



MONASH University

Department of Econometrics and Business Statistics

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

**Forecasting time series with
complex seasonal patterns using
exponential smoothing**

Alysha M De Livera and Rob J Hyndman

December 2009

Working Paper 15/09

Forecasting time series with complex seasonal patterns using exponential smoothing

Alysha M De Livera

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

Email: alysha.delivera@buseco.monash.edu

Division of Mathematics, Informatics and Statistics,
Commonwealth Scientific and Industrial Research Organisation,
Clayton, VIC 3168
Australia.

Email: alysha.delivera@csiro.au

Rob J Hyndman

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

Email: rob.hyndman@buseco.monash.edu

12 December 2009

JEL classification: C22,C53

Forecasting time series with complex seasonal patterns using exponential smoothing

Abstract

A new innovations state space modeling framework, incorporating Box-Cox transformations, Fourier series with time varying coefficients and ARMA error correction, is introduced for forecasting complex seasonal time series that cannot be handled using existing forecasting models. Such complex time series include time series with multiple seasonal periods, high frequency seasonality, non-integer seasonality and dual-calendar effects. Our new modelling framework provides an alternative to existing exponential smoothing models, and is shown to have many advantages. The methods for initialization and estimation, including likelihood evaluation, are presented, and analytical expressions for point forecasts and interval predictions under the assumption of Gaussian errors are derived, leading to a simple, comprehensible approach to forecasting complex seasonal time series. Our trigonometric formulation is also presented as a means of decomposing complex seasonal time series, which cannot be decomposed using any of the existing decomposition methods. The approach is useful in a broad range of applications, and we illustrate its versatility in three empirical studies where it demonstrates excellent forecasting performance over a range of prediction horizons. In addition, we show that our trigonometric decomposition leads to the identification and extraction of seasonal components, which are otherwise not apparent in the time series plot itself.

Keywords: exponential smoothing, Fourier series, prediction intervals, seasonality, state space models, time series decomposition.

1 Introduction

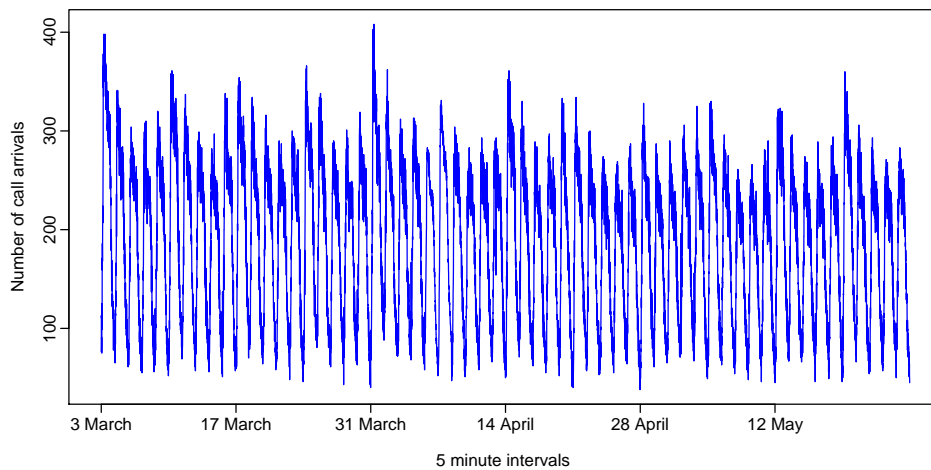
Many time series exhibit complex seasonal patterns. For example, Figure 1(a) shows the number of retail banking call arrivals per 5-minute interval between 7:00am and 9:05pm each weekday. There is a daily seasonal pattern with frequency 169 and a weekly seasonal pattern with frequency $169 \times 5 = 845$. If a longer series of data were available, there may also be an annual seasonal pattern. Such multiple seasonal patterns are becoming more common with high frequency data recording. Further examples where multiple seasonal patterns can occur include daily hospital admissions, requests for cash at ATMs, electricity and water usage, and access to computer web sites.

Other time series (most commonly weekly data) have patterns with a non-integer frequency. Figure 1(b) shows the weekly United States finished motor gasoline products in thousands of barrels per day. The time series clearly exhibits an annual seasonal pattern with frequency $365.25/7 \approx 52.179$.

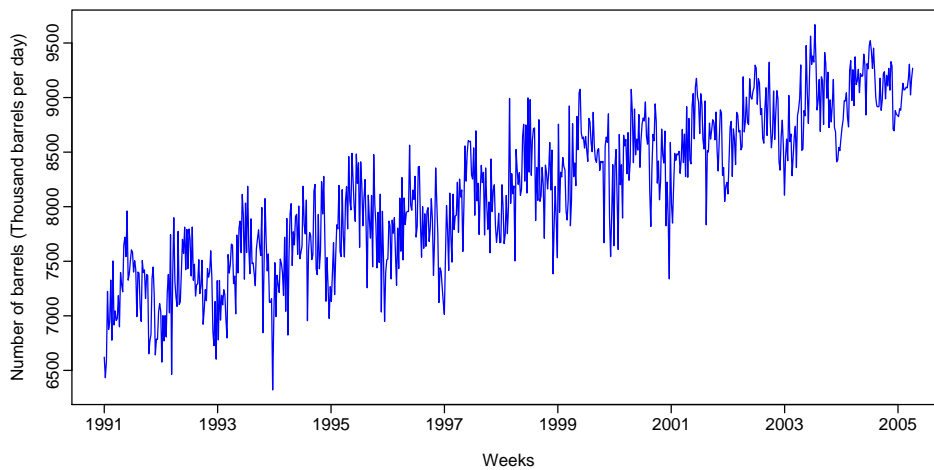
In addition, some time series may have dual-calendar seasonal effects. Figure 1(c) shows the daily electricity demand in Turkey over nine years, from 1 January 2000 to 31 December 2008. A clear weekly seasonal pattern and an annual seasonal pattern can be observed in the time series. The annual seasonality consists of two separate seasonal patterns with frequencies of 354.37 and 365.25, following the Hijri and Gregorian calendars respectively. 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 & Liu 2002, Riazuddin & Khan 2005). Unlike the multiple periodicities 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 periodicity (such as 12 for monthly data or 4 for quarterly data). There are a few models which attempt to deal with more complex seasonal patterns (e.g., Harvey & Koopman (1993), Harvey et al. (1997), Pedregal & Young (2006), Taylor (2003b), Gould et al. (2008), Taylor & Snyder (2009), Taylor (2009)), but none that is able to handle all of the complexities above.

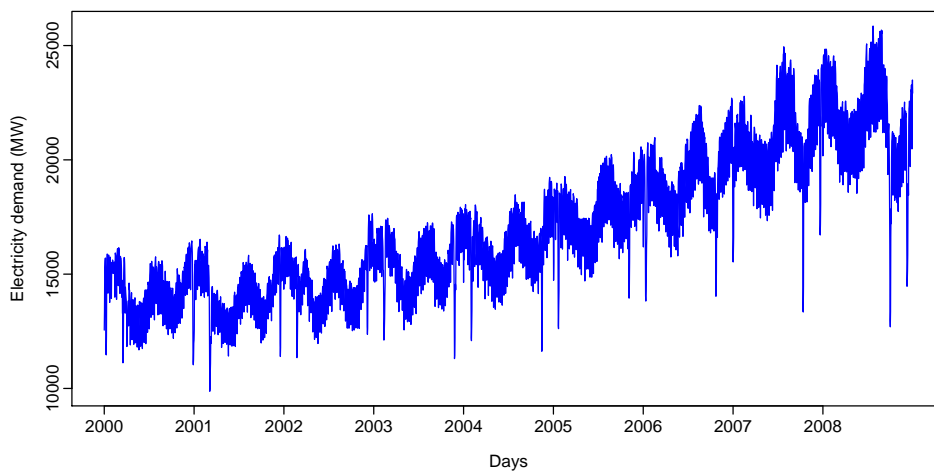
In this paper we introduce a new innovations state space modeling framework based on a trigonometric formulation which is capable of tackling all of these seasonal complexities. Using the above time series, we show that these trigonometric exponential smoothing models provide exceptional



(a) Number of call arrivals handled on weekdays between 7am and 9:05pm from March 3, 2003, to May 23, 2003 in a large North American commercial bank.



