U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

INAGBE
GOVERNO DE ANGOLA

Contribuições para o desenho de modelos de previsão da procura: Aplicação no planeamento energético para a cidade de Cabinda

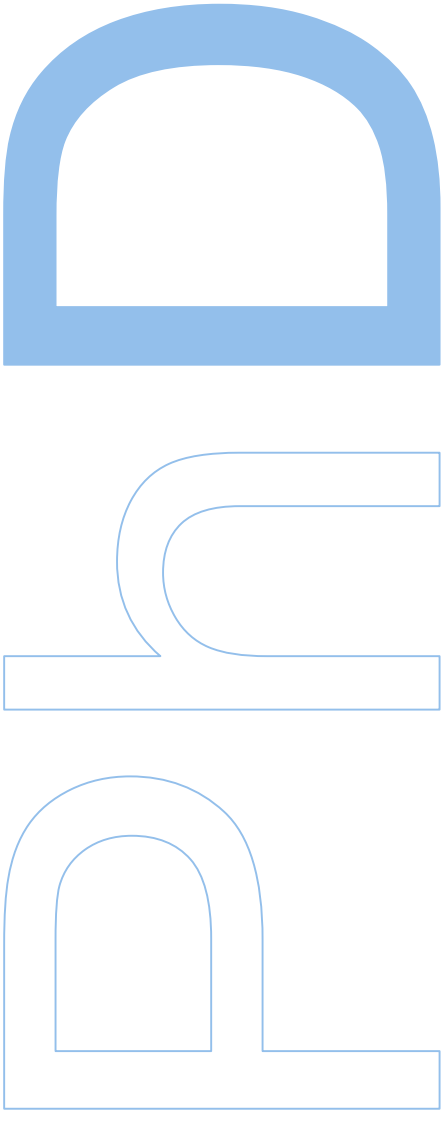
António Casimiro Puindi

Tese de Doutoramento apresentada à
Faculdade de Ciências da Universidade do Porto.
Matemática Aplicada
2018

Contribuições para o desenho de modelos de previsão da procura: Aplicação no planeamento energético para a cidade de Cabinda

António Casimiro Puindi
Doutoramento em Matemática Aplicada
Departamento de Matemática
2018

Orientadora
Prof. Dra Maria Eduarda Silva
Professora da Faculdade de Economia (FEP - UP) e membro do Centro de Investigação e Desenvolvimento da Matemática da Universidade do Porto (CIDMA)



Agradecimentos

À Ti meu Deus, pela Tua bondade e poder. Tu que me ensinaste que nada é impossível, que perante qualquer dificuldade quem acredita no Teu amor, na Tua graça e infinita generosidade encontrará o caminho da superação [*Salmos 37:5*].

À minha orientadora, Prof. Dra. Maria Eduarda Silva, pela sua paciência, generosidade, instruções, comentários construtivos, incentivos e apoio contínuo e incondicional ao longo da formação e de todo o curso dessa tese. A minha gratidão é também extensiva aos professores Sílvio Gama e Slava Razbash e aos companheiros de doutoramento por toda a sua ajuda.

À querida mãe, que muito segurou a minha mão e se dedicou e abdicou seu tempo para que eu tivesse a oportunidade de estudar; à ela devo a honra.

À Cecília, minha amada esposa e nossos filhos, cujo amor por eles proporcionado foi a força que impulsionou-me a superar não poucas as dificuldades enfrentadas durante o doutoramento.

Finalmente, ao Instituto Nacional da Gestão de Bolsas de Estudo de Angola (INAGBE) pela bolsa de estudo concedida.

De todo meu coração, muito obrigado.

ANTÓNIO PUINDI

Declaração do Autor

Eu, *António Casimiro Puindi*, declaro que essa tese intitulada, "Contribuições para o Desenho de Modelos de Previsão da Procura: Aplicação no Planeamento Energético para a Cidade de Cabinda", e a abordagem nela apresentada são da minha autoria. Confirmo ainda que:

- Essa tese foi feita enquanto aluno na Faculdade de Ciências da Universidade do Porto e candidato à pesquisa científica de doutoramento.
- Todas as consultas feitas nos trabalhos publicados, quer sejam artigos, dissertações ou teses dos outros, foram de forma clara e honestamente referenciados.
- Reconheço todas as fontes de ajuda consultadas, e estão todas indicadas.
- Onde essa tese baseou–se no trabalho dos outros, deixei claro exatamente o que foi feito por mim como contribuição e o que os outros fizeram.

Assinado:

_____

Data:

_____

Resumo

Padrões sazonais complexos são um fenómeno comum em vários tipos de séries temporais, incluindo as de consumo de energia elétrica. Tais séries, algumas exibem padrões sazonais múltiplos com frequência alta, outras, mais comumente em dados semanais, possuem padrões com uma frequência não inteira. Outras ainda exibem padrões sazonais com o efeito duplo de calendário. As formulações de modelos estruturais em espaço de estados são bastante potentes para acomodar esse tipo de séries temporais. Mas, quando o assunto é previsão, sabe-se que a informação adicional, que pode melhorar as previsões, pode estar disponível na forma de variáveis de influência externa. Nesse contexto, trabalhos sobre modelos de previsão de séries temporais com sazonalidade complexa e que integram os efeitos das covariáveis são praticamente inexistentes. Uma das razões desse vazio se prende com o facto de, até ao momento, existir apenas um modelo formulado para lidar com séries temporais de sazonalidade complexa, trata-se do modelo TBATS. Este modelo está implementado no pacote forecast do R e é um modelo automático no sentido que não requer interação do utilizador, pelo que a integração das covariáveis é improvável.

Esta tese tem como objetivo contribuir para colmatar esse vazio através da formulação de um quadro de modelos estruturais dinâmicos com a integração dos efeitos das covariáveis. Desse modo constroem-se dois modelos estruturais baseados na formulação de múltiplas fontes de aleatoriedade. O primeiro, designado por SCov, é uma redefinição dos métodos tradicionais de suavização exponencial sazonal simples. O segundo modelo, denominado por TSCov é uma extensão do modelo TBATS formulado para acomodar as séries temporais com sazonalidade complexa. Ambos os modelos são formulados através das três componentes não observáveis: nível, tendência e a sazonalidade que são consideradas aleatórias e variantes no tempo. Para a extração dos sinais da série temporal, introduz-se um filtro de Kalman com matrizes de covariância calculadas recursivamente. Para a estimação dos modelos, um procedimento computacional baseado no filtro de Kalman usando a abordagem de estimativa de máxima verossimilhança é projetado. Trata-se de um procedimento automático que congrega o filtro de Kalman e o método de seleção do número adequado de harmônicas para os termos trigonométricos. A otimização é realizada com base no método de *Newton-Rapson* usando as rotinas de otimização. Este procedimento permite calcular, não só, os estados completos do sistema e as estimativas dos parâmetros, como também permite calcular o erro-padrão de cada estimativa dos parâmetros do modelo.

Os modelos são experimentados com dados reais. Previsões pontuais, intervalares e probabilísticas sob a forma de densidades preditivas são calculadas. O procedimento de reamostragem por *bootstrap* apresentado aprimora o método de previsão baseado nas recursões de filtro de Kalman. Em todos os casos de estudo os resultados obtidos mostram a boa classificação dos modelos e sugerem que o quadro de modelos propostos nesta tese é promissor para futuros estudos. Um estudo específico sobre Cabinda é realizado. O objetivo é estimar um modelo de procura diária de energia elétrica e obter previsões. Ademais, estimar as regiões de maior densidade no consumo e o máximo de carga elétrica consumida. Os resultados obtidos são sa-

tisfatórios, que permitem assegurar o ótimo desempenho do modelo estrutural trigonométrico (TSCov).

Palavras-chave: Bootstrap, Filtro de Kalman, Intervalos de previsão, Modelos estruturais de séries temporais, Regiões de maior densidade, Sazonalidade complexa, Séries temporais.

Abstract

Complex seasonal patterns are a common phenomenon in various types of time series, including those of electricity consumption. Some of these time series exhibit multiple seasonal patterns with high-frequency, some, most commonly weekly series, have patterns with a non-integer period. Other series may have dual-calendar seasonal effects. Formulations of state-space structural models are powerful enough to accommodate such time series. But when the subject is forecasting, it is known that, the additional information which can improve forecasts can be provided by covariates. In this context, works on models with covariates for forecasting of time series with complex seasonal patterns is practically non-existent. One reason for this gap is linked to the fact that, untill now, there is only one automatic formulated model - the TBATS model, dealing with time series with with complex seasonal patterns, such as those time series mentioned above. Being an automatic model, it does not allow to incorporate covariates.

This thesis aims to contribute to fill this gap by formulating a framework of dynamic structural models with covariates. Two structural models are constructed based on the formulation of multiple sources of randomness. The first one designated by SCov is a redefinition of traditional methods of simple exponential smoothing. The second model named by TSCov is an extension of the TBATS model formulated to deal with time series with complex seasonal patterns. Both models are formulated through the three main unobservable components, level, trend and seasonality which are allowed to be random and time varying. 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. For time series signals extraction, a Kalman filter with covariance matrices calculated recursively. For the estimation process, a computational procedure based on the Kalman filter using the maximum likelihood estimation approach is designed. It is an automatic procedure that brings together the Kalman filter and the selecting method of the appropriate number of harmonics for the trigonometric terms. The optimization is performed with Newton-Rapson method using the optimization routines. This procedure allows to calculate, not only, the complete system states and the parameter estimates, but also allows to obtaining the standard errors for the unknown estimate parameters.

The proposed models are empirically explored with real time series. Prediction intervals and probabilistic forecasts in predictive densities form are calculated. The bootstrap procedure presented enhances the forecasting method based on Kalman filter recursions. In all cases tested, the obtained results show the good classification of the models and suggest that the proposed framework in this thesis is promising for future studies. A specific study on Cabinda is carried out. The goal is to estimate a daily electricity demand model and obtainig forecasts. In addition, estimating the highest density regions in consumption and the load maximum consumed. The results obtained are satisfactory, which allow assuring the optimal performance of the Trigonometric Structural model with Covariates (TSCov).

Keywords: Bootstrap, Complex seasonality, Highest density regions, Kalman filter, Prediction intervals, Structural time series models. Time series.

ÍNDICE GERAL

LISTA DE TABELAS

LISTA DE FIGURAS

Lista de Símbolos e Abreviaturas

t	Índice que representa o tempo
$\dot{a}$	Parâmetro de transformação Box-Cox
$\alpha, \beta, \gamma, \rho$	Parâmetros de suavização
$\{y_1, \cdots, y_n\}$	Série temporal
y_t	Valor observado da série temporal inicial no instante t
n	representa o comprimento da série temporal
$\hat{y}_{t+h\|t}$	Previsão no instante t a h passos ($h \geq 1$)
$\{\hat{y}_1^*, \cdots, \hat{y}_n^*\}$	Réplica bootstrap da série inicial
ℓ_t	Nível da série temporal no período t
b	Tendência de longo prazo da série temporal
b_t	Tendência de curto prazo da série temporal no período t
s_t	Componente sazonal da série temporal no período t
ϕ	Parâmetro de transição
$\{\psi_1, \cdots, \psi_p\}$	Parâmetros do processo Autorregressivo (AR)
$\{\theta_1, \cdots, \theta_q\}$	Parâmetros do processo Média Móvel (MA)
m_i	i-ésimo período sazonal existente na série temporal
$\dot{k}_t$	Número de harmônicos para os termos trigonométricos
T	T número de padrões sazonais da série temporal
$\boldsymbol{A}_t$	Matriz do modelo de medição
$\boldsymbol{\Phi}$	Matriz do modelo de transição
$\boldsymbol{\Gamma}$	Matriz de transição das entradas
$\{\beta_1^*, \cdots, \beta_K^*\}$	Coeficientes de regressão
$\boldsymbol{z}_t$	Vetor de covariáveis
$\boldsymbol{x}_t$	Vetor dos estados não observados

$\{\boldsymbol{R}_t, \boldsymbol{Q}_t\}$	Matrizes de covariância associadas às equações de medida e de estado, respetivamente
$\{\sigma^2_{\varepsilon}, \sigma^2_{\xi}, \sigma^2_{\zeta}, \sigma^2_{w}\}$	Variâncias associadas à observação, nível, tendência e sazonalidade, respetivamente, dos modelos estruturais com covariáveis
$\tilde{\boldsymbol{q}}^{(i)}$	Ruído da componente sazonal
$\boldsymbol{\varepsilon}_t$	Inovações a priori de medição
$\boldsymbol{\eta}_t$	Inovações a posteriori do vetor dos estados
$\boldsymbol{\Sigma}_t$	Matriz de covariância das inovações de medição
$\boldsymbol{K}_t$	Ganho do filtro de Kalman
$\boldsymbol{P}_t$	Matriz de covariância de estado
δ	Fator de esquecimento aplicado às covariâncias do sistema
$\boldsymbol{\Omega}$	Vetor dos parâmetros desconhecidos
ϱ	Número de estados estimados
$\boldsymbol{v}^s_t$	Inovações padronizadas do modelo inicial
$\boldsymbol{v}^{*s}_t$	Amostra bootstrap das inovações padronizadas
h	Horizonte de previsão
B	Número de réplicas bootstrap
AIC	Critério de Informação Akaike
SCov	Modelo estrutural básico com covariáveis
TSCov	Modelo estrutural trigonométrico com covariáveis
Boot.TSCov	Modelo TSCov bootstrap
MSE	Erro quadrático médio

Parte I

Introdução e Revisão da Literatura

Capítulo 1

INTRODUÇÃO

1.1 Antecedentes e Justificativa da Pesquisa

A projeção de um modelo em espaço de estados para problemas de previsão de séries temporais tem merecido, até o momento, muita atenção da parte dos pesquisadores, entre os quais, (Harvey, 1989; Harvey and Koopman, 1993; Ord et al., 1997; Hyndman et al., 2002; Ord et al., 2005; Hyndman et al., 2008; Durbin and S.J.Koopman, 2011; Koehler et al., 2012; Ahmad and Maxwell, 2015; Shumway and Stoffer, 2017). Essa atenção está justificada pelo facto desses modelos serem flexíveis para a classe de modelos de suavização exponencial, Hyndman et al. (2008), e para incorporar os efeitos das covariáveis, (Wang, 2006; Dordonnat et al., 2008; Gob et al., 2013; Shumway and Stoffer, 2017), assim como para incluir a metodologias de reamostragem, como a metodologia *bootstrap*, (Stoffer and Wall, 2004; Menezes et al., 2006; Rodriguez and Ruiz, 2009; Cordeiro and Neves, 2011; Bergmeir et al., 2015; Shumway and Stoffer, 2017).

No quadro da formulação de modelos estruturais em espaço de estados para séries temporais sazonais, a literatura existente permite distinguir dois tipos: (i) modelos para séries temporais sazonais que não integram os efeitos das covariáveis; (ii) modelos para séries temporais sazonais que integram os efeitos das covariáveis. Do tipo (i), os modelos mais comumente empregados na formulação de espaço de estados incluem aqueles que subjazem os conhecidos métodos aditivos e multiplicativos de *Holt-Winters*, Taylor (2003); Hyndman et al. (2008); dos quais, Taylor and R.Buizza (2003) estenderam a versão linear do método *Holt-Winters* para incorporar uma segunda componente sazonal. Outras formulações foram propostas para lidar com a sazonalidade múltipla em séries temporais, por exemplo, as abordagens de Pedregal and Young (2006) e Harvey and Koopman (1993). As formulações propostas por estes permitem acomodar séries temporais com sazonalidade dupla, mas não são suficientemente estruturadas para acomodar as séries temporais com mais de dois padrões sazonais; igualmente não são capazes de acomodar a não-linearidade encontrada em muitas séries temporais na prática. As formulações propostas por Taylor (2003); Gould et al. (2008); Taylor (2010); Taylor and Snyder (2012), tentam lidar com os padrões sazonais complexos, mas, também sofrem de várias fraquezas, como a sobre-parametrização e a incapacidade de acomodar tanto o período não inteiro quanto os efeitos duplos de calendário. Para o tipo (ii), as propostas desses modelos podem ser vistas em Brockwell and Davis (2002) com o SARIMA (*Seasonal Auto-Regressive Integrated Moving Average*) e em Wang (2006); Dordonnat et al. (2008);

Koehler et al. (2012); Gob et al. (2013); Ahmad and Maxwell (2015). Apesar de integrarem os efeitos das covariáveis na previsão e tentarem lidar com os padrões sazonais complexos, também são incapazes de lidar tanto com as séries temporais com período não inteiro, com frequência alta quanto às séries temporais com o efeito duplo de calendário.

Todos os modelos referenciados acima, têm como foco a previsão. São úteis em muitos campos, no entanto, todos têm a limitação que se prende com a capacidade de lidar com as séries temporais com frequência alta, sazonalidade não-inteira e efeito duplo de calendário. Para resolver essas limitações e melhorar os métodos tradicionais de suavização exponencial sazonal, De Livera et al. (2011) introduziu duas estruturas, BATS (como acrônimo para os principais recursos do modelo: *Box-Cox transform, ARMA errors, Trend, and Seasonal components*) e TBATS (com o T inicial conotando *Trigonometric*), apresentadas em (1.1) e (1.2) e adota o método de suavização exponencial para o processo de estimação dos parâmetros dos modelos.

1.2 Objetivo

O principal objetivo dessa tese é fornecer contribuições inerentes a modelos de previsão de séries temporais, em especial para séries temporais com sazonalidade complexa[1], Figuras 1.1, 1.2, 1.3.

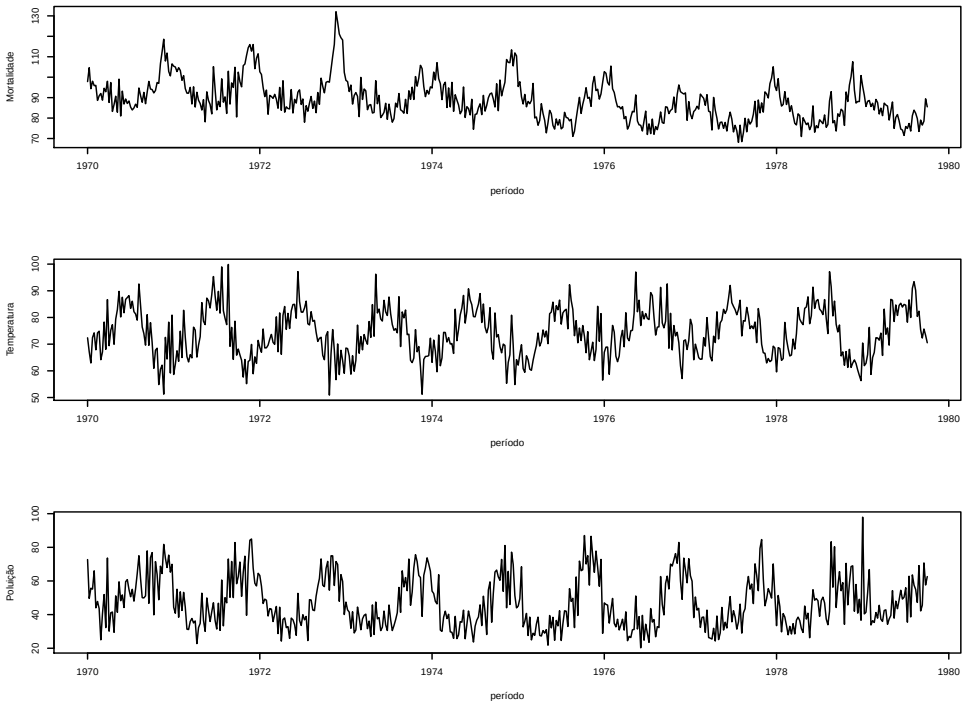


Figura 1.1: Dados semanais sobre mortalidade cardiovascular por temperatura e poluição em Los Angels–Estados Unidos de América entre 1970-1979 (Stoffer, 2016).

Para concretizar o nosso objetivo:

[1]São séries temporais com sazonalidade complexa, àquelas séries temporais com vários períodos sazonais, com sazonalidade de frequência alta, sazonalidade de frequência não-inteira e sazonalidade com efeito duplo de calendário. A descrição mais objetiva dessas séries pode ser encontrada em De Livera (2010), ver também apêndice A.

