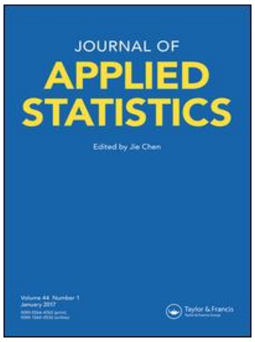


Journal of Applied Statistics

ISSN: 0266-4763 (Print) 1360-0532 (Online) Journal homepage: https://www.tandfonline.com/loi/cjas20

Dynamic structural models with covariates for short-term forecasting of time series with complex seasonal patterns

António Casimiro Puindi & Maria Eduarda Silva

JOURNAL OF APPLIED STATISTICS
https://doi.org/10.1080/02664763.2020.1748178

Dynamic structural models with covariates for short-term forecasting of time series with complex seasonal patterns

António Casimiro Puindi[a] and Maria Eduarda Silva[b]

[a]CIDMA & Faculdade de Ciências, Universidade do Porto, Porto, Portugal; [b]CIDMA & Faculdade de Economia, Universidade do Porto, Porto, Portugal

ABSTRACT

This work presents a framework of dynamic structural models with covariates for short-term forecasting of time series with complex seasonal patterns. The framework is based on the multiple sources of randomness formulation. A noise model is formulated to allow the incorporation of randomness into the seasonal component and to propagate this same randomness in the coefficients of the variant trigonometric terms over time. A unique, recursive and systematic computational procedure based on the maximum likelihood estimation under the hypothesis of Gaussian errors is introduced. The referred procedure combines the Kalman filter with recursive adjustment of the covariance matrices and the selection method of harmonics number in the trigonometric terms. A key feature of this method is that it allows estimating not only the states of the system but also allows obtaining the standard errors of the estimated parameters and the prediction intervals. In addition, this work also presents a non-parametric bootstrap approach to improve the forecasting method based on Kalman filter recursions. The proposed framework is empirically explored with two real time series.

ARTICLE HISTORY

Received 6 November 2018
Accepted 22 March 2020

KEYWORDS

Bootstrap; Kalman filter; prediction intervals; structural time series models; seasonal time series.

1. Introduction

In modern management operations, forecasting plays a key role. Researches on forecast generally consider three main categories: long-term, medium-term and short-term forecasts [15,17]. An efficient prediction, for example, can allow a company commit its resources with greater security to make long-term profits (long-term forecast), since it helps to identify future demand patterns and facilitates the new products development. The short-term forecast, for example, is important for studying the balance of national power grid, which requires a balance between the electricity produced and consumed at any moment in the day [13].

Forecasting an event depends on how well we understand the factors that contribute to its occurrence and how much unexplained variability is involved, as well as the factors determining actual outcomes, types of data patterns, and so on. As for the method (quantitative or qualitative), the choice depends on what data are available and the predictability

CONTACT António Casimiro Puindi acpuindi@gmail.com

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

of the quantity to be forecast. Since the numerical information is available about the past of the phenomenon, it is reasonable to assume that some aspects of the past patterns will continue into the future, the quantitative forecasting can be applied.

There is a wide range of quantitative forecasting models for specific purposes, such as the models designed in the state-space form, which has deserved much attention by researchers, among them [2,14,21,22,24,25,27,30,31,37]. This attention is justified, on the one hand, by the fact that state-space structures are optimal and flexible for class of exponential smoothing models [25]; on the other hand, such models are very flexible to incorporate covariates effects [13,16,37,44], as well as to accommodate resampling methods such as bootstrap methodology [12,28,37,38].

Regardless of how these methods model a time series, implicitly some of these models assume that past observations of a time series contain all information required for forecasting its future; that is, they forecast the future of a time series using only its past observations [44]. For this class of models without covariates (with reference to short-term forecast), the most common models of state-space structure include those underlying the well-known additive and multiplicative methods of Holt-Winters [18,22,25,33,39–42].

It is obvious that the history of a time series certainly contains information about its future. However, other information beyond what is available in a time series history can also shed light on the series movements over time, therefore, lead to more accurate forecasting of its future if incorporated [44]. Such other information can be provided by so-called external influence variables (or covariates). Proposals of this type of models can be seen in [8] with SARIMA (Seasonal Auto-Regressive Integrated Moving Average) and [2,13,16,27,44].

As Wang [44] also found, the projection of forecast models with covariates has two fundamental advantages: (i) to take into account the history of the time series of interest and the information hidden in quantifiable covariates, may lead to a more accurate forecast of the time series of interest; (ii) to forecast, one only needs to know how the time series of interest move over time. But, in order to make the series to move in the desired directions, for example, to estimate (in order to reduce or increase) the maximum production of electric power due to demand, we need to understand the reasons behind this demand. Such knowledge can often be learned from its relation to other variables. In this context, the essential methodological challenge is the ability to relate the history of the time series to exogenous factors or covariates [16]. Still within the advantages (or importance) of using covariates for short-term forecasting, we highlight the research works of [7,13,15,16,23,32,36].

The focus of this work is primarily to explore the use of covariates on short-term forecasting of time series with complex seasonal patterns. The framework proposed is inspired by De Livera *et al.* [1], who introduced two structures, BATS as an acronym for the main features of the model: Box-Cox transformation, ARMA errors, Trend, and Seasonal components and TBATS with the initial T connoting Trigonometric. The authors estimate the parameters of these models by exponential smoothing. Among the advantages presented by TBATS we highlight the following: (i) the ability to accommodate data with non-integer seasonal periods, high-frequency data and dual calendar effects data; (ii) the Box-Cox transformation that allows to deal with the non-linearity of the data; (iii) the ARMA process on residuals to solve the autocorrelation problem. However, we notice that TBATS models are related to ETS models, tbats() is fully automatic, is unlikely to over include covariates.

The fact that TBATS does not allow for covariates which may be important for short-term forecasting may be pointed out as a disadvantage of these models.

Our proposed framework differs from TBATS approach in four key aspects, namely: (i) the framework is a state-space formulation with multiple sources of randomness (MSR); (ii) the models do not incorporate smoothing parameters; (iii) the models deal with covariates, but they can work without the covariates; (iv) in the estimation procedure, we used the Kalman filter (with recursive adjustment of the covariance matrices) to obtain one-step-ahead forecasting errors and associated variances needed for evaluating fitting criteria for given trial values of the parameters. For the proposed framework, the Kalman filter is appropriate in particular. The choice of MSR formulation is justified by the fact that it is more flexible way to treat the covariates in the proposed framework and using Kalman filter as an algorithm for statistical treatment. In addition, once our proposed framework does not incorporate the smoothing parameters, this fact limits the possibility of having many parameters to estimate. From the forecast point of view, there is a growing interest among researchers in combining the resampling methods with the state-space models [4,11,12,19,28,37,38,46]. With these ideas, we formulated a bootstrap procedure based on the residuals trigonometric structural model with covariates.

Our subsequent study has three objectives: (i) to explore the use of covariates in short-term forecasting of time series with complex seasonal patterns; (ii) to construct a Kalman filter with recursive adjustment of covariance matrices and to project a computational procedure for estimation of the proposed models; (iii) to construct a bootstrap procedure for forecasting. The rest of the work is organized as follows: In Section 2, we provide a brief review of TBATS models. We then introduce in this Section the new proposed structural model with covariates (TSCov–*Trigonometric Structural models with Covariates*) including its state-space representation and the Kalman filter with recursively covariances computed. Section 3 contains the empirical fitting of proposed model. Specifically addresses the maximum likelihood estimation, the new computational procedure and the model selection criterion. Section 4 provides the forecasting strategies. Two procedures are presented: the first is related to the direct use of Kalman filter recursions, the second is based on bootstrap method. The proposed model is then applied in Section 5. Conclusion and future direction are drawn in Section 6.

2. Models for time series with complex seasonal patterns

2.1. TBATS *models*

TBATS models which is BATS model plus Trigonometric Seasonal models was proposed by De Livera *et al.* [1] in order to overcome some weaknesses found on the traditional seasonal exponential smoothing models. The BATS model is the most obvious generalization of the traditional seasonal innovations models to allow for multiple seasonal periods [1], formulated as follows:

$$y_t^{(\omega)} = \begin{cases} \frac{y_t^{\omega}-1}{\omega} & \text{se } \omega \neq 0 \\ \log y_t & \text{se } \omega = 0 \end{cases} \tag{1a}$$

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

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

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

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \boldsymbol{\gamma}_i d_t \tag{1e}$$

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

Furthermore, De Livera *et al.* [1] found that BATS model cannot accommodate non-integer seasonality. In the quest for a more flexible parsimonious approach, the authors introduced the following trigonometric representation of seasonal components based on Fourier series:

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)} \tag{2a}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \boldsymbol{\gamma}_1^{(i)} d_t \tag{2b}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \boldsymbol{\gamma}_2^{(i)} d_t \tag{2c}$$

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 $\alpha, \beta, \boldsymbol{\gamma}^{(i)}$, $\boldsymbol{\gamma}_1^{(i)}$ and $\boldsymbol{\gamma}_2^{(i)}$ for $i = 1, \ldots, T$, and $\lambda_j^{(i)} = 2\pi j/m_i$. The stochastic level of the ith seasonal component is described 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 is described by $s_{j,t}^{*(i)}$. The number of harmonics required for the ith seasonal component is denoted by k_i. So the new class designated by TBATS, is obtained by replacing the seasonal component $s_t^{(i)}$ in Equation (1) by the trigonometric seasonal formulation, and the measurement equation by

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

From a practical point of view, as stated above, tbats() is fully automatic and is unlikely to over include covariates. In contrast, we introduce a new state-space model with multiple sources of randomness and covariates, as shown in the next section.

2.2. *The* TSCov *model*

Let $\{Y_t\} = \{y_1, \ldots, y_n\}$ be the observed time series and $z_{\kappa,t}$ ($\kappa = 1, \ldots, r$) the set of regressor variables. We notice that our study is limited for cases of additive trends and seasonality.