(b) US finished motor gasoline products supplied (thousands of barrels per day), from February 1991 to July 2005.



(c) Turkish electricity demand data from January 1, 2000, to December 31, 2008.

Figure 1: Examples of complex seasonality showing (a) multiple nested seasonal periods, (b) non-integer seasonal periods and (c) multiple non-nested and non-integer seasonal periods.

out-of-sample forecasting performances and offer an elegant decomposition of complex seasonal time series.

In Section 2 we discuss the existing exponential smoothing models, their weaknesses and their inadequacy in handling complex seasonal patterns, and present a modified, generalized modeling framework in order to overcome these problems. We then introduce in Section 3 the new trigonometric innovations state space modeling framework, which is capable of handling complex seasonal patterns, as well as the usual single seasonal patterns, in a straightforward manner with fewer parameters. Section 4 describes both analytical and simulated prediction distributions, as well as point and interval predictions for the models. Section 5 presents the methods used for initialization and estimation, including the derivation of maximum likelihood estimators and the methodology used in applying the models. In Section 6, we explain the trigonometric formulation as a way of decomposing complex seasonal time series, which cannot be decomposed using any of the existing decomposition techniques. The proposed models are then applied to the US gasoline products data, the call center data and the Turkey electricity demand data in Section 7, and it is shown that these new trigonometric exponential smoothing models provide outstanding forecasting performances over a range of forecasting horizons, compared to the existing models. Furthermore, using these applications, we demonstrate the decomposition of complex seasonal time series using our trigonometric approach. Some conclusions are drawn in Section 8.

2 Exponential smoothing models for seasonal data

2.1 Traditional models

Single seasonal exponential smoothing methods are among the most widely used forecasting procedures in practice (Snyder et al. 2002, Makridakis et al. 1982, Makridakis & Hibon 2000). These methods have been shown to be optimal for a class of innovations state space models (Ord et al. 1997, Hyndman et al. 2002), thus allowing a stochastic modelling framework for exponential smoothing including likelihood calculation, prediction intervals, model selection, and so on. The single source of error or innovations approach is known to be simple yet robust, and has been shown to have several advantages over the multiple source of error models (Ord et al. 2005, Hyndman et al. 2008).

The most commonly employed seasonal models in the innovations state space modeling framework include the underlying models for the well-known Holt-Winters' additive and multiplicative seasonal exponential smoothing methods. However, these models are inadequate for handling complex

seasonal time series such as multiple seasonality, non-integer seasonality and dual-calendar effects. [Taylor \(2003b\)](#) extended the single seasonal Holt-Winters' model to accommodate a second seasonal component in order to handle time series with two seasonal patterns. This requires a large number of values to be estimated for the initial seasonal components, especially when the frequencies of the seasonal patterns are high, which may lead to over-parameterization. [Gould et al. \(2008\)](#) attempted to reduce the over-parameterization of this model by dividing the longer seasonal length into sub-seasonal cycles that have similar patterns. However, this model is relatively complex and can only be used in modeling double seasonal patterns when one seasonality is a multiple of the other. Using six years of British and French electricity demand data, [Taylor \(2009\)](#) illustrated that extended additive seasonal versions of the above models to handle a third seasonal pattern can outperform the double seasonal exponential smoothing models. However, none of these models can be used to model complex seasonal patterns such as non-integer seasonality and calendar effects, or time series with more than two non-nested seasonal patterns.

In addition, the non-linear versions of exponential smoothing models, although widely used, suffer from some important weaknesses. [Akram et al. \(2009\)](#) showed that most non-linear seasonal exponential smoothing models can be unstable, having infinite forecast variances beyond a certain forecasting horizon. Of the multiplicative error models which do not have this flaw, [Akram et al. \(2009\)](#) proved that sample paths will converge almost surely to zero even when the error distribution is non-Gaussian. Furthermore, for non-linear exponential smoothing models, analytical results for the prediction distributions are not available.

The models used for exponential smoothing assume that the error process is serially uncorrelated. However, the assumption of an uncorrelated error process does not always hold. In an empirical study, using the Holt-Winters' method for multiplicative seasonality, [Chatfield \(1978\)](#) showed that the error process is correlated and can be described by an AR(1) process. This was further illustrated by [Taylor \(2003b\)](#) in a study of electricity demand forecasting using a double-seasonal Holt-Winters' multiplicative method. [Gardner \(1985\)](#), [Reid \(1975\)](#), and [Gilchrist \(1976\)](#) have also mentioned this issue of correlated errors, in improving the forecast accuracy.

2.2 Modified models

We now consider various modifications to the standard exponential smoothing models to enable them to handle a wider variety of seasonal patterns, and to deal with the problems raised above.

Extending non-linear exponential smoothing models to handle more than two seasonal patterns may make these models unnecessarily complex, and the estimation and model selection procedure may become cumbersome. Also, the problems with non-linear models that are noted above are also a problem in any extended versions. Consequently, rather than allow non-linear forms, we restrict attention to linear homoscedastic models but allow some types of non-linearity using Box-Cox transformations (Box & Cox 1964). 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 exponential smoothing models to accommodate T seasonal patterns as follows.

$$\begin{aligned}
 y_t^{(\omega)} &= \begin{cases} \frac{y_t^\omega - 1}{\omega}; & \omega \neq 0 \\ \log y_t & \omega = 0 \end{cases} \\
 y_t^{(\omega)} &= \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^T s_{t-m_i}^{(i)} + d_t \\
 \ell_t &= \ell_{t-1} + \phi b_{t-1} + \alpha d_t \\
 b_t &= \phi b_{t-1} + \beta d_t \\
 s_t^{(i)} &= s_{t-m_i}^{(i)} + \gamma_i d_t \\
 d_t &= \sum_{i=1}^p \varphi_i d_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t,
 \end{aligned} \tag{1}$$

where m_1, \dots, m_T denote the seasonal periods, ℓ_t and b_t represent the level and trend components of the series at time t , respectively, $s_t^{(i)}$ represents the i th 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, γ_i for $i = 1, \dots, T$, and ϕ is the dampening parameter, which gives more control over trend extrapolation when the trend component is damped (Hyndman et al. 2002, Taylor 2003a).

The notation BATS($p, q, m_1, m_2, \dots, m_T$) is used for these models. B stands for the Box-Cox transformation, A stands for the ARMA residuals, T stands for the trend component in the model and S stands for the seasonal components. The arguments indicate the ARMA parameters (p and q) and the seasonal periods (m_1, \dots, m_T). For example, BATS(0, 0, m_1) with $\phi = 1$ and $\omega = 1$ 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 (2003b) is given by BATS(0, 0, m_1, m_2) with $\phi = 1$ and $\omega = 1$, and that with the residual AR(1) adjustment in the model

of Taylor (2003b, 2008) is given by $BATS(1, 0, m_1, m_2)$. The Holt-Winters' additive triple seasonal model with AR(1) adjustment in Taylor (2009) is given by $BATS(1, 0, m_1, m_2, m_3)$.

The BATS model is the most obvious generalization of the traditional exponential smoothing models to allow for multiple seasonal periods. However, it is not capable of handling non-integer seasonality, and it suffers from a very large number of parameters that require estimation; the initial seasonal component alone contains $m_1 + m_2 + \dots + m_T$ parameters. This becomes a huge number of values when the frequencies of the seasonal patterns are high. For example, for the call center data shown in Figure 1(a), $169 + 845 = 1014$ initial seasonal values must be estimated.

3 Trigonometric exponential smoothing models for seasonal data

Consequently, we introduce a new trigonometric representation of seasonal components based on Fourier series. We could replace the equation for $s_t^{(i)}$ in the BATS model with

$$s_t^{(i)} = \sum_{j=1}^{k_i} \alpha_{j,t}^{(i)} \cos(\lambda_j^{(i)} t) + \beta_{j,t}^{(i)} \sin(\lambda_j^{(i)} t) \quad (2a)$$

$$\alpha_{j,t}^{(i)} = \alpha_{j,t-1}^{(i)} + \kappa_1^{(i)} d_t \quad (2b)$$

$$\beta_{j,t}^{(i)} = \beta_{j,t-1}^{(i)} + \kappa_2^{(i)} d_t, \quad (2c)$$

where $\kappa_1^{(i)}$ and $\kappa_2^{(i)}$ are the smoothing parameters and $\lambda_j^{(i)} = 2\pi j/m_i$. This is an extended, modified single source of error version of a single seasonal multiple source of error representation suggested by Hannon et al. (1970), and 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 . But most seasonal terms will require much smaller values of k_i , thus reducing the number of parameters to be estimated.