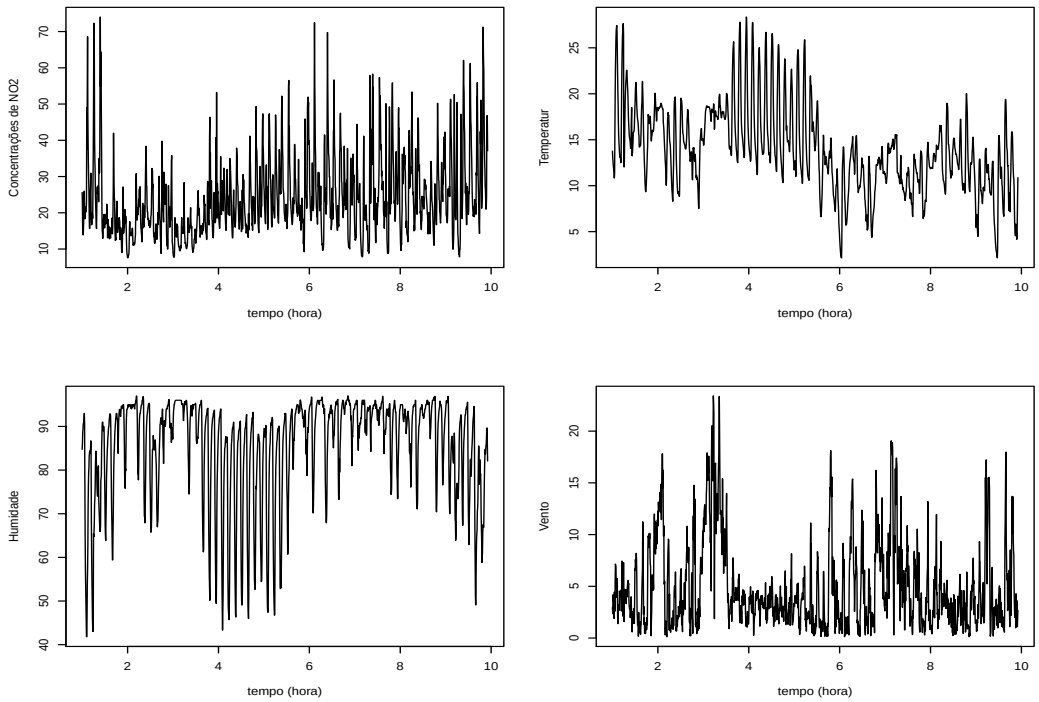
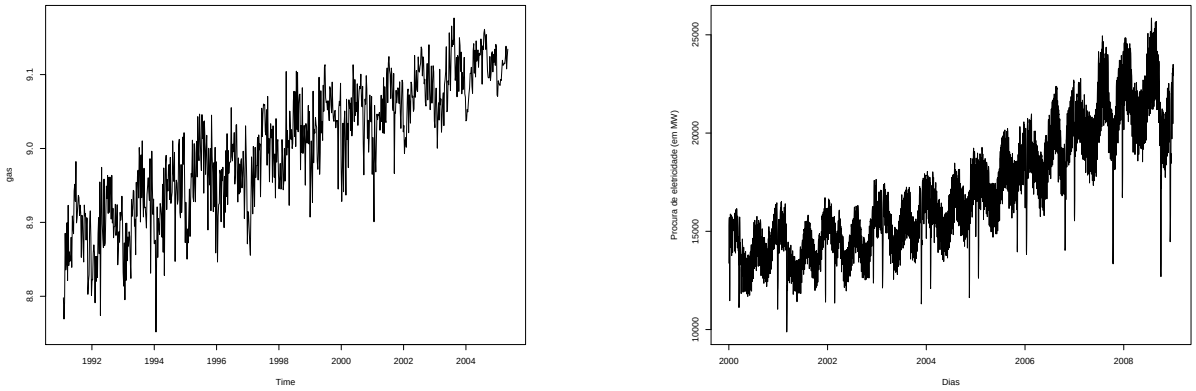


Figura 1.2: Dados horários de NO_2 observados em 2014 entre 1 de Outubro e 31 de Dezembro na estação de Paredes, Portugal. A temperatura, Humidade e Vento são as covariáveis, observadas igualmente de hora em hora (QualAr, 2015).



(a) (b)

Figura 1.3: (a) Dados sobre produção de gasolina a motor dos EUA (milhares de barris por dia), entre Fevereiro de 1991 e Julho de 2005; (b) Dados da procura diária de eletricidade na Turquia a partir de 01 de Janeiro de 2000 até 31 de Dezembro de 2008. Esse conjunto de dados foi utilizado também no trabalho de De Livera et al. (2011) e pode ser encontrado na página web do professor *Rob J. Hyndman*: https://robjhyndman.com/publications/complex-seasonality/

- Explora-se os efeitos das covariáveis na previsão e desenvolve-se um quadro de modelos estruturais com efeitos das covariáveis que descrevem adequadamente esse tipo de dados e calculam não só as previsões pontuais, como também os intervalos de previsão. A escolha do quadro de modelos estruturais justifica-se pelo fato de serem modelos muito

fáceis de generalizar, por exemplo, adicionar covariáveis. Ademais, quando o assunto é lidar com valores omissos, os modelos estruturais comportam-se melhor;

- No âmbito da estimação, constrói-se um procedimento baseado nas inovações (a priori e a posteriori) que permite a correta configuração das matrizes covariância dos ruídos do sistema projetado no filtro de Kalman. A adoção do filtro de Kalman justifica-se pelas seguintes razões:

  - Proporciona um conjunto de equações muito gerais para qualquer modelo no formato de espaço de estados;

  - Proporciona estimadores ótimos no sentido do menor MSE (*Mean Squared Error*);

  - É fácil de lidar com valores faltantes;

  - A verossimilhança é calculada com facilidade.

- Combina-se o método *bootstrap* e o filtro de Kalman para previsão de curto prazo.

1.3 Motivação

Do ponto de vista social, a motivação desse estudo é dupla: (1) os recentes projetos no setor energético implementados pelo Governo Angolano, fornecem uma boa justificativa para se começar com pesquisas sobre previsão da procura de energia elétrica, pois, não existem até ao momento trabalhos científicos nesse âmbito com foco às cidades angolanas. (2) Angola vive uma era de políticas de projetos de inovação, dentre elas, disposição de fontes de energia eléctrica fiáveis e de qualidade, centrais térmicas e barragens, em quantidade suficiente para cobrir a procura existente. Os contínuos défices registados no fornecimento de energia elétrica nas cidades Angolanas, em particular Cabinda, provocados pela falta de capacidade de resposta à crescente procura da carga elétrica, impõem à necessidade de compreensão da dinâmica dessa procura. Porém, a quantificação das previsões da variabilidade da procura de energia elétrica é um processo complexo, devido a vários fatores que influenciam no comportamento do consumidor. A par disso, TRÊS questões de interesse local precisam-se responder:

1. Como modelar e prever o comportamento da procura de energia elétrica na cidade Cabinda?

2. Em quais regiões do espaço amostral a densidade de consumo de energia elétrica é alta?

3. Como estimar o fluxo máximo de carga elétrica consumida num dado período?

A motivação matemática para esse estudo resulta, primeiro, das formulações dos modelos TBATS[2], De Livera et al. (2011), que incluem uma transformação Box-Cox, erros ARMA e T padrões sazonais, conforme apresentado a seguir:

$$y_t^{(\dot{a})} = \begin{cases} \frac{y_t^{\dot{a}}-1}{\dot{a}} & \text{se } \dot{a} \neq 0 \\ \log y_t & \text{se } \dot{a} = 0 \end{cases} \tag{1.1a}$$

$$y_t^{(\dot{a})} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t \tag{1.1b}$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha d_t \tag{1.1c}$$

$$b_t = (1-\phi)b + \phi b_{t-1} + \beta d_t \tag{1.1d}$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t \tag{1.1e}$$

$$d_t = \sum_{i=1}^{p} \varphi_i d_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{1.1f}$$

onde y_t representa a série temporal observada no instante t e $\dot{a}$ é o parâmetro de transformação Box-Cox; $i = 1, \cdots, T$ (T número de padrões sazonais da série temporal); $m_1, \cdots, m_T$ denotam os períodos sazonais da série; ℓ_t e b_t são o nível estocástico local e a tendência de curto prazo, respetivamente; b é a tendência de longo prazo e d_t é um processo ARMA(p,q) com ruído branco Gaussiano de média zero e variância constante σ^2; $s_t^{(i)}$ representa a i-ésima componente sazonal no instante t. Os parâmetros de suavização são dados por α, β e γ_i. Os dois modelos, BATS e TBATS estão acoplados com o erro de fonte única d_t. Ademais, De Livera et al. (2011) introduziu a representação trigonométrica para componente sazonal baseada em séries de Fourier:

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)} \tag{1.2a}$$

$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t \tag{1.2b}$$

$$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t \tag{1.2c}$$

onde, $\lambda_j^{(i)} = \frac{2\pi j}{m_i}$ ($j = 1, 2, \cdots, k_i$) sendo k_i o número de harmônicas necessário para os termos trigonométricos na i–ésima componente sazonal, onde $\gamma_1^{(i)}$ e $\gamma_2^{(i)}$ são parâmetros de suavização, com $i = 1, \cdots, T$. Dessa forma, o modelo TBATS é definido substituindo (1.2) em (1.1e) e a equação de medida (1.1b) por

$$y_t^{(\dot{a})} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + d_t \tag{1.3a}$$

A formulação trigonométrica do seu modelo sazonal permite dar ao modelo TBATS mais flexibilidade para lidar com os padrões sazonais complexos presentes na série temporal. É um

[2] A existência de algumas séries temporais exibindo padrões sazonais complexos deu lugar a formulação dos modelos tbats – junção de dois modelos: bats e tbats. O bats é a generalização dos modelos sazonais tradicionais de suavização exponencial projetado para lidar com as séries temporais com múltiplos períodos sazonais. Ao ser incapaz de lidar com séries temporais de sazonalidade não-inteira, é introduzido o modelo tbats através da representação trigonométrica de componentes sazonais baseada em séries de Fourier.

modelo automático e útil para previsão de séries temporais com padrões sazonais complexos. Dentre as vantagens que apresenta, destacam-se:

1. A capacidade de acomodar dados com períodos sazonais não-inteiros, dados com frequência alta e dados com efeitos duplo de calendário, incluindo os padrões sazonais múltiplos e simples;
2. A transformação Box-cox pode lhe permitir lidar com a não-linearidade dos dados;
3. O processo ARMA sobre os resíduos pode resolver o problema da autocorrelação;
4. Pode obter não apenas a previsão pontual, como também os intervalos de previsão.

Apesar dessas vantagens, também sofre da fraqueza que se prende, fundamentalmente, com a integração dos efeitos das covariáveis na previsão. Em problemas de previsão, sabe-se que a informação adicional pode estar disponível na forma de variáveis de entrada (variáveis de influência externa). No entanto,

- Os modelos TBATS são totalmente automáticos, a inclusão de covariáveis nesses modelos é improvável[3]. Ademais,
- A suposição de $\varepsilon_t \sim iidN(0, \sigma^2)$ pode não ser válida. Pois, quando $p = q = 0$ os modelos TBATS são equivalentes aos modelos de suavização exponencial aditivos denominados TETS (*Trigonometric ExponenTial Smoothing*).

A segunda motivação se prende com a aplicação dos métodos de reamostragem na previsão; concretamente o método *bootstrap*. Até o momento há um vazio na literatura existente sobre modelos de previsão para séries temporais com sazonalidade complexa aplicando o método *bootstrap*.

Motivados por esses impasses, um quadro de modelos estruturais que integra os efeitos das covariáveis e que tenha a vantagem da formulação dinâmica para lidar com as séries temporais de sazonalidade complexa é apresentado. Admite-se que o modelo estrutural pode incluir a transformação Box–Cox.

A nossa abordagem difere da abordagem TBATS em quatro aspectos, fundamentalmente:

1. A estrutura do modelo é baseada na formulação de múltiplas fontes de aleatoriedade;
2. A estrutura do modelo não integra parâmetros de suavização;
3. A estrutura do modelo lida com covariáveis, mas, pode funcionar sem a integração das covariáveis;
4. O procedimento de estimação dos parâmetros é baseado no filtro de Kalman com as matrizes de covariância do sistema calculadas recursivamente, conforme as subsecções 1.3.1 e 1.4. Um processo automático que agrega o filtro de Kalman e o método de seleção dos harmônicos necessários para os termos trigonométricos é construído para a otimização das estimativas dos parâmetros através do método de Newton-Rapson.

[3]São modelos implementados no pacote de Hyndman and Khandakar (2008), e as funções genéricas bats() e tbats() que permitem executar os modelos bats e tbats não disponibilizam algum parâmetro que permite o usuário incorporar as covariáveis.

1.4 Estrutura da Tese

A primeira parte da tese que inclui a introdução também apresenta a revisão da literatura que se subdivide em dois Capítulos:

Capítulo 2 revisa os modelos básicos de séries temporais, fundamentalmente os modelos lineares.

Capítulo 3 tem como foco os conceitos fundamentais sobre modelos estruturais de séries temporais e modelos lineares dinâmicos convencionais. Um olhar para as diferentes estratégias de estimação de modelos em espaço de estados, critérios de seleção dos modelos e avaliação do desempenho dos modelos são dentre outros aspetos também revisados nesse Capítulo.

A segunda parte é reservada para as contribuições da tese.

Capítulo 1 introduz o quadro dos modelos estruturais com a integração das covariáveis e sua descrição em espaço de estados. O primeiro modelo é designado por SCov como acrônimo de *Structural Models with Covariates*. O segundo modelo é designado por TSCov como iniciais de *Trigonometric Structural Models with Covariates* é uma extensão do primeiro. Este segundo integra o modelo sazonal trigonométrico no quadro da formulação de múltiplas fontes de aleatoriedade. Esse capítulo é também dedicado à construção do procedimento computacional de estimação dos parâmetros do modelo, que é baseado no filtro de Kalman com ajusto recursivo das matrizes de covariâncias do sistema.

Capítulo 2 apresenta os resultados da análise empírica dos modelos propostos. Discute-se o procedimento de inicialização dos parâmetros assim como a capacidade preditiva dos modelos estimados. Previsões pontuais e intervalares são calculadas. A adequação dos modelos ajustados é avaliada usando ferramentas para diagnóstico e validação de modelos com base na análise de resíduos.

Capítulo 3 aborda a implementação da metodologia de reamostragem por *bootstrap* na previsão dos modelos estruturais desenvolvidos. Uma breve resenha sobre *bootstrap* em modelos de espaço de estado é apresentado. Em seguida, desenvolve-se o procedimento geral que é descrito em etapas de implementação. Os delineamentos computacionais associados ao algoritmo *bootstrap* são igualmente apresentados. A bordagem do capítulo termina com alguns experimentos usando dados reais.

Capítulo 4 é dedicado à modelação e previsão da procura diária de eletricidade na cidade de Cabinda. Com a aplicação do modelo TSCov, previsões pontuais e probabilísticas sob a forma de densidades preditivas são calculadas.

A terceira parte desta tese enumera as principais contribuições e resultados da tese e alguns desafios que requerem maior investigação para o futuro. Esses aspetos constituem o Capítulo 1 da tese.

Finalmente, a quarta parte tem a ver com os apêndices e a bibliografia consultada.

Todas as partes deste trabalho que apresentam resultados do tratamento computacional, foram conseguidos mediante o uso do programa R Core Team (2017). As razões do seu uso resultam no facto de ser uma linguagem criada pelos estatísticos para estatísticos, e não uma linguagem de informáticos para estatísticos.

Capítulo 2

MODELOS DE SÉRIES TEMPORAIS: CONCEITOS BÁSICOS

Índice do Capítulo

2.1 Processos Estocásticos

Uma série temporal $\{y_1, \cdots, y_n\}$ é um conjunto de observações, sendo cada uma observação registada cronologicamente num tempo específico t (Brockwell and Davis, 2002). Dado esse conjunto de observações, o objetivo é analisar e conhecer o comportamento dos fenômenos associados à série observada e fundamentalmente prever ocorrências futuras.

Um processo estocástico é um fenômeno estatístico que evolui no tempo de acordo com as leis probabilísticas. A ideia básica do processo estocástico é extensão do conceito de variável aleatória. Se, por exemplo, considerarmos a temperatura T_t de uma cidade no período diurno, é óbvio que T_t é uma variável aleatória, pois toma valores diferentes a cada hora ou a cada dia de observação. A obtenção das estatísticas completas de T_t, implica observar e armazenar os valores de T_t durante vários dias, e a partir desses dados podemos determinar as funções densidade e de distribuição de probabilidade da variável aleatória T_t. Ademais, a temperatura é também função do tempo. Num dado período temporal, a temperatura pode ter uma distribuição diferente daquela obtida para o outro período temporal.

Dado um espaço de probabilidade $(\Xi, \mathscr{A}, \mathscr{P})$, um espaço mensurável (E, ϵ) e um certo conjunto de índices $\mathscr{I}$,

Definição 1 *Um processo estocástico é uma coleção de variáveis aleatórias* $\{\boldsymbol{X}_t : t \in \mathscr{I}\}$ *definidos sobre* $(\Xi, \mathscr{A}, \mathscr{P})$ *e com valores em* (E, ϵ)*; podendo ser representado por*$\{\boldsymbol{X}(t, \zeta_i) : \zeta_i \in \Xi, t \in \mathscr{I}\}$ *ou como uma aplicação* $\boldsymbol{X} : \mathscr{I} \times \Xi \to \epsilon$ *(Brockwell and Davis, 2002; Cordeiro and Neves, 2011; Douc et al., 2014).*

Sobre o exemplo de temperatura, precisaríamos armazenar temperaturas diárias, por exemplo, para cada valor de t (cada hora do dia) durante um período determinado. Este processo fornece-nos uma forma de onda $\{\boldsymbol{X}(t, \zeta_i)\}$, tal como indicada na Figura (2.1), onde i indica o dia em que foi feita a observação.

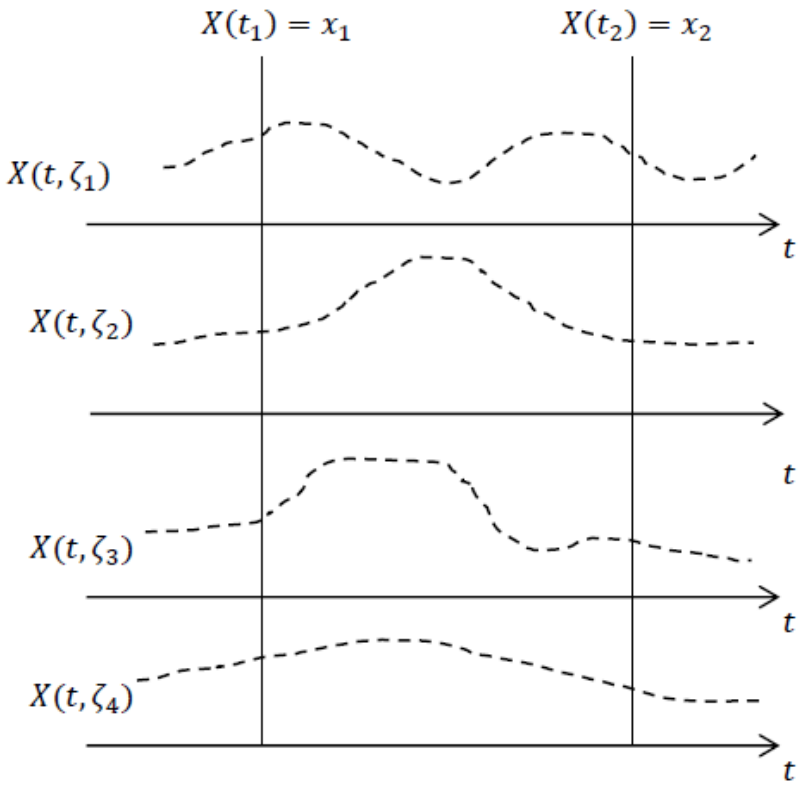


Figura 2.1: Exemplo de um processo estocástico que representa a temperatura de uma cidade.

Para todo $\zeta_i \in \Xi$ fixo, $\boldsymbol{X}(t, \zeta_i)$ é uma função de parâmetro t com domínio em $\mathscr{I}$ ($\mathscr{I} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$), a qual designamos por trajetória ou realização do processo e dizemos tratar-se de um processo estocástico de tempo discreto. No caso de $\mathscr{I} = R$ ou $\mathscr{I} = [0, +\infty[$, dizemos que é um processo estocástico de tempo contínuo.

Considerando $\{t_1, t_2, \ldots, t_n\} \subset \mathscr{I}$ e sendo $(\boldsymbol{X}_{t_1}, \ldots, \boldsymbol{X}_{t_n})$ um conjunto finito de n variáveis aleatórias de um processo estocástico discreto, com observações em intervalos regulares, uma maneira de caracterizar o processo estocástico é através da distribuição de probabilidade conjunta das suas n variáveis aleatórias $\boldsymbol{X}_t$. Mas essa maneira de caracterizar o processo estocástico é bastante complicada e de certo modo não viável na prática.

Usualmente define-se o processo estocástico $\{\boldsymbol{X}_t : t = 0, \pm 1, \pm 2, \pm 3, \cdots\}$ como função de variáveis aleatórias cuja distribuição de probabilidade é conhecida. Outra forma de caracterizar um processo estocástico é a partir das funções determinísticas a ele associadas que são particularmente importantes para a descrição e caracterização do comportamento do processo, que são os momentos do processo, em especial o primeiro e o segundo momentos, que designamos por *média*, *variância* e *covariância*, respectivamente (Jonathan and Chan, 2008;

Cordeiro and Neves, 2011; Shumway and Stoffer, 2011; Douc et al., 2014). Assim, para um processo estocástico $\{\boldsymbol{X}_t : t = 0, \pm1, \pm2, \pm3, \ldots\}$, a *função média* é definida por

$$\mu_t \quad = \quad E[\boldsymbol{X}_t] \qquad para \quad t = 0, \pm1, \pm2, \pm3, \ldots \tag{2.1}$$

Nesse caso, μ_t é apenas o valor esperado do processo no instante t. Em geral, μ_t pode tomar valores diferentes em cada instante de tempo t. A *função auto-covariância* $\gamma_{t,s}$ é definida como

$$\gamma_{t,s} \quad = \quad Cov(\boldsymbol{X}_t, \boldsymbol{X}_s) \qquad para \quad t, s = 0, \pm1, \pm2, \pm3, \ldots \tag{2.2}$$

onde $Cov(\boldsymbol{X}_t, \boldsymbol{X}_s) = E[(\boldsymbol{X}_t - \mu_t)(\boldsymbol{X}_s - \mu_s)] = E(\boldsymbol{X}_t\boldsymbol{X}_s) - \mu_t\mu_s$. A *função de auto-correlação*, $\rho_{t,s}$, do processo é dada por

$$\rho_{t,s} \quad = \quad Corr(\boldsymbol{X}_t, \boldsymbol{X}_s) \qquad para \quad t, s = 0, \pm1, \pm2, \pm3, \ldots \tag{2.3}$$

onde

$$Corr(\boldsymbol{X}_t, \boldsymbol{X}_s) \quad = \quad \frac{Cov(\boldsymbol{X}_t, \boldsymbol{X}_s)}{\sqrt{Var(\boldsymbol{X}_t)Var(\boldsymbol{X}_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}\gamma_{s,s}}} \tag{2.4}$$

2.2 Estacionaridade

Da inferência estatística sobre a estrutura de um processo estocástico baseada nos valores observados, usualmente faz-se algumas suposições, provavelmente razoáveis, acerca dessa mesma estrutura. A mais importante dessas suposições é a estacionaridade. A ideia básica da estacionaridade é que a lei de probabilidade que governa o comportamento do processo não muda ao longo do tempo. De acordo com Jonathan and Chan (2008),

Definição 2 *Um processo estocástico $\{\boldsymbol{X}_t\}$ é dito ser estritamente estacionário se a distribuição conjunta de $\boldsymbol{X}_{t_1}, \boldsymbol{X}_{t_2}, \ldots, \boldsymbol{X}_{t_n}$ é a mesma com a distribuição conjunta de $\boldsymbol{X}_{t_1-k}, \boldsymbol{X}_{t_2-k}, \ldots, \boldsymbol{X}_{t_n-k}$ para todos os pontos temporais $t_1, t_2, \ldots t_n$ escolhidos e todos os lag k temporais escolhidos.*

Na prática, é difícil avaliar a estacionaridade estrita dos dados. No entanto, é usual aplicar a definição mais branda que impõe condições apenas nos dois primeiros momentos da série temporal, a denominada estacionaridade fraça.