In addition, to deal with non-linearity problems, it is assumed that TSCov model is applicable to a Box-Cox transformation. Using the same notation as in (1) and (2), the regressor variables may be incorporated into the TSCov model as follows:

$$y_t^{(\omega)} = \begin{cases} \frac{y_t^{(\omega)}-1}{\omega} & \text{se } \omega \neq 0 \\ \ln y_t & \text{se } \omega = 0 \end{cases} \tag{4a}$$

$$y_t^{(\omega)} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + \sum_{\kappa=1}^{r} \beta_\kappa^* z_{\kappa,t} + \varepsilon_t; \quad \varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2) \tag{4b}$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \xi_t; \quad \xi_t \sim iid\mathcal{N}(0, \sigma_\xi^2) \tag{4c}$$

$$b_t = (1-\phi)b + \phi b_{t-1} + \zeta_t; \quad \zeta_t \sim iid\mathcal{N}(0, \sigma_\zeta^2) \tag{4d}$$

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

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

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

We assume that $e_{j,t}^{(i)} = e_{j,t}^{*(i)} \sim iid\mathcal{N}(0, \sigma_e^{2(i)})$ and $\varepsilon_t, \xi_t, \zeta_t, e_{j,t}^{(i)}$ are independent processes. As well as in (1) and (2), $m_1, \ldots, m_T$ represent the seasonal periods and T denotes the seasonal patterns; $\lambda_j^{(i)} = 2\pi j/m_i$ ($j = 1, 2, \ldots, k_i$ and $i = 1, \ldots, T$); ℓ_t and b_t are the local level and the short-term trend in period t; b is the long-term trend. $s_t^{(i)}$ and $s_{j,t}^{(i)}$ denotes also the seasonal component in period t and the stochastic level of the i*th* seasonal component and $s_{j,t}^{*(i)}$ is the stochastic growth in the level of the i*th* seasonal component that is needed to describe the change in the seasonal component over time. k_i is also the number of harmonics required for the i*th* seasonal component, whose 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. The transition parameter is bounded by $0 < \phi \leq 1$ to prevent a negative coefficient being applied to b_t. In case $\phi = 0$, it would indicate the absence of the trend in time series.

2.2.1. State-space representation

The linear state-space models can be extended to incorporate the fixed-effects regression. Such regression effects can be included in one or two different ways: (i) the first approach is including exogenous or predetermined variables in the signal equation; (ii) the second approach is including the exogenous or predetermined variables in the state equation. The above model with fixed-effects regression may be written in a state-space formulation. We adopt $\boldsymbol{z}_t$ to represent the vector containing any control inputs (meteorological variables, for example, air temperature, wind-speed, relative humidity or predetermined variables, it may also contain the indicator variables) and $\boldsymbol{\Gamma}$ to represent the control input matrix (coefficient matrix formed by regression coefficients $\boldsymbol{\beta}_\kappa^*$ which applies the effect of each control input parameter in $\boldsymbol{z}_t$ on the observation vector, for example, applies the effect of temperature on the electricity consumption. The vector $\boldsymbol{\Gamma}$ contains unknown parameters but these do not affect the stochastic properties of the model, only enter the model in a deterministic way, that is, the parameters appearing in $\boldsymbol{\Gamma}$ only affect the expected value of the observations $\boldsymbol{y}_t$ in a deterministic way. This distinction can become blurred, for example, if $\boldsymbol{\Gamma}$ is a function of a lagged value of $\boldsymbol{y}_t$. If $\boldsymbol{\Gamma}$ is a linear function of unknown parameters, these parameters can be treated as state variables [21]. So, we wrote the Gaussian linear model in state-space representation as

$$\boldsymbol{y}_t^{(\omega)} = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{\Gamma}\boldsymbol{z}_t + \boldsymbol{v}_t \quad t = 1, 2, \ldots, n \tag{5a}$$

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{\eta}_t \quad t = 1, 2, \ldots, n \tag{5b}$$

where $\boldsymbol{A}_t$ is a $q \times p$ *measurement* or *observation matrix*; (5a) is called the *observation equation*. The observed data vector, $\boldsymbol{y}_t$, is q-dimensional, which can be larger than or smaller than p, the state dimension. (5b) is called the *state equation*, $\boldsymbol{\Phi}$ is a $p \times p$ transition matrix. We suppose we have an $r \times 1$ vector of inputs $\boldsymbol{z}_t$ and $\boldsymbol{\Gamma}$ is $q \times r$ matrix. $\boldsymbol{v}_t$ and $\boldsymbol{\eta}_t$ are white noises, being that:

$$E(\boldsymbol{v}_t\boldsymbol{v}_t') = \boldsymbol{R} \tag{6a}$$

$$E(\boldsymbol{\eta}_t\boldsymbol{\eta}_t') = \boldsymbol{Q} \tag{6b}$$

Furthermore, $\boldsymbol{v}_t$ and $\boldsymbol{\eta}_t$ are assumed to be uncorrelated:

$$E(\boldsymbol{v}_t\boldsymbol{\eta}_t) = 0 \tag{7}$$

$\boldsymbol{x}_t$ represents the unobserved state vector. The observation and transition models are represented by $\boldsymbol{A}_t$ and $\boldsymbol{\Phi}$ matrices, respectively. Given (7), Equation (5b) is typically used to describe a finite time series of observations $y_1, y_2, \ldots, y_n$ and for which assumptions about the initial value of the state vector are necessary [20]. It is assumed that $\boldsymbol{x}_t$ is uncorrelated with any realization of $\boldsymbol{v}_t$ or $\boldsymbol{\eta}_t$:

$$E(\boldsymbol{v}_t\boldsymbol{x}_t') = 0 \quad \text{for } t = 1, 2, \ldots, n \tag{8a}$$

$$E(\boldsymbol{\eta}_t\boldsymbol{x}_t') = 0 \quad \text{for } t = 1, 2, \ldots, n \tag{8b}$$

The statement that $\boldsymbol{z}_t$ is exogenous or predetermined means that $\boldsymbol{z}_t$ does not provide information about the $\boldsymbol{x}_{t+h}$ or $\boldsymbol{\eta}_{t+h}$ for $h = 1, 2, \ldots$ beyond the information contained in $y_{t-1}, y_{t-2}, \ldots, y_1$. So, the system matrices (5) for TSCov model can be obtained by defining: (i) the state vector, $\boldsymbol{x}_t = \{\ell_t, b_t, s_{1,t}^{(i)}, s_{2,t}^{(i)}, \ldots, s_{k_i,t}^{(i)}, s_{1,t}^{*(i)}, s_{2,t}^{*(i)}, \ldots, s_{k_i,t}^{*(i)}\}$, (ii) the replica vector of 0 and 1 (which depends on the seasonal component) defined as $\boldsymbol{a} = (\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(T)})$ with $\boldsymbol{a}^{(i)} = (\mathbf{1}_{k_i}, \mathbf{0}_{k_i})$ and $\tau = 2\sum_{i=1}^{T} k_i$. (iii) we also need to define the block matrix, $\boldsymbol{B}$, resulting from the direct sum, $\bigoplus$, of the matrices $\boldsymbol{B}_i$, that is, $\boldsymbol{B} = \bigoplus_{i=1}^{T} \boldsymbol{B}_i$,

$$\boldsymbol{B}_i = \begin{bmatrix} \boldsymbol{C}^{(i)} & \boldsymbol{S}^{(i)} \\ -\boldsymbol{S}^{(i)} & \boldsymbol{C}^{(i)} \end{bmatrix}$$

where, $\boldsymbol{C}^{(i)}$ and $\boldsymbol{S}^{(i)}$ are diagonal matrices of size $k_i \times k_i$ with elements $\cos(\lambda_j^{(i)})$ and $\sin(\lambda_j^{(i)})$, for $j = 1, 2, \ldots, k_i$.

$$\boldsymbol{B} = \bigoplus_{i=1}^{T} \boldsymbol{B}_i = \begin{bmatrix} \boldsymbol{B}_1 & \cdots & & \boldsymbol{0} \\ & \boldsymbol{B}_2 & & \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \cdots & & \boldsymbol{B}_T \end{bmatrix} = \begin{bmatrix} \cos\lambda_j^{(1)} & \sin\lambda_j^{(1)} & \cdots & \boldsymbol{0} \\ -\sin\lambda_j^{(1)} & \cos\lambda_j^{(1)} & & \\ \vdots & \vdots & \ddots & \vdots \\ & & \cos\lambda_j^{(T)} & \sin\lambda_j^{(T)} \\ \boldsymbol{0} & \cdots & -\sin\lambda_j^{(T)} & \cos\lambda_j^{(T)} \end{bmatrix}$$

The covariance matrices for measurement and the state are given by

$$\boldsymbol{R} = \sigma_\varepsilon^2 \text{ and } \boldsymbol{Q}^{(i)} = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \tilde{\boldsymbol{q}}^{(i)} \end{bmatrix}$$

where σ_ε^2 is the noise variance in the measurement equation, σ_ξ^2 and σ_ζ^2, the model variances corresponding to the level and trend components, $\tilde{\boldsymbol{q}}^{(i)}$ given in (9) is the model variance corresponding the seasonal component.

$$\tilde{\boldsymbol{q}}^{(i)} = \{\tilde{\boldsymbol{\sigma}}_e^{2(1)}, \ldots, \tilde{\boldsymbol{\sigma}}_e^{2(T)}\}, \quad \text{where} \quad \tilde{\boldsymbol{\sigma}}_e^{2(i)} = \{\boldsymbol{\sigma}_e^{2(i)}, \boldsymbol{\sigma}_{e*}^{2(i)}\} \quad \text{and} \tag{9a}$$

$$\boldsymbol{\sigma}_e^{2(i)} = \sigma_e^{2(i)} \mathbf{1}_{k_i} \tag{9b}$$

$$\boldsymbol{\sigma}_{e*}^{2(i)} = \sigma_{e*}^{2(i)} \mathbf{1}_{k_i} \tag{9c}$$

By modeling in this way, we allow the seasonal component noise to have a dual function: (i) to be the source of randomness for seasonal component; (ii) propagate the randomness effect on the stochastically variant trigonometric coefficients terms over time. This way of modeling the seasonal component variances it is similar to the methodology used by De Livera *et al.* [1] to model the smoothing parameter of seasonal component.