In the single seasonal multiple source of error setting (Harvey 1989), an alternative, but equivalent formulation of representation (2) is preferred (Durbin & Koopman 2001), which can be obtained by re-parameterizing the single seasonal multiple source of error version of (2) using

$$\alpha_{j,t}^{(i)} = s_{j,t}^{(i)} \cos(\lambda_j^{(i)} t) - s_{j,t}^{*(i)} \sin(\lambda_j^{(i)} t)$$

and

$$\beta_{j,t}^{(i)} = s_{j,t}^{(i)} \sin(\lambda_j^{(i)} t) + s_{j,t}^{*(i)} \cos(\lambda_j^{(i)} t).$$

For our modified multiple seasonal single source of error formulation, it can be shown (see Appendix A) that the above re-parametrization leads to the following:

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)}, \quad (3)$$

$$\text{where } s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \left[\kappa_1^{(i)} \cos(\lambda_j^{(i)} t) + \kappa_2^{(i)} \sin(\lambda_j^{(i)} t) \right] d_t$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \left[\kappa_2^{(i)} \cos(\lambda_j^{(i)} t) - \kappa_1^{(i)} \sin(\lambda_j^{(i)} t) \right] d_t$$

This then gives rise to a heteroscedastic error process. However, to be consistent with the homoscedastic nature of the traditional additive innovations state space models, the following representation is employed in our models:

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

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

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

In the single seasonal multiple source of error setting, (2) is equivalent to (4) (Proietti 2000). In our multiple seasonal single source of error setting, the seasonal representations (2) and (4) are equivalent only when the smoothing parameters are equal to zero. This then reduces to a standard Fourier series representation when the initial state values $s_{j,0}^{(i)}$ and $s_{j,0}^{*(i)}$ are set to Fourier series coefficients for the i th seasonal component. Then, following Koopman & Lee (2005), it can be shown that $s_{j,t}^{*(i)} = \frac{1}{\lambda_j^{(i)}} \frac{ds_{j,t}^{(i)}}{dt}$. Thus $s_{j,t}^{*(i)}$ is proportional to the rate of growth in $s_{j,t}^{(i)}$ for each seasonal component. Hence, in the representation (4), the term $s_{j,t}^{(i)}$ could be defined as the stochastic *level* term of the i th seasonal component and the term $s_{j,t}^{*(i)}$ as a stochastic *auxiliary* variable describing the growth in the *level* of the i th seasonal component that is needed to describe the change in the seasonal component over time.

Replacing the seasonal component $s_t^{(i)}$ in (1) with that given by (4), a new class of exponential smoothing models, termed *trigonometric exponential smoothing models*, is obtained. The notation TBATS($p, q, \{m_1, k_1\}, \{m_2, k_2\}, \dots, \{m_T, k_T\}$) is used for these trigonometric models. These models require the estimation of $2(k_1 + k_2 + \dots + k_T)$ initial seasonal values, which is expected to greatly reduce the number of parameters required compared to BATS models. The use of trigonometric functions also allows the modelling of non-integer seasonal frequencies.

4 Point forecasts, prediction distributions and interval predictions

The BATS and TBATS models can be written in the following linear innovations state space form:

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

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

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 .

To obtain the matrices in (5a), we first define $\tilde{\mathbf{s}}_t^{(i)} = (s_{1,t}^{(i)}, s_{2,t}^{(i)}, \dots, s_{k_i,t}^{(i)})$, $\tilde{\mathbf{s}}_t^{*(i)} = (s_{1,t}^{*(i)}, s_{2,t}^{*(i)}, \dots, s_{k_i,t}^{*(i)})$, $\mathbf{s}_t^{(i)} = (\tilde{\mathbf{s}}_t^{(i)}, \tilde{\mathbf{s}}_t^{*(i)})'$; let $\mathbf{1}_r = (1, 1, \dots, 1)$ and $\mathbf{0}_r = (0, 0, \dots, 0)$ be row vectors of length r ; let $\boldsymbol{\gamma}_1^{(i)} = \gamma_1^{(i)} \mathbf{1}_{k_i}$, $\boldsymbol{\gamma}_2^{(i)} = \gamma_2^{(i)} \mathbf{1}_{k_i}$, $\boldsymbol{\gamma}^{(i)} = (\boldsymbol{\gamma}_1^{(i)}, \boldsymbol{\gamma}_2^{(i)})'$, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_p)$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$; 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 $\mathbf{a}^{(i)} = (\mathbf{1}_{k_i}, \mathbf{0}_{k_i})'$. We shall also need the matrices

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\gamma}^{(1)} \boldsymbol{\varphi} \\ \vdots \\ \boldsymbol{\gamma}^{(T)} \boldsymbol{\varphi} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \boldsymbol{\gamma}^{(1)} \boldsymbol{\theta} \\ \vdots \\ \boldsymbol{\gamma}^{(T)} \boldsymbol{\theta} \end{bmatrix}, \quad \mathbf{A}_i = \begin{bmatrix} \mathbf{C}^{(i)} & \mathbf{S}^{(i)} \\ -\mathbf{S}^{(i)} & \mathbf{C}^{(i)} \end{bmatrix}, \quad \tilde{\mathbf{A}}_i = \begin{bmatrix} \mathbf{O}_{m_i-1} & 1 \\ \mathbf{I}_{m_i-1} & \mathbf{O}'_{m_i-1} \end{bmatrix},$$

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 $j = 1, 2, \dots, k_i$ and $i = 1, \dots, T$, and where \bigoplus denotes the direct sum of the matrices. Let $\tau = 2 \sum_{i=1}^T k_i$.

Then the TBATS model can be written in the form (5a) with

$$\begin{aligned} \mathbf{x}_t &= \left(\ell_t, b_t, \mathbf{s}_t^{(1)'}, \dots, \mathbf{s}_t^{(T)'}, d_t, d_{t-1}, \dots, d_{t-p+1}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1} \right)' \\ \mathbf{w} &= \left(1, \boldsymbol{\phi}, \mathbf{a}^{(1)'}, \dots, \mathbf{a}^{(T)'}, \boldsymbol{\varphi}, \boldsymbol{\theta} \right)' \\ \mathbf{g} &= \left(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}^{(1)'}, \dots, \boldsymbol{\gamma}^{(T)'}, 1, \mathbf{0}_{p-1}, 1, \mathbf{0}_{q-1} \right)' \end{aligned}$$

and

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

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

The state space form of the BATS model can be obtained by letting $\mathbf{s}_t^{(i)} = (s_t^{(i)}, s_{t-1}^{(i)}, \dots, s_{t-(m_i-1)}^{(i)})'$, $\mathbf{a}^{(i)} = (\mathbf{0}_{m_i-1}, 1)'$, $\boldsymbol{\gamma}^{(i)} = (\gamma_i, \mathbf{0}_{m_i-1})'$, $A = \bigoplus_{i=1}^T \tilde{A}_i$, and by replacing $2k_i$ with m_i in the matrices presented above for the TBATS models.

Let $\boldsymbol{\theta}$ be a vector of all parameters to be estimated in the model, consisting of the smoothing parameters and the Box-Cox parameter, and let n be the length of the time series, h the length of the forecast horizon, and $y_{n+h|n} \equiv y_{n+h} | \mathbf{x}_n$, $\boldsymbol{\theta}$ the prediction distribution of a future value of the series given the model, its estimated smoothing parameters and the state vector at the last observation. A Gaussian assumption for the errors implies that $y_{n+h|n}^{(\omega)}$ is also normally distributed, with mean $E(y_{n+h|n}^{(\omega)})$ and variance $V(y_{n+h|n}^{(\omega)})$ given by the equations (Hyndman et al. 2005):

$$E(y_{n+h|n}^{(\omega)}) = \mathbf{w}' \mathbf{F}^{h-1} \mathbf{x}_n \quad (6a)$$

$$V(y_{n+h|n}^{(\omega)}) = \begin{cases} \sigma^2 & \text{if } h = 1; \\ \sigma^2 \left[1 + \sum_{j=1}^{h-1} c_j^2 \right] & \text{if } h \geq 2; \end{cases} \quad (6b)$$

where $c_j = \mathbf{w}' \mathbf{F}^{j-1} \mathbf{g}$. Point forecasts and forecast intervals may be obtained using the inverse Box-Cox transformation. Point forecasts obtained this way are equal to the median of the conditional probability density of $y_{n+h|n}$. Forecast intervals retain the required probability coverage under back-transformation because the Box-Cox transformation is monotonic.