Definição 3 *Um processo estocástico $\{\boldsymbol{X}_t\}$ é dito ser fracamente (segunda ordem) estacionário se:*

(i) $E(\boldsymbol{X}_t) = \mu(t) = \mu$ *é constante ao longo do tempo,*

(ii) $E(\boldsymbol{X}^2) < \infty$ *para todo t,*

(ii) $\gamma_x(s,t) = cov(\boldsymbol{X}_s, \boldsymbol{X}_t) = E[(\boldsymbol{X}_s - \mu_s)(\boldsymbol{X}_t - \mu_t)]$ *é uma função de* $|s - t|$.

2.3 Modelos Lineares para Séries Temporais Estacionárias

Um importante exemplo de processo estacionário é o chamado processo Ruído branco, $\{w_t : t = 1, 2, 3, \cdots, n\}$, o qual é definido como a sequência de variáveis aleatórias independentes e identicamente distribuídas; o que implica que todas as variáveis tenham a mesma

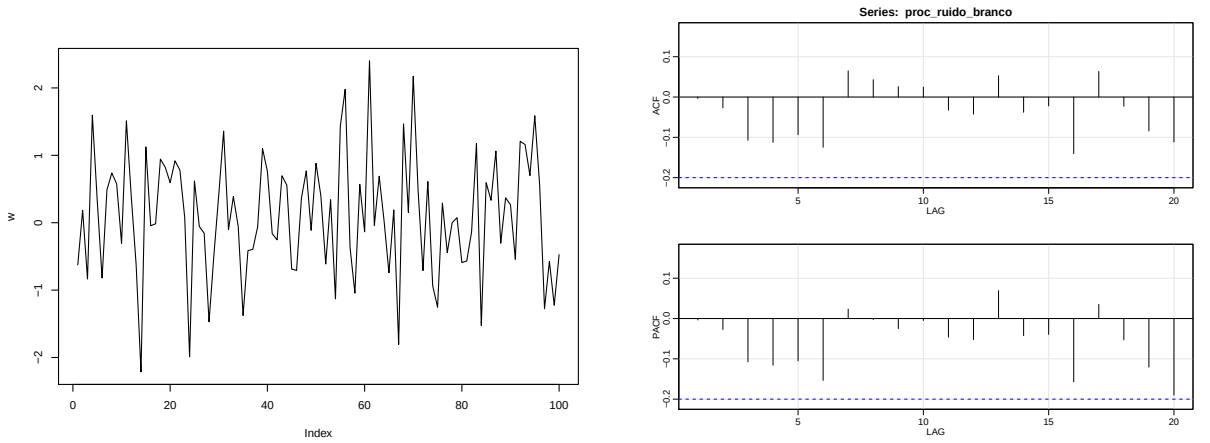


Figura 2.2: Simulação de um ruído branco gaussiano e o seu respetivo correlograma

variância, σ^2 e $Cor(w_i, w_j) = 0$ para todo $i \neq j$. Adicionalmente, se as variáveis seguem uma distribuição normal, isto é, $w_t \sim N(0, \sigma^2)$, então o processo $\{w_t\}$ é chamado de *ruido branco Gaussiano*, Figura 2.2, (Jonathan and Chan, 2008; Cowpertwait and Metcalfe, 2009; Douc et al., 2014). Este processo é muitas vezes designado de ruído branco, que por vezes é útil para a construção de processos estocásticos mais complexos. No caso de um processo discreto, o processo é chamado de *puramente* aleatório se este consiste de uma sequência de variáveis aleatórias $\{w_t\}$ independentes e identicamente distribuídas, com as seguintes propriedades:

1. $E(w_t) = E(w_t|w_{t-1}, w_{t-2}, \cdots) = \mu$

2. $Var(w_t) = Var(w_t|w_{t-1}, w_{t-2}, \cdots) = \sigma_w^2$

3. $\gamma(k) = Cov(w_t, w_{t+k}) = 0, \quad k = \pm 1, \pm 2, \cdots$

Como a média e a função de auto-covariância não dependem do tempo, o processo é estacionário em segunda ordem. Assim, a função de auto-correlação é, portanto, definida por

$$\rho(k) = \begin{cases} 1 & , k = 0 \\ 0 & , k = \pm 1, \pm 2, \cdots \end{cases}$$

Já para os processos estocásticos auto-regressivos, que se caracterizam no vasto número de coeficientes de auto-correlação distintos de zero e decrescendo com o retardo, são processos com memória relativamente longa, já que o valor atual da série é correlacionado com todos os valores anteriores. Todavia, esses processos podem representar séries de memória muito curta, onde o valor atual da série somente está correlacionado com um número pequeno de valores anteriores, de maneira que a função de auto-correlação simples tenha apenas poucas auto-correlações distintas de zero. A família de processos que tem esta propriedade de memória muito curta são os processos de Média Móvel, designados por simplicidade *MA* de seu nome em inglês, *Moving Average*.

2.3.1 Processos Média Móvel (MA)

Seja w_t um processo ruido branco. De acordo com Brockwell and Davis (2002); Jonathan and Chan (2008); Cowpertwait and Metcalfe (2009),

Definição 4 *Um processo* $\{\boldsymbol{X}_t\}$ *é média móvel de ordem* q*, ou* $MA(q)$*, se*

$$\boldsymbol{X}_t = w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{q-1}, \tag{2.5}$$

sendo $\theta_i \in \mathscr{R}, i = 1, \ldots, q$*. A média e a variância do processo são definidas por*

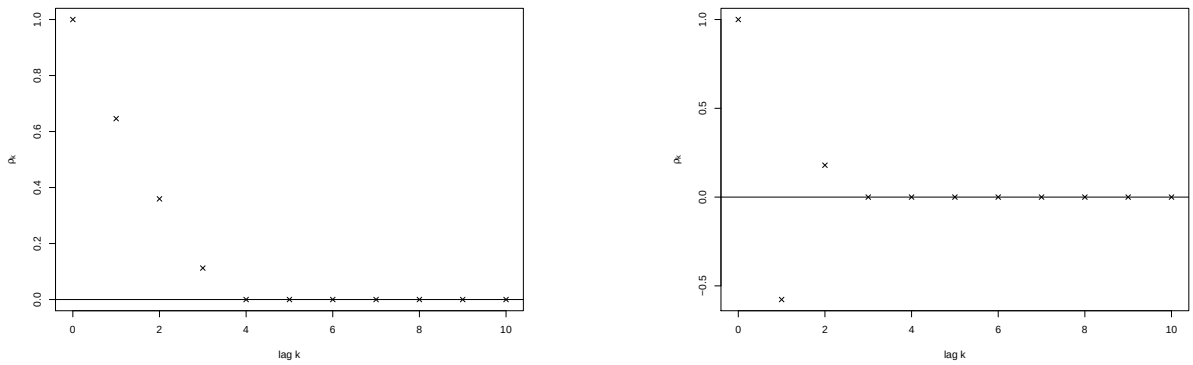
$$\begin{aligned} E(\boldsymbol{X}_t) &= E(w_t) + \sum_{j=1}^{q} \theta_j E(w_{t-j}) = 0 \\ Var(\boldsymbol{X}_t) &= Var(w_t) + \sum_{j=1}^{q} \theta_j^2 Var(w_{t-j}) = (1 + \theta_1^2 + \ldots + \theta_q^2)\sigma_w^2. \end{aligned}$$

Como a média e a variância são constantes, a função de auto-correlação para $k \geq 0$*, é dada por*

$$\rho(k) = \begin{cases} 1 & , k = 0 \\ \sum_{i=0}^{q-k} \theta_i \theta_{i+k} / \sum_{i=0}^{q} \theta_i^2, \qquad k = 1, \cdots, q & \\ 0 & , k > q \end{cases} \tag{2.6}$$

onde β_0 *é* 1.

A função é zero quando $K > q$ porque $\boldsymbol{X}_t$ e $\boldsymbol{X}_{t+k}$ consistem das somas dos termos de ruido branco independentes e têm covariância zero. Ademais, tem um ponto de corte no *lag* q, isto é, $\rho(k) = 0$ para $k > q$, Figura 2.3. Esta é uma característica específica de processos médias móveis e torna-se útil na especificação do valor de q na prática. Os gráficos da Figura 2.3 são igualmente as funções de autocorrelação para os dois processos $MA(3)$: o primeiro com correlações positivas nos *lag* 1 e 3, o segundo com correlações negativas nos *lag* 1 e 3. A Figura 2.4 mostra o processo média móvel de ordem $q = 1$.



(a) (b)

Figura 2.3: Funções de autocorrelação para dois processos $MA(3)$: (a) $\theta_1 = 0.7, \theta_2 = 0.5, \theta_3 = 0.2$; (b) $\theta_1 = -0.7, \theta_2 = 0.5, \theta_3 = -0.2$ (Cowpertwait and Metcalfe, 2009).

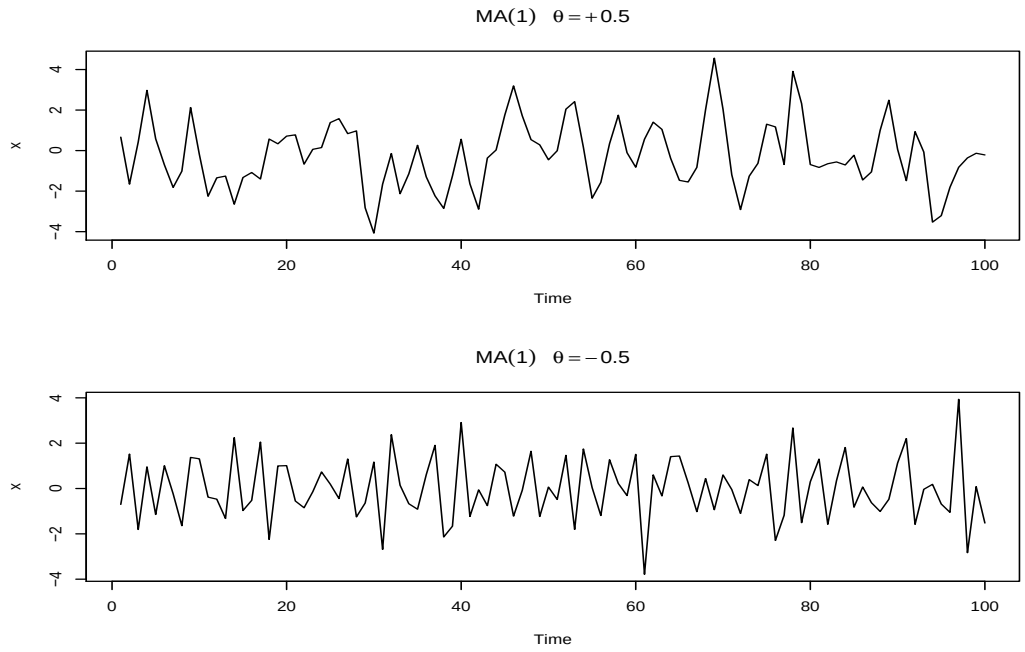


Figura 2.4: Simulação de um modelo MA(1): $\theta = .9$ (primeiro painel); $\theta = -.9$ (segundo painel).

2.3.2 Modelos Autorregressivos

Modelos autorregressivos são baseados na ideia de que o valor atual da série temporal, $\{\boldsymbol{X}_t\}$, pode ser explicado por p valores passados, $X_{t-1}, X_{t-2}, \cdots, X_{t-p}$. São modelos que resultam da imposição da dependência linear entre as variáveis do processo, similar a uma equação de regressão. Podem ser usados como modelos, sendo razoável, assumir que o valor atual da série temporal depende do seu passado imediato mais um erro aleatório, conforme a definição 5.

Definição 5 *Um processo estacionário $\{\boldsymbol{X}_t\}$ é autorregressivo de ordem p, usualmente denotado por $AR(p)$, se para todo $t = 0, \pm 1, \pm 2, \ldots$,*

$$\boldsymbol{X}_t \quad = \quad \psi_1 X_{t-1} + \psi_2 X_{t-2} + \ldots + \psi_p X_{t-p} + w_t, \tag{2.7}$$

onde w_t é um processo de ruído branco e ψ_i são constantes reais. Utilizando o operador de atraso, L, a equação (2.7) pode escrever-se por

$$\psi_p(L)\boldsymbol{X}_t \quad = \quad w_t, \tag{2.8}$$

onde $\psi_p(L) = 1 - \psi_1 L - \psi_2 L^2 - \ldots - \psi_p L^p$ (chamado polinômio autorregressivo). A Figura 2.5 mostra o processo autorregressivo de ordem $p = 1$.

$\psi_p(L)$ em (2.8), é chamado de polinômio característico do processo e as suas raízes determinam quando o processo é estacionário ou não; ou seja, todas as suas raízes devem exceder a unidade em valor absoluto para que o processo seja estacionário. Por exemplo,

- O processo AR(1) dado por $x_t = 0.5x_{t-1} + w_t$ é estacionário, porque a raíz de $1 - 0.5L = 0$ é $L = 2$, que é maior que 1;

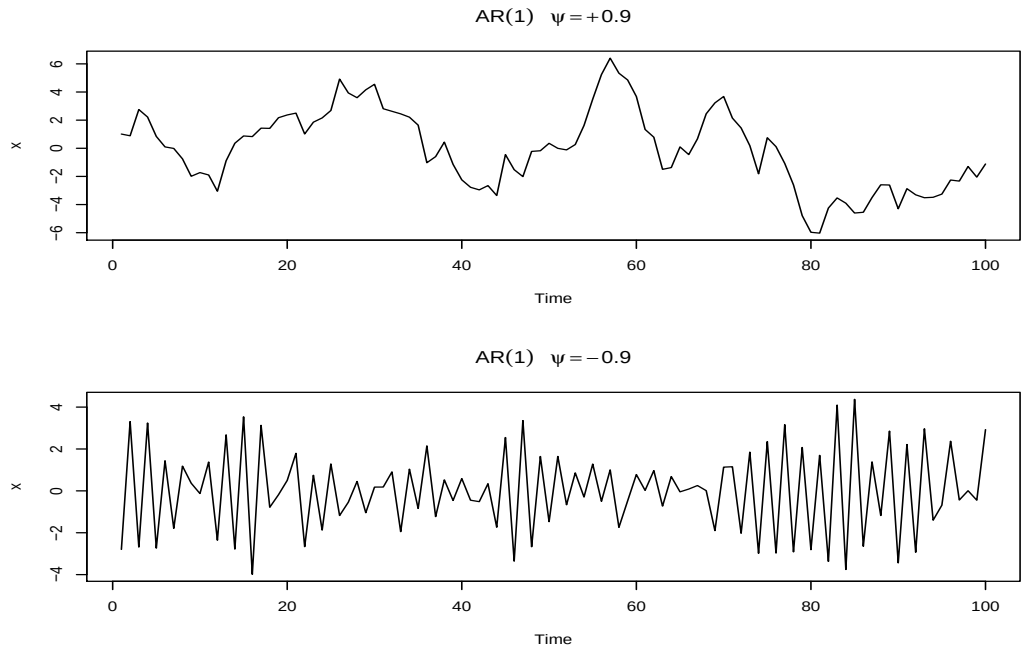


Figura 2.5: Simulação de um modelo Ar(1): $\psi = .9$ (primeiro painel); $\psi = -.9$ (segundo painel).

- O processo AR(2) dado por $x_t = 0.5x_{t-1} + 0.5x_{t-2} + w_t$ é não estacionário, pois, uma das raízes é unidade. Prova: ao expressar o modelo como $-0.5(L^2 + L - 2)x_t = w_t \Leftrightarrow -0.5(L-1)(L+2)x_t = w_t$, o polinômio $\psi(L) = -0.5(L-1)(L+2)$ tem raízes $L = 1$ e -2. Como há uma raiz unitária $L = 1$, o processo não é estacionário. $L = -2$ excede a unidade em valor absoluto.

2.3.3 Modelos Autorregressivos Média Móvel

A combinação das propriedades dos processos $AR(p)$ e $MA(q)$ permitem definir os processos $ARMA(p,q)$, que dão lugar a uma família muito ampla e flexível de processos estocásticos estacionários úteis e parcimoniosos para descrever dados de séries temporais. Os processos AR e MA aparecem como casos particulares desta representação geral $ARMA(p,q)$.

Definição 6 *O processo $\{\boldsymbol{X}_t\}$ é tido como um processo $ARMA(p,q)$ se $\{\boldsymbol{X}_t\}$ é estacionário e se para cada t, (Cowpertwait and Metcalfe, 2009),*

$$X_t - \psi_1 X_{t-1} - \ldots - \psi_p X_{t-p} \quad = \quad w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q} \tag{2.9}$$

sendo $w_t \sim N(0, \sigma^2)$, $\{\psi_i\}_{i=1}^p$ e $\{\theta_i\}_{i=1}^q$ são constantes reais e os polinômios $\psi_p(w) = 1 - \psi_1 w - \cdots - \psi_p w^p$ e $\theta_q(w) = 1 - \theta_1 w - \ldots - \theta_q w^q$ não têm fatores comuns. De forma compacta, o modelo (2.9) se expressa por

$$\psi_p(L)\boldsymbol{X}_t = \theta_q(L)w_t$$

Assim,

- Um processo $ARMA(p,q)$ é estacionário quando as raízes de ψ todas excedem a unidade em valor absoluto;
- É invertível quando as raízes de θ todas excedem a unidade em valor absoluto.

Existindo alguma raiz comum nos dois polinômios, então, o modelo (2.9) estaria sobre parametrizado desnecessariamente, já que o modelo poderia escrever-se como um $ARMA(p-1, q-1)$. Tal como dito antes, os processos AR e MA são casos particulares do processo $ARMA$, ou seja, o processo $AR(p)$ é um caso particular de $ARMA(p,q)$ quando $q = 0$; o processo $MA(q)$ corresponde ao processo $ARMA(p,q)$ quando $p = 0$.

2.3.4 Modelos Autorregressivos Integrados de Médias Móveis

Os processos discutidos acima são flexíveis para séries temporais estacionárias. Pois, para ajustar tais modelos a uma série temporal, basta remover as fontes de variação não estacionárias. Por exemplo, sendo a série não estacionária na média, poder-se-ia tentar remover a tendência tomando-se uma ou mais diferenças. Outra situação comum é da variabilidade dos dados que pode crescer ou diminuir com o nível da série. Neste caso diríamos que a série não é estável na variância. Também pode ocorrer que a série seja não estacionária nas auto-correlações quando estas se modificam com o tempo. Estas situações são melhor abordadas com o uso dos modelos auto-regressivos integrados e de média móvel. Um modelo $ARMA$ no qual $\{\boldsymbol{X}_t\}$ é substituído pela sua d-ésima diferença $\bigtriangledown^d \boldsymbol{X}_t$ é capaz de descrever alguns tipos de séries não estacionárias.

Sendo a série das diferenças denotada por $\boldsymbol{Y}_t = \bigtriangledown^d \boldsymbol{X}_t = (1-L)^d \boldsymbol{X}_t$, conforme Velasco and Garcia (2009),

Definição 7 *Um processo $\{\boldsymbol{X}_t\}$ é autorregressivo integrado e de média móvel de ordem (p,d,q), denotado por $ARIMA(p,d,q)$, se o processo $\boldsymbol{Y}_t = \bigtriangledown^d \boldsymbol{X}_t$ é um modelo $ARMA(p,q)$ estacionário e invertível. Isto é, para $t = 0, \pm 1, \pm 2, \ldots,$*

$$(1 - \psi_1 L - \ldots - \psi_p L^p)\boldsymbol{Y}_t \quad = \quad (1 - \theta_1 L - \ldots - \theta_q L^q) w_t \tag{2.10}$$

ou equivalente

$$\psi_p(L)(1-L)^d \boldsymbol{X}_t \quad = \quad \theta_q(L) w_t \tag{2.11}$$

com $\psi_p(L) = 1 - \psi_1 L - \ldots - \psi_p L^p$ e $\theta_q(L) = 1 - \theta_1 L - \ldots - \theta_q L^q$. Da equação (2.11) pode-se notar que o modelo para X_t é claramente não estacionário, uma vez que o polinômio auto-regressivo $\psi_p(L)(1-L)^d$ tem exatamente d raízes unitárias. Portanto, um processo que se torna estacionário após d diferenças é dito ser integrado de ordem d. O nome $ARIMA$ provém das iniciais em inglês dos processos autorregressivos integrados de média móvel (Autoregressive Integrated Moving Average), donde *integrado* indica que, chamando $w_t = \bigtriangledown^d \boldsymbol{X}_t$, ao processo estacionário, $\{\boldsymbol{X}_t\}$ se obtém como soma (integrada) de w_t. Esta classe de processos está formada por todos os processos que podem ser transformados em estacionários mediante a aplicação do operador diferença de uma determinada ordem. A estes modelos denomina-se também *processos não estacionários homogêneos* (Velasco and Garcia, 2009).