The homoscedastic exponential smoothing model (BATS) with covariates can be obtained by letting $\boldsymbol{x}_t = \{l_t, b_t, s_t^{(i)}, s_{t-1}^{(i)}, \ldots, s_{t-(m_i-1)}^{(i)}\}$, $\boldsymbol{a}^{(i)} = (\mathbf{0}_{m_i-1}, 1)$, $\boldsymbol{B} = \bigoplus_{i=1}^{T} \tilde{\boldsymbol{D}}_i$, and by replacing $2k_i$ with m_i in the matrices presented above for the TSCov model, that is, $\tau = \sum_{i=1}^{T} m_i$,

$$\tilde{\boldsymbol{D}}_i = \begin{bmatrix} \mathbf{0}_{m_i-1} & \mathbf{1} \\ \boldsymbol{I}_{m_i-1} & \mathbf{0}'_{m_i-1} \end{bmatrix}$$

where $\tilde{\boldsymbol{q}}^{(i)} = (\tilde{\boldsymbol{\sigma}}_w^{2(i)}, \mathbf{0}_{m_i-1})$ and $\tilde{\boldsymbol{\sigma}}_w^{2(i)} = \sigma_w^{2(i)} \mathbf{1}_{m_i}$.

2.2.2. Kalman filter with recursively computed covariances

Given (5), according to the principles featuring a state-space model, the Kalman predictor used when $t > s$ and the Kalman filter applied when $t = s$ are given by (10). For more details, see [37].

$$\hat{\boldsymbol{x}}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{x}_{t-1|t-1} \tag{10a}$$

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{P}_{t-1|t-1}\boldsymbol{\Phi}' + \boldsymbol{Q} \tag{10b}$$

$$\boldsymbol{x}_{t|t} = \hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{K}_t\boldsymbol{\epsilon}_t \tag{10c}$$

$$\boldsymbol{P}_{t|t} = \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{A}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{A}_t\boldsymbol{P}_{t|t-1} \tag{10d}$$

where

$$\boldsymbol{K}_t = \boldsymbol{P}_{t|t-1}\boldsymbol{A}_t'\left[\boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t' + \boldsymbol{R}_t\right]^{-1} \tag{11a}$$

$$\boldsymbol{\epsilon}_t = \boldsymbol{y}_t^{(\omega)} - \boldsymbol{A}_t\hat{\boldsymbol{x}}_{t|t-1} - \boldsymbol{\Gamma}\boldsymbol{z}_t \tag{11b}$$

$$\boldsymbol{\Sigma}_t = \mathrm{Var}(\boldsymbol{\epsilon}_t) = \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t' + \boldsymbol{R} \tag{11c}$$

When performing a Kalman filter, the correct setting of the covariance matrices is critical, since the Kalman filter performance is highly affected by the system covariance matrices. Its inadequate choice can significantly degrade the performance of Kalman filter and even make the filter divergent [29]. It is quite common to use ad-hoc procedures to determine the system of covariance matrices, such as conventional filters [8,14,37], among others, in which $\boldsymbol{Q}$ and $\boldsymbol{R}$ are constant matrices and adjusted manually by trial and error. To address this challenge, a Kalman filter with recursively computed covariance is constructed with reference to [3] and [43] approaches. Now we denote the covariance matrices of measurement and state equations variants over time as R_t and Q_t. The procedure that we apply is based on the innovations (a priori and a posteriori) of the model; these will influence the adjustment of covariance matrices recursively to improve the accuracy of state estimate. Given (10c), first, set up the a posteriori innovations as

$$\boldsymbol{\varsigma}_t = \boldsymbol{y}_t^{(\omega)} - \boldsymbol{A}_t\boldsymbol{x}_{t|t} - \boldsymbol{\Gamma}\boldsymbol{z}_t, \tag{12}$$

then, define the covariance estimates $\boldsymbol{R}_t$ and $\boldsymbol{Q}_t$ that participate in the recursive process by synchronizing them with a priori ($\boldsymbol{\epsilon}_t$) and a posteriori ($\boldsymbol{\varsigma}_t$) innovations.

For measurement covariance estimation, $\boldsymbol{R}_t$. Given (11c), the measurement covariance can be given by

$$\boldsymbol{R}_t = \boldsymbol{\Sigma}_t - \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t' \tag{13}$$

Theoretically $\boldsymbol{\Sigma}_t$ should be set positive. However, the Equation (13) does not guarantee the positivity of the estimated matrix, $\boldsymbol{R}_t$, because it results from the subtraction of two positive definite matrices. According to (6a), it is ensured that $\boldsymbol{R}_t$ is a positive definite matrix by combining covariance with a posteriori innovations, $\boldsymbol{\varsigma}_t$, as in (14).

$$\begin{aligned}\boldsymbol{\Sigma}_t^* &= E[\boldsymbol{\varsigma}_t\boldsymbol{\varsigma}_t'] = E[\boldsymbol{\nu}_t\boldsymbol{\nu}_t'] - \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t' \\ \boldsymbol{R}_t &= E[\boldsymbol{\varsigma}_t\boldsymbol{\varsigma}_t'] + \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t' \end{aligned} \tag{14}$$

where, the operation on $\boldsymbol{\varsigma}_t\boldsymbol{\varsigma}_t'$ is usually approximated by averaging $\boldsymbol{\varsigma}_t\boldsymbol{\varsigma}_t'$ over time t [29]. Instead, we apply the procedure used by Akhlaghi *et al.* [3], which consists of applying a forgetting factor, $0 < \delta \leq 1$, to estimate the covariance adaptively, as defined in (15).

$$\boldsymbol{R}_t = \delta \boldsymbol{R}_{t-1} + (1-\delta)(\boldsymbol{\varsigma}_t \boldsymbol{\varsigma}_t' + \boldsymbol{A}_t \boldsymbol{P}_{t|t-1} \boldsymbol{A}_t') \tag{15}$$

For state covariance estimation, $\boldsymbol{Q}_t$. Given (5b), the state covariance estimate, $\boldsymbol{Q}_t$ can be obtained by doing $\boldsymbol{\eta}_t = \boldsymbol{x}_t - \boldsymbol{\Phi x}_{t-1}$. Since $\boldsymbol{x}_{t|t}$ is the estimator of $\boldsymbol{x}_{t|t-1}$, given (10c), the estimated state error can be given by $\hat{\boldsymbol{\eta}}_t = \boldsymbol{x}_{t|t} - \boldsymbol{\Phi} \hat{\boldsymbol{x}}_{t|t-1} = \boldsymbol{K}_t \boldsymbol{\epsilon}_t$. Its covariance is

$$E(\hat{\boldsymbol{\eta}}_t \hat{\boldsymbol{\eta}}_t') = E[\boldsymbol{K}_t(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t')\boldsymbol{K}_t'] = \boldsymbol{K}_t E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t')\boldsymbol{K}_t' \tag{16}$$

From (16), given (11c), the covariance estimate of state is given by $\hat{\boldsymbol{Q}}_t = \boldsymbol{K}_t \boldsymbol{\Sigma}_t \boldsymbol{K}_t'$. Using the same procedure, the estimate of $\boldsymbol{Q}_t$ over time is given by (17)

$$\boldsymbol{Q}_t = \delta \boldsymbol{Q}_{t-1} + (1-\delta)(\boldsymbol{K}_t \boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' \boldsymbol{K}_t') \tag{17}$$

3. Empirical fitting of TSCov model

3.1. Maximum likelihood estimation

The approach used was obtaining the conditional distribution $p(\boldsymbol{x}_t | y_t^{(\omega)})$ of the state $\boldsymbol{x}_t$ for a set of observations $\boldsymbol{Y}_{t-1}$, [26,47,48]. We calculate the conditional densities and the classical maximum likelihood theory based on the situation by which the n transformed observations, $y_1^{(\omega)}, y_2^{(\omega)}, \ldots, y_n^{(\omega)}$, are independent and identically distributed with $\boldsymbol{\Omega}$ a vector of unknown parameters, allowed us defining the joint density function as

$$\mathcal{L}(\boldsymbol{y}_t^{(\omega)}, \boldsymbol{\Omega}) = \prod_{t=1}^{n} p(\boldsymbol{y}_t^{(\omega)} | \boldsymbol{Y}_{t-1}) \tag{18}$$

where $p(\boldsymbol{y}_t^{(\omega)} | \boldsymbol{Y}_{t-1})$ describes the distribution of $\boldsymbol{y}_t^{(\omega)}$ conditioned on the established information in the period $t-1$, that is, $\boldsymbol{Y}_{t-1} = \{y_{t-1}, y_{t-2}, \ldots, y_1\}$ [21]. Since $\boldsymbol{x}_0 \sim iid\mathcal{N}(\mu_0, \boldsymbol{\Sigma}_0)$, the distribution of $\boldsymbol{y}_t^{(\omega)}$ conditioned to $\boldsymbol{Y}_{t-1}$ is in itself normal. The mean and covariance of this distribution are given by Kalman filter from the derivations and the likelihood is calculated using (11b) and (11c). Then, the likelihood of the model (5) at time t for time series possibly transformed[1] is given as

$$\mathcal{L}(\boldsymbol{\Omega}) = \prod_{t=1}^{n} g_t(y_t^{(\omega)} | y_1, \ldots, y_{t-1}, \boldsymbol{\Omega}) = \prod_{t=1}^{n} g_t(y_t^{(\omega)} | \mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \quad \text{where}$$

$$g_t(y_t^{(\omega)} | \mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{\kappa} \left|\Sigma_t\right|^{-(1/2)} exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\epsilon}_t\right\}, \quad \text{then}$$

$$g_t(y_t | \mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) = g_t(y_t^{(\omega)} | \mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \left|\det\left(\frac{\partial y_t^{(\omega)}}{\partial y_t}\right)\right| = g_t(y_t^{(\omega)} | \mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \prod_{t=1}^{n} y_t^{\omega-1}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^{\kappa} \left|\Sigma_t\right|^{-(1/2)} \exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\epsilon}_t\right\} \prod_{t=1}^{n} y_t^{\omega-1}$$

The log-likelihood is given by

$$\mathcal{L}(\boldsymbol{\Omega}) = -\frac{\kappa n}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Sigma}_t| - \frac{1}{2}\sum_{t=1}^{n}\boldsymbol{\epsilon}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\epsilon}_t + (\omega-1)\sum_{t=1}^{n}\log y_t$$

