption may conflict with the underlying structure of the data generating process, our approach can be readily adapted to non-Gaussian situations. Being based on exponential smoothing where the conditioning process ensures that successive state vectors become fixed quantities, any suitable distribution can substitute for the role of the Gaussian distribution. Thus, if the innovations have a t-distribution, the prediction error form of the likelihood can be formed directly from the product of t-distributions. The analytical form of successive prediction distributions is unknown, but they can be simulated from successive t-distributions using means obtained from the application of the equations of the innovations state space model. This can be contrasted with a Kalman filter approach, which must usually be adapted in the presence of non-Gaussian distributions, to a form which necessitates the use of computationally intensive simulation methods.

The proposed frameworks can also be extended to exploit the potential inter-series dependencies of a set of related time series, providing an alternative to the existing vector exponential smoothing framework (de Silva, Hyndman, and Snyder 2007), but with several advantages (see De Livera 2010a, chap. 7).

Our experiences suggest that the use of our proposed estimation procedure for complex seasonal time series requires relatively less computational time than some of the heuristic estimation procedures employed in recent empirical studies (e.g., Taylor 2010a). Further, it should be noted that when the estimation scheme is strictly stable, in large samples, the successive powers of the matrix $\mathbf{D}$ in Equation (10) converge to the null matrix. Thus, in handling complex seasonal time series, the sparse structure of this matrix can be utilized for further gain in computational efficiency (see Koenker and Ng 2003).

The R code for the methods implemented in this article will be available in the *forecast* package for R (Hyndman 2010).

[Received December 2009. Revised June 2011.]

REFERENCES

Abraham, B., and Box, G. E. P. (1978), "Deterministic and Forecast-Adaptive Time-Dependent Models," *Journal of the Royal Statistical Society, Ser. C*, 27 (2), 120–130. [1516]

Akram, M., Hyndman, R., and Ord, J. (2009), "Exponential Smoothing and Non-Negative Data," *Australian and New Zealand Journal of Statistics*, 51 (4), 415–432. [1515]

Anderson, B. D. O., and Moore, J. B. (1979), *Optimal Filtering*, Englewood Cliffs: Prentice-Hall. [1517]

Ansley, C. F., and Kohn, R. (1985), "Estimation, Filtering, and Smoothing in State Space Models With Incompletely Specified Initial Conditions," *The Annals of Statistics*, 13, 1286–1316. [1518]

Archibald, B. C., and Koehler, A. B. (2003), "Normalization of Seasonal Factors in Winters' Methods," *International Journal of Forecasting*, 19, 143–148. [1520]

Billah, B., Hyndman, R. J., and Koehler, A. B. (2005), "Empirical Information Criteria for Time Series Forecasting Model Selection," *Journal of Statistical Computation and Simulation*, 75, 831–840. [1519]

Box, G. E. P., and Cox, D. R. (1964), "An Analysis of Transformations," *Journal of the Royal Statistical Society, Ser. B*, 26 (2), 211–252. [1516]

Box, G. E. P., and Jenkins, G. M. (1970), *Time Series Analysis: Forecasting and Control* (1st ed.), San Francisco: Holden-Day. [1517]

Burnham, K. P., and Anderson, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.), New York: Springer-Verlag. [1519]

Chatfield, C. (1978), "The Holt–Winters Forecasting Procedures," *Applied Statistics*, 27, 264–279. [1515]

Chatfield, C., and Yar, M. (1991), "Prediction Intervals for Multiplicative Holt–Winters," *International Journal of Forecasting*, 7, 31–37. [1517]

Cleveland, R., Cleveland, W., McRae, J., and Terpenning, I. (1990), "STL: A Seasonal-Trend Decomposition Procedure Based on Loess," *Journal of Official Statistics*, 6 (1), 3–73. [1525]

De Jong, P. (1991), "The Diffuse Kalman Filter," *The Annals of Statistics*, 19, 1073–1083. [1518]

De Livera, A. (2010a), "Automatic Forecasting With a Modified Exponential Smoothing State Space Framework," Working Paper 10/10, Dept. of Econometrics and Business Statistics, Monash University. [1520,1525, 1526]

——— (2010b), "Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing," Ph.D. thesis, Dept. of Econometrics and Business Statistics, Monash University. [1517,1519,1520]

de Silva, A., Hyndman, R., and Snyder, R. (2007), "The Vector Innovations Structural Time Series Framework: A Simple Approach to Multivariate Forecasting," *International Journal of Statistical Modelling*, 10 (4), 353–374. [1526]