5 Estimation and model selection

5.1 Maximum likelihood estimation

The Box-Cox parameter and smoothing parameters, $\boldsymbol{\vartheta}$, and the initial states, \mathbf{x}_0 , can be estimated from the observed data $\mathbf{y} = (y_1, \dots, y_n)$ by maximizing the likelihood. If $\varepsilon_t \sim N(0, \sigma^2)$ and \mathbf{x}_0 and $\boldsymbol{\vartheta}$ are known, then $y_t^{(\omega)} \sim N(\mathbf{w}'\mathbf{x}_{t-1}, \sigma^2)$. Thus,

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

and so

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

Therefore, the log likelihood is given by

$$\mathcal{L} = \frac{-n}{2} \log(2\pi\sigma^2) - \frac{\sum_{t=1}^n \varepsilon_t^2}{2\sigma^2} + (\omega - 1) \sum_{t=1}^n \log y_t. \quad (7)$$

The maximum likelihood estimate of the error variance is obtained by letting the partial derivative of \mathcal{L} with respect to σ^2 be equal to zero, giving

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^n \varepsilon_t^2 \quad (8)$$

Then, substituting (8) into (7), multiplying by -2 , and omitting constant terms, we obtain

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

This quantity can be minimized to obtain maximum likelihood estimates.

To start the optimization, we need some initial estimates of \mathbf{x}_0 and $\boldsymbol{\vartheta}$. If a Box-Cox transformation is required for the data, an initial value for the Box-Cox parameter ω has to be approximated first. This could be done by inspection of the data under different transformations or simply by letting $\omega = 0$ to start with. (We found this choice led to better model selection and forecasting than other initial values for ω .)

Let $m^* = \lfloor \max(m_1, \dots, m_T) \rfloor$. Then we compute a $2 \times m^*$ moving average through the first few seasons of transformed data. Denote this by $\{f_t\}$ ($t = m^*/2 + 1, m^*/2 + 2, \dots$), and let $z_t = y_t^{(\omega)} - f_t$. For TBATS models, the seasonal component is then approximated using

$$z_t \approx \sum_{i=1}^T \sum_{j=1}^{k_i} a_j^{(i)} \cos(\lambda_j^{(i)} t) + b_j^{(i)} \sin(\lambda_j^{(i)} t), \quad (10)$$

where $\hat{a}_j^{(i)}$ and $\hat{b}_j^{(i)}$ are estimated by regressing z_t against the trigonometric terms. The initial seasonal states for the i th seasonal component can then be set to $\hat{a}_j^{(i)}$ and $\hat{b}_j^{(i)}$. For BATS models, we follow the same procedure but replace k_i with $m_i/2$ for even frequencies and $(m_i - 1)/2$ for odd frequencies in equation (10), in order to obtain $\hat{a}_j^{(i)}$ and $\hat{b}_j^{(i)}$. Using these values, we define the initial seasonal state estimates as $\hat{z}_t^{(i)} = \sum_{j=1}^{k_i} [\hat{a}_j^{(i)} \cos(\lambda_j^{(i)} t) + \hat{b}_j^{(i)} \sin(\lambda_j^{(i)} t)]$.

The initial level and trend components can be approximated by computing a linear regression on the first m^* deseasonalized values, against a time variable $t = 1, \dots, m^*$. Then set the initial level to be the intercept of this regression and let the initial trend be equal to the slope. For BATS models with $T = 1$, this is equivalent to the procedure proposed by [Hyndman et al. \(2002\)](#).

These initial state values can then be optimized along with the Box-Cox parameter and the smoothing parameters. In practice, the initial seasonal states can only be optimized when there are a small number of values. For this reason, when handling high frequency data, the initial seasonal values can not be optimized in the BATS model. One of the advantages of the TBATS model is that there are fewer initial seasonal values, and hence the initial seasonal states may be optimized.

For the BATS model, the seasonal values are constrained when optimizing, so that each seasonal component sums to zero. For both models, using the matrices and vectors \mathbf{D} , \mathbf{g} and \mathbf{w} from equation (5a), the smoothing parameters are restricted to the *forecastability region* given by [Hyndman et al. \(2007\)](#). Restricting the parameters in this way, rather than restricting them to the usual parameter region of $[0, 1]$, serves two purposes. First, it guarantees stable forecasts. Second, if the usual parameter region lies within the forecastability region, then restricting the parameters to the usual region may lead to inferior forecasts as the maximum likelihood parameters may lie outside that region. In addition, ARMA coefficients are restricted to the causality and invertibility regions.

5.2 Model selection

In this paper, the AIC = $\mathcal{L}^*(\hat{\boldsymbol{\theta}}, \hat{\mathbf{x}}_0) + 2K$ is used for choosing between the models, where K is the total number of parameters in $\boldsymbol{\theta}$ plus the number of free states in \mathbf{x}_0 , and $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{x}}_0$ denote the estimates

of θ and x_0 . When one of the smoothing parameters takes the boundary value 0, the value of K is reduced by one 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 one in each instance to account for the resulting simplified model.

In an empirical study, [Billah et al. \(2005\)](#) indicated that information criterion approaches, such as the AIC, provide the best basis for automated model selection, relative to other methods such as the prediction-validation method. Alternative information criteria such as the AICc or BIC ([Burnham & Anderson 2002](#)) may also be used.

Selecting the number of harmonics k_i in the trigonometric models

The forecasts from the TBATS model depend on the number of harmonics k_i used for the i th seasonal component. It would be impractical to consider all possible combinations when searching for the optimal k_i combination. Instead, the following procedure was used to obtain the number of harmonics for the applications in this paper. In practice we have found that this approach leads to good models and further computational effort rarely leads to much improvement.

Using the first few seasons, we use multiple linear regression applied to (10) to find the trigonometric coefficients. Starting with a single harmonic, we 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 i th seasonal component. We then fit the required model to the data with $k_i = k_i^*$, and compute the AIC. Considering one seasonal component at a time, we 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.

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 & 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. The ARMA component is only retained if the resulting model has lower AIC than the model with no ARMA component.

6 Decomposition of complex seasonal time series

Decomposing a time series into constituent latent subseries is a vital aspect in practical, retrospective time series analysis such as deseasonalisation, analyzing seasonal effects and isolating latent, quasi-cyclical components (West & Harrison 1997, Pole et al. 1994). None of the existing decomposition methods are capable of handling all of the seasonal complexities described in this paper, such as multiple seasonality, non-integer seasonality, and dual-calendar effects. Our trigonometric formulation offers an elegant way of decomposing complex seasonal time series into trend, seasonal and irregular components. In particular, it leads to the identification and extraction of one or more seasonal components, which may not be apparent in the time series plots themselves.

Using our trigonometric models for multiple seasonal time series, the overall seasonal component can be decomposed into several individual seasonal components with different frequencies. Each of these individual seasonal components are obtained by $s_t^{(i)}$ of the model 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 the model.

In decomposing time series, TBATS approach has several important advantages over BATS. First, seasonal components obtained from BATS model are not normalized. 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 & 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. 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.

In addition, BATS model cannot be used to decompose time series with non-integer seasonality and dual calendar effects.

7 Empirical analysis

7.1 Application to weekly US 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. These data were part of the Transportation competition. (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 data consist of 745 observations and were split into two segments: an estimation sample period (484 observations) and test sample (261 observations). The estimation sample was used to estimate the initial values, select the appropriate number of harmonics, and estimate the smoothing parameters. Following the procedure for finding the number of harmonics to start with, it was found that only one harmonic was highly significant. Hence, starting with $k_1^* = 1$, the initial values were estimated using the heuristic method. The model was then fitted to the whole estimation sample of 484 values and optimized to minimize equation (9). The values of the AIC decreased until $k_1 = 7$, then started to increase.