Multiplying this expression by -1 and omitting constant terms, we get

$$-\mathcal{L}(\boldsymbol{\Omega}) = \frac{1}{2}\log|\boldsymbol{\Sigma}_t| + \frac{1}{2}\sum_{t=1}^{n}\boldsymbol{\epsilon}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\epsilon}_t - (\omega-1)\sum_{t=1}^{n}\log y_t \tag{19}$$

If there is no need the Box–Cox transformation, the log-likelihood is given by

$$-\mathcal{L}(\boldsymbol{\Omega}) = \frac{1}{2}\log|\boldsymbol{\Sigma}_t| + \frac{1}{2}\sum_{t=1}^{n}\boldsymbol{\epsilon}_t'\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\epsilon}_t \tag{20}$$

3.2. Computational procedure

The functions (19) and (20) are highly nonlinear and complicated functions of the unknown parameters. The procedure is to fix the initial state vector $\boldsymbol{x}_0$ and develop a recursive process for log-likelihood function and successively apply the Newton–Raphson algorithm to update the parameter estimates until the log-likelihood is minimized. The optimization process is combined with the Kalman filter and conditioned according to the necessity or not of the Box–Cox transformation. The adopted seasonal formulation for TSCov model requires the estimation of seasonal initial values $2(k_1, k_2, \ldots, k_T)$. In this work, we apply the method described by De Livera *et al.* [1], based on the multiple linear regression for selecting the appropriate number of harmonics in the trigonometric terms.

The parameters are incorporated in the Kalman filter by the following procedure: (i) construct a vector of unknown parameters, $\boldsymbol{\Omega} = \{\sigma_\varepsilon, \sigma_\xi, \sigma_\zeta, \phi, \boldsymbol{\sigma}_e^{2(i)}, \boldsymbol{\sigma}_{e^*}^{2(i)}\}$ conditioned to seasonal patterns, Box–Cox transformation (indicate whether to use the Box-Cox transformation or no) and damping parameter (indicate whether to include a damping parameter in the trend or not) and incorporate it into the log-likelihood function of the model, (ii) incorporate the log-likelihood function in the Kalman filter to ensure step 2 described below. According to the steps 3 to 5, the optimal values of parameters are determined by minimizing the MSE (Mean Squared Error) of the one-step-ahead forecasting errors, through the Newton–Rapson method using the *optim* r function under the L-BFGS-B method. We formed a unique, recursive and systematic process combining the Kalman filter and the multiple linear regression to the Newton–Rapson method. We summarize the iterative procedure as follows:

(i) Select the initial values of the parameters, $\boldsymbol{\Omega}^{(0)}$. In this step, the transition parameter is configured with TRUE/FALSE to indicate whether the final model should or not include damping in the trend. Being set to NULL, the previous two cases are tried and from AIC the best fitted is selected;
(ii) Run the Kalman filter using the initial values of the parameters, $\boldsymbol{\Omega}^{(0)}$, to obtain the set of innovations and covariance, that is, $\{\boldsymbol{\epsilon}_t^{(0)}; t = 1, \ldots, n\}$ and $\{\boldsymbol{\Sigma}_t^{(0)}; t = 1, \ldots, n\}$;

(iii) Perform one iteration of the Newton-Rapson procedure by taking $-\ln[\mathcal{L}(\boldsymbol{\Omega})]$ as a criterion function to obtain the new set of estimates, therefore, $\boldsymbol{\Omega}^{(1)}$. In this step, the selection procedure of harmonics for seasonal component comes into play;
(iv) At iteration j, ($j = 1, 2, \ldots$), repeat step (*ii*) using $\boldsymbol{\Omega}^{(j)}$ in place of $\boldsymbol{\Omega}^{(j-1)}$ to obtain a new set of innovations values $\{\boldsymbol{\epsilon}_t^{(j)}\}$ and $\{\boldsymbol{\Sigma}_t^{(j)}\}$. Then, step (*iii*) is repeated to obtain new estimates, $\boldsymbol{\Omega}^{(j+1)}$.
(v) While run step (*iii*) in (*iv*), the Kalman filter is updated with new estimates $\boldsymbol{\Omega}^{(j+1)}$. The process ends when the estimates or the likelihood stabilize.

Standard errors of parameters estimates (St.Error). Since we use Newton's procedure, the Hessian matrix at the time of convergence may be used as an estimate to obtain the standard errors estimates, that is, we included a numerical evaluation of the Hessian matrix of $-\ln[\mathcal{L}(\hat{\boldsymbol{\Omega}})]$, where $\hat{\boldsymbol{\Omega}}$ is the vector of estimated parameters at the time of convergence.

3.3. Model selection

Let $\boldsymbol{x}_0$ be the initial state vector and $\boldsymbol{\Omega}$ the vector of unknown parameters. The *Information Criterion*

$$AIC = \mathcal{L}^*(\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{x}}_0) + 2(\Delta + \varrho) \tag{21}$$

is used for choosing between the models, where Δ is the number of parameters in $\boldsymbol{\Omega}$ and ϱ is the number of estimated states, and $\hat{\boldsymbol{\Omega}}$ and $\hat{x}_0$ denote the estimates of $\boldsymbol{\Omega}$ and $\boldsymbol{x}_0$.

4. Forecasting

4.1. Empirical forecasting under TSCov model without bootstrap

The Kalman filter equations can deal with missing observations in a natural way. We used as first forecast strategy the so-called *Increasing Horizon Prediction of the State* [26]. By extending the sample data $y_1^{(\omega)}, \ldots, y_n^{(\omega)}$ as missing values for $\boldsymbol{y}_t^{(\omega)}$ with $t = n+1, n+2, \ldots$, and applying the Kalman filter to the extended sample, the predictions are produced. Since $\boldsymbol{z}_t$ does not contain information about $\boldsymbol{x}_t$ beyond that contained in $\mathbf{y}_{1:t-1}$, $E(\boldsymbol{x}_t|\boldsymbol{z}_t, \mathbf{y}_{1:t-1}) = E(\boldsymbol{x}_t|\mathbf{y}_{1:t-1}) = \hat{\boldsymbol{x}}_{t|t-1}$.

Let $\hat{\boldsymbol{y}}_{t|t-1}^{(\omega)} \equiv E(\hat{\boldsymbol{y}}_t^{(\omega)}|\boldsymbol{x}_t, \mathbf{y}_{1:t-1})$ be the forecast of $\boldsymbol{y}_t$. From (5a), $E(\boldsymbol{y}_t|\boldsymbol{x}_t, \boldsymbol{z}_t) = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{\Gamma}\boldsymbol{z}_t$ and applying the law of iterated projections, one obtains

$$\hat{\boldsymbol{y}}_{t|t-1}^{(\omega)} = \boldsymbol{A}_t E(\boldsymbol{x}_t|\mathbf{y}_{1:t-1}, \boldsymbol{z}_t) + \boldsymbol{\Gamma}\boldsymbol{z}_t = \boldsymbol{A}_t\hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{\Gamma}\boldsymbol{z}_t$$

and its MSE is given as 11c.

Covariates forecast strategy. According to Hyndman *et al.* [25], if the covariates consist of indicator variables, their values are known to a certain future point in time. In addition, if such indicator variables reflect the effect of known future interventions that have also occurred in the past, then these values are also known. However, when they are unknown, predictions of future values of covariates are needed. In this work, we adopt the exponentially weighted moving average approach to predict covariates. The covariates are recursively smoothed by calculating the exponentially weighted moving average. In this way, the forecast of $\boldsymbol{z}_t$ in time $t+1$ is equal to a weighted average of the most recent observation $\boldsymbol{z}_t$ and the previous forecast $\hat{\boldsymbol{z}}_{t|t-1}$, that is

$$\hat{\boldsymbol{z}}_{t+1|t} = \rho \boldsymbol{z}_t + (1-\rho)\hat{\boldsymbol{z}}_{t|t-1} \tag{22}$$

where $0 \leq \rho \leq 1$ is the smoothing parameter that is typically close to 1. This strategy is similar to that applied by Dordonnat *et al.* [13].

Due to Markov structure in the state dynamics and assumptions about conditional independence of observations, the predictive distributions can be recursively calculated. The Kalman filter update in time t provide $\boldsymbol{x}_{t+1|t}$ and $\boldsymbol{P}_{t+1|t}$, used to obtain the one-step-ahead forecast of $\boldsymbol{y}^{(\omega)}_{t+1}$. The h-steps-ahead forecast is given by

$$\hat{\boldsymbol{y}}^{(\omega)}_{t+h|t} \equiv E(\boldsymbol{y}^{(\omega)}_{t+1}|\mathbf{y}_{1:t-1}, \boldsymbol{z}_t) = \boldsymbol{A}_{t+1}\hat{\boldsymbol{x}}_{t+1|t} + \boldsymbol{\Gamma}\hat{\boldsymbol{z}}_{t+h|t} \tag{23a}$$

$$MSE(\hat{\boldsymbol{y}}^{(\omega)}_{t+h|t}) = \boldsymbol{A}_{t+h}\boldsymbol{P}_{t+h|t}\boldsymbol{A}'_{t+h} + \boldsymbol{R}_{t+h} \tag{23b}$$

where

$$\hat{\boldsymbol{x}}_{t+h|t} = \boldsymbol{\Phi}\hat{\boldsymbol{x}}_{t+h-1|t} \tag{24a}$$

$$MSE(\hat{\boldsymbol{x}}_{t+h|t}) = \boldsymbol{P}_{t+h|t} = \boldsymbol{\Phi}\boldsymbol{P}_{t+h-1|t}\boldsymbol{\Phi}' + \boldsymbol{Q}_{t+h} \tag{24b}$$

The prediction intervals can be obtained directly. Since the h-steps-ahead forecast errors are Gaussian, we generate the prediction intervals (PI) of the nominal coverage rate of 95% as,

$$PI = \hat{y}^{(\omega)}_{t+h} \pm 1.96\sqrt{MSE(\hat{y}^{(\omega)}_{t+h|t})}, \tag{25}$$

4.2. Bootstrap procedure

The combination of state-space models and bootstrap methods currently has received much attention from researchers [11,12,19,28,37,38]. The results of these works shows that the combination of bootstrap methods with state-space models is valuable for time series forecasting, once it can provide more accurate forecasts than individual methods. The basic approaches of existing bootstrap methods, one of these approaches is the residual bootstrap in state-space models. Theoretically, if the model is correctly fitted, the residuals of the model would be independent and identically distributed. So, it is possible to sample these residuals with replacement to obtain a replica of the original sample. Then fit the model to the original sample replica and repeat the process, see [10].