Durbin, J., and Koopman, S. (2001), *Time Series Analysis by State Space Methods*, New York: Oxford University Press. [1518]

Findley, D., and Soukup, R. (2000), "Modeling and Model Selection for Moving Holidays," in *Proceedings of the Business and Economic Statistics Section*, Alexandria, VA: American Statistical Association. [1526]

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B.-C. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program," *Journal of Business & Economic Statistics*, 16 (2), 127–152. [1525]

Gardner, E. S., Jr. (1985), "Exponential Smoothing: The State of the Art," *Journal of Forecasting*, 4, 1–28. [1515]

Gardner, E. S., Jr., and McKenzie, E. (1985), "Forecasting Trends in Time Series," *Management Science*, 31 (10), 1237–1246. [1516]

Gardner, G., Harvey, A., and Phillips, G. (1980), "Algorithm AS 154: An Algorithm for Exact Maximum Likelihood Estimation of Autoregressive–Moving Average Models by Means of Kalman Filtering," *Journal of the Royal Statistical Society, Ser. C*, 29 (3), 311–322. [1517]

Gilchrist, W. (1976), *Statistical Forecasting*, Chichester: Wiley. [1515]

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Downloaded by [University of Washington Libraries] at 13:41 30 March 2016

Gómez, V., and Maravall, A. (1998), "Programs TRAMO and SEATS, Instructions for the Users," Working Paper 97001, Direcciòn General de Anilisis y Programaciòn Presupuestaria, Ministerio de Economia y Hacienda. [1525]

Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J., and Vahid-Araghi, F. (2008), "Forecasting Time Series With Multiple Seasonal Patterns," *European Journal of Operational Research*, 191 (1), 207–222. [1513,1515]

Hannan, E. J., Terrel, R., and Tuckwell, N. (1970), "The Seasonal Adjustment of Economic Time Series," *International Economic Review*, 11, 24–52. [1516]

Harrison, P. J., and Stevens, C. F. (1976), "Bayesian Forecasting," *Journal of the Royal Statistical Society, Ser. B*, 38 (3), 205–247. [1520]

Harvey, A. (1989), *Forecasting Structural Time Series Models and the Kalman Filter*, New York: Cambridge University Press. [1515,1516]

Harvey, A., and Koopman, S. J. (1993), "Forecasting Hourly Electricity Demand Using Time-Varying Splines," *Journal of the American Statistical Association*, 88, 1228–1236. [1513]

Harvey, A., Koopman, S. J., and Riani, M. (1997), "The Modeling and Seasonal Adjustment of Weekly Observations," *Journal of Business & Economic Statistics*, 15, 354–368. [1513]

Hyndman, R., and Khandakar, Y. (2008), "Automatic Time Series Forecasting: The Forecast Package for R," *Journal of Statistical Software*, 26 (3), 1–22. [1519]

Hyndman, R. J. (2010), "forecast: Forecasting Functions for Time Series," R package version 2.09. [1519,1526]

Hyndman, R. J., Akram, M., and Archibald, B. C. (2007), "The Admissible Parameter Space for Exponential Smoothing Models," *Annals of the Institute of Statistical Mathematics*, 60, 407–426. [1519]

Hyndman, R. J., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2005), "Prediction Intervals for Exponential Smoothing Using Two New Classes of State Space Models," *Journal of Forecasting*, 24, 17–37. [1519]

——— (2008), *Forecasting With Exponential Smoothing: The State Space Approach*, Berlin: Springer-Verlag. Available at *www.exponentialsmoothing.net*. [1513,1515,1516,1518-1520,1526]

Hyndman, R. J., Koehler, A. B., Snyder, R. D., and Grose, S. (2002), "A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods," *International Journal of Forecasting*, 18 (3), 439–454. [1515]

Johnson, N. L. (1949), "Systems of Frequency Curves Generated by Methods of Translation," *Biometrika*, 36, 149–176. [1526]

Koenker, R., and Ng, P. (2003), "Sparsem: A Sparse Matrix Package for R," *Journal of Statistical Software*, 8 (6), 1–9. [1526]

Kohn, R., and Ansley, C. (1985), "Computing the Likelihood and Its Dierivatives for a Gaussian ARMA Model," *Journal of Statistical Computation and Simulation*, 22 (3), 229–263. [1518]