Sabe-se que os processos *MA* finitos são sempre estacionários, mas os processos *AR* são estacionários somente se as raízes de $\psi(L) = 0$ estiverem fora do círculo unitário. Por exemplo, para o processo *AR(1)*,

$$\boldsymbol{X}_t \quad = \quad c + \psi X_{t-1} + w_t \tag{2.12}$$

se $|\psi| < 1$, o processo é estacionário. Se $|\psi| > 1$, os valores da variável crescem sem limite até o infinito; obtém-se, portanto, um processo explosivo, tal como mostrado no segundo painel da Figura 2.6, obtida com $c = 0$ e $\psi = 1.03$. Estes são processos menos frequentes na prática e menos úteis para representar séries reais. A terceira simulação do passeio aleatório (2.6) foi obtida com $c = \psi = 1$. Trata-se de um processo com desvio igual 1, e o processo mostra uma tendência linear de declive c.

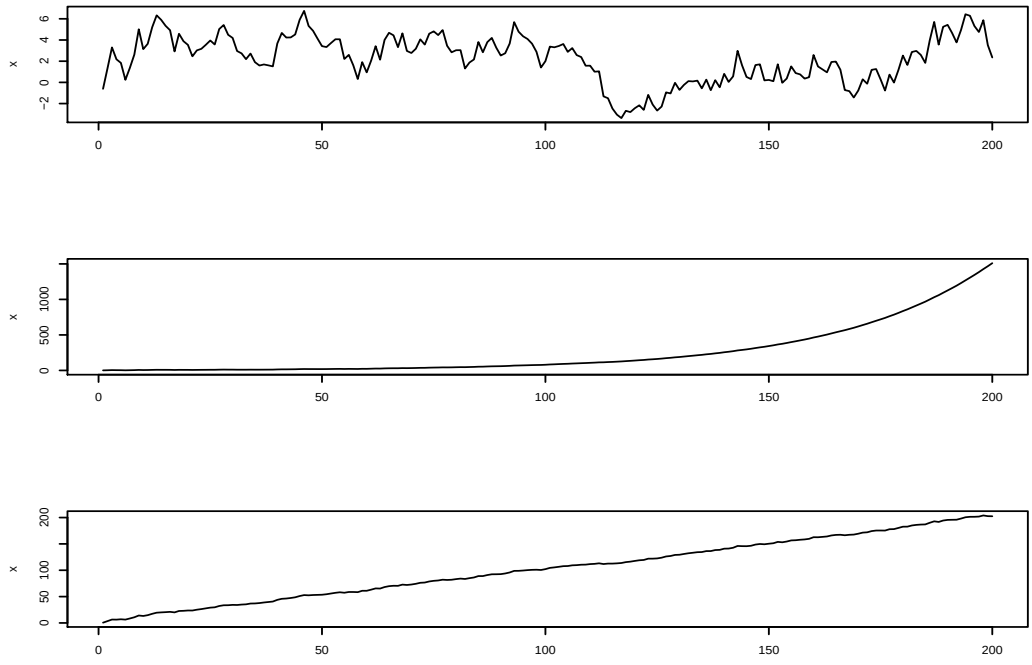


Figura 2.6: Simulação de três realizações do passeio aleatório (2.12). A primeira foi feita com $c = 0$ e $\psi = 1$. A segunda com $c = 0$ e $\psi = 1.03$. A terceira com $c = \psi = 1$.

Quando $\psi = 1$, Figura 2.6, o processo não é estacionário, tão pouco explosivo e pertence a classe de processos integrados de ordem 1. Trata-se de um processo com desvio, e o nível da série oscila ao longo do tempo. Portanto, o passeio aleatório dado em (2.6) é um modelo *ARIMA(0,1,0)*.

Capítulo 3

MODELOS ESTRUTURAIS DE SÉRIES TEMPORAIS

Índice do Capítulo

Nesse Capítulo o objetivo principal é descrever as formulações dos modelos estruturais convencionais e sua representação em espaço de estados, com destaque para os modelos dinâmicos lineares e Gaussianos. É também apresentada uma síntese comparativa e contraste entre as duas formulações, a de múltiplas fontes de aleatoriedade e a de única fonte de aleatoriedade. Não constitui objetivo desse trabalho abordar sobre os modelos não lineares e não-Gaussianos, no entanto, importa observar que a literatura sobre modelos estruturais não-lineares e não-Gaussianos tem crescido bastante durante a última década, juntamente com os avanços na inferência computacional que envolve técnicas de simulação estocástica. Mais detalhes sobre os temas, os materiais encontrados em Harvey (1989); Kitagawa and Gersch (1996); West and Harrison (1997), são bastante úteis.

Tendo em conta as semelhanças que os modelos estruturais possuem com os métodos de suavização exponencial, uma síntese sobre esses métodos é apresentada a seguir.

3.1 Introdução à Suavização Exponencial

A análise de dados experimentais observados em diferentes pontos ou intervalos temporais, leva a problemas novos e únicos de modelação estatística e inferências. O objetivo principal da análise é desenvolver modelos matemáticos que forneçam descrições plausíveis para dados de amostra, a fim de fornecer um cenário estatístico para identificar e descrever o padrão do comportamento dos dados que possibilitem fazer previsões (Makridakis et al., 1998; Shumway and Stoffer, 2006; Kitagawa, 2010). A esse respeito, vários modelos de previsão (de curto, médio ou longo prazos) têm sido propostos na literatura. Muito desses modelos são

formulados para lidarem somente com a tendência, outros para lidarem simultaneamente com a tendência e sazonalidades (diárias, semanais ou anuais) exibidas pela série temporal (Weron, 2006; Taylor and P.McSharry, 2008; Rebennack et al., 2010).

Na construção desses modelos, duas principais abordagens são evidenciadas: a abordagem que consiste em usar um modelo de equação única ou várias e diferentes equações para descrever o fenômeno observado. A outra abordagem que contempla modelos mais sofisticadas incorpora formulações vetoriais com equações interligadas para diferentes cenários (Taylor, 2003; Cottet and M, 2003; Ramanathan et al., 1997; Soares and Medeiros, 2008; Dordonnat et al., 2008). A maioria desses modelos assume que o comportamento da série no passado ou a sua relação com outras séries continuará no futuro. Sendo por isso importante a escolha do modelo que melhor descreva o comportamento da série em estudo, pois este terá importantes implicações nos procedimentos de estimação e na obtenção das previsões, (Cordeiro and Neves, 2011).

Métodos de Suavização Exponencial. Os métodos de suavização exponencial têm sido desde a década de 1950 até hoje os métodos de previsão mais populares utilizados para formulação de modelos de aplicação em negócios e na indústria. São técnicas estatísticas que visam descrever a série baseando-se no estudo das mudanças que se produzem na mesma, e não buscando ou construindo um modelo matemático explícito que tenha gerado os dados que a constituem. O propósito da aplicação desses métodos se centra na eliminação das flutuações aleatórias presentes nos dados, porém, aproveitar qualquer comportamento evidente da série com a finalidade de prever novos valores. Destacam-se pela sua versatilidade na vasta opção de modelos que integram; robustos e simples de aplicar. A sua ampla divulgação faz deles dos métodos mais utilizados na modelação e previsão em séries temporais (Hyndman et al., 2008; Cordeiro and Neves, 2011; Velasco and Garcia, 2009).

Historicamente, a suavização exponencial descreve uma classe de métodos de previsão. Há uma variedade de métodos que se encorporam na família de suavização exponencial. O nome *suavização exponencial* reflete o facto de que no cálculo da previsão, os pesos das observações passadas diminuem exponencialmente (Hyndman et al., 2008). O método de suavização exponencial mais básico é o *método de suavização exponencial simples* para o qual necessita-se estimar apenas um parâmetro. Outros métodos de suavização mais complexos são o *método linear de Holt* que faz o uso de dois parâmetros e o *método de Holt-Winters* com três parâmetros. O método de suavização exponencial simples é um método adequado para prognosticar os valores futuros das séries temporais não sazonais e sem tendência. Se a série temporal em estudo tem alguma dessas componentes (tendência ou sazonalidade), são preferíveis os métodos linear de *Holt* ou de *Holt-Winters*, respetivamente, (Hyndman et al., 2008; Velasco and Garcia, 2009).

No processo de suavização exponencial considera-se primeiro a componente tendência, que é uma combinação do nível (ℓ_t) e do declive (b_t). Sendo $\hat{y}_{t+h|t}$ a previsão no instante t a h passos de y_t, a tendência da previsão, T_h, e o parâmetro de amortecimento $0 < \phi < 1$, a Tabela 3.1 apresenta as diferentes combinações possíveis dos termos nível (ℓ_t) e declive (b_t) (Cordeiro and Neves, 2011).

Implicitamente o método de suavização exponencial simples assume que os dados

Classificação	Tendência
Nenhum	$T_h = \ell$
Aditivo	$T_h = \ell + bh$
Aditivo amortecido	$T_h = \ell + (\phi + \phi^2 + \cdots + \phi^h)$
Multiplicativo	$T_h = \ell b^h$
Multiplicativo amortecido	$T_h = \ell b^{(\phi+\phi^2+\cdots+\phi^h)}$

Tabela 3.1: Padrão da tendência para previsão (Hyndman et al., 2008)

proveem de um modelo com média localmente constante, ou seja, que os dados tenham sido gerados por um modelo da forma

$$y_t = \mu_t + \varepsilon_t \tag{3.1}$$

onde ε_t é o processo ruido branco. O nível, μ_t, pode mudar lentamente com o tempo, no entanto, em qualquer segmento de tempo local, a constante μ fornece um modelo razoavelmente bom da série temporal. Dado ℓ_{t-1}, com ℓ_t o estimador do nível no tempo t e y_t disponível, o método de suavização exponencial simples atualiza o estimador do nível através da equação de recorrência

$$\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1} \tag{3.2}$$

onde $0 \leq \alpha \leq 1$ é o parâmetro de suavização que determina o peso dado a cada uma das componentes para gerar as previsões. Pode ser usado para ajustar a sensibilidade do estimador a mudanças no nível da série. Quanto maior o seu valor, mais peso recebe a observação y_t, mais sensível o estimador a mudanças no nível. O estimador, ℓ_t, é a média ponderada da última observação, y_t, e ℓ_{t-1} é o estimador no instante $t-1$.

A previsão de h passos à frente feita no instante $t+h$ é dada por

$$\hat{y}_{t+h|t} = \ell_t, \quad h > 0 \tag{3.3}$$

Ao substituirmos recursivamente e sucessivamente l_{t-1} na expressão anterior, as previsões são uma soma ponderada de todas as observações anteriores e do valor inicial do nível ℓ_0 (Cordeiro and Neves, 2011).

$$\hat{y}_{t+h|t} = (1-\alpha)^t \ell_0 + \alpha \sum_{i=0}^{t-1} (1-\alpha)^i y_{t-i} \tag{3.4}$$

Se tomamos α próximo de 1, as previsões para os distintos períodos são muitos similares e pouco variam com a nova informação. Contrariamente ao α muito pequeno, próximo de zero, a previsão adequa-se em função do último valor observado (Wang, 2006; Velasco and Garcia, 2009).

Quando se trata de dados com tendência, o método de suavização exponencial simples

não proporciona bons resultados. Para resolver esse problema, é introduzido o método linear de *Holt* que atualiza os estimadores usando as equações de recorrência

$$\begin{aligned} \text{Nível:} \quad \ell_t &= \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1}) && (3.5a) \\ \text{Tendência :} \quad b_t &= \beta(\ell_t - \ell_{t-1}) + (1-\beta) b_{t-1} && (3.5b) \\ \text{Previsão :} \quad \hat{y}_{t+h|t} &= \ell_t + h b_t && (3.5c) \end{aligned}$$

onde α e β são parâmetros de suavização que tomam valores do intervalo $(0, 1]$.

Quando existe sazonalidade nos dados, os métodos descritos anteriormente não descrevem bem o comportamento da componente sazonal. Para esse tipo de dados, o método de *Holt-Winters* comporta-se melhor. O método sazonal *Holt-Winters* compreende a equação de previsão e três equações de suavização – uma para o nível, uma para o declive e outra para a sazonalidade. Dependendo se a sazonalidade é combinada com a tendência linear aditiva ou multiplicativa, existem duas versões do método de *Holt-Winters*. Na síntese a apresentar sobre o método, far-se-á uma distinção entre o efeito sazonal aditivo e/ou multiplicativo. A preferência pelo método aditivo eleva-se quando as variações sazonais são praticamente constantes pela série, enquanto o método multiplicativo é preferido quando as variações sazonais vão se alterando proporcionalmente ao nível da série.

- Holt-Winters aditivo

$$\begin{aligned} \hat{y}_{t+h|t} &= \ell_t + h b_t + s_{t-m+h_m^+} && (3.6a) \\ \ell_t &= \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1}) && (3.6b) \\ b_t &= \beta(\ell_t - \ell_{t-1}) + (1-\beta) b_{t-1} && (3.6c) \\ s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma) s_{t-m} && (3.6d) \end{aligned}$$

  onde $0 \leq \gamma \leq 1 - \alpha$, e $h_m^+ = \lfloor (h-1) mod\ m \rfloor + 1$, o qual garante que as estimativas dos índices sazonais utilizados para a previsão provenham do último ano da amostra. A notação $\lfloor u \rfloor$ significa (o maior inteiro contido em u). A forma de correção do erro das equações de suavização é dada por:

$$\begin{aligned} \ell_t &= \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t && (3.7a) \\ b_t &= b_{t-1} + \alpha\beta\varepsilon_t && (3.7b) \\ s_t &= s_{t-m} + \gamma\varepsilon_t && (3.7c) \end{aligned}$$

- Holt-Winters multiplicativo

$$\begin{aligned} \hat{y}_{t+h|t} &= (\ell_t + h b_t) s_{t-m+h_m^+} && (3.8a) \\ \ell_t &= \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1}) && (3.8b) \\ b_t &= \beta(\ell_t - \ell_{t-1}) + (1-\beta) b_{t-1} && (3.8c) \\ s_t &= \gamma \frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1-\gamma) s_{t-m} && (3.8d) \end{aligned}$$

Para mais detalhes, ver Velasco and Garcia (2009); Cowpertwait and Metcalfe (2009); Hyndman (2014).

3.2 Modelos Lineares Dinâmicos

Os modelos estruturais de séries temporais (ou modelos lineares dinâmicos) são casos especiais de modelos de espaço de estados, considerados como generalizações dos modelos tratados até agora. Foram usados extensivamente na teoria de sistemas, nas ciências físicas, e na engenharia. A terminologia é, em grande parte, a partir desses campos (Helmut, 2005). A motivação para o estudo e a aplicação dos modelos estruturais em espaço de estados, pode ser desenvolvida na análise de séries temporais como as apresentadas na Secção 1.3. De acordo com Stoffer and Wall (2004), as formulações de espaço de estados são bastante gerais e podem assumir assumir casos especiais de interesse. Embora esses modelos tenham sidos originalmente desenvolvidos como métodos principalmente para o uso em pesquisas relacionadas o aeroespaço, eles foram amplamente aplicados a modelos de dados nos diferentes campos do conhecimento humano; e um dos excelentes tratamentos modernos da análise de séries temporais com base a modelos de espaço de estados é o trabalho desenvolvido por Durbin and S.J.Koopman (2011).

Nas últimas três décadas assiste-se um crescente interesse no estudo de modelos de espaço de estados e sua aplicação na análise de séries temporais, a exemplo dos trabalhos de Harvey (1989); Wang (2006); Hyndman et al. (2008); Petris et al. (2009); Durbin and S.J.Koopman (2011); De Livera et al. (2011); Gob et al. (2013), as recentes revisões de Ahmad and Maxwell (2015) e Shumway and Stoffer (2017) e suas referências. Tais modelos consideram uma série temporal como a saída de um sistema dinâmico perturbado por ruídos aleatórios. Possuem uma estrutura probabilística elegante e poderosa que oferece uma estrutura flexível para uma ampla gama de aplicações. Os cálculos podem ser implementados por algoritmos recursivos. Os problemas de estimação e de previsão são resolvidos através da computação recursiva da distribuição condicional das quantidades de interesse, dada a informação disponível, (Petris et al., 2009).

Como a tarefa inicial na análise de séries temporais é a pesquisa da evolução dinâmica das observações ao longo do tempo, sabe-se que as propriedades dinâmicas não podem ser observadas diretamente dos dados. O processo dinâmico não observável no instante t é referido por **estado** da série temporal e pode integrar várias componentes (Cordeiro and Neves, 2011; Douc et al., 2014). Assume-se, por isso, que há uma cadeia de Markov não observável, $\boldsymbol{x}_t$, a que se denomina por *processo de estado*, e $\boldsymbol{y}_t$ a medida de $\boldsymbol{x}_t$ caracterizada por dois princípios, Figura 3.1, (Petris et al., 2009; Shumway and Stoffer, 2017):

1. Existência de um processo latente denominado *processo de estado*, assumido ser um processo de *Markov*–significa que o futuro $\{\boldsymbol{x}_n; n > t\}$ e o passado $\{\boldsymbol{x}_n; n < t\}$ são condicionalmente independentes dado o presente;

2. As observações são independentes dados os estados–significa que a dependência entre as observações é gerada por estados;

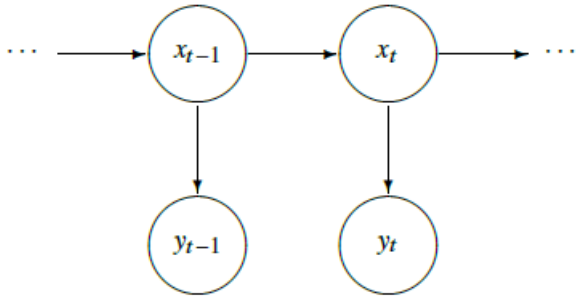


Figura 3.1: Diagrama de um modelo em espaço de estado (Shumway and Stoffer, 2017)

A primeira classe importante de modelos de espaço de estados é dada por modelos lineares Gaussianos, também chamados de modelos lineares dinâmicos (Petris et al., 2009). De acordo com Douc et al. (2014),

Definição 8 *Um processo* $\{\boldsymbol{X}_t, t \in \mathcal{Z}\}$ *é linear se é definida como uma combinação linear de ruídos brancos* $w_t \sim \mathcal{N}(0, \sigma^2)$*, e é dado por*

$$\boldsymbol{X}_t = \boldsymbol{\mu} + \sum_{j=-\infty}^{\infty} \Psi_j \boldsymbol{w}_{t-j}, \quad \sum_{j=-\infty}^{\infty} \Psi_j^2 < \infty. \tag{3.9}$$

Definição 9 *Um processo* $\{\boldsymbol{X}_t, t \in \mathcal{Z}\}$ *é dito ser Gaussiano se os vetores n-dimensionais* $\boldsymbol{X} = (X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ *para qualquer coleção de pontos temporais distintos* $t_1, t_2, \ldots, t_n$*, e qualquer inteiro positivo* n*, tem uma distribuição normal multivariada não-singular.*

Sendo $E(\boldsymbol{X}) \equiv \boldsymbol{\mu} = (\mu_{t_1}, \mu_{t_2}, \ldots, \mu_{t_n})$ o vetor média de dimensão $n \times 1$ e a matriz variância-covariância $n \times n$ dada por $Cov(\boldsymbol{X}) \equiv \boldsymbol{\Lambda} = \{\gamma(t_i, t_j); i, j = 1, \ldots, n\}$, a qual é assumida ser positiva, a função densidade normal multivariada é definida por

$$f(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\Lambda}) = (2\pi)^{-n/2} |\boldsymbol{\Lambda}|^{-1/2} exp\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})' \boldsymbol{\Lambda}^{-1} \{\boldsymbol{x} - \boldsymbol{\mu}\}\} \tag{3.10}$$

onde $|\cdot|$ denota o determinante e $\boldsymbol{x} \in \mathcal{R}^n$. Essa distribuição forma o básico para resolver os problemas que envolvem inferências estatísticas para séries temporais lineares Gaussianas. A definição da estrutura de modelos de espaço de estados com múltiplas fontes de aleatoriedade pode ser encontrada em Brockwell and Davis (2002); Ord et al. (2005); Tsay (2005); Wang (2006); Durbin and S.J.Koopman (2011); Hyndman et al. (2008); Shumway and Stoffer (2017). Segundo Brockwell and Davis (2002),

Definição 10 *Um modelo em espaço de estado para uma série temporal (possivelmente multivariada)* $\{\boldsymbol{y}_t, t = 1, 2, \ldots, n\}$ *consiste em duas equações. A primeira, conhecida de* equação de observação *ou* equação de medida*, expressa a observação* $\boldsymbol{y}_t$ *como uma função linear de uma variável de estado* $\boldsymbol{x}_t$ *mais o ruido, ou seja*

$$\boldsymbol{y}_t = \boldsymbol{A}_t \boldsymbol{x}_t + \boldsymbol{\nu}_t, \quad t = 1, 2, \ldots, n \tag{3.11}$$

onde $\boldsymbol{y}_t$ *é um vetor de observações de dimensões* $k \times 1$*;* $\boldsymbol{A}_t$ *é uma matriz* $(k \times p)$ *que supõe-se conhecida para todo* t*;* $\boldsymbol{x}_t$ *é o vetor de variáveis de estado não observável de dimensão* $(p \times 1)$ *e* $\boldsymbol{\nu}_t$ *é um processo de ruído branco que supõe-se ter a distribuição* $\mathcal{N}(0, \boldsymbol{R})$*.*

A descrição do sistema inclui uma equação que descreve a evolução dinâmica das variáveis de

estado, $\boldsymbol{x}_t$, denominada equação de estado ou equação de transição, que determina o estado $\boldsymbol{x}_t$ no instante t em termos do estado anterior $\boldsymbol{x}_{t-1}$ e um termo de ruído, ou seja

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{\Theta}\boldsymbol{w}_t, \quad t = 1, 2, \ldots, n \tag{3.12}$$

onde $\boldsymbol{\Phi}$ é uma matriz (de transição) conhecida de dimensão $(p \times p)$, $\boldsymbol{\Theta}$ é uma matriz de parâmetros e w_t é outro processo de ruído branco, independente do anterior, que tem distribuição $\mathcal{N}(0, \boldsymbol{Q})$. $\boldsymbol{A}_t$, $\boldsymbol{x}_t$ e $\boldsymbol{\Phi}$ contêm parâmetros desconhecidos e que devem ser estimados. ν_t e w_t podem ser referidos como o erro transitório e erro permanente, respetivamente. O erro transitório apenas afeta a observação atual y_t, enquanto o efeito do erro permanente persiste ao longo do tempo.