Our proposed bootstrap procedure, which we call henceforward Boot.TSCov, is inspired by Cordeiro and Neves [12] and Rodrigues and Ruiz [38]. The first goal of our bootstrap procedure is to improve the forecasting method described in section 4.1. The second goal is to provide contributions for short-term forecasting of time series with complex seasonal patterns using bootstrap method, once it is a paradigm that in the scope for forecasting of time series with complex seasonal patterns still presents a void in the scant existing literature.

4.2.1. Boot.TSCov *general procedure*

First, use the procedure presented in section 3.2 to estimate the initial model TSCov given in (4) to obtain the sequence of innovations, $\boldsymbol{\epsilon}_t$ (which must be uncorrelated), and the fitted values $\{\hat{y}_1, \ldots, \hat{y}_n\}$ of the initial estimated model. Second, the standardized innovations, (26a), with the assurance that these innovations have, at least, the same first two moments, are resampled with replacement b times to obtain a bootstrap sample of standardized innovations, $\boldsymbol{v}_t^{*s}$ (step 3.1). Then, according to Cordeiro and Neves [12] the bootstrap replica of original time series can be obtained using (26b).

$$\boldsymbol{v}_t^s = \boldsymbol{\Sigma}_t^{-1/2}\boldsymbol{\epsilon}_t \tag{26a}$$

$$\boldsymbol{y}_t^* = \hat{\boldsymbol{y}}_t + \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{v}_t^{*s} \tag{26b}$$

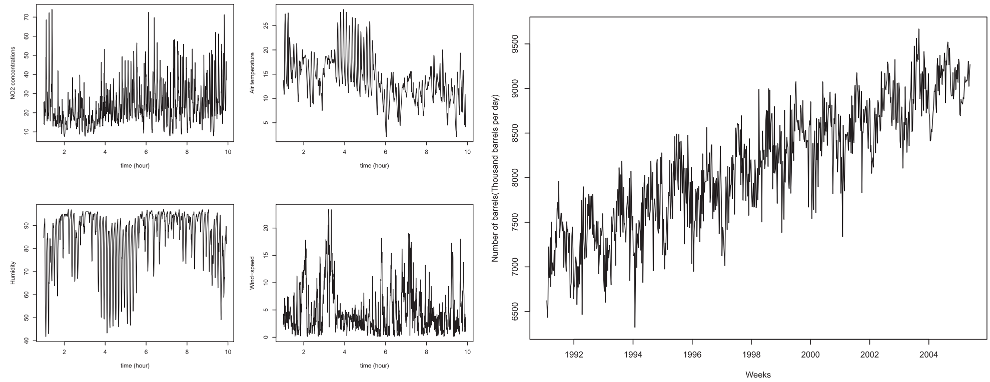
Then, use bootstrap replica, $\boldsymbol{y}_t^*$, estimate the bootstrap model to obtain bootstrap estimates, $\hat{\boldsymbol{\Omega}}^*$, a priori ($\hat{\boldsymbol{\epsilon}}_t^*$) and a posteriori ($\hat{\boldsymbol{\varsigma}}_t^*$) innovations, the state vector and other Kalman filter derivatives. The forecast up to h-steps-ahead is obtained using the Kalman filter recursions with bootstrap estimates. Next, we summarize the main steps of our proposed bootstrap procedure.

4.2.2. *Procedure steps*

1. Use the iterative procedure described in section 3.2 and estimate the model defined by (5) to obtain the sequence of innovations, $\boldsymbol{\epsilon}_t$;
2. Compute the standardized innovations using (26a);
3. For each replica B,
   - 3.1 Resample with replacement b times the standardized innovations $\{\boldsymbol{v}_1^s, \boldsymbol{v}_2^s, \ldots, \boldsymbol{v}_n^s\}$ to obtain the bootstrap sample of standardized innovations $\{\boldsymbol{v}_1^{*s}, \boldsymbol{v}_2^{*s}, \ldots, \boldsymbol{v}_n^{*s}\}$;
   - 3.2 Compute a bootstrap replicate, $\boldsymbol{y}_t^*$, by Equation (26b) using $\boldsymbol{v}_t^{*s}$. From the iterative procedure described in section 3.2, estimate the corresponding bootstrap parameters, $\hat{\boldsymbol{\Omega}}^*$, and the Kalman filter derivatives, such as a priori ($\hat{\boldsymbol{\epsilon}}_t^*$) and a posteriori ($\hat{\boldsymbol{\varsigma}}_t^*$) innovations, the state vector at time t and others;
   - 3.3 Obtain conditional bootstrap h-step-ahead forecast, $\hat{\boldsymbol{y}}_{t+h|t}^*$, from the following expressions:

$$\begin{aligned}
\hat{\boldsymbol{x}}_{t+h|t}^* &= \hat{\boldsymbol{\Phi}}^*\hat{\boldsymbol{x}}_{t+h-1|t}^* \\
\hat{\boldsymbol{P}}_{t+h|t}^* &= \hat{\boldsymbol{\Phi}}^*\hat{\boldsymbol{P}}_{t+h-1|t}^*\hat{\boldsymbol{\Phi}}^{*'} + \hat{\boldsymbol{Q}}_{t+h|t}^* \\
\hat{\boldsymbol{y}}_{t+h|t}^* &= \hat{\boldsymbol{A}}_{t+h}^*\hat{\boldsymbol{x}}_{t+h|t}^* + \hat{\boldsymbol{\Gamma}}^*\boldsymbol{z}_{t+h} \\
\hat{\boldsymbol{\Sigma}}_{t+h|t}^* &= \hat{\boldsymbol{A}}_{t+h}^*\hat{\boldsymbol{P}}_{t+h|t}^*\hat{\boldsymbol{A}}_{t+h}^{*'} + \hat{\boldsymbol{R}}_{t+h|t}^* \\
\hat{\boldsymbol{Q}}_{t+h|t}^* &= \delta\hat{\boldsymbol{Q}}_{t+h-1}^* + (1-\delta)(\hat{\boldsymbol{K}}_t^*\hat{\boldsymbol{\epsilon}}_t^*\hat{\boldsymbol{\epsilon}}_t^{*'}\hat{\boldsymbol{K}}_t^{*'}) \\
\hat{\boldsymbol{R}}_{t+h|t}^* &= \delta\hat{\boldsymbol{R}}_{t+h-1}^* + (1-\delta)(\hat{\boldsymbol{\varsigma}}_t^*\hat{\boldsymbol{\varsigma}}_t^{*'} + \hat{\boldsymbol{A}}_{t+h}^*\hat{\boldsymbol{P}}_{t+h-1|t}^*\hat{\boldsymbol{A}}_{t+h}^*)
\end{aligned}$$

where, $\hat{\boldsymbol{\varsigma}}_t^* = \boldsymbol{y}_t - \hat{\boldsymbol{A}}_t^*\hat{\boldsymbol{x}}_{t|t}^* - \hat{\boldsymbol{\Gamma}}_t^*\boldsymbol{z}_t$ and $\hat{\boldsymbol{\epsilon}}_t^* = \boldsymbol{y}_t - \hat{\boldsymbol{A}}_t^*\hat{\boldsymbol{x}}_{t|t-1}^* - \hat{\boldsymbol{\Gamma}}_t^*\boldsymbol{z}_t$. The hat on the matrices means that they are matrices estimated by bootstrap and used for forecast. By taking the



(a) (b)

Figure 1. Complex seasonality showing: (a) Hourly NO_2 concentrations levels (with multiple seasonal periods) measured from October 1^{st} and December 31^{st} in 2014 in Paredes/Portugal, including Temperature, Relative humidity and Wind-speed as covariates also observed at hourly intervals; (b) Weekly US finished motor gasoline products in thousands (with non-integer seasonal periods), from February 1991 to July 2005.

average ($\hat{y}_H$) of the empirical distribution of $\hat{y}^*_{t+h|t}$, the prediction intervals are generated as (25).

5. Applications to real time series

The results obtained from the application of TSCov and TBATS to the two complex time series in Figure 1 are reported in this Section. Each one of time series is splitted into two parts: fitting set and validation set.

$$\underbrace{y_1, y_2, \ldots, y_{n-h},}_{\text{Fitting set}} \quad \overbrace{y_{n-h+1}, \ldots, y_n}^{\text{Validation set}}$$

The TSCov model is configured to run with or without covariates. If there is no need to include covariates in the model, we only need to set input $= 0$ and $\boldsymbol{\Gamma} = 0$ from the generic function which projects the Kalman filter. The initialization of parameters is not automatic, it depends on the features of each dataset used, except the transition parameter, ϕ. Details about the initialization for each estimated models are presented in subsections 5.1 and 5.2.

We reported some results of TBATS model to compare with the TSCov model and Boot. TSCov procedure, for example, using the same nominal coverage rate to generate the prediction intervals and compare the performance of each model. To assess forecasting performance, we used the mean absolute percentage forecast error (MAPE) and the root mean squared forecast error (RMSE).

5.1. Application to multiple seasonal patterns data

Figure 1(a) shows the hourly NO_2 concentrations levels, obtained from the online database on air quality [34]. The dataset under analysis concern 49 stations located over Portugal (mainland) from October 1st and December 31st in 2014. The selected period corresponds to the highest NO_2 levels along the year, according to Andreia *et al.* [5]. The time series of NO_2 denoted by N_t has a daily pattern with period 24 and a weekly seasonal pattern with period 168. The covariates considered are Temperature, T_t, Relative humidity, H_t, and Wind-speed, W_t. We estimated two TSCov models: one model with real covariates and other with predicted covariates. The series, which consists of 2208 observations, was split into two segments: an estimation sample period (1488 observations) and a test sample (720 observations). The point forecasts are obtained only using the *Increasing Horizon Prediction of the State* strategy.