In order to investigate the out-of-sample performance, we computed the Root Mean Square Error (RMSE), defined as

$$\text{RMSE} = \sqrt{\frac{1}{p-h+1} \sum_{t=n}^{n+p-h} (y_{t+h} - \hat{y}_{t+h|t})^2}, \quad (11)$$

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. Hence, the TBATS(0, 0, {365.25/7, 7}) model with $\omega = \phi = 1$ was fitted. 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, 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, 1, 52) model with $\omega = \phi = 1$ 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 clearly performs better for all lead times.

This result can be explained by some of the advantages the trigonometric representation has over the traditional seasonal representation. First, the BATS model cannot handle the non-integer frequency, and hence it had to be rounded off to the nearest integer. Second, the BATS model may be over-parameterized, as 52 initial seasonal values have to be estimated for the model. Third, these initial values cannot be optimized along with the smoothing parameters, due to the high dimensionality of the optimization space. These problems are overcome in the trigonometric formulation, which allows for non-integer frequencies and requires fewer initial values to be estimated for the initial seasonal component, hence allowing them to be optimized.

Decomposition of the Gasoline time series obtained using the fitted TBATS model is shown in Figure 3. The vertical bars at the right side of each plot are of equal heights but plotted on different scales; thus providing a comparison of the size of each component. The trigonometric formulation in TBATS models allows for more randomness to be eliminated from the seasonal component, yet allowing for those influential bumps to be revealed.

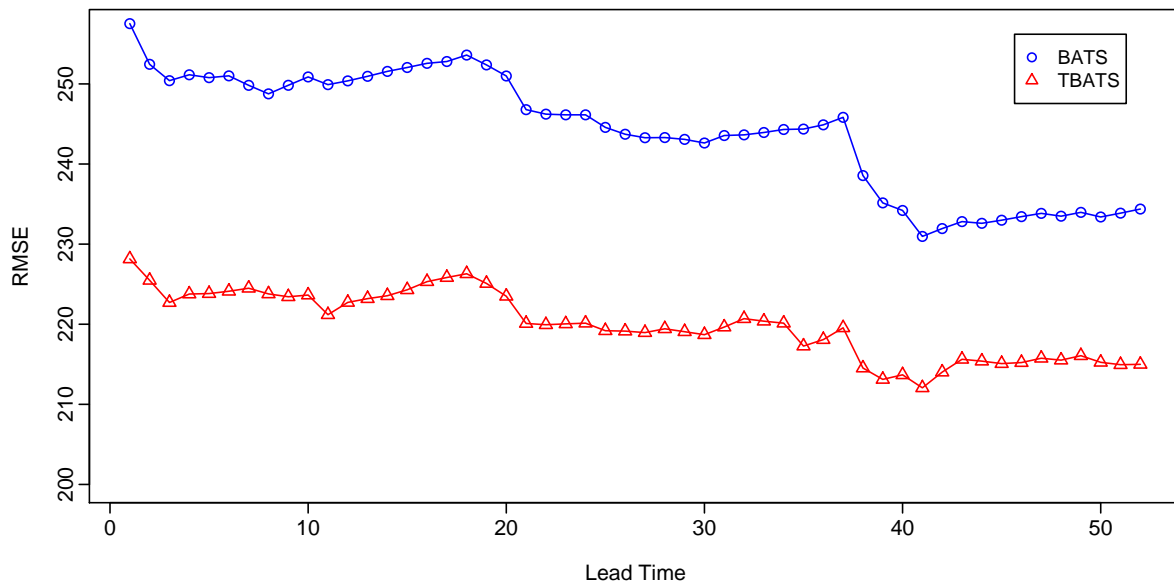


Figure 2: Out-of-sample results for the US gasoline data using BATS(0, 1, 52) and TBATS(0, 1, {365.25/7, 7}).

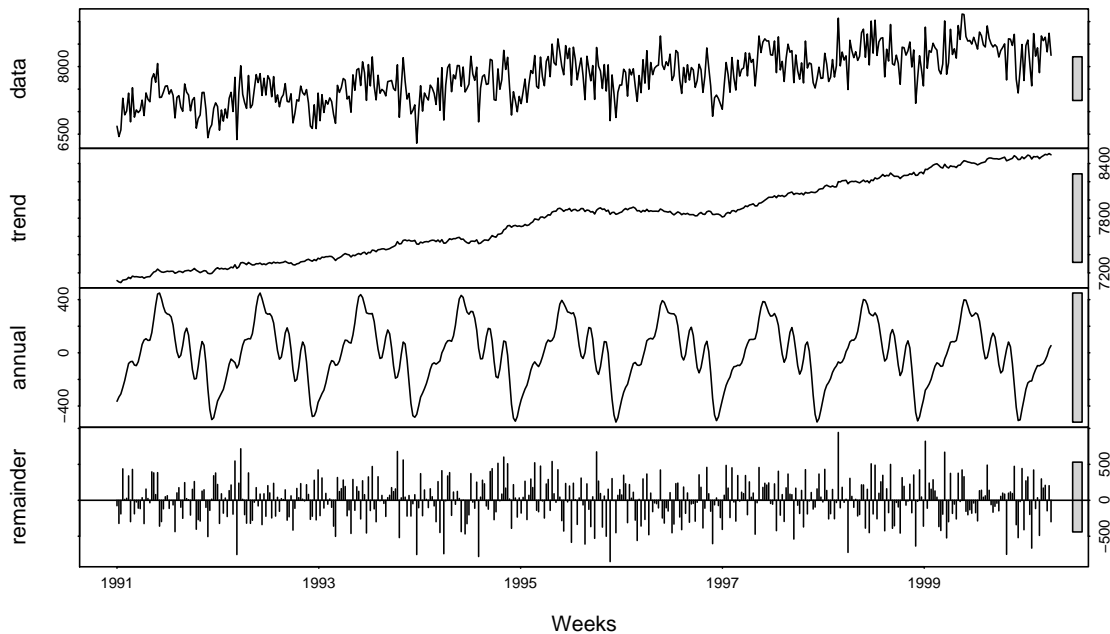


Figure 3: *Trigonometric decomposition of the US gasoline data.*

7.2 Application to call center data

The call center data shown in Figure 1(a) consist of 10,140 observations, that is 12 weeks of data starting from 3 March 2003 (Weinberg et al. 2007). They contain a daily seasonal pattern with frequency 169 and a weekly seasonal pattern with frequency $169 * 5 = 845$. The fitting sample consists of 7,605 observations (9 weeks). The trend appears to be close to zero, and hence the growth rate b_t was omitted from the models.

The model selection procedure led to the model TBATS(3, 1, {169, 29}, {845, 15}) with $\omega = \phi = 1$ and $\beta = b_0 = 0$. The BATS model, with $m_1 = 169$ and $m_2 = 845$, was then considered. The model selection procedure led to the BATS(3, 0, 169, 845) model with $\omega = 0.317$, $\phi = 1$ and $\beta = b_0 = 0$ was chosen. (Other BATS models with $\omega = 1$ were also tried, but their forecasting performance was worse.)

Figure 4 compares the post-sample forecasting accuracies of the selected BATS and TBATS models, and it can be seen that the TBATS model, which requires the estimation of 102 values, provides much greater forecasting accuracy than the BATS model, which requires the estimation of 1025 values, for all forecast horizons.