A ideia geral por trás desses modelos é que uma série temporal $\{y_1, ..., y_n\}$ depende de um estado possivelmente não observado $\boldsymbol{x}_t$ que é conduzido por um processo estocástico (Helmut, 2005). São modelos muito usados para modelar séries temporais univariadas ou multivariadas, também na presença de não-estacionaridade, mudanças estruturais e padrões irregulares (Petris et al., 2009). Várias terminologias têm sido usadas e tipos de notação, o que justifica a sua larga aplicação tanto por engenheiros como por estatísticos. Para os engenheiros, alguns dos termos usados são *filtros de Kalman* e *modelos lineares dinâmicos*, para os estatísticos *regressão linear dinâmica*. De igual modo, significados diferentes são atribuidos às designações; ou seja, usa-se a designação de *espaço de estados* no contexto da formulação de modelos e filtro de Kalman no contexto da técnica de previsão recursiva usada no quadro desses modelos (Cordeiro and Neves, 2011).

Covariáveis em modelos de espaço de estados. Em muitas aplicações a informação adicional do problema em estudo pode estar disponível na forma de variáveis de entrada (variáveis de influência externa) e que pode representar as chamadas variáveis explicativas ou variáveis de intervenção. Entende-se de variável explicativa àquela variável que fornece ao analista informações adicionais. O termo *explicativa* nesse contexto não exige a relação de causalidade entre a entrada e as variáveis dependentes, apenas representa uma série que está disponível em tempo oportuno para melhorar o processo de previsão. A intervenção é frequentemente representada por uma variável indicadora, tomando valores de 0 e 1 ou outras formas mais gerais. Essas variáveis podem representar mudanças planeadas ou eventos incomuns (por exemplo, condições climáticas extremas). Também podem ser utilizadas para sinalizar observações ou valores extremos incomuns (Herman, 1994; Hyndman et al., 2008; Koehler et al., 2012). Elas podem ser incluídas no modelo para capturar efeitos exógenos e vários tipos de intervenções. Se $\boldsymbol{\Gamma}$ e $\boldsymbol{\Upsilon}$ denotam as matrizes (fixas e conhecidas) dos coeficientes de regressão e $\boldsymbol{z}_t$ a matriz das covariáveis, o modelo de espaço de estados pode ser escrito na forma

$$\boldsymbol{y}_t = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{\Gamma}\boldsymbol{z}_t + \boldsymbol{\nu}_t, \quad t = 1, 2, \ldots, n \tag{3.13a}$$

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{\Upsilon}\boldsymbol{z}_t + \boldsymbol{w}_t, \quad t = 1, 2, \ldots, n \tag{3.13b}$$

onde,

- $\boldsymbol{A}_t$ é a matriz de medição, variante ao longo do tempo;
- $\boldsymbol{\Phi}$ é a matriz de transição (ou do sistema);

- $\boldsymbol{\Gamma}$ e $\boldsymbol{\Upsilon}$ são matrizes de entradas formadas a partir dos coeficientes de regressão para as equações de medida e do estado, respetivamente;
- Os ruídos $\boldsymbol{\nu}_t$ e $\boldsymbol{w}_t$ são não correlacionados. As matrizes do sistema têm as seguintes dimensões:

$$\begin{aligned}
\boldsymbol{y}_{(q\times 1)} &= \boldsymbol{A}_{(q\times p)}\boldsymbol{x}_{(p\times 1)} + \boldsymbol{\Gamma}_{(q\times r)}\boldsymbol{z}_{(r\times 1)} + \boldsymbol{\nu}_t; \quad \boldsymbol{\nu}_t \sim iid\mathcal{N}_q(0, \boldsymbol{R}) \\
\boldsymbol{x}_{(p\times 1)} &= \boldsymbol{\Phi}_{(p\times p)}\boldsymbol{x}_{(p\times 1)} + \boldsymbol{\Upsilon}_{(p\times r)}\boldsymbol{z}_{(r\times 1)} + \boldsymbol{w}_t; \quad \boldsymbol{w}_t \sim iid\mathcal{N}_m(0, \boldsymbol{Q})
\end{aligned}$$

A formulação de um modelo estrutural em espaço de estados é bastante simples. Por exemplo, o modelo de tendência linear local com o declive variando no tempo conforme a dinâmica de ℓ_t, definido por

$$\begin{aligned}
y_t &= \ell_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2) && \text{(3.14a)} \\
\ell_t &= \ell_{t-1} + \nu_{t-1} + \xi_t, \qquad \xi_t \sim iid\mathcal{N}(0, \sigma_\xi^2) && \text{(3.14b)} \\
\nu_t &= \nu_{t-1} + \zeta_t, \qquad \zeta_t \sim iid\mathcal{N}(0, \sigma_\zeta^2) && \text{(3.14c)}
\end{aligned}$$

a sua estrutura matricial de espaço de estados é dada por

$$y_t = [1, 0].\begin{bmatrix} \ell_{t-1} \\ \nu_{t-1} \end{bmatrix} + \varepsilon_t$$

$$\begin{bmatrix} \ell_t \\ \nu_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.\begin{bmatrix} \ell_{t-1} \\ \nu_{t-1} \end{bmatrix} + \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix}$$

A adição da componente sazonal (modelo aditivo de período m), s_t, a equação de medida defini-se por

$$y_t = \ell_t + s_{t-m} + w_t, \qquad w_t \sim iid\mathcal{N}(0, \sigma_w^2) \qquad \text{(3.15)}$$

onde s_t tem a seguinte dinâmica $s_t = -s_{t-1} + \ldots + s_{t-m+1} + w_t$. No caso de um modelo estrutural com nível local, tendência e sazonalidade, com inclusão do parâmetro de amortecimento na tendência, ideia introduzida por Gardner and McKenzie (1985), pode ser formulado como

$$\begin{aligned}
y_t &= \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t; \qquad \varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2) && \text{(3.16a)} \\
\ell_t &= \ell_{t-1} + \phi b_{t-1} + \xi_t; \qquad \xi_t \sim iid\mathcal{N}(0, \sigma_\xi^2) && \text{(3.16b)} \\
b_t &= \phi b_{t-1} + \zeta_t; \qquad \zeta_t \sim iid\mathcal{N}(0, \sigma_\zeta^2) && \text{(3.16c)} \\
s_t &= s_{t-m} + \omega_t \qquad \omega_t \sim iid\mathcal{N}(0, \sigma_\omega^2) && \text{(3.16d)}
\end{aligned}$$

O modelo de tendência em (3.16c) pode ser extensivo adicionando uma constante,

$$b_t = (1 - \phi)b + \phi b_{t-1} + \zeta_t \qquad \text{(3.17)}$$

onde, b denota a tendência de longo prazo e b_t a tendência de curto prazo. Em espaço de estados, o modelo é expresso pela seguinte estrutura matricial:

$$\boldsymbol{A}_t = \begin{bmatrix} 1 \\ \phi \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}; \boldsymbol{x}_t = \begin{bmatrix} \ell_t \\ b_t \\ s_t \\ s_{t-1} \\ \vdots \\ s_{t-m+1} \end{bmatrix}; \boldsymbol{\Phi} = \begin{bmatrix} 1 & \phi & 0 & 0 \dots 0 & 0 \\ 0 & \phi & 0 & 0 \dots 0 & 0 \\ 0 & 0 & 0 & 0 \dots 0 & 1 \\ 0 & 0 & 1 & 0 \dots 0 & 0 \\ 0 & 0 & 0 & 1 \dots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \ddots \vdots & \vdots \\ 0 & 0 & 0 & 0 \dots 1 & 0 \end{bmatrix} \text{ e } \boldsymbol{\Theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix}.$$

As matrizes de variância-covariância dos ruídos w_t e ν_t são dadas por

$$\boldsymbol{Q} = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \sigma_\omega^2 \end{bmatrix} \quad \text{e} \quad \boldsymbol{R} = Var(\varepsilon_t)$$

3.3 Representação dos Modelos Estruturais Convencionais

Os modelos estruturais de séries temporais são formulados diretamente em termos de componentes não observáveis, que possuem uma interpretação natural e representam as principais características da série em análise. Esses modelos possuem muitas semelhanças com métodos de suavização exponencial, mas possuem múltiplas fontes de aleatoriedade. Por exemplo, o modelo estrutural básico é semelhante ao método de Holt-Winters para dados sazonais e inclui um nível, tendência e componente sazonal, (Tommaso, 1991; Jan and Hyndman, 2005). Os modelos estruturais são facilmente estendidos para lidar com qualquer tipo de frequência (semanal, diária, horária, etc) e padrões das séries que são difíceis de lidar, por exemplo, na estrutura ARIMA (heteroscedasticidade, não-linearidade, não-Gaussianidade). Ademais, são configurados como modelos de regressão nos quais as variáveis explicativas são funções do tempo e os coeficientes podem variar ao longo do tempo, abrangendo a decomposição tradicional de uma série temporal em componentes determinísticos. A extensão dessa classe de modelos com uma abordagem não bayesiana de estimação pode ser vista em Harvey (1989), e as principais ideias e aspectos metodológicos subjacentes a esses modelos podem ser vistos em Harvey (1989); West and Harrison (1997); Kitagawa and Gersch (1996).

Especificar um modelo estrutural de série temporal depende dos recursos exibidos pela série em análise e outros conhecimentos prévios. O modelo estrutural mais elementar lida com uma série cujo nível subjacente varia ao longo do tempo e o processo gerador de dados pode resultar de uma tendência, ℓ_t, evoluindo de acordo com um passeio aleatório com um componente irregular, ε_t:

$$y_t = \ell_t + \varepsilon_t \tag{3.18a}$$

$$\ell_t = \ell_{t-1} + \eta_t \tag{3.18b}$$

onde $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$ e $\eta_t \sim iid\mathcal{N}(0, \sigma_\eta^2)$ são duas séries independentes, para $t = 1, 2, \dots, n$. O valor inicial ℓ_0 ou é dado ou segue uma distribuição conhecida, e é independente de ε_t e η_t. Na literatura ℓ_t é referido como *tendência* da série, que não é diretamente observável, e y_t são os

dados observados com ruído observacional ε_t. Esse modelo é chamado de *modelo de tendência local*. O modelo é também chamado de *modelo de nível local* em Durbin and S.J.Koopman (2011), que é um caso simples do modelo estrutural de séries temporais de Harvey (1989), e é um modelo especial de espaço de estados Gaussiano linear.

As flutuações sazonais são responsáveis por grande parte da variação de um amplo espectro de fenômenos econômicos, sociais e ambientais. Descrevem movimentos sistemáticos não necessariamente regulares, causados pelas mudanças do clima, calendário e tempo das decisões (Tommaso, 1991). Existem várias especificações de uma componente sazonal, geralmente denotado por St, satisfazendo este requisito. Por exemplo, no domínio da frequência, um padrão sazonal fixo é modelado pela soma dos ciclos $[k/2]$ definidos nas frequências sazonais $\lambda_j = 2\pi j/k$, com $j = 1, 2, \cdots, [k/2]$, onde $[k/2] = k/2$ para k par e $(k-1)/2$ se k ímpar, isto é:

$$S_t = \sum_{j=1}^{[k/2]} s_{jt}, \quad s_{jt} = \alpha_j cos\lambda_j + \alpha_j^* sin\lambda_j \tag{3.19}$$

Quando k é par o termo *sin* desaparece para $j = k/2$, então o número de termos trigonométricos é sempre $k-1$. Uma possível extensão estocástica do modelo sazonal trigonométrico é tal que o efeito sazonal no instante t resulte da combinação de $[k/2]$ ciclos estocásticos formulados como em (3.20),

$$\begin{aligned} S_t &= \sum_{j=1}^{[k/2]} s_{j,t} \\ s_{j,t} &= s_{j,t-1} cos\lambda_j + s^*_{j,t-1} sin\lambda_j + w_{j,t} \\ s^*_{j,t} &= -s_{j,t-1} sin\lambda_j + s^*_{j,t-1} cos\lambda_j + w^*_{j,t} \end{aligned} \tag{3.20}$$

Para k par, o último componente, definido em $\lambda_{k/2} = \pi$, reduz-se para $S_{k/2,t+1} = -S_{k/2,t} + w_{k/2,t}$. Os ruídos $w_{j,t}$ e $w^*_{j,t}$ são assumidos como sendo distribuídos normal e independentemente com variância comum $\sigma^2_{w_j}$ (Tommaso, 1991).

A Tabela 3.2 mostra os modelos estruturais padrão correspondentes às duas formulações: a formulação de fonte única de aleatoriedade (designado em inglês por *single source of error - SSOE*) e a formulação de múltiplas fontes de aleatoriedade (designado em inglês por *multi-disturbance or multiple source of error - MSOE*). Os símbolos ℓ_t, b_t e ε_t são comumente usados para representar o nível, a tendência e a inovação. No entanto, seus valores e significados diferem entre as duas estruturas, ver Hyndman et al. (2008). Por exemplo, o vetor dos estados não observados para ambas as estruturas, é composto pelos componentes nível, ℓ_t, tendência, b_t, e sazonalidade, s_t. No caso dos modelos de suavização exponencial (ETS–*Exponential SmooThing State Space model*),

- $\boldsymbol{y}_t$ depende de $\boldsymbol{x}_{t-1}$.
- O mesmo processo do erro afeta $\boldsymbol{x}_t|\boldsymbol{x}_{t-1}$ e $\boldsymbol{y}_t|\boldsymbol{x}_t$.

Para os modelos estruturais,

- $\boldsymbol{y}_t$ depende de $\boldsymbol{x}_t$.
- Processos diferentes do erro afetam $\boldsymbol{x}_t|\boldsymbol{x}_{t-1}$ e $\boldsymbol{y}_t|\boldsymbol{x}_t$.

Um outro exemplo tem a ver com a previsão um passo à frente para o modelo de suavização exponencial simples dada por $\hat{y}_{t+1|t} = \hat{\ell}_t$, onde $\hat{\ell}_t = \ell_{t-1} + \hat{\alpha}\varepsilon_t$ é a estimativa da média ou nível de y_t feito no instante t, e $\hat{\alpha}$ é um parâmetro de suavização estimado. Para formulações de múltiplas fontes de aleatoriedade, o modelo que corresponde à suavização exponencial simples é dado por (3.18a), onde ℓ_t é a variável de estados não observados (o nível ou a média das séries temporais que podem ser estimadas) quando os valores de $\{y_1, ..., y_n\}$ são conhecidos. As previsões de suavização exponencial simples são as previsões quadráticas médias mínimas para o modelo de múltiplas fontes de aleatoriedade. A aplicação do filtro de Kalman ao modelo, permite verificar que a equação do estado estacionário para o nível corresponde a estimativa da média em suavização exponencial simples (Ord et al., 2005).

As duas formulações		
Modelo	Modelos de múltiplas fontes de aleatoriedade	Modelos de inovações
Nivel	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \xi_t$	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$
Tendência	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \xi_t$ $b_t = b_{t-1} + \zeta_t$	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$
Sazonalidade	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \xi_t$ $b_t = b_{t-1} + \zeta_t$ $s_t = s_{t-m} + w_t$	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$

Tabela 3.2: Modelos lineares dinâmicos convencionais (Hyndman et al., 2008)

A par dos aspectos realçados acima, ambas as formulações são igualmente gerais. As vantagens dessas duas formulações são amplamente discutidas na literatura, tal é o caso do trabalho apresentado em Jerez et al. (2015). Nesse trabalho que coleta os pontos de vistas de vários autores, *Jerez* fornece também evidências do seu estudo particular intitulado *Single and multiple error state-space models for signal extraction*, que a seguir transcrevemos:

1. As formulações de fonte única de aleatoriedade possuem vantagens de velocidade e estabilidade numérica para o cálculo da verossimilhança de um modelo de espaço de estados com coeficientes fixos. Fornece, igualmente, uma decomposição estrutural que segue a dinâmica do sistema e evita revisão, o que o torna mais adequado quando se deseja remover uma determinada componente, por exemplo, a sazonalidade da série;

2. As formulações de múltiplas fontes de aleatoriedade têm duas vantagens para o cálculo do indicador de tendência. Primeiro, cada valor na componente tendência combina informações tanto passadas como futuras, de modo que os valores atuais da tendência conduzem a tendência do modelo de única fonte de aleatoriedade. Segundo, a formulação de múltiplas fontes de aleatoriedade permite impor restrições de relação ruído-variância e, portanto, escolher a *suavidade* da componente tendência, sendo este recurso muito conveniente para detectar pontos de rotação do ciclo comercial, por exemplo.

3. As duas formulações podem ser combinadas, pois, têm vantagens específicas para

diferentes aplicações. Os sinais extraídos a partir dessas duas formulações são muito semelhantes; no entanto, suas variações são drasticamente diferentes porque os estados para a formulação de única fonte de aleatoriedade colapsam para valores com variância zero.

3.4 Filtro de Kalman: Extração do sinal e previsão

O objetivo do filtro de Kalman é atualizar o conhecimento da variável de estado recursivamente quando um novo ponto de dados se torna disponível. Dado o modelo (3.13), o intuito da análise é inferir propriedades do estado $\boldsymbol{x}_t$ a partir dos dados $\{y_1, \cdots, y_n\}$ e do modelo. Três tipos de inferência são comumente discutidos na literatura, são eles: a filtragem, previsão e suavização (Shumway and Stoffer, 2017).

- **Filtragem**. Filtrar significa recuperar a variável de estado $\boldsymbol{x}_t$ dado $\{y_1, \cdots, y_n\}$, isto é, remover os erros de medição a partir dos dados, onde $n = t$.
- **Previsão**. Previsão significa prever $\boldsymbol{x}_{t+h}$ ou $\boldsymbol{y}_{t+h}$ dado $\{y_1, \cdots, y_n\}$, onde t é a origem da previsão e $n < t$.
- **Suavização**. Suavizar significa estimar $\boldsymbol{x}_t$ dado $\{y_1, \cdots, y_n\}$, onde $n > t$.

Alguns resultados elementares que fornecem a base para o tratamento do filtro de Kalman, segundo Durbin and S.J.Koopman (2011), são apresentados a seguir, e as observações serão tradadas como multivariadas. Para simplificar a descrição, as variáveis x e y são usadas para designar dois vetores aleatórios normalmente distribuídos em conjunto, isto é

$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \qquad Var\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy} & \Sigma_{yy} \end{bmatrix} \tag{3.21}$$

onde Σ_{yy} é assumido ser uma matriz não singular.

Lema 1. A distribuição condicional de x dado y é normal com vetor média

$$E(x|y) = \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y) \tag{3.22}$$

e a matriz de variância

$$Var(x|y) = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^{'} \tag{3.23}$$

Considera-se a estimação de x quando x é desconhecido e y é conhecido. Pelo **Lema 1** pode-se tomar a estimativa de x como a expectativa condicional $\hat{x} = (x|y)$, isto é

$$\hat{x} = \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y) \tag{3.24}$$

sendo $\hat{x} - x$ o erro da estimativa, então x é condicionalmente não enviesado no sentido de que $E(\hat{x} - x|y) = \hat{x} - E(x|y) = 0$. Quando a suposição de que (x, y) é normalmente distribuído é descartada, dá-se lugar o **Lema 2**.

Lema 2. Se (x, y) é normalmente distribuída ou não, a estimativa $\hat{x}$ definida por (3.24) é uma estimativa linear não enviesada de variância mínima de x dado y, e sua matriz de variância do erro é dada por

$$Var(\hat{x} - x) = Var[\Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y) - (x - \mu_x)] = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^{'} \tag{3.25}$$

A propriedade de estimativa linear não enviesada da estimativa do vetor $\hat{x}$ implica que as funções lineares arbitrárias dos elementos de $\hat{x}$ são estimativas lineares não enviesadas de variância mínima das correspondentes funções lineares dos elementos de x. O **Lema 2** é significativo para aquelas abordagens cujos pesquisadores preferem não assumir a normalidade como base para a análise de séries temporais, uma vez que muitas séries temporais reais têm distribuições que parecem estar longe de ser normal. Outros pesquisadores preferem abordar os problemas de inferência na análise de séries temporais de espaço de estados desde o ponto de vista Bayesiano, em vez do ponto de vista clássico, para o qual os **Lema 1, 2** são apropriados.

A extração do sinal de uma série temporal é uma tarefa que trata da definição e estimativa dos sinais interessantes ocultados na série temporal, como a tendência, os ciclos ou componentes sazonais. Para esse propósito e por duas razões óbvias, os modelos de espaço de estados são uma escolha natural para o efeito. A primeira dessas razões tem a ver com o facto de esses modelos configurarem as séries temporais observadas como uma combinação linear de variáveis latentes, os chamados *estados*, que podem ter propriedades dinâmicas e estocásticas muito flexíveis. Em segundo lugar podemos referir que os modelos de espaço de estados permitem construir algoritmos robustos e eficientes para se obter estimativas dos estados e covariâncias correspondentes ao sistema em análise (Jerez et al., 2015).