A preliminary analysis suggests the following considerations: fit a TBATS model to NO_2 to determine the cross-correlation function (CCF) between the NO_2 residuals, temperature series, humidity series and wind-speed series. The results (no output is shown) indicates that the strongest correlations occur at 12-h lag with air temperature (T_{t-12}), 2-h lag with relative humidity (H_{t-2}) and current Wind-speed (W_t).

Initialization of TSCov *model with real covariates.* The initial mean was fixed at $\boldsymbol{x}_0 = 0.7$ with uncertainty modeled by the diagonal covariance matrix $\boldsymbol{\Sigma}_{0ii} = 4$, for $i = 1, \ldots, 16$. Initial state covariance values were taken as $\boldsymbol{Q}_0 = \text{diag}\{\sigma^2_\xi, \sigma^2_\zeta, \sigma_e^{2(i)}\} = \{0.004, 0.17, 0.0001, 0.0001\}$ $(i = 4)$. The measurement error covariance was started at $\boldsymbol{R}_0 = \sigma^2_\varepsilon = 10^{-8}$. Initial regression coefficients values was fixed at $\beta^*_1 = \beta^*_2 = \beta^*_3 = 0.1$, and the forgetting factor, $\delta = 0.999$.

Initialization of TSCov *model with predicted covariates.* The initial mean and its uncertainty (modeled by the diagonal covariance matrix) was fixed at $\boldsymbol{x}_0 = 0.7$ and $\boldsymbol{\Sigma}_{0ii} = 3.1$ $(i = 1, \ldots, 17)$, respectively. Initial state covariance values were taken as $\boldsymbol{Q}_0 = \text{diag}\{\sigma^2_\xi, \sigma^2_\zeta, \sigma_e^{2(i)}\} = \{0.004, 0.045, 0.001, 0.001\}$ $(i = 4)$. The measurement error covariance was started at $\boldsymbol{R}_0 = \sigma^2_\varepsilon = 10^{-6}$. Initial regression coefficients values was fixed at $\{\beta^*_1, \beta^*_2, \beta^*_3\} = 0.01$, and the forgetting factor, $\delta = 0.8$.

The TBATS models are implemented in the *forecast r package* [45]. Therefore, its initialization process is automatic.

Figure 2 shows the residual analysis. For TSCov model with real covariates, Figure 2(a), the lag 18 is outside the interval, but we view this value as acceptable, since we do not find any pattern that leads us to believe that we have missed a structural dynamic feature in the time series. The Ljung-Box test provides a Chi-square value equal to 29.154, with 19 degrees of freedom and p-value $= .164$, allowing not to reject the null hypothesis that the residuals are independent. The required number of significant harmonics for the trigonometric terms of daily seasonal pattern with periodicity 24 was $k_1 = 5$ and $k_2 = 3$ for weekly seasonal pattern with periodicity 168. For TBATS model, Figure 2(b), the correlation in lags 23 and 24 shows that the model does not capture all the dynamics well. However, the Ljung-Box test provides a Chi-square value equal to 19.443 with 18 degrees of freedom and p-value $= .265$, we also do not reject the null hypothesis.

Table 1 reports the estimated parameters of TSCov model with real covariates and TBATS model. Almost all of the estimates are significant. The estimated $\hat{\phi}$ from TBATS model does not suggest the damping effect in the trend component, unlike TSCov model that suggests a damping effect in the trend component. The measurement uncertainty was, in general, large at $\sigma^2_\varepsilon = 2.294$, compared with the model uncertainties of the level, trend and seasonal.

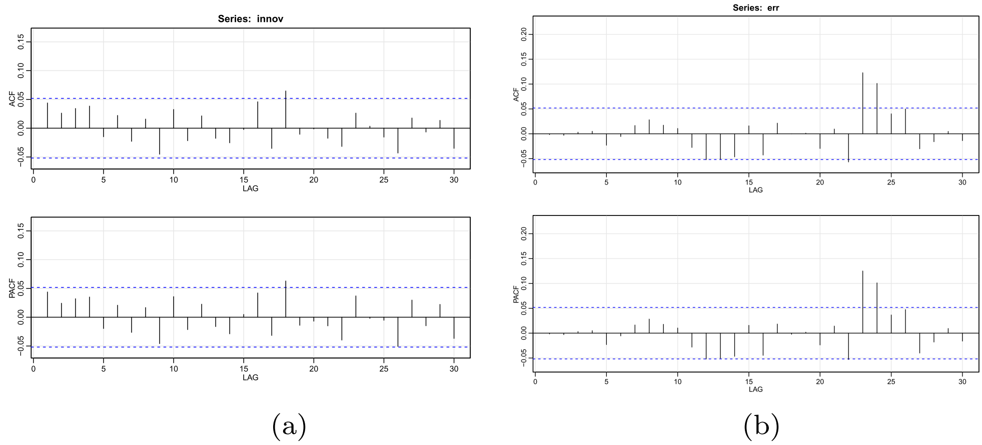


Figure 2. Autocorrelation plots of the one-step-ahead forecasting errors under the: (a) TSCov model (with real covariates) and (b) TBATS model for NO_2 levels data. Dashed: confidence intervals of level 95%.

Table 1. Estimated parameters and their standard errors obtained from TSCov model (with real covariates): NO_2 levels in Paredes/Portugal.

	Estimated models for NO_2 concentrations		
Parameter	MLE (TSCov)	St.Error	MLE(TBATS)
β_1^*	−0.724	0.437	–
β_2^*	−0.007	0.013	–
β_3^*	0.257	0.035	–
ω	–	–	–
α	–	–	0.039
β	–	–	0.0002
ϕ	0.947	0.175	1
σ_ε^2	2.294	0.161	–
σ_ξ^2	1.596	0.073	–
σ_ζ^2	0.091	0.036	–
σ_e^2	{0.002; 0.019}	{0.012; 0.022}	–
σ_{e*}^2	{0.012; −0.048}	{0.164; 0.011}	–
γ_1	–	–	$\{-10^{-4}; -2.10^{-6}\}$
γ_2	–	–	$\{-3.10^{-5}; 3.10^{-5}\}$

Note: The table also shows the parameters obtained from TBATS model.

We may also consider forecasting the NO_2 series, and the result of a 24-steps-ahead forecast is shown in Figure 3(a). Using the same nominal coverage rate, 95%, we generate the prediction intervals for both models, Figure 3(b). As it can be seen the prediction intervals based on the proposed TSCov model are narrower and more regular over the forecast horizon and cover all future values than those obtained with TBATS model.

Table 2 shows the forecasts accuracy computed using real and predicted covariates for TSCov model. It should be noted that, for TSCov model (with real or predicted covariates), the RMSE provided values that we consider small in almost all forecast horizons, unlike the TBATS model that provided higher values for higher forecast horizons. In addition,

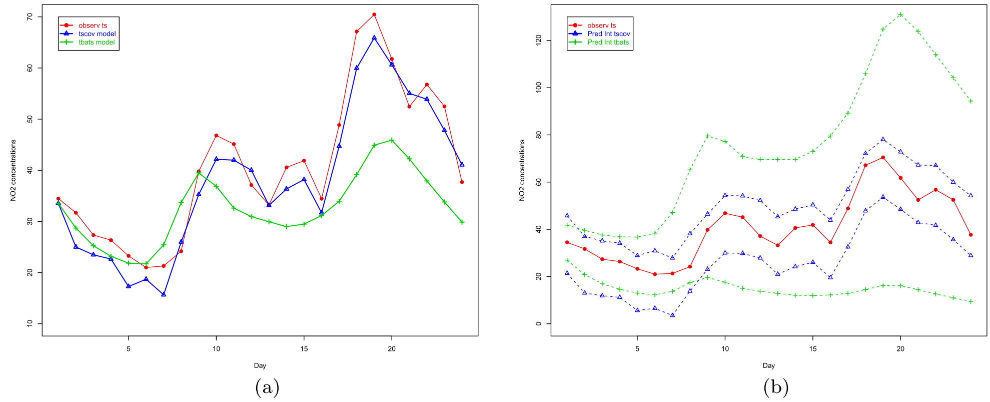


Figure 3. Forecasts up to 24-steps-ahead of NO_2 levels in Paredes/Portugal under the: (a) TSCov model (with real covariates) and TBATS model; (b) 95% prediction intervals obtained from TSCov and TBATS models.

Table 2. Forecasting accuracy up to 24-steps-ahead of NO_2 levels in Paredes/Portugal.

	TSCov[a]		TSCov[b]		TBATS	
Horizon	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1 – 3	4.127	3.643	4.717	5.217	2.179	6.819
1 – 6	4.067	3.646	4.727	5.387	2.111	6.699
1 – 9	4.206	3.787	4.833	5.287	3.866	9.326
1 – 12	4.246	3.607	5.303	6.513	5.973	11.213
1 – 15	4.293	3.827	5.081	6.571	6.962	12.404
1 – 18	4.361	3.827	5.356	6.581	6.969	15.243
1 – 21	4.388	3.848	5.369	7.748	9.840	15.657
1 – 24	4.495	3.911	5.903	7.804	11.448	17.574

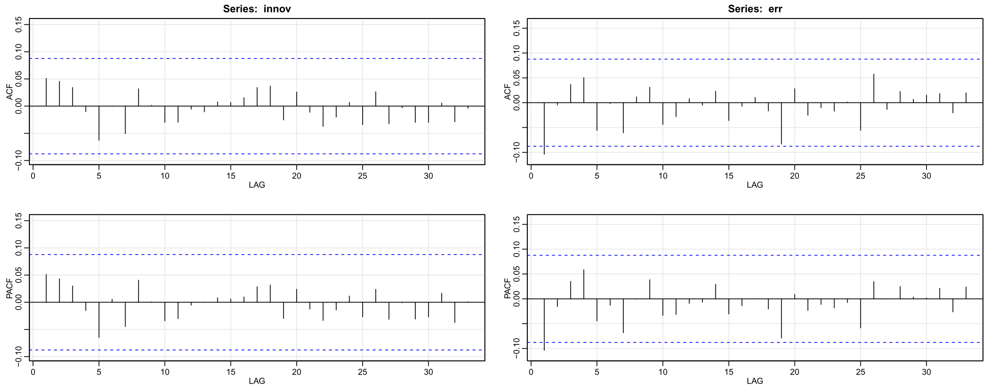
[a] Forecast with real covariates.
[b] Forecast with predicted covariates.

both graphically and in terms of accuracy measures, both are in agreement that the TSCov model got a good rating.