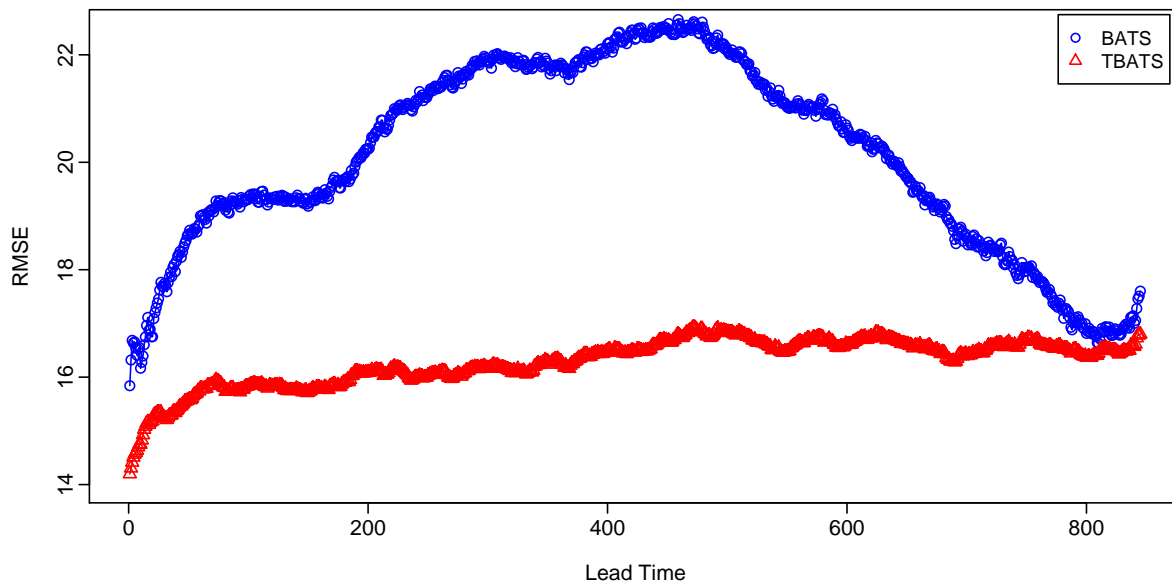


Figure 4: Out-of-sample results for the call center data using $BATS(3,0,169,845)$ and $TBATS(3,1,\{169,29\},\{845,15\})$

The decomposition obtained from TBATS is shown in Figure 5, which clearly exhibits strong daily and weekly seasonal components. The weekly seasonal pattern seems to be evolving rapidly with time while the daily seasonal pattern stays relatively more constant. As is seen from the time series plot itself, the trend component is very small in magnitude compared to the seasonal components.

7.3 Application to the Turkey electricity demand data

The Turkey electricity demand data shown in Figure 1(c) have 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 and Kurban 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.

In this study, the data, which cover a period of 9 years, are 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 order selection procedure was followed and the $TBATS(3,2,\{7,3\},\{354.37,23\},\{365.25,3\})$ model with $\omega = 0.00165$ and $\phi = 1$ was selected.

None of the existing exponential smoothing models can handle this time series; however, the BATS model was applied with the frequencies rounded off to the nearest integer (i.e., $m_1 = 7, m_2 =$

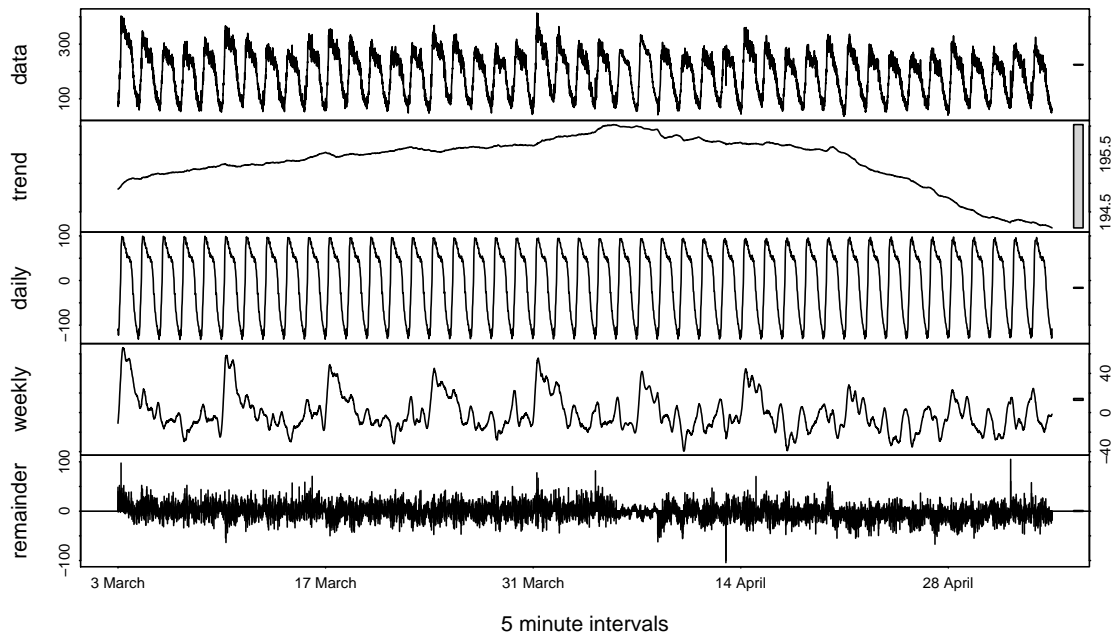


Figure 5: *Trigonometric decomposition of the call center data.*

354, $m_3 = 365$), which is expected to affect the accuracy of the forecasts. In addition, this model requires the estimation of $354 + 365 + 7 = 726$ values for the seasonal component alone.

Figure 6 shows the exceptional post-sample forecasting performance of the TBATS model compared to the BATS model.

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 which does not seem to have much variation with time. The fourth panel shows the overall annual component, which is composed of a seasonal component based on the Hijri calendar with frequency 354.37 and a seasonal component based on the Gregorian calendar with frequency 365.25, which are shown separately in the fifth and the sixth panels respectively.

The rather wiggly seasonal components has probably occurred 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 assume that the stochastic seasonal component is augmented by deterministic holiday effects, we can reduce the number of harmonics required by using dummy variables in handling those sharp drops which occur due to holidays. Table 1 gives the dates of the holidays from the Hijri and Gregorian calendars. Seker is a three-day festival when sweets are eaten to

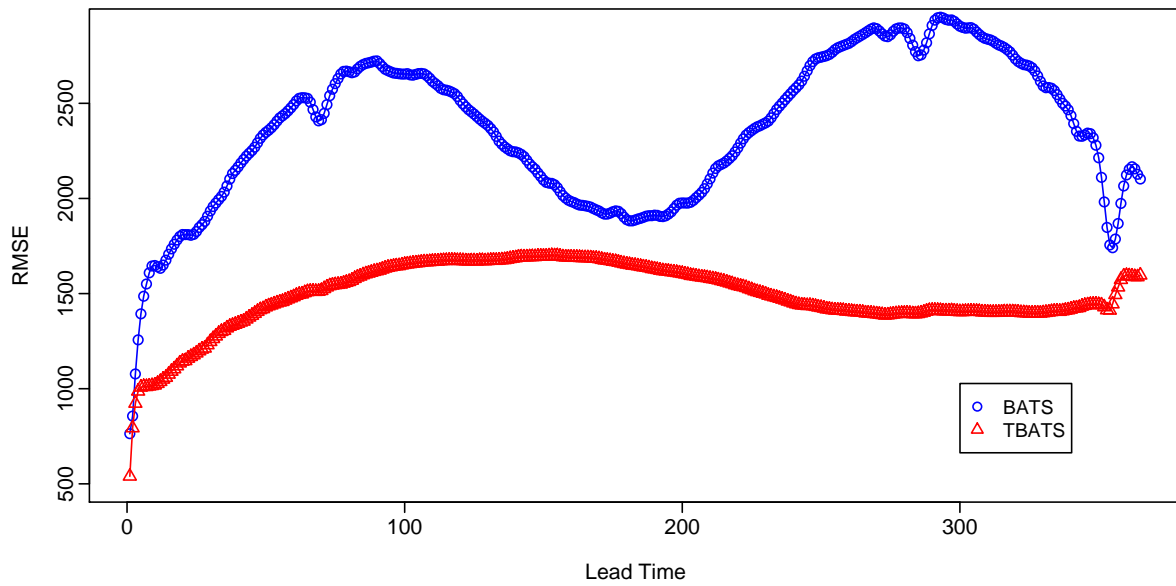


Figure 6: Out-of-sample results for the Turkey electricity demand data using $BATS(0, 0, 7, 354, 365)$ and $TBATS(3, 2, \{7, 3\}, \{354.37, 23\}, \{365.25, 3\})$