A seguir apresenta-se a recursão do filtro de Kalman do modelo (3.13) para o caso em que o estado inicial $\boldsymbol{x}_0 \sim \mathcal{N}_p\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0\right)$ para $t = 1, \cdots, n$ e onde $\boldsymbol{x}_0 = \boldsymbol{\mu}_0$ e $\boldsymbol{P}_0 = \boldsymbol{\Sigma}_0$ são conhecidos. A derivação apresentada é baseada na inferência clássica usando o **Lema 1**. Considera-se $\boldsymbol{x}_{t|t-1} = E(\boldsymbol{x}_t|\mathbf{y}_{1:t-1})$, $\boldsymbol{P}_{t|t-1} = Var(\boldsymbol{x}_t|\mathbf{y}_{1:t-1})$ e $\boldsymbol{x}_{t|t}$ o estimador do estado $\boldsymbol{x}_{t|t-1}$. Esta estimativa pode ser conseguida tomando as esperanças de (3.13b) condicionadas a $\mathbf{y}_{1:t-1}$, isto é, $E(\boldsymbol{x}_t|\mathbf{y}_{1:t-1})$. Denotamos $\boldsymbol{P}_{t|t-1}$ como a matriz de covariância a priori da estimativa do estado a priori, $\boldsymbol{x}_{t|t-1}$, e $\boldsymbol{P}_{t|t}$ a matriz de covariância a posteriori do estado atualizado $\boldsymbol{x}_{t|t}$ (ou a posteriori). Considera-se as seguintes suposições:

1. Os ruídos Gaussianos da medida e do estado com covariâncias $\boldsymbol{R} = E(\boldsymbol{\nu}_t\boldsymbol{\nu}_t^{'})$ e $\boldsymbol{Q} = E(\boldsymbol{w}_t\boldsymbol{w}_t^{'})$, respetivamente, são independentes e mutuamente não correlacionados;
2. Os ruídos têm matrizes de covariâncias constantes;
3. O estado e a medida têm a mesma frequência da amostragem.

Assim, o filtro de Kalman para o modelo (3.13) é composto das etapas de previsão e atualização, que a seguir resumimos.

Etapa de previsão:

$$\boldsymbol{x}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{x}_{t-1|t-1} + \boldsymbol{\Upsilon}\boldsymbol{z}_t \tag{3.26a}$$

$$\boldsymbol{P}_{t|t-1} = \boldsymbol{\Phi}\boldsymbol{P}_{t-1|t-1}\boldsymbol{\Phi}^{'} + \boldsymbol{Q} \tag{3.26b}$$

Etapa de atualização

$$
\begin{aligned}
\boldsymbol{x}_{t|t} &= \boldsymbol{x}_{t|t-1} + \boldsymbol{K}_t\boldsymbol{\varepsilon}_t && \text{(3.27a)}\\
\boldsymbol{P}_{t|t} &= \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{A}_t\boldsymbol{P}_{t|t-1} && \text{(3.27b)}\\
\boldsymbol{K}_t &= \boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'}[\boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'} + \boldsymbol{R}]^{-1} && \text{(3.27c)}\\
\boldsymbol{\varepsilon}_t &= \boldsymbol{y}_t - \boldsymbol{A}_t\boldsymbol{x}_{t|t-1} - \boldsymbol{\Gamma}\boldsymbol{z}_t && \text{(3.27d)}\\
\boldsymbol{\Sigma}_t &= \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'} + \boldsymbol{R} && \text{(3.27e)}
\end{aligned}
$$

onde $\boldsymbol{K}_t$ é o ganho de Kalman, $\boldsymbol{\varepsilon}_t$ são as inovações (erros de previsão) e $\boldsymbol{\Sigma}_t$ é a matriz de variância-covariância das inovações. Os detalhes das derivações do filtro de Kalman dado em (3.26) podem ser vistos em (Harvey, 1989; Petris et al., 2009; Kitagawa, 2010; Durbin and S.J.Koopman, 2011; Shumway and Stoffer, 2017), entre outros.

A previsão é muitas vezes a tarefa principal de todo processo, e a estimação do estado completo do sistema é um passo importante e fundamental para a previsão das observações futuras. Em geral, primeiro estima-se a distribuição (ou previsão um passo à frente) do estado do sistema usando as recursões do filtro de Kalman, e em seguida, estima-se a distribuição preditiva da observação usando o estado previsto. Muitas vezes o pesquisador tem especial interesse na estimativa da evolução do sistema para horizontes superiores a 1, isso implica fazer previsões de h passos à frente das observações usando, igualmente, a distribuição preditiva do estado. Esta é uma das vantagens dos modelos de espaço de estados – a de estimar as distribuições preditivas recursivamente de forma direta a partir do filtro de Kalman. Como todas as distribuições relevantes do sistema são Gaussianas, estas são completamente determinadas através das suas médias e variâncias, (Cottet and M, 2003; Gould et al., 2008; Dordonnat et al., 2008; Kitagawa, 2010; Durbin and S.J.Koopman, 2011; Shumway and Stoffer, 2017).

3.5 Estimação por Máxima Verossimilhança

Usa-se $\boldsymbol{\Omega}$ para representar o vetor de parâmetros desconhecidos na média e covariância iniciais $\boldsymbol{x}_0$ e $\boldsymbol{\Sigma}_0$, na matriz de transição $\boldsymbol{\Phi}$, nas matrizes de covariância de estado e da observação $\boldsymbol{R}$ e $\boldsymbol{Q}$ e nas matrizes de coeficientes de entradas $\boldsymbol{\Gamma}$ e $\boldsymbol{\Upsilon}$.

Há, em geral, duas maneiras possíveis de estimar os parâmetros de um modelo de espaço de estados com ou sem covariáveis: (i) a referente via frequentista – abordagem *MLP* (Maximum Likelihood Procedure); (ii) a abordagem Bayesiana. Na perspetiva frequentista, o objeto primário para *MLP* é a função de verossimilhança. Se o modelo contém parâmetros e sua distribuição pode ser expressa como $f(\boldsymbol{y}) = f(\boldsymbol{y}|\boldsymbol{\Omega})$, a log-verossimilhança $\mathcal{L}$ pode ser considerada como uma função de parâmetros em $\boldsymbol{\Omega}$. Explicitamente o vetor $\boldsymbol{\Omega}$ expresso como

$$
\mathcal{L}(\boldsymbol{\Omega}) = \begin{cases} \sum_{t=1}^{n} \log f(\boldsymbol{y}|\boldsymbol{\Omega}) & \text{para dados independentes} \\ \log f(y_1, \cdots, y_n|\boldsymbol{\Omega}) & \text{caso contrário} \end{cases}
$$

é chamado de log-verossimilhança de $\boldsymbol{\Omega}$.

Como a função log-verossimilhança $\mathcal{L}(\boldsymbol{\Omega})$ avalia a qualidade do ajuste do modelo especificado pelo vetor de parâmetros $\boldsymbol{\Omega}$, selecionando $\boldsymbol{\Omega}$ de modo a maximizar $\mathcal{L}(\boldsymbol{\Omega})$, podemos determinar os valores ótimos dos parâmetros do modelo, $f(\boldsymbol{y}|\boldsymbol{\Omega})$. O método de estimação de parâmetros obtido pela maximização da função de verossimilhança ou da função log-verossimilhança é chamado de *método de máxima verossimilhança*, os valores obtidos para os parâmetros dizem-se as *estimativas da máxima verossimilhança* e são denotados por $\hat{\boldsymbol{\Omega}}$.

Definição 11 *A **função de verossimilhança** é a função de densidade considerada como uma função de $\boldsymbol{\Omega}$.*

$$\mathcal{L}(\boldsymbol{\Omega}|\mathbf{x}) = f(\mathbf{x}|\boldsymbol{\Omega})$$

*O **estimador de máxima verossimilhança**,*

$$\hat{\boldsymbol{\Omega}}(\mathbf{x}) = \underbrace{argmax}_{\boldsymbol{\Omega}} \mathcal{L}(\boldsymbol{\Omega}|\mathbf{x})$$

Em determinadas condições de regularidade, a estimativa de máxima verossimilhança tem as seguintes propriedades:

1. **Consistência:** a estimativa $\hat{\boldsymbol{\Omega}}$ de máxima verossimilhança é assintoticamente consistente ($n \to \infty$). Para valores finitos de n pode haver um viés.
2. **Normalidade:** a estimativa $\hat{\boldsymbol{\Omega}}$ é, sob condições muito gerais, assintótico normalmente distribuido com variância mínima $V(\hat{\boldsymbol{\Omega}})$.
3. **Invariância:** a solução de máxima verossimilhança é invariante sob a mudança dos parâmetros – se $\hat{\boldsymbol{\Omega}}$ é o estimador de máxima verossimilhança de $\boldsymbol{\Omega} \in \Delta$ e $h(\boldsymbol{\Omega})$ é uma função bijetiva, $h : \Delta \to \boldsymbol{C}$, então $h(\hat{\boldsymbol{\Omega}})$ é o estimador de máxima verossimilhança de $h(\boldsymbol{\Omega}) \in \boldsymbol{C}$.

   Para mais detalhes, ver Kitagawa (2010).

Existem para tal três formas de funções de verossimilhança dependendo das condições iniciais, Durbin and S.J.Koopman (2011): (i) decomposição do erro de previsão - quando as condições iniciais são conhecidas; (ii) log-verosimilhança difusa - quando alguns dos elementos do vetor de estado são difusos; (iii) verosimilhança - quando elementos do vetor de estado inicial são fixos mas desconhecidos. No caso de um modelo de espaço de estados com ou sem covariáveis e o vetor de variáveis de estado é fixo e de parâmetros desconhecidos, vários autores como Hyndman et al. (2008); De Livera et al. (2011) e Koehler et al. (2012) sugerem a utilização da suavização exponencial, que implica tratar o vetor de estados como um vetor fixo contendo os coeficientes de regressão (caso estejam presentes no modelo) e os demais parâmetros do modelo para optimizar a verossimilhança. As estimativas dos coeficientes de regressão são, nesse contexto, obtidas a partir do valor optimizado do vetor de estado estimado. No entanto, para além de ser um processo computacionalmente complexo, é ainda preferível que o modelo tenha uma única fonte de aleatoriedade.

Para formulações de múltiplas fontes de aleatoriedade e envolvendo as covariáveis, o filtro de Kalman é o algoritmo mais flexível para o tratamento estatístico. Pois, oferece a facilidade de integração dos modelos de espaço de estado simplesmente aumentando o

vetor de estados ou ambos os modelos, o de medição e o de transição. Sob a suposição gaussiana, produz o estimador do vetor de estados juntamente com o seu erro quadrático médio condicionado à informação passada, e é usado para construir a previsão um passo à frente da medição e obter o seu respetivo erro quadrático médio. Devido à independência dos erros de previsão um passo à frente, a verossimilhança pode ser avaliada através da decomposição do erro de previsão, Tommaso (1991).

Considera-se o modelo de espaço de estados dado em (3.13). Quando a série temporal $\{y_1, \cdots, y_n\}$ é dada, a função de densidade conjunta de $\{y_1, \cdots, y_n\}$ especificada pelo modelo (3.13) é denotada por

$$\mathcal{L}(\boldsymbol{\Omega}) = f_n(y_1, \cdots, y_n|\boldsymbol{\Omega})$$

Aplicando repetidamente esta relação

$$f_n(y_1, \cdots, y_n|\boldsymbol{\Omega}) = f_{n-1}(y_1, \cdots, y_{n-1}|\boldsymbol{\Omega}) g_n(y_n|y_1, \cdots, y_{n-1}, \boldsymbol{\Omega}),$$

obtenho a verossimilhança do modelo (3.13) no instante t, que pode ser expressa como um produto de funções de densidade condicional

$$\mathcal{L}(\boldsymbol{\Omega}) = \prod_{t=1}^{n} g_t(y_t|y_1, \cdots, y_{t-1}, \boldsymbol{\Omega}) = \prod_{t=1}^{n} g_t(y_t|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \tag{3.28}$$

Por simplicidade de notação, assume-se que $\mathbf{y}_0 = \emptyset$ (conjunto vazio) e $f_1(y_1|\boldsymbol{\Omega}) \equiv g_1(y_1|\mathbf{y}_0, \boldsymbol{\Omega})$. Tomando logaritmos de $\mathcal{L}(\boldsymbol{\Omega})$, a log-verossimilhança do modelo é obtida como

$$\mathcal{L}^*(\boldsymbol{\Omega}) = \log \mathcal{L}(\boldsymbol{\Omega}) = \sum_{t=1}^{n} \log g_t(y_t|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \tag{3.29}$$

Como $g_t(y_t|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega})$ é uma distribuição condicional de y_t dada a observação $\mathbf{y}_{1:t-1}$, $g_t(y_t|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega})$ é de facto, uma distribuição normal com média $y_{t|t-1}$ e covariância $\boldsymbol{\Sigma}_{t|t-1}$ (para $t = 2, \cdots, n$). Assim, a verossimilhança pode ser expressa em relação a (3.27d) e (3.27e). Isto é

$$g_t(y_t|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{\kappa} \left|\Sigma_t\right|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t\right\} \tag{3.30}$$

Substituindo essa função de densidade em (3.29), a log-verossimilhança do modelo é obtida como

$$\mathcal{L}^*(\boldsymbol{\Omega}) = -\frac{1}{2}\left\{\kappa \mathrm{n} \log(2\pi) + \sum_{t=1}^{n} \log|\boldsymbol{\Sigma}_t| + \sum_{t=1}^{n} \boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t\right\} \tag{3.31}$$

ou seja, ignorando a constante, podemos escrever a verossimilhança como

$$-\mathcal{L}^*(\boldsymbol{\Omega}) \quad = \quad \frac{1}{2}\sum_{t=1}^{n} \log|\boldsymbol{\Sigma}_t| + \frac{1}{2}\sum_{t=1}^{n} \boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t \tag{3.32}$$

As quantidades, $\boldsymbol{\Sigma}_t$ e $\boldsymbol{\varepsilon}_t$, são calculadas recursivamente a partir do filtro de Kalman.

Mais detalhes, ver Harvey (1989); Brockwell and Davis (2002); Helmut (2005); Kitagawa (2010); Tommaso and L.Alessandra (2012); Shumway and Stoffer (2017).

A seguir mostra-se exemplo de um modelo estrutural com dados reais[1]. Trata-se do lucro trimestral por ação da companhia dos Estados Unidos Johnson & Johnson e contemplam 84 trimestres (21 anos) medidos desde o primeiro trimestre de 1960 até o último trimestre de 1980. Para mais detalhes sobre o exemplo, ver o Capítulo 6 de Shumway and Stoffer (2017).

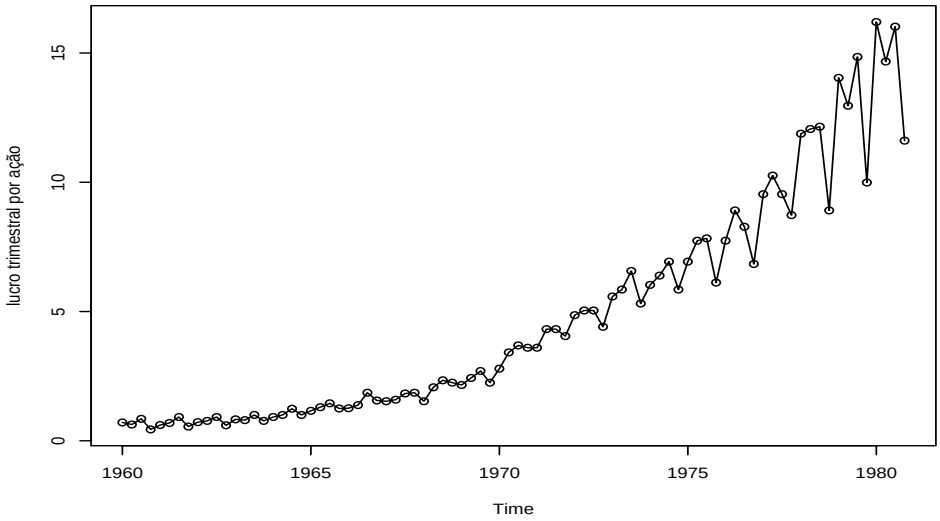


Figura 3.2: Lucro trimestral por ação da companhia Johnson & Johnson, 84 trimestres, primeiro trimestre de 1960 – último trimestre de 1980. Para mais detalhes sobre o exemplo, ver o Capítulo 6 de Shumway and Stoffer (2017).

Conforme o padrão da série temporal y_t, Figura 3.2, o modelo matemático que descreve esse comportamento pode resultar da soma das componentes tendência b_t, sazonalidade s_t e ruído ε_t, ou seja

$$y_t = b_t + s_t + \varepsilon_t$$

A tendência com um aumento exponencial é modelada como

$$b_t = \phi b_{t-1} + \xi_{t1}$$

onde $\phi > 1$ é o parâmetro que caracteriza o aumento na componente tendência. A componente sazonal é definida como $s_t + s_{t-1} + s_{t-2} + s_{t-3} = \xi_{t2}$. Na forma de espaço de estados, as equações de medida e de estado são dadas por

$$\boldsymbol{y}_t = [1, 1, 0, 0].\begin{bmatrix} b_t \\ s_t \\ s_{t-1} \\ s_{t-2} \end{bmatrix} + \varepsilon_t,$$

[1]Estes dados podem ser encontrados no pacote de Stoffer (2016). Os resultados desse exemplo é uma réplica do exemplo que pode ser encontrado no livro de Shumway and Stoffer (2017), Cap.7.

$$\begin{bmatrix} b_t \\ s_t \\ s_{t-1} \\ s_{t-2} \end{bmatrix} = \begin{bmatrix} \phi & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} . \begin{bmatrix} b_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \end{bmatrix} + \begin{bmatrix} \xi_{t1} \\ \xi_{t2} \\ 0 \\ 0 \end{bmatrix},$$

e as perturbações para a observação e o estado, respetivamente, são dadas por

$$\boldsymbol{R} = Var(\varepsilon_t) \quad \text{e} \quad \boldsymbol{Q} = \begin{bmatrix} \sigma^2_{\xi_{t1}} & 0 & 0 \\ 0 & \sigma^2_{\xi_{t2}} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

A Figura 3.3 mostra as componentes da série da companhia Johnson & Johnson estimadas através do erro quadrático médio usando o procedimento de Newton-Rapson através das rotinas de otimização implementadas no R.

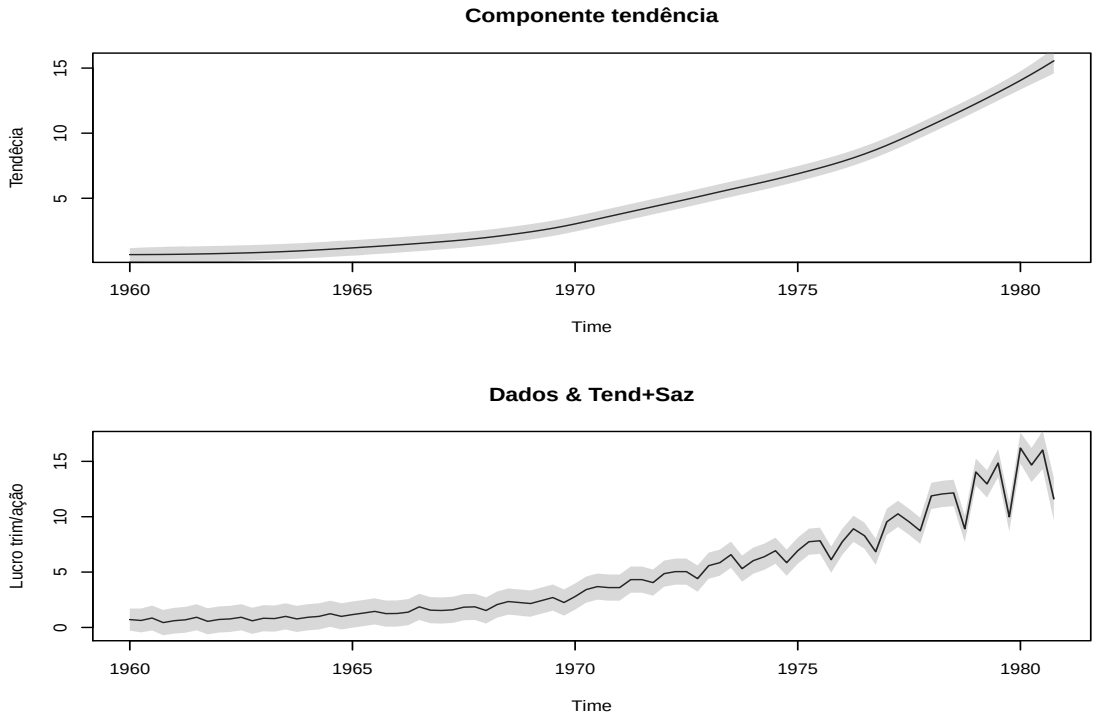


Figura 3.3: As componentes b_t e s_t estimadas sobre o lucro trimestral por ação da companhia Johnson & Johnson, com intervalos de previsão (área em cinza). Para mais detalhes sobre o exemplo, ver o Capítulo 6 de Shumway and Stoffer (2017).

3.6 Critérios de Seleção de Modelos

No âmbito das aplicações, vários modelos são julgados adequados em termos do comportamento dos seus resíduos. Para isso, uma forma de discriminar os modelos competidores é utilizar os critérios de informação, por exemplo, o critério de informação AIC, que leva em conta não apenas a qualidade do ajuste do modelo mas também penaliza a inclusão de parâmetros extras. Significa que um modelo com mais parâmetros pode ter um ajuste melhor mas não

necessariamente ser preferível em termos de critério de informação. A regra básica consiste em selecionar o modelo cujo critério de informação calculado seja mínimo, isto é

$$AIC = \mathscr{L}^*(\hat{\Omega}, \hat{\boldsymbol{x}}_0) + 2p$$

onde $\mathscr{L}^*(\hat{\Omega}, \hat{\boldsymbol{x}}_0)$ é o valor da verossimilhança calculada nas estimativas e para o vetor dos estados iniciais $\hat{\boldsymbol{x}}_0$; $\hat{\Omega}$ é o vetor que contém os parâmetros a estimar e $\hat{\boldsymbol{x}}_0$ é o vetor dos estados iniciais; p denota o número de parâmetros contidos em $\boldsymbol{\Omega}$, (Hyndman et al., 2002; Billah et al., 2005; Anthanasopoulos et al., 2006). Um segundo critério, também bastante utilizado é o critério de informação Bayesiano – BIC. Este critério de informação Bayesiano (BIC), proposto por Schwarz (1978) é dado por:

$$BIC = -\log f(\boldsymbol{x}_n|\boldsymbol{\theta}) + p \log n$$

em que $f(\boldsymbol{x}_n|\boldsymbol{\theta})$ é o modelo escolhido, p é igualmente o número de parâmetros a serem estimados e n é o número de observações da amostra.