5.2. Application to non-integer seasonal periods data

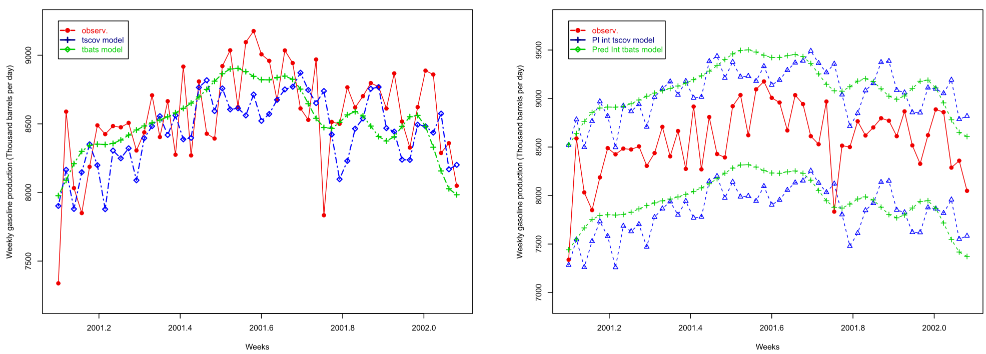
Figure 1(b) shows the number of barrels of motor gasoline product supplied in the United States, in thousands of barrels per day, from February 1991 to July 2005. This dataset was also used by De liver *et al.* [1] (see https://robjhyndman.com/publications/complex-seasonality/). The data are observed weekly and show a strong annual seasonal pattern, where the length of seasonality of the time series is $365.25/7 \approx 52.179$. According to De Livera *et al.* [1], 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.

For this case, the TSCov model and Boot.TSCov procedure are applied without the inclusion of covariates. The time series consists of 745 observations and was split into two segments: an estimation sample period (520 observations) and a test sample (225 observations). The initial model for Boot.TSCov procedure is fitted using the same procedure



(a) (b)

Figure 4. Autocorrelation plots of the one-step-ahead forecasting errors under the: (a) TSCov model without covariates for weekly U.S. Gasoline data; (b) TBATS model for weekly U.S. Gasoline data. Dashed: confidence intervals of level 95%



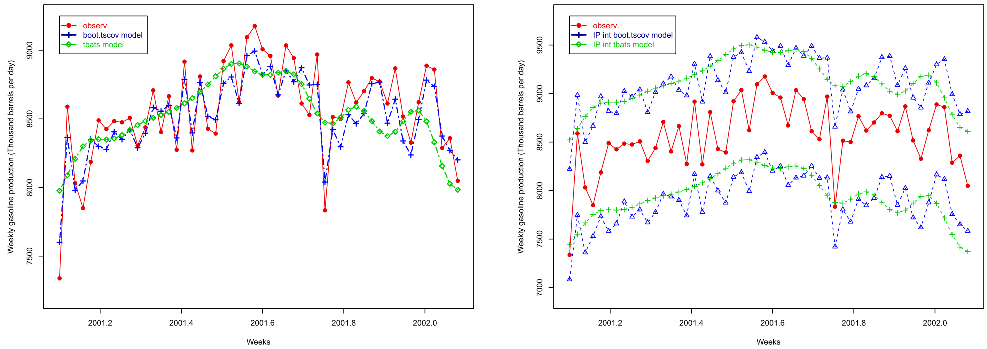
(a) (b)

Figure 5. Increasing State Horizon Strategy – Forecast up to 52-steps-ahead of weekly U.S. Gasoline data under: (a) TSCov and TBATS models; (b) 95% prediction intervals obtained from TSCov and TBATS models.

described in section 3.2. Thus, the initial mean was fixed at $\hat{\boldsymbol{x}}_0 = 0$ with uncertainty modeled by the diagonal covariance matrix $\boldsymbol{\Sigma}_{0ii} = 6.5$ $(i = 1, \ldots, 18)$, the forgetting factor was fixed at $\delta = 0.999$. The measurement error covariance was started at $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 10^{-8}$. The initial state covariance values were taken as $\boldsymbol{Q}_0 = \text{diag}\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^2\} = \{0.0005, 0.0005, 0, 0\}$.

The point forecasts are obtained using the *Increasing Horizon Prediction of the State* and *Bootstrap* strategies.

Figure 4 shows the residual analysis. We noted that, although the Box–Cox transformation, the empirical autocorrelation of estimated TBATS model shows a negative correlation for lag 1, Figure 4(b), which is unlikely to be due to random sampling variation. However, the Ljung-Box test provides a Chi-square value equal to 19,392, with 24 degrees of freedom and p-value $= .731$, allowing not reject the null hypothesis. For TSCov model, there is no significant correlations, Figure 4(a). They are all within the 95% confidence interval, which



(a) (b)

Figure 6. Bootstrap Strategy–Forecast up to 52-steps-ahead of weekly production of weekly U.S. Gasoline data under: the (a) Boot.TSCov and TBATS models; (b) 95% prediction intervals obtained from Boot.TSCov and TBATS models.

Table 3. Estimated parameters and their standard errors of the TSCov model for weekly U.S. Gasoline data.

	Estimated models for weekly U.S. Gasoline data				
Parameter	MLE(TSCov)	St.Error	Boot.TSCov	St.Error	MLE(TBATS)
ω	–	–	–	–	0.709
α	–	–	–	–	−0.063
β	–	–	–	–	0.031
ϕ	0.829	0.154	0.801	0.457	0.834
σ_ε^2	220.69	1.751	222.25	1.916	–
σ_ξ^2	1094.29	9.597	1143.19	9.947	–
σ_ζ^2	860.39	3.819	991.51	4.022	–
σ_w^2	7.055	0.489	1.249	0.604	–
σ_{w*}^2	2.892	0.216	0.029	0.263	–
γ_1	–	–	–	–	−0.003
γ_2	–	–	–	–	0.002

Note: The table also shows the parameters obtained from TBATS model.

is satisfactory. In addition, the Ljung-Box test provides a Chi-square value equal to 11,485, with 24 degrees of freedom and p-value $= .985$; we also do not reject the null hypothesis.

Table 3 shows the estimated parameters and both models including Boot.TSCov procedure suggests a damping effect in the trend component. Figure 5(a) shows the 52-steps-ahead forecasts obtained from TSCov model using *Increasing State Horizon* strategy and Figure 6(a) shows 52-step-ahead forecasts obtained from Boot.TSCov procedure. The forecasts accuracy are shown in Table 4. As it can be seen, TSCov model competes with TBATS model, but Boot.TSCov procedure performs better for all lead times than TBATS model. The fitted values of the two models indicate that the forecasts generated by the TSCov model are closer to the validation series and TBATS model offers smoother fitted values and forecasts than those obtained from the TSCov model.

We may also consider the prediction intervals, see Figures 5(b) and 6(b). Using the same nominal coverage rate, 95%, the prediction intervals based on the Boot.TSCov procedure got a good rating compared to those obtained with the TSCov and TBATS models.

Table 4. Forecasting accuracy up to 52-steps-ahead of weekly U.S. Gasoline data.

	TSCov[a]		Boot.TSCov[b]		TBATS	
Horizon	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1 – 7	204.524	2.834	213.342	2.451	269.031	2.446
1 – 14	253.443	3.116	219.713	2.643	269.543	2.691
1 – 21	273.329	3.203	226.368	2.669	272.326	2.687
1 – 28	270.323	3.250	255.145	2.725	275.282	2.711
1 – 35	291.233	3.314	263.405	2.776	278.056	2.745
1 – 42	297.547	3.425	267.096	2.780	281.356	2.760
1 – 49	323.651	3.671	293.612	2.860	296.847	2.936
1 – 52	358.524	3.993	311.634	2.973	359.863	3.906

[a] Forecast (without covariates) obtained using the Increasing Horizon Prediction of the State strategy.
[b] Forecast (without covariates) obtained using Bootstrap strategy.

These results allow to conclude that the TSCov model and Boot.TSCov procedure are also pertinent in handling the non-integer seasonality.

6. Conclusion and future direction

The main contribution of this work is to explore the use of covariates in short-term forecasting of time series with complex seasonal patterns, as an extension of the TBATS model. Our proposed framework responds two interesting problems: (i) in the field of forecasting, the proposed framework deals with covariates that may be important for the short-term forecasting; (ii) from the point of view of formulation, the proposed framework is a new tool for structural models for modeling and forecasting of time series with complex seasonal patterns. The answer to these ”two” problems is a valuable contribution to limited existing literature on structural models for predicting time series with complex seasonal patterns. The formulated bootstrap procedure has as main objective to improve the short-term forecasting obtained from the usual procedure with Kalman filter recursions. The procedure was applied satisfactorily and the results show that the proposed bootstrap procedure provides good results for the used dataset.

The empirical study shows the potential of our proposals as promising methodologies for short-term forecasting of time series with complex seasonal patterns. The used covariates had a significant impact on the forecast and, as expected, the forecasts obtained were more accurate under the use of covariables. Our estimation procedure not only obtains point forecasts and prediction intervals, but also allows to obtain the standard errors of each estimated parameter. However, our proposed methodologies can be improved in several ways and we can see some lines of future work listed below:

- *Automatic selection of covariates*. Some study can be done in this line for the automatic selection of candidate covariates to the estimated final model;
- *Estimator of the coefficient matrix*, $\boldsymbol{\Gamma}$. The state space model given in (5) involves covariates in the measurement equation. However, the Kakman filter constructed for this model does not provide the estimator of the coefficient matrix $\boldsymbol{\Gamma}$. An updated estimate of this matrix can be obtained by applying the Expectation Maximization algorithm (EM) [6];

- *Multivariate analysis*. When significant dependencies between individual time series can not be ignored, multivariate time series need to be introduced. A projection of the proposed framework to deal with multivariate time series is necessary. Our study was done for the univariate case. But the proposed framework can "easily" be reformulated to the multivariate case.