Lin, J., and Liu, T. (2002), "Modeling Lunar Calendar Holiday Effects in Taiwan," *Taiwan Economic Forecast and Policy*, 33 (1), 1–37. [1513]

Makridakis, S., and Hibon, M. (2000), "The M3-Competition: Results, Conclusions and Implications," *International Journal of Forecasting*, 16, 451–476. [1515]

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., and Winkler, R. (1982), "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition," *International Journal of Forecasting*, 1, 111–153. [1515]

Makridakis, S., Wheelwright, S. C., and Hyndman, R. J. (1998), *Forecasting: Methods and Applications* (3rd ed.), New York: Wiley. [1525]

McKenzie, E. (1984), "General Exponential Smoothing and the Equivalent ARMA Process," *Journal of Forecasting*, 3, 333–344. [1517]

Melard, G. (1984), "Algorithm AS 197: A Fast Algorithm for the Exact Likelihood of Autoregressive–Moving Average Models," *Journal of the Royal Statistical Society, Ser. C*, 33 (1), 104–114. [1517]

Nelder, J., and Mead, R. (1965), "A Simplex Method for Function Minimization," *The Computer Journal*, 7 (4), 308. [1519]

Ord, J. K., Koehler, A. B., and Snyder, R. D. (1997), "Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models," *Journal of the American Statistical Association*, 92, 1621–1629. [1515]

Pankratz, A., and Dudley, U. (1987), "Forecasts of Power-Transformed Series," *Journal of Forecasting*, 6, 239–248. [1519]

Pedregal, D. J., and Young, P. C. (2006), "Modulated Cycles, an Approach to Modelling Periodic Components From Rapidly Sampled Data," *International Journal of Forecasting*, 22, 181–194. [1513]

Proietti, T. (2000), "Comparing Seasonal Components for Structual Time Series Models," *International Journal of Forecasting*, 16, 247–260. [1516]

Proietti, T., and Riani, M. (2009), "Transformations and Seasonal Adjustment," *Journal of Time Series Analysis*, 30, 47–69. [1519]

Reid, D. J. (1975), "A Review of Short-Term Projection Techniques," in *Practical Aspects of Forecasting*, ed. H. A. Gordon, London: Operational Research Society, pp. 8–25. [1515]

Riazuddin, R., and Khan, M. (2005), "Detection and Forecasting of Islamic Calendar Effects in Time Series Data," *State Bank of Pakistan Research Bulletin*, 1 (1), 25–34. [1513]

Snyder, R. (1985a), "Estimation of Dynamic Linear Model—Another Approach," technical report, Dept. of Econometrics and Operations Research, Monash University. [1518]

——— (1985b), "Recursive Estimation of Dynamic Linear Models," *Journal of the Royal Statistical Society, Ser. B*, 47 (2), 272–276. [1518]

——— (2006), "Discussion," *International Journal of Forecasting*, 22 (4), 673–676. [1516]

Snyder, R., Koehler, A., and Ord, J. (2002), "Forecasting for Inventory Control With Exponential Smoothing," *International Journal of Forecasting*, 18 (1), 5–18. [1515]

Taylor, J. (2008), "An Evaluation of Methods for Very Short-Term Load Forecasting Using Minute-by-Minute British Data," *International Journal of Forecasting*, 24 (4), 645–658. [1516]

——— (2010a), "Exponentially Weighted Methods for Forecasting Intraday Time Series With Multiple Seasonal Cycles," *International Journal of Forecasting*, 26, 627–646. [1526]

——— (2010b), "Triple Seasonal Methods for Short-Term Electricity Demand Forecasting," *European Journal of Operational Research*, 204, 139–152. [1513,1515,1516]

Taylor, J. W. (2003), "Short-Term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing," *Journal of the Operational Research Society*, 54, 799–805. [1513,1515,1516]

Taylor, J. W., and Snyder, R. D. (2009), "Forecasting Intraday Time Series With Multiple Seasonal Cycles Using Parsimonious Seasonal Exponential Smoothing," Technical Report 09/09, Dept. of Econometrics and Business Statistics, Monash University. [1513]

Weinberg, J., Brown, L., and Stroud, J. (2007), "Bayesian Forecasting of an Inhomogeneous Poisson Process With Applications to Call Center Data," *Journal of the American Statistical Association*, 102 (480), 1185–1198. [1521]

West, M., and Harrison, J. (1997), *Bayesian Forecasting and Dynamic Models* (2nd ed.), New York: Springer-Verlag. [1516]