3.7 Avaliação do Desempenho do Modelo

No quadro da avaliação do desempenho do modelo estatístico, o modelo é considerado adequado se os erros gerados pelo processo de estimação são não-correlacionados e cumprem com o pressuposto da normalidade. Existem vários diagnósticos que podem ser utilizados quando a questão é verificar a adequação do modelo estimado. Se o foco é a previsão, a análise das características dinâmicas dos erros de previsão é suficiente para avaliar a adequação do modelo, e o gráfico da função de autocorrelação da amostra dos erros de previsão, incluindo o histograma com curva normal são utilizados para a requerida análise. A Tabela 3.3 mostra as métricas mais usuais.

Sigla	Designação	Fórmula
ME	Erro médio	$\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)$
RMSE	Raiz do erro quadrático médio	$\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$
MAE	Erro absoluto médio	$\frac{1}{n}\sum_{t=1}^{n}\lvert y_t - \hat{y}_t\rvert$
MPE	Erro percentual médio	$\frac{100}{n}\sum_{t=1}^{n}\frac{y_t - \hat{y}_t}{y_t}$
MAPE	Erro percentual absoluto médio	$\frac{100}{n}\sum_{t=1}^{n}\left\lvert\frac{y_t - \hat{y}_t}{y_t}\right\rvert$
sMAPE	MAPE simétrico	$\frac{200}{n}\sum_{t=1}^{n}\left\lvert\frac{y_t - \hat{y}_t}{y_t + \hat{y}_t}\right\rvert$

Tabela 3.3: Medidas de precisão da previsão (Hyndman et al., 2008).

As três primeiras medidas pertencem à categoria de medidas dependentes da escala dos

dados. São medidas que se baseiam na variabilidade da previsão quando comparadas com as observações reais. São úteis quando se pretende comparar diferentes métodos aplicados a mesma série temporal. A mais popular e preferida, das três primeiras, é *RMSE* mas, é também a mais sensível a *outliers* em relação a *MAE*. Motivo suficiente que leva muitos autores não recomendarem o seu uso na avaliação de previsões. Quanto as medidas *MPE* e *MAPE*, estas pertencem à categoria de erros percentuais com a vantagem de não dependerem da escala dos dados. No entanto, estas também apresentam problemas quando $y_t = 0$ ou muito próximo de zero, e no caso dos erros serem positivos estes têm uma maior penalização do que para o caso dos erros negativos, por exemplo, *MAPE*. Facto que conduziu ao aparecimento e uso da medida *sMAPE* (Hyndman et al., 2008).

As partes que seguem apresentam as contribuições dessa pesquisa, cuja perspetiva teórica se fundamenta nos modelos de espaço de estados com múltiplas fontes de aleatoriedade e o objeto de estudo tem a ver com os modelos estruturais para a previsão de séries temporais com sazonalidade complexa. Portanto, as abordagens das Secções e Capítulos seguintes estão baseados nesse tipo de modelos.

Parte II

Contribuições para Previsão de Séries Temporais com Sazonalidade Complexa

Capítulo 1

OS MODELOS ESTRUTURAIS COM A INTEGRAÇÃO DAS COVARIÁVEIS

Índice do Capítulo

O principal objetivo desse Capítulo é definir um quadro de modelos estruturais dinâmicos com a integração das covariáveis e que pode admitir a transformação Box–Cox definida por Box and Cox (1964). A nossa proposta é inspirada nos modelos TBATS (Exponential Smoothing State Space Model with Box-Cox Transformation, ARMA Errors, Trend And Seasonal Components), mas, baseada na formulação de múltiplas fontes de aleatoriedade (MSOE), Tabela 1.1. Primeiro, defini-se o modelo estrutural básico que denominamos (SCov) como acrônimo de *Structural model with Covariates*. Este modelo é uma versão do modelo de suavização exponencial homocedástico (BATS). Segundo, adota-se o modelo sazonal trigonométrico proposto por De Livera (2010), mas, na versão de múltiplas fontes de aleatoriedade e desenvolve-se o modelo estrutural trigonométrico que denominamos por TSCov – iniciais de *Trigonometric Structural model with Covariates*. O procedimento de estimação dos parâmetros e a sua implementação computacional são outros propósitos deste Capítulo.

1.1 Especificação do Quadro dos Modelos Estruturais

1.1.1 O Modelo Estrutural Básico com Covariáveis

Seja $\{y_t\}$ a série temporal observada no instante t e z_{kt}, com $\{k = 1, \cdots, K\}$, o conjunto das variáveis de influência externa também observadas no instante t. Adota-se β_k^* para designar os coeficientes de regressão para K covariáveis. Para lidar com os problemas da não linearidade, admite-se que o modelo estrutural é aplicável para uma transformação Box–Cox $y_t^{(\grave{a})}$ da série original y_t, dependendo de um único parâmetro de transformação, $\grave{a}$, feita por Box and Cox (1964), então, o modelo estrutural com covariáveis (SCov) é definido com as componentes não observáveis que usualmente se designam por nível (ℓ_t), tendência (b_t) conforme (3.17) e a sazonalidade (s_t), tal como se apresenta em seguida.

$$y_t^{(\grave{a})} = \begin{cases} \frac{y_t^{\grave{a}}-1}{\grave{a}} & \text{se } \grave{a} \neq 0 \\ \ln y_t & \text{se } \grave{a} = 0 \end{cases} \tag{1.1a}$$

$$y_t^{(\grave{a})} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + \sum_{k=1}^{K} \beta_k^* z_{kt} + \varepsilon_t; \quad \varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2) \tag{1.1b}$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \xi_t; \qquad \xi_t \sim iid\mathcal{N}(0, \sigma_\xi^2) \tag{1.1c}$$

$$b_t = (1-\phi)b + \phi b_{t-1} + \zeta_t; \qquad \zeta_t \sim iid\mathcal{N}(0, \sigma_\zeta^2) \tag{1.1d}$$

$$s_t^{(i)} = s_{t-m_i}^{(i)} + w_{i,t}; \qquad w_{i,t} \sim iid\mathcal{N}(0, \sigma_w^2) \tag{1.1e}$$

onde, $\varepsilon_t, \xi_t, \zeta_t, w_{i,t}$ são processos independentes; $m_1, \cdots, m_T$ representam os períodos sazonais e T padrões sazonais, respetivamente, com $i = 1, \cdots, T$. Os componentes ℓ_t e b_t são o nível local e a tendência de curto-prazo no instante t, respetivamente; b é a tendência de longo prazo e $s_t^{(i)}$ denota a componente sazonal no instante t. Tal como BATS, o modelo SCov é, igualmente, a generalização dos modelos de inovações sazonais tradicionais para permitir vários períodos sazonais, mas, com a vantagem de lidar com os efeitos das covariáveis. No entanto, este modelo não pode lidar com a sazonalidade não-inteira e séries temporais sazonais com efeito duplo de calendário.

1.1.2 O Modelo Estrutural Trigonométrico com Covariáveis

Ao adotar a formulação sazonal trigonométrica da classe de modelos de suavização exponencial (TETS) na versão de múltiplas fontes de aleatoriedade, o modelo estrutural trigonométrico com a integração das covariáveis é definido conforme (1.2). A tabela 1.1 mostra o modelo TSCov como uma extensão do modelo TBATS.

$$y_t^{(\grave{a})} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + \sum_{k=1}^{K} \beta_k^* z_{kt} + \varepsilon_t; \quad \varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2) \tag{1.2a}$$

$$\ell_t = \ell_{t-1} + \phi b_{t-1} + \xi_t; \qquad \xi_t \sim iid\mathcal{N}(0, \sigma_\xi^2) \tag{1.2b}$$

$$b_t = (1-\phi)b + \phi b_{t-1} + \zeta_t; \qquad \zeta_t \sim iid\mathcal{N}(0, \sigma_\zeta^2) \tag{1.2c}$$

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)} \tag{1.2d}$$

$$
\begin{aligned}
s_{j,t}^{(i)} &= s_{j,t-1}^{(i)} cos\lambda_j^{(i)} + s_{j,t-1}^{*(i)} sin\lambda_j^{(i)} + w_{j,t}^{(i)} \\
s_{j,t}^{*(i)} &= -s_{j,t-1}^{(i)} sin\lambda_j^{(i)} + s_{j,t-1}^{*(i)} cos\lambda_j^{(i)} + w_{j,t}^{*(i)}
\end{aligned} \tag{1.2e}
$$

onde, $\varepsilon_t, \xi_t, \zeta_t, w_{j,t}^{(i)}$ são também processos independentes. O parâmetro de amortecimento[1] é limitado por $0 < \phi \leq 1$ para evitar que um coeficiente negativo seja aplicado em b_t. No caso de $\phi = 0$, indicaria a inexistência da tendência na série temporal. Tal como em (1.2), $\lambda_j^{(i)} = \frac{2\pi j}{m_i}$ ($j = 1, 2, \cdots, \dot{k}_i$ e $i = 1, \cdots, T$). Também adotamos $s_{j,t}^{(i)}$ para descrever o nível estocástico da componente sazonal e o crescimento estocástico no nível da i-ésima componente sazonal que é necessário para descrever a mudança na componente sazonal ao longo do tempo t por $s_{j,t}^{*(i)}$. $\dot{k}_i$ é o número de harmônicas necessário para os termos trigonométricos na i–ésima componente sazonal cuja abordagem é equivalente as abordagens de índices sazonais quando $k_i = m_i/2$ para valores pares de m_i e quando $k_i = (m_i - 1)/2$ para valores ímpares de m_i.

	Formulações	
Modelo	Modelo TSCov	Modelo TBATS
Obs.	$y_t^{(\grave{a})} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + \sum_{k=1}^{K} \beta_k^* z_{kt} + \varepsilon_t$	$y_t^{(\grave{a})} = \ell_{t-1} + \phi b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + d_t$
Niv.	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \xi_t$	$\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha d_t$
Tend.	$b_t = (1-\phi) + \phi b_{t-1} + \zeta_t$	$b_t = (1-\phi) + \phi b_{t-1} + \beta d_t$
Saz.	$s_t^{(i)} = \sum_{j=1}^{\dot{k}_i} s_{j,t}^{(i)}$	$s_t^{(i)} = \sum_{j=1}^{\dot{k}_i} s_{j,t}^{(i)}$
	$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} cos\lambda_j^{(i)} + s_{j,t-1}^{*(i)} sin\lambda_j^{(i)} + w_{j,t}^{(i)}$	$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} cos\lambda_j^{(i)} + s_{j,t-1}^{*(i)} sin\lambda_j^{(i)} + \gamma_1^{(i)} d_t$
	$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} sin\lambda_j^{(i)} + s_{j,t-1}^{*(i)} cos\lambda_j^{(i)} + w_{j,t}^{*(i)}$	$s_{j,t}^{*(i)} = -s_{j,t-1}^{(i)} sin\lambda_j^{(i)} + s_{j,t-1}^{*(i)} cos\lambda_j^{(i)} + \gamma_2^{(i)} d_t$

Tabela 1.1: Modelo estrutural trigonométrico com covariáveis e o modelo tbats.

Partimos da suposição de que os termos trigonométricos estocasticamente variantes ao longo do tempo têm uma única fonte de aleatoriedade, isto é, $w_{j,t}^{(i)} = w_{j,t}^{*(i)} \sim iid\mathcal{N}(0, \sigma_w^{2(i)})$ e a variação dinâmica nessa componente (sazonal) é incorporada no sistema de evolução do erro a partir da matriz covariância do estado do sistema, cujo modelo do ruído dessa componente está dado em (1.4).

1.2 Formulação em Espaço de Estados

O processo de estimação implica, primeiro, formular os modelos em espaço de estados, que requer a definição das matrizes dos seus sistemas. O estado é uma descrição de todos os parâmetros ou componentes que precisamos para descrever o sistema atual e realizar a previsão. Assim, utilizamos a seguinte notação (Shumway and Stoffer, 2017): $\boldsymbol{x}_t$ para representar o vetor dos estados não observados no instante t. As matrizes de medição e de

[1] Também chamado de "parâmetro de transição".

transição[2] dos estados, respetivamente, as representamos por $\boldsymbol{A}_t$ e $\boldsymbol{\Phi}$. A matriz $\boldsymbol{\Gamma}$ é a matriz de transição de entradas formada pelos coeficientes de regressão ($\boldsymbol{\beta}_k^*$); $\boldsymbol{z}_t$ é a matriz de controlo de entrada das covariáveis. A relação entre a observação e os estados é descrita pela equação (1.3a). A transição dos estados no instante $t-1$ para o instante t é descrita pela equação (1.3b). O termo $\boldsymbol{\Phi x}_{t-1}$ representa o efeito do passado no estado atual $\boldsymbol{x}_t$. Assim, a formulação em espaço de estados é dada pelas equações do sistema.

$$\boldsymbol{y}_t^{(\grave{a})} = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{\Gamma z}_t + \boldsymbol{\nu}_t \qquad t = 1, 2, \cdots, n \tag{1.3a}$$

$$\boldsymbol{x}_t = \boldsymbol{\Phi x}_{t-1} + \boldsymbol{w}_t \qquad t = 1, 2, \cdots, n \tag{1.3b}$$

onde, $\boldsymbol{\nu}_t$ e $\boldsymbol{w}_t$ são os vetores que contêm os termos de ruído do processo para cada parâmetro no vetor de estado. Supõe-se que os ruídos do processo sejam retirados de uma distribuição normal multivariada de média zero com covariâncias dadas pelas matrize de covariância $\boldsymbol{R}$ e $\boldsymbol{Q}$. Assim, os modelos podem então acomodar qualquer variável regressora, como por exemplo, as variáveis de intervenção representadas por indicadores que usam valores de 0 e 1 e que na prática são usadas para representar determinadas circunstâncias, que inclui mudanças planeadas, eventos incomuns e *outliers*.

A matriz do modelo de medição, $\boldsymbol{A}_t$, contém o parâmetro de transição e a matriz do modelo de transição, $\boldsymbol{\Phi}$, inclui não só o parâmetro de transição, ϕ, como também os ruídos do nível, da tendência e da sazonalidade.

1.2.1 Modelo TSCov em espaço de estados

Como em De Livera et al. (2011), a obtenção das matrizes do sistema (1.3) para o modelo TSCov requer primeiro definir:

O vetor de réplicas de 1's e 0's, $\boldsymbol{a} = (\boldsymbol{a}^{(1)}, \cdots, \boldsymbol{a}^{(T)})$ com $\boldsymbol{a}^{(i)} = (\boldsymbol{1}_{k_i}, \boldsymbol{0}_{k_i})$ e $\tau = 2\sum_{i=1}^{T} \dot{k}_i$. Precisamos definir também a matriz de blocos, $\boldsymbol{B}$, que resulta da soma direta, $\bigoplus$, das matrizes $\boldsymbol{B}_i$, ou seja, $\boldsymbol{B} = \bigoplus_{i=1}^{T} \boldsymbol{B}_i$,

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

onde, $\boldsymbol{C}^{(i)}$ e $\boldsymbol{S}^{(i)}$ são matrizes diagonais de dimensão $\dot{k}_i \times \dot{k}_i$ com os elementos $\cos(\lambda_j^{(i)})$ e $\sin(\lambda_j^{(i)})$, respectivamente, para $j = 1, 2, \cdots, \dot{k}_i$.

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

O vetor do estado do sistema é composto por

[2] Matriz de transição de estado que aplica o efeito de cada parâmetro de estado do sistema no instante $t-1$ para o estado do sistema no instante t (por exemplo, a posição e a velocidade no instante $t-1$ afetam a posição no instante t).

$\boldsymbol{x}_t = \{\ell_t, b_t, s_{1,t}^{(i)}, s_{2,t}^{(i)}, \cdots, s_{\dot{k}_i,t}^{(i)}, s_{1,t}^{*(i)}, s_{2,t}^{*(i)}, \cdots, s_{\dot{k}_i,t}^{*(i)}\}$. Dessa forma, o sistema (1.3) é configurado na forma

$$\boldsymbol{y}_t^{(\grave{\mathbf{a}})} = [1, \phi, \boldsymbol{a}] \cdot \begin{bmatrix} \ell_{t-1} \\ b_{t-1} \\ s_{t-1}^{(i)} \end{bmatrix} + [\beta_1^*, \cdots, \beta_K^*] \cdot \begin{bmatrix} z_{1t} \\ \vdots \\ z_{Kt} \end{bmatrix} + \varepsilon_t,$$

$$\begin{bmatrix} \ell_t \\ b_t \\ s_{j,t}^{(i)} \end{bmatrix} = \begin{bmatrix} 1 & \phi & \mathbf{0}_\tau \\ 0 & \phi & \mathbf{0}_\tau \\ \mathbf{0}_\tau' & \mathbf{0}_\tau' & \mathbf{B} \end{bmatrix} \cdot \begin{bmatrix} \ell_{t-1} \\ b_{t-1} \\ s_{j,t-1}^{(i)} \end{bmatrix} + \begin{bmatrix} \xi_t \\ \zeta_t \\ \tilde{w}_t \end{bmatrix}$$

onde, $\tilde{w}_t$ representa os ruídos w_t e w_t^* da componente sazonal;

$\boldsymbol{R} = \sigma_\varepsilon^2$ e $\boldsymbol{Q}^{(i)} = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \tilde{\boldsymbol{q}}^{(i)} \end{bmatrix}$ onde $\tilde{\boldsymbol{q}}^{(i)}$ é dado por

$$\tilde{\boldsymbol{q}}^{(i)} = \{\tilde{\boldsymbol{\sigma}}_w^{2(1)}, \ldots, \tilde{\boldsymbol{\sigma}}_w^{2(T)}\}, \quad \text{sendo} \quad \tilde{\boldsymbol{\sigma}}_w^{2(i)} = \{\boldsymbol{\sigma}_w^{2(i)}, \boldsymbol{\sigma}_{w^*}^{2(i)}\} \tag{1.4a}$$

e (1.4b)

$$\boldsymbol{\sigma}_w^{2(i)} = \sigma_w^{2(i)} \mathbf{1}_{\dot{k}_i} \tag{1.4c}$$

$$\boldsymbol{\sigma}_{w^*}^{2(i)} = \sigma_{w^*}^{2(i)} \mathbf{1}_{\dot{k}_i} \tag{1.4d}$$

Dessa forma permitimos que o ruído na componente sazonal tenha dupla função: (i) ser a fonte de aleatoriedade para a componente sazonal; (ii) estabelecer a extensão do efeito da aleatoriedade nos coeficientes dos termos trigonométricos estocasticamente variantes ao longo do tempo. Essa forma de modelar o ruído da componente sazonal é similar à metodologia utilizada pelo De Livera et al. (2011) para modelar o parâmetro de suavização, $\boldsymbol{\gamma}$, da componente sazonal.

Suponhamos, por exemplo, disponível uma série temporal com um padrão sazonal, $T = 1$, e o número de harmônicos necessário para os termos trigonométricos seja $\dot{k}_i = 2$. Um modelo estrutural baseado em TSCov com a integração de uma covariável z_{1t} se define como,

$$\begin{aligned}
y_t^{(\grave{\mathbf{a}})} &= \ell_{t-1} + \phi b_{t-1} + s_t^{(1)} + \beta_1^* z_{1t} + \varepsilon_t \\
\ell_t &= \ell_{t-1} + \phi b_{t-1} + \xi_t \\
b_t &= (1-\phi)b + \phi b_{t-1} + \zeta_t \\
s_{1,t}^{(1)} &= s_{1,t-1}^{(1)} cos\lambda_1^{(1)} + s_{1,t-1}^{*(1)} sin\lambda_1^{(1)} + w_{1,t}^{(1)} \\
s_{1,t}^{*(1)} &= -s_{1,t-1}^{(1)} sin\lambda_1^{(1)} + s_{1,t-1}^{*(1)} cos\lambda_1^{(1)} + w_{1,t}^{*(1)} \\
s_{2,t}^{(1)} &= s_{2,t-1}^{(1)} cos\lambda_2^{(1)} + s_{2,t-1}^{*(1)} sin\lambda_2^{(1)} + w_{2,t}^{(1)} \\
s_{2,t}^{*(1)} &= -s_{2,t-1}^{(1)} sin\lambda_2^{(1)} + s_{2,t-1}^{*(1)} cos\lambda_2^{(1)} + w_{2,t}^{*(1)}
\end{aligned}$$

Na forma de espaço de estados, o vetor de estados é dado como

$\boldsymbol{x}_t = (\ell_t, b_t, s_{1,t}^{(1)}, s_{2,t}^{(1)}, s_{1,t}^{*(1)}, s_{2,t}^{*(1)})$

As matrizes do sistema são definidas como se segue:

$$\boldsymbol{A}_t = [1, \phi, 1, 1, 0, 0]; \quad \boldsymbol{\Phi} = \begin{bmatrix} 1 & \phi & 0 & 0 & 0 & 0 \\ 1 & \phi & 0 & 0 & 0 & 0 \\ 0 & 0 & \cos\lambda_1^{(1)} & \sin\lambda_1^{(1)} & 0 & 0 \\ 0 & 0 & -\sin\lambda_1^{(1)} & \cos\lambda_1^{(1)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos\lambda_2^{(1)} & \sin\lambda_2^{(1)} \\ 0 & 0 & 0 & 0 & -\sin\lambda_2^{(1)} & \cos\lambda_2^{(1)} \end{bmatrix};$$

$$\boldsymbol{\Gamma} = \beta_1^*; \quad \boldsymbol{R} = \sigma_\varepsilon^2 \quad \text{e} \quad \boldsymbol{Q}^{(1)} = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_\zeta^2 & 0 \\ 0 & 0 & \tilde{\boldsymbol{q}}^{(1)} \end{bmatrix}, \quad \text{onde,} \quad \tilde{\boldsymbol{q}}^{(1)} = (\sigma_{w_1}^2, \sigma_{w_2}^2).$$

Este modelo é sempre observável, no entanto, só é controlável e estável se $\sigma_\varepsilon^2, \sigma_\xi^2$ e σ_ζ^2 são estritamente positivos, ou seja a matriz Q é definida positiva. De igual modo para a estabilidade do modelo sazonal requer que a variância de cada harmônica seja estritamente positiva, isto é, $\sigma_j^2 > 0$. Essa condição de controlabilidade e estabilidade é fundamental para assegurar que o filtro de Kalman convirja para o estado estável e exponencialmente rápido (Harvey, 1989).