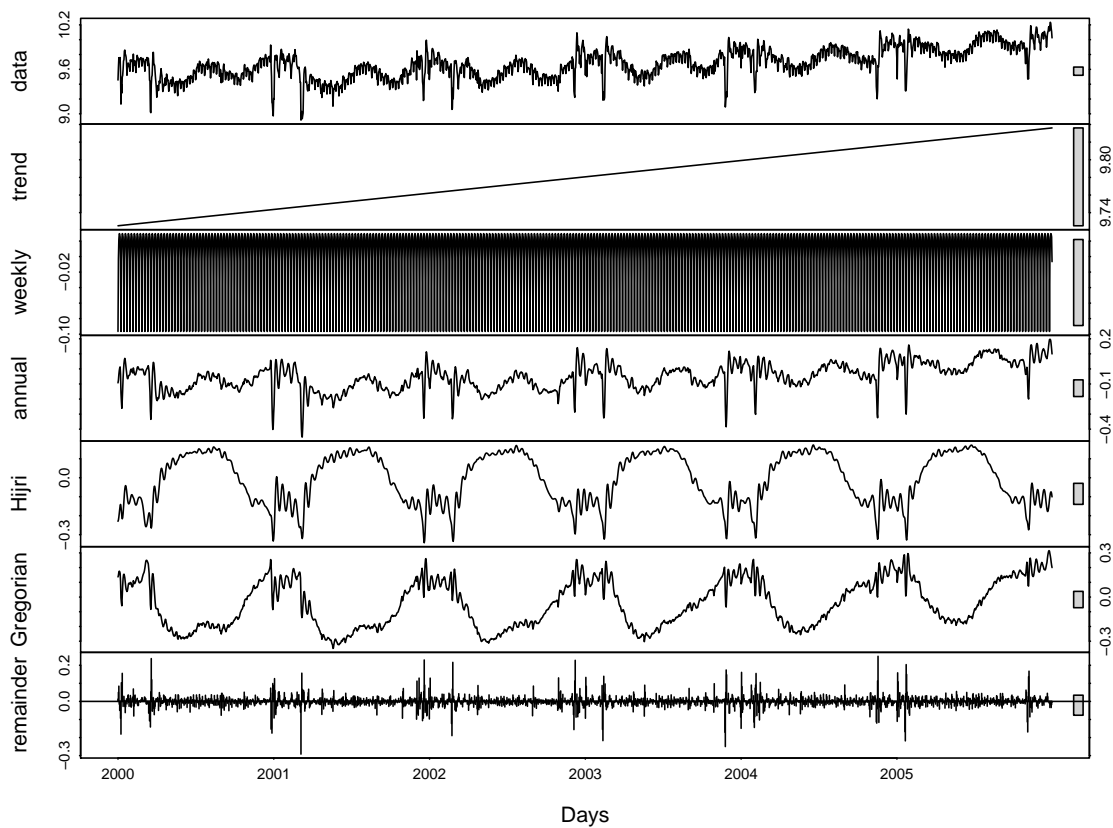


Figure 7: Trigonometric decomposition of the Turkey electricity demand data.

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. 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 in capturing the multiple seasonality with $k_1 = 3, k_2 = k_3 = 1$. The estimated holiday effect was then removed from the series and the remainder was decomposed using TBATS to achieve the decomposition shown in Figure 8.

| Year | Religious holidays | | National holidays |
|------|--------------------------------|-------------------------|--|
| | Seker holiday | Kurban holiday | |
| 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 |

Table 1: The dates of Turkish holidays between 1 January 2000 to 31 December 2006.

The fourth panel shows the overall annual component, which comprises the Hijri seasonal effect (fifth panel) and the Gregorian seasonal effect (sixth panel). It is seen that this now provides a much smoother seasonal decomposition and only one harmonic is required for each seasonal component (i.e., $k_2 = k_3 = 1$) in capturing seasonal calendar effects. This analysis demonstrates the capability of our trigonometric decomposition in extracting those seasonal components which are otherwise not apparent in graphical displays. Existing decomposition techniques are unable to handle such complex seasonalities.

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.

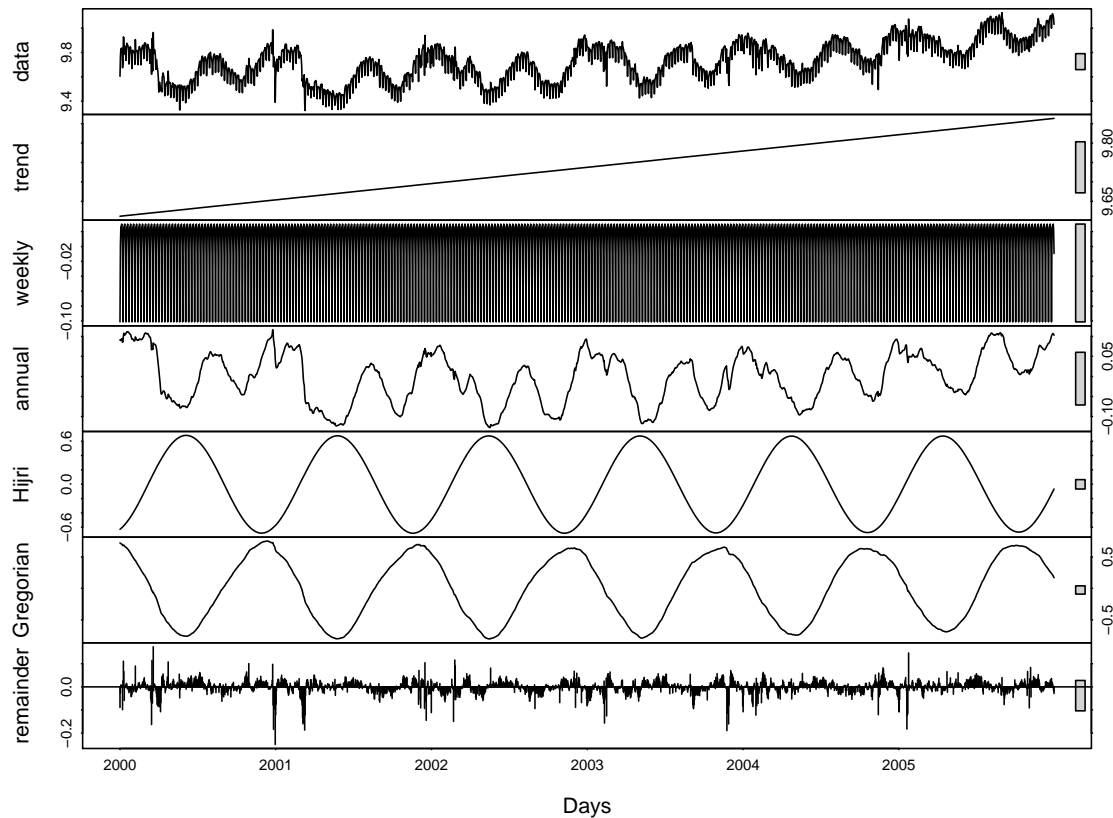


Figure 8: *Trigonometric decomposition of the regressed Turkey electricity demand data.*

8 Concluding remarks

A new state space modeling framework, based on the innovations approach, is developed for forecasting time series with complex seasonal patterns which cannot be forecast using any current forecasting approaches. The new approach not only offers an alternative to both linear and non-linear traditional exponential smoothing models, but also has many advantages and provides additional options, some of which are not seen in any of the existing forecasting approaches. The new approach is also capable of decomposing and forecasting time series with multiple seasonality, high frequency seasonality, non-integer seasonality and dual calendar effects.

The superiority of the new modeling framework in handling such seasonal patterns is illustrated in three empirical studies, where it is shown to greatly improve the out-of-sample forecasting accuracy. It is also shown that our trigonometric approach offers an elegant way of decomposing complex seasonal time series, which cannot be decomposed using any of the existing methods. Table 2 demonstrates that the trigonometric approach requires substantially fewer values to be estimated

than traditional seasonal exponential smoothing models in all three applications. Moreover, these models also overcome the weaknesses of the existing exponential smoothing models, such as the instability of non-linear models and correlated errors.

Table 2: *Number of estimated parameters for each model in each application.*

| Data | Model | No.parameters |
|-------------------------|--|----------------------|
| Gasoline data | BATS(0, 1, 52) | 59 |
| | TBATS(0, 1, {365.25/7, 7}) | 22 |
| Call center data | BATS(3, 0, 169, 845) | 1025 |
| | TBATS(3, 1, {169, 29}, {845, 15}) | 102 |
| Electricity demand data | BATS(0, 0, 7, 354, 365) | 734 |
| | TBATS(3, 2, {7, 3}, {354.37, 23}, {365.25, 3}) | 79 |

Acknowledgement

We thank Ralph Snyder and Peter Toscas for helpful comments that improved this paper. We are also grateful to Derek K. Baker and Merih Aydinalp Koksal for providing the Turkish electricity data.