Finally, with this work, a question can be made here: which models are preferred to use, TSCov and TBATS? It should, however, be noted from the results that there is little distinction between the two models, therefore, the immediate response is, the TSCov approach which is preferable if there are covariates that are useful predictors since they can be added as regressors and improve the forecasts.

All computational results of this work were obtained with the R software environment [35].

Note

1. Transformation Box-Cox [9]. If Box–Cox transformation is required, the point forecasts and forecast intervals may be obtained using the inverse Box–Cox transformation of appropriate quantiles of the distribution of $y^{(\omega)}_{t+h|t}$. Moreover, the prediction intervals retain the probability of coverage required by the back-transformation because the Box–Cox transformation is monotone increasing [1].

Acknowledgements

The authors thank the Portuguese Environment Agency and Professor Robin John Hyndman for providing the data used in this work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was partially supported by the Center for Research and Development in Mathematics and Applications (CIDMA) through the Portuguese Foundation for Science and Technology (FCT – Fundação para a Ciência e a Tecnologia), references UIDB/04106/2020 and UIDP/04106/2020 and ENAGBE (National Institute of Managment of Scholarships) – Angola.

ORCID

Maria Eduarda Silva http://orcid.org/0000-0003-2972-2050

References

[1] Alysha M. De Livera, Rob J. Hyndman, and Ralph D. Snyder, *Forecasting time series with complex seasonal patterns using exponential smoothing*, J. Am. Stat. Assoc. 106 (2011), pp. 1513–1527. doi:10.1198/jasa.2011.tm09771.

[2] F.O. Ahmad and L.K. Maxwell, *Exponential smoothing with regressors: estimation and initialization*, Model. Assist. Stat. Appl. 10 (2015), pp. 253–263.

[3] S. Akhlaghi, N. Zhou, and Z. Huang, *Adaptive adjustment of noise covariance in Kalman filter for dynamic state estimation*, 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, 2017, pp. 1–5.
[4] A.M. Alonso, C. Garcia-Martos, J. Rodriguez, and M.J. Sanchez, *Seasonal dynamic factor analysis and bootstrap inference: application to electricity market forecasting*, Technometrics 53(2) (2011), pp. 137–151. doi:10.1198/TECH.2011.09050.
[5] M. Andreia, R. Menezes, and M. Eduarda Silva, *Modelling spatio-temporal data with multiple seasonalities: the NO2 Portuguese case*, Spat. Stat. 22 (2017), pp. 1–25.
[6] H.N. Arlene, *State space models with exogenous variables and missing data*, PhD thesis, University of Florida, 2007.
[7] A. Ba, M. Sinn, Y. Goude, and P. Pompey, *Adaptive learning of smoothing functions: application to electricity load forecasting*, in Advances in Neural Information Processing Systems. 25, P. Bartlett, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, eds. MIT Press: Cambridge, MA, 2012, pp. 2519–2527.
[8] P.J. Brockwell and R.A. David, *Introduction to Time Series and Forecasting*, 2nd ed. Springer-Verlang, New York, 2002.
[9] G. Box and D. Cox, *An analysis of transformations*, J. R. Stat. Soc.; Ser. B 26 (1964), pp. 211–252.
[10] M.R. Chernick and R.A. LaBudde, *An Introduction to Bootstrap Methods with Applications to R*, Wiley, 2011. pp. 3–129.
[11] B. Christoph, R.J. Hyndman, and J.M. Benitez, *Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation*, Int. J. Forecast. 32 (2015), pp. 2–18.
[12] C. Cordeiro and M.M. Neves, *Forecasting with exponential smoothing methods and bootstrap*, REVSTAT–Stat. J. 7(2) (2009), pp. 135–149.
[13] V. Dordonnat, S.J. Koopman, M. Ooms, A. Dessertaine, and J. Collet, *An hourly periodic state space model for modelling French national electricity load*, Int. J. Forecast. 24 (2008), pp. 566–587.
[14] J. Durbin and S.J. Koopman, *Time Series Analysis by State Space Methods*, Oxford University Press, 2011.
[15] S. Fan and R.J. Hyndman, *Short-term load forecasting based on a semi-parametric additive model*, IEEE Trans. Power Syst. 27 (2012), pp. 134–141.
[16] R. Gob, K. Lurz, and A. Pievatolo, *Electrical load forecasting by exponential smoothing with covariates*, Appl. Stoch. Model. Bus. Ind. 29 (2013), pp. 629–645.
[17] H.L. Shang, *Functional time series approach for forecasting very short-term electricity demand*, J. Appl. Stat. 40 (2013), pp. 152–168. doi:10.1080/02664763.2012.740619.
[18] P.G. Gould, A.B. Koehler, F. Vahid-Araghi, R.D. Snyder, J.K. Ord, and R.J. Hyndman, *Forecasting time-series with multiple seasonal patterns*, Eur. J. Oper. Res. 191 (2008), pp. 207–222.
[19] G. Hafida and F. Hamdi, *Bootstrapping periodic state-space models*, Commun. Stat. Simul. Comput. 44 (2015), pp. 374–401.
[20] J.D. Hamilton, *Time Series Analysis*. Vol. 41, William St. Princeton ed. Princeton University Press, New Jersey, 1994. p. 08540, 1994.
[21] A.C. Harvey, *Forecasting, structural time series models and the Kalman Filter*, Cambridge University Press, Cambridge, 1989.
[22] A.C. Harvey and S.J. Koopman, *Forecasting hourly electricity demand using timevarying splines*, J. Am. Stat. Assoc. 88 (1993), pp. 1228–1236.
[23] J. Hinman and E. Hickey, *Modeling and forecasting short-term electricity load using regression analysis*.University of Illinois Research Report, Chicago, IL, 2009.
[24] R.J. Hyndman, A.B. Koehler, R.D. Snyder, and S. Grose, *A state space framework for automatic forecasting using exponential smoothing methods*, Int. J. Forecast. 18 (2002), pp. 439–454.
[25] R.J. Hyndman, A.B. Koehler, J.K. Ord, and R.D. Snyder, *Forecasting with Exponential Smoothing: The State Space Approach*, Springer-Verlang, 2008.
[26] G. Kitagawa, *Introduction to Time Series Modeling*, CRC Press, Boca Raton, 2010.
[27] A.B. Koehler, R.D. Snyder, J.K. Ord, and A. Beaumont, *A study of outliers in the exponential smoothing approach to forecasting*, Int. J. Forecast. 28 (2012), pp. 477–484.

[28] J.C. Menezes, V.V. Lopes, and C.C. Pinheiro, *Determination of state-space model uncertainty using bootstrap techniques*, J. Process. Control. 16 (2006), pp. 685–692.
[29] A. Mohamed and K. Schwarz, *Adaptive Kalman Filtering for INS/GPS*, J. Geod. 73 (1999), pp. 193–203. doi:10.1007/s001900050236.
[30] J. Ord, A.B. Koehler, and R.D. Snyder, *Estimation and prediction for a class of dynamic nonlinear statistical models*, J. Am. Stat. Assoc. 92 (1997), pp. 1621–1629.
[31] J. Ord, R.D. Snyder, A.B. Koehler, R.J. Hyndman, and M. Leeds, *Time series forecasting: the case for the single source of error state space approach*, Unpublished manuscript, Monash University, 2005, pp. 2–33.
[32] A.D. Papalexopoulos and T.C. Hesterberg, *A regression-based approach to short-term system load forecasting*, IEEE Trans. Power Syst. 5 (1990), pp. 1535–1547. doi:10.1109/59.99410.
[33] D.J. Pedregal and P.C. Young, *Modulated cycles, an approach to modelling periodic components from rapidly sampled data*, Int. J. Forecast. 22 (2006), pp. 181–194.
[34] QualAr: *Online database on air quality*, 2015. Available at https://qualar.apambiente.pt/qualar/index.phpSEP.
[35] R Core Team: *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017; software available at https://www.R-project.org.
[36] R. Ramanathan, R. Engle, C.W.J. Granger, F. Vahid-Araghi, and C. Brace, *Shorte-run forecasts of electricity loads and peaks*, Int. J. Forecast. 13 (1997), pp. 161–174. doi:10.1016/S0169-2070(97)00015-0.
[37] H.S. Robert and David S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, 4th ed., Springer, New York, 2017.
[38] A. Rodriguez and E. Ruiz, *Bootstrap prediction intervals in state-space models*, J. Time Ser. Anal. 30 (2009), pp. 167–178.
[39] J.W. Taylor, *Short-term electricity demand forecasting using double seasonal exponential smoothing*, J. Operat. Res. Soc. 54 (2003), pp. 799–805.
[40] J.W. Taylor and R. Buizza, *Using weather ensemble predictions in electricity demand forecasting*, IEEE Trans. Power Syst. 19 (2003), pp. 57–70.
[41] J.W. Taylor, *Triple seasonal methods for short-term electricity demand forecasting*, Eur. J. Operat. Res. 204 (2010), pp. 139–152.
[42] J.W. Taylor and R.D. Snyder, *Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing*, Omega 40 (2012), pp. 748–757.
[43] J. Wang, *Stochastic modeling for real-time kinematic gps/glonass positioning*, Navigation 46 (2000), pp. 297–305.
[44] S. Wang, *Exponential smoothing for forecasting and Bayesian validation of computer models*, PhD thesis, Georgia Institute of Technology, 1, 2006, pp. 96–126.
[45] S. Razbash and R.J. Hyndman, *Forecasting functions for time series and linear models*, cran.r-project.org, Package forecast, 2018.
[46] K.D. Wall and D.S. Stoffer, *A state space approach to bootstrapping conditional forecasts in ARMA models*, J. Time Ser. Anal. 23 (2002), pp. 733–751.
[47] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, Chapel Hill, NC, 2001. 27599:3175, Unpublished manuscript.
[48] P. Zarchan and H. Musoff, *Fundamentals of Kalman Filtering: A Practical Approach*, 3rd ed., American Institute of Aeronautics and Astronautics Inc, 2009.