Appendix A

Equation (2) can be written in the following form:

$$s_t = \sum_{j=1}^k (\cos \lambda_j t, \sin \lambda_j t)' \mathbf{x}_{j,t} \quad (12)$$

$$\mathbf{x}_{j,t} = \mathbf{x}_{j,t-1} + \boldsymbol{\kappa}_t, \quad (13)$$

where $\mathbf{x}_{j,t} = (\alpha_{j,t}, \beta_{j,t})'$ and $\boldsymbol{\kappa}_t = (k_1 \varepsilon_t, k_2 \varepsilon_t)'$. We re-parameterize using

$$\mathbf{x}_{j,t} = \mathbf{A}_{j,t} \mathbf{s}_{j,t},$$

where $\mathbf{s}_{j,t} = (s_{j,t}, s_{j,t}^*)'$ and $\mathbf{A}_{j,t} = \begin{bmatrix} \cos(\lambda_j t) & -\sin(\lambda_j t) \\ \sin(\lambda_j t) & \cos(\lambda_j t) \end{bmatrix}$. Then, substituting this into equation (13), we obtain

$$\mathbf{A}_{j,t} \mathbf{s}_{j,t} = \mathbf{A}_{j,t-1} \mathbf{s}_{j,t-1} + \boldsymbol{\kappa}_t$$

$$\mathbf{s}_{j,t} = \mathbf{A}_{j,t}^{-1} \mathbf{A}_{j,t-1} \mathbf{s}_{j,t-1} + \mathbf{A}_{j,t}^{-1} \boldsymbol{\kappa}_t,$$

where $\mathbf{A}_{j,t}^{-1} = \begin{bmatrix} \cos(\lambda_j t) & \sin(\lambda_j t) \\ -\sin(\lambda_j t) & \cos(\lambda_j t) \end{bmatrix}$. Using well-known trigonometric identities, $\mathbf{A}_{j,t}^{-1} \mathbf{A}_{j,t-1} = \begin{bmatrix} \cos(\lambda_j) & \sin(\lambda_j) \\ -\sin(\lambda_j) & \cos(\lambda_j) \end{bmatrix}$, and so we obtain the required result (3).

References

- Akram, M., Hyndman, R. J. & Ord, J. K. (2009), 'Exponential smoothing and non-negative data', *Australian & New Zealand Journal of Statistics* **51**(4), 415–432.
- Archibald, B. C. & Koehler, A. B. (2003), 'Normalization of seasonal factors in Winters' methods', *International Journal of Forecasting* **19**, 143–148.
- Billah, B., Hyndman, R. J. & Koehler, A. B. (2005), 'Empirical information criteria for time series forecasting model selection', *Journal of Statistical Computation and Simulation* **75**, 831–840.
- Box, G. E. P & Cox, D. R. (1964), 'An analysis of transformations', *Journal of the Royal Statistical Society, Series B* **26**(2), 211–252.
- Burnham, K. P. & Anderson, D. R. (2002), *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn, Springer-Verlag.
- Chatfield, C. (1978), 'The Holt-Winters forecasting procedures', *Applied Statistics* **27**, 264–279.
- Durbin, J. & Koopman, S. (2001), *Time series analysis by state space methods*, Oxford University Press.
- Gardner, Jr, E. S. (1985), 'Exponential smoothing: The state of the art', *Journal of Forecasting* **4**, 1–28.
- Gilchrist, W. (1976), *Statistical Forecasting*, Wiley: Chichester.
- Gould, P. G., Koehler, A. B., Ord, J. K., Snyder, R. D., Hyndman, R. J. & Vahid-Araghi, F. (2008), 'Forecasting time series with multiple seasonal patterns', *European Journal of Operational Research* **191**(1), 207–222.
- Hannon, E. J., Terrel, R. & Tuckwell, N. (1970), 'The seasonal adjustment of economic time series', *International Economic Review* **11**, 24–52.
- Harvey, A. (1989), *Forecasting structural time series models and the Kalman filter*, Cambridge University Press.
- Harvey, A. & Koopman, S. J. (1993), 'Forecasting hourly electricity demand using time-varying splines.', *Journal of the American Statistical Association* **88**, 1228–1236.
- Harvey, A., Koopman, S. J. & Riani, M. (1997), 'The modeling and seasonal adjustment of weekly observations', *Journal of Business and Economic Statistics* **15**, 354–368.

- Hyndman, R. J., Akram, M. & Archibald, B. C. (2007), 'The admissible parameter space for exponential smoothing models', *Annals of the Institute of Statistical Mathematics* **60**, 407–426.
- Hyndman, R. J., Koehler, A. B., Ord, J. K. & Snyder, R. D. (2005), 'Prediction intervals for exponential smoothing using two new classes of state space models', *Journal of Forecasting* **24**, 17–37.
- Hyndman, R. J., Koehler, A. B., Ord, J. K. & Snyder, R. D. (2008), *Forecasting with exponential smoothing: the state space approach*, Springer-Verlag, Berlin.
URL: www.exponentialsMOOTHING.net
- Hyndman, R. J., Koehler, A. B., Snyder, R. D. & Grose, S. (2002), 'A state space framework for automatic forecasting using exponential smoothing methods', *International Journal of Forecasting* **18**(3), 439–454.
- Hyndman, R. & Khandakar, Y. (2008), 'Automatic time series forecasting: The forecast package for R', *Journal of Statistical Software* **26**(3).
- Koopman, S. J. & Lee, K. M. (2005), 'Measuring asymmetric stochastic cycle components in U.S. macroeconomic time series', (05-081/4).
URL: <http://ideas.repec.org/p/dgr/uvatin/20050081.html>
- Lin, J. & Liu, T. (2002), 'Modeling lunar calendar holiday effects in Taiwan', *Taiwan Economic Forecast and Policy* **33**(1), 1–37.
- Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E. & Winkler, R. (1982), 'The accuracy of extrapolation (time series) methods: results of a forecasting competition', *International Journal of Forecasting* **1**, 111–153.
- Makridakis, S. & Hibon, M. (2000), 'The M3-competition: Results, conclusions and implications', *International Journal of Forecasting* **16**, 451–476.
- Ord, J. K., Koehler, A. B. & Snyder, R. D. (1997), 'Estimation and prediction for a class of dynamic nonlinear statistical models', *Journal of the American Statistical Association* **92**, 1621–1629.
- Ord, J. K., Snyder, R. D., Koehler, A. B., Hyndman, R. J. & Leeds, M. (2005), Time series forecasting: the case for the single source of error state space approach, Working paper 7/05, Department of Econometrics & Business Statistics, Monash University.
- Pedregal, D. J. & Young, P. C. (2006), 'Modulated cycles, an approach to modelling periodic components from rapidly sampled data', *International Journal of Forecasting* **22**, 181–194.

- Pole, A., West, M. & Harrison, J. (1994), *Applied Bayesian forecasting and time series analysis*, Chapman & Hall/CRC.
- Proietti, T. (2000), 'Comparing seasonal components for structural time series models', *International Journal of Forecasting* **247-260**, 16.
- Reid, D. J. (1975), *A review of short-term projection techniques. Practical aspects of forecasting*, Operational Research Society: London.
- Riazuddin, R. & Khan, M. (2005), 'Detection and forecasting of Islamic calendar effects in time series data', *State Bank of Pakistan Research Bulletin* **1(1)**, 25–34.
- Snyder, R., Koehler, A. & Ord, J. (2002), 'Forecasting for inventory control with exponential smoothing', *International Journal of Forecasting* **18(1)**, 5–18.
- 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.
- Taylor, J. (2009), 'Triple seasonal methods for short-term electricity demand forecasting', *European Journal of Operational Research* .
- Taylor, J. W. (2003a), 'Exponential smoothing with a damped multiplicative trend', *International Journal of Forecasting* **19**, 715–725.
- Taylor, J. W. (2003b), 'Short-term electricity demand forecasting using double seasonal exponential smoothing', *Journal of the Operational Research Society* **54**, 799–805.
- Taylor, J. W. & Snyder, R. D. (2009), 'Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing', (9/09).
URL: <http://ideas.repec.org/p/msh/ebswps/2009-9.html>
- Weinberg, J., Brown, L. & 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.
- West, M. & Harrison, J. (1997), *Bayesian forecasting and dynamic models*, 2nd edn, Springer-Verlag, New York.