1.2.2 Modelo SCov em espaço de estados

A sua forma de espaço de estados é obtida usando o exposto acima. O estado do sistema é defini-se por $\boldsymbol{x}_t = \{\ell_t, b_t, s_t^{(i)}, s_{t-1}^{(i)}, \cdots, s_{t-(m_i-1)}^{(i)}\}$. Precisamos também um vetor de réplicas de 1's e 0's que definimos como $\boldsymbol{a}^{(i)} = (\boldsymbol{0}_{m_i-1}, 1)$ e $\boldsymbol{B} = \bigoplus_{i=1}^{T} \tilde{\boldsymbol{D}}_i$. Substituindo $2k_i$ por m_i nas matrizes do modelo TSCov, obtém-se:

$$\tau = \sum_{i=1}^{T} m_i, \quad \text{e} \quad \tilde{\boldsymbol{D}}_i = \begin{bmatrix} \boldsymbol{0}_{m_i-1} & \boldsymbol{1} \\ \boldsymbol{I}_{m_i-1} & \boldsymbol{0}'_{m_i-1} \end{bmatrix}$$

onde $\boldsymbol{I}$ é uma matriz rectangular diagonal com elemento 1 na diagonal. O ruído é modelado por $\tilde{\boldsymbol{q}}^{(i)} = (\tilde{\boldsymbol{\sigma}}_w^{2i}, \boldsymbol{0}_{m_i-1})$, sendo $\tilde{\boldsymbol{\sigma}}_w^{2(i)} = \sigma_w^{2(i)} \boldsymbol{1}_{m_i}$.

Os parâmetros a serem estimados, para os dois modelos, inclui: o parâmetro de transição, ϕ, os coeficientes de regressão, $\beta_1^*, \cdots, \beta_K^*$ e as matrizes de covariância da observação e do estado, $\boldsymbol{R}$ e $\boldsymbol{Q}$ cujas componentes são as variâncias associadas à observação, nível, tendência e sazonalidade, respetivamente.

1.3 Filtro de Kalman e Estimação

Por conveniência, antes de se construir o filtro de Kalman com matrizes de covariância ajustadas recursivamente, reproduzimos aqui o modelo de espaço de estados dado em (1.3).

$$\boldsymbol{y}_t^{(\grave{a})} = \boldsymbol{A}_t \boldsymbol{x}_t + \boldsymbol{\Gamma} \boldsymbol{z}_t + \boldsymbol{\nu}_t \qquad t = 1, 2, \cdots, n \tag{1.6a}$$

$$\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{w}_t \qquad t = 1, 2, \cdots, n \tag{1.6b}$$

onde $\boldsymbol{\nu}_t$ e $\boldsymbol{w}_t$ são ruídos brancos assumidos serem não correlacionados,

$$E(\boldsymbol{\nu}_t \boldsymbol{w}_t) = 0 \tag{1.7}$$

e

$$E(\boldsymbol{\nu}_t \boldsymbol{\nu}_t^{'}) = \boldsymbol{R} \tag{1.8a}$$
$$E(\boldsymbol{w}_t \boldsymbol{w}_t^{'}) = \boldsymbol{Q} \tag{1.8b}$$

A matriz $\boldsymbol{z}_t$ contém as variáveis exógenas ou predeterminadas. Pode igualmente conter as variáveis indicadoras. A declaração de que $\boldsymbol{z}_t$ é exógena ou predeterminada significa que $\boldsymbol{z}_t$ não fornece informação sobre o $\boldsymbol{x}_{t+h}$ ou $\boldsymbol{w}_{t+h}$ para $h = 1, 2, \cdots$ além da informação contida em $y_{t-1}^{(\dot{a})}, y_{t-2}^{(\dot{a})}, \cdots, y_1^{(\dot{a})}$. Assim, por exemplo, $\boldsymbol{z}_t$ poderia incluir valores defasados de $\boldsymbol{y}_t^{(\dot{a})}$ ou variáveis que não são correlacionadas com $\boldsymbol{x}_t$ e $\boldsymbol{w}_t$.

Dado (1.7), a equação (1.6b) é tipicamente usada para descrever uma série temporal finita de observações $\{y_1, y_2, \cdots, y_n\}$ e para quais suposições sobre o valor inicial do vetor de estados são necessárias (Hamilton, 1994). Assume-se que $\boldsymbol{x}_t$ é não correlacionado com qualquer realização de $\boldsymbol{\nu}_t$ ou $\boldsymbol{w}_t$:

$$E(\boldsymbol{\nu}_t \boldsymbol{x}_t^{'}) = 0 \quad \text{para } t = 1, 2, \cdots, n \tag{1.9a}$$
$$E(\boldsymbol{w}_t \boldsymbol{x}_t^{'}) = 0 \quad \text{para } t = 1, 2, \cdots, n \tag{1.9b}$$

Importante notar que o estado (verdadeiro) do sistema $\boldsymbol{x}_t$ não é observado diretamente, e o filtro de Kalman fornece um algoritmo para determinar a estimativa $\hat{\boldsymbol{x}}_{t|t-1}$ combinando modelos do sistema e medições ruidosas de certos parâmetros ou funções lineares de parâmetros. As estimativas dos parâmetros de interesse no vetor de estado são, portanto, fornecidas agora por funções de densidade de probabilidade (pdfs), em vez de valores discretos. O filtro de Kalman é baseado em pdfs Gaussianas. Para descrever completamente as funções Gaussianas, precisamos conhecer suas variâncias e covariâncias, e estas são armazenadas na matriz de covariância[3] dada em (1.10b). A subsecção seguinte dedica-se a projeção do filtro de Kalman com matrizes de covariância calculadas recursivamente.

1.3.1 O Filtro de Kalman com Matrizes de Covariância Calculadas Recursivamente

Dado o modelo (1.6) e $y_{1:n} = \{y_1, \cdots, y_n\}$ o conjunto das observações (possivelmente com transformação Box-Cox). Baseado na inferência clássica usando o **Lema 1** conforme a secção 3.4 e de acordo com os princípios que caracterizam um modelo de espaço de estados,

[3]Os termos ao longo da diagonal principal da matriz de covariância são as variações associadas aos termos correspondentes no vetor de estado. Os termos fora da diagonal de (1.10b) fornecem as covariâncias entre os termos no vetor de estado. No caso de um sistema linear unidimensional bem modelado com erros de medição retirados de uma distribuição Gaussiana de média zero, o filtro de Kalman tem se mostrado como o melhor estimador.

Figura 3.1, o preditor de Kalman usado quando $n < t$ e o filtro de Kalman aplicado quando $n = t$ são dados por

$$\begin{aligned}
\hat{\boldsymbol{x}}_{t|t-1} &= \boldsymbol{\Phi}\boldsymbol{x}_{t-1|t-1} && \text{(1.10a)}\\
\boldsymbol{P}_{t|t-1} &= \boldsymbol{\Phi}\boldsymbol{P}_{t-1|t-1}\boldsymbol{\Phi}^{'} + \boldsymbol{Q} && \text{(1.10b)}\\
\boldsymbol{x}_{t|t} &= \hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{K}_t(\boldsymbol{y}_t^{(\grave{a})} - \boldsymbol{A}_t\hat{\boldsymbol{x}}_{t|t-1} - \boldsymbol{\Gamma}\boldsymbol{z}_t) && \text{(1.10c)}\\
\boldsymbol{P}_{t|t} &= \boldsymbol{P}_{t|t-1} - \boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'}\boldsymbol{\Sigma}_t^{-1}\boldsymbol{A}_t\boldsymbol{P}_{t|t-1} && \text{(1.10d)}
\end{aligned}$$

onde

$$\begin{aligned}
\boldsymbol{K}_t &= \boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'}\left[\boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'} + \boldsymbol{R}\right]^{-1} && \text{(1.11a)}\\
\boldsymbol{\varepsilon}_t &= \boldsymbol{y}_t^{(\grave{a})} - \boldsymbol{A}_t\hat{\boldsymbol{x}}_{t|t-1} - \boldsymbol{\Gamma}\boldsymbol{z}_t && \text{(1.11b)}\\
\boldsymbol{\Sigma}_t &= Var(\boldsymbol{\varepsilon}_t) = \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t^{'} + \boldsymbol{R} && \text{(1.11c)}
\end{aligned}$$

O ganho de Kalman é dado por $\boldsymbol{K}_t$, as inovações (ou erros de previsão um passo) são dadas por $\boldsymbol{\varepsilon}_t$ e $\boldsymbol{\Sigma}_t$ é a matriz de variância-covariância das inovações.

Um dos problemas importantes na execução de um filtro de Kalman tem a ver com a correta configuração das matrizes de covariância dos ruídos do sistema, pois o desempenho do filtro de Kalman é altamente afetado pelas matrizes de covariância do sistema. A sua escolha inadequada pode degradar significativamente o desempenho do filtro de Kalman e inclusive tornar o filtro divergente (Mohamed, 1999; Akhlaghi et al., 2017). É bastante comum o uso de procedimentos *ad-hoc* para determinar as matrizes de covariância do sistema, como os filtros convencionais vistos em (Brockwell and Davis, 2002; Durbin and S.J.Koopman, 2011; Shumway and Stoffer, 2017), entre outros, nos quais $\boldsymbol{Q}$ e $\boldsymbol{R}$ entram no filtro de Kalman como matrizes invariantes ao longo do tempo, o que pode ser muito desafiador para o pesquisador.

Para enfrentar esse desafio, um filtro de Kalman com matrizes de covariância calculadas recursivamente é construído com referência às abordagens de Wang (2000); Akhlaghi et al. (2017). Passaremos agora representar as matrizes de covariância da medida e do estado variantes ao longo do tempo como $\boldsymbol{R}_t$ e $\boldsymbol{Q}_t$, respetivamente. O procedimento que aplicamos é baseado nas inovações (a priori e a posteriori) do modelo; estes é que vão influenciar o ajustamento das matrizes de covariância recursivamente até melhorar a precisão da estimativa do estado.

Dado (1.10c), as inovações a posteriori são definidas como

$$\boldsymbol{\eta}_t = \boldsymbol{y}_t^{(\grave{a})} - \boldsymbol{A}_t\boldsymbol{x}_{t|t} - \boldsymbol{\Gamma}\boldsymbol{z}_t \qquad \text{(1.12)}$$

e as estimativas das covariâncias do sistema $\boldsymbol{R}_t$ e $\boldsymbol{Q}_t$ que participam do processo recursivo são obtidas sincronizando estas matrizes às inovações a priori $\boldsymbol{\varepsilon}_t$ e a posteriori $\boldsymbol{\eta}_t$ e projetá-las no mesmo processo recursivo do filtro de Kalman.

Estimativa da matriz de covariância, $\boldsymbol{R}_t$.

De acordo com (1.11c), a matriz de covariância $\boldsymbol{R}_t$ pode ser calculada como

$$\boldsymbol{R}_t = \boldsymbol{\Sigma}_t - \boldsymbol{A}_t \boldsymbol{P}_{t|t-1} \boldsymbol{A}_t^{'} \tag{1.13}$$

onde teoricamente $\boldsymbol{\Sigma}_t$ deve ser definida positiva. No entanto, a equação (1.13) não garante a positividade da matriz estimada, $\boldsymbol{R}_t$, uma vez que resulta da subtração de duas matrizes definidas positivas. Como $\boldsymbol{R}_t = E(\boldsymbol{\nu}_t \boldsymbol{\nu}_t^{'})$, conforme (1.8a), garante-se que $\boldsymbol{R}_t$ seja uma matriz definida positiva ao combinar a covariância com as inovações a posteriori, $\boldsymbol{\eta}_t$, como em (1.14).

$$\begin{aligned} \boldsymbol{\Sigma}_t^* &= E[\boldsymbol{\eta}_t \boldsymbol{\eta}_t^{'}] = E[\boldsymbol{\nu}_t \boldsymbol{\nu}_t^{'}] - \boldsymbol{A}_t \boldsymbol{P}_{t|t-1} \boldsymbol{A}_t^{'} \\ \boldsymbol{R}_t &= E[\boldsymbol{\eta}_t \boldsymbol{\eta}_t^{'}] + \boldsymbol{A}_t \boldsymbol{P}_{t|t-1} \boldsymbol{A}_t^{'} \end{aligned} \tag{1.14}$$

onde $E[\boldsymbol{\eta}_t \boldsymbol{\eta}_t^{'}]$ é usualmente aproximada pela média de $\boldsymbol{\eta}_t \boldsymbol{\eta}_t^{'}$ ao longo do tempo t, média dentro da janela de estimativa móvel (Mohamed, 1999). Em vez disso, um fator de esquecimento $0 < \delta \leq 1$ é usado para estimar de forma adaptativa a covariância, $\boldsymbol{R}_t$.

$$\boldsymbol{R}_t = \delta \boldsymbol{R}_{t-1} + (1-\delta)(\boldsymbol{\eta}_t \boldsymbol{\eta}_t^{'} + \boldsymbol{A}_t \boldsymbol{P}_{t|t-1} \boldsymbol{A}_t^{'}) \tag{1.15}$$

Estimativa da matriz de covariância, $\boldsymbol{Q}_t$.

Conforme a equação (1.6b), a estimativa da covariância de estado, $\boldsymbol{Q}_t$, resulta fazendo

$$w_t = \boldsymbol{x}_t - \boldsymbol{\Phi} \boldsymbol{x}_{t-1} \tag{1.16}$$

A equação (1.10c) é equivalente à

$$\boldsymbol{K}_t \boldsymbol{\varepsilon}_t = \boldsymbol{x}_{t|t} - \hat{\boldsymbol{x}}_{t|t-1} \tag{1.17}$$

Dado que $\boldsymbol{x}_{t|t}$ é o estimador de $\boldsymbol{x}_{t|t-1}$, a partir de (1.17) a estimativa do erro do estado pode ser dada por

$$\hat{\boldsymbol{w}}_t = \boldsymbol{x}_{t|t} - \boldsymbol{\Phi} \hat{\boldsymbol{x}}_{t|t-1} = \boldsymbol{K}_t \boldsymbol{\varepsilon}_t \tag{1.18}$$

e a sua covariância é

$$E(\hat{\boldsymbol{w}}_t \hat{\boldsymbol{w}}_t^{'}) = E[\boldsymbol{K}_t (\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^{'}) \boldsymbol{K}_t^{'}] = \boldsymbol{K}_t E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^{'}) \boldsymbol{K}_t^{'} \tag{1.19}$$

De (1.19), dado (1.11c), a estimativa da covariância do estado é dada por

$$\hat{\boldsymbol{Q}}_t = \boldsymbol{K}_t \boldsymbol{\Sigma}_t \boldsymbol{K}_t^{'} \tag{1.20a}$$

e usamos também um fator de esquecimento δ para ponderar essa estimativa,

$$\boldsymbol{Q}_t = \delta \boldsymbol{Q}_{t-1} + (1-\delta)(\boldsymbol{K}_t \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^{'} \boldsymbol{K}_t^{'}) \tag{1.21}$$

O fluxograma da Figura 1.1 apresenta o processo recursivo do filtro de Kalman. Uma vez

aplicado (o filtro de Kalman) no procedimento computacional descrito na secção (1.4), permite obter as previsões no instante t a um passo dos estados e das observações, incluindo os seus erros quadráticos médios de previsão, o filtro do estado, erro quadrático médio do filtro, a log-verossimilhança negativa, a série de inovações, etc. A condição de os ruídos serem estritamente positivos é necessária para a controlabilidade e estabilidade do modelo; o que assegura o filtro de Kalman convergir para o estado estável.

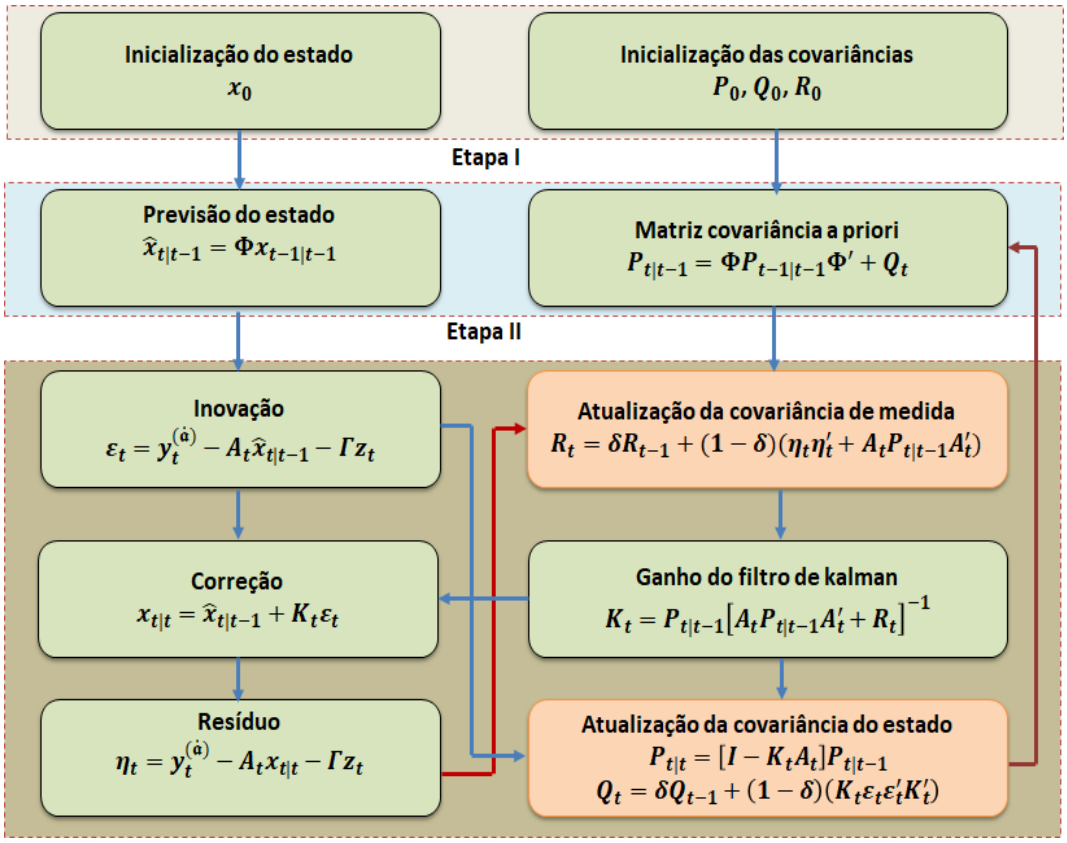


Figura 1.1: Fluxograma do filtro de kalman com efeitos das covariáveis e matrizes de covariância calculadas recursivamente. Figura adaptada conforme Akhlaghi et al. (2017).

1.3.2 Estimativa de Máxima Verossimilhança

Para a estimação do vetor de estados completo do sistema, a abordagem utilizada foi a de obtenção da distribuição condicional $p(\boldsymbol{x}_t|y_t^{(\dot{a})})$ do estado $\boldsymbol{x}_t$ para o conjunto das observações $\mathbf{y}_{1:t-1}$, (Welch and Bishop, 2001; Zarchan and Musoff, 2009; Kitagawa, 2010). Conforme a secção 3.5, a verossimilhança do modelo (1.3) no instante t para a série temporal possivelmente transformada[4] é dada como

$$\mathcal{L}(\boldsymbol{\Omega}) = \prod_{t=1}^{n} g_t(y_t^{(\dot{a})}|y_1, \cdots, y_{t-1}, \boldsymbol{\Omega}) = \prod_{t=1}^{n} g_t(y_t^{(\dot{a})}|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega})$$

[4]Transformação Box-Cox, Box and Cox (1964). Sendo necessária a transformação Box-Cox, as previsões pontuais e os intervalos de previsão podem ser obtidos usando a transformação inversa Box-Cox de quantis apropriados da distribuição de $\hat{\boldsymbol{y}}_{t+h|t}$. Ademais, os intervalos de previsão mantêm a cobertura de probabilidade exigida pela transformação de volta, porque a transformação Box-Cox é monótona crescente.

sendo

$$g_t(y_t^{(\dot{a})}|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{\kappa} \left|\Sigma_t\right|^{-\frac{1}{2}} exp\left\{ -\frac{1}{2}\boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t \right\}$$

e então

$$\begin{aligned} g_t(y_t|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) &= g_t(y_t^{(\dot{a})}|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \left| det\left(\frac{\partial y_t^{(\dot{a})}}{\partial y_t}\right)\right| = g_t(y_t^{(\dot{a})}|\mathbf{y}_{1:t-1}, \boldsymbol{\Omega}) \prod_{t=1}^{n} y_t^{\dot{a}-1} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^{\kappa} \left|\Sigma_t\right|^{-\frac{1}{2}} exp\left\{ -\frac{1}{2}\boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t \right\} \prod_{t=1}^{n} y_t^{\dot{a}-1} \end{aligned}$$

Portanto, a log-verossimilhança é dada por

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

Ao multiplicar essa expressão por -1 e omitir o termo constante, obtenho

$$-\mathcal{L}(\boldsymbol{\Omega}) = \frac{1}{2}\log|\boldsymbol{\Sigma}_t| + \frac{1}{2}\sum_{t=1}^{n} \boldsymbol{\varepsilon}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{\varepsilon}_t - (\dot{a}-1)\sum_{t=1}^{n}\log y_t \tag{1.22}$$

Para o caso em que a transformação Box - Cox não se faz necessária, a verossimilhança é dada como em (3.32).

1.4 Procedimento Computacional de Estimação do Modelo

Como (1.22) e (3.32) são funções não-lineares de parâmetros desconhecidos, o procedimento de estimação consiste em fixar o vetor dos estados iniciais, $\boldsymbol{x}_0$, e depois construir o processo recursivo para a função de log-verossimilhança e aplicar sucessivamente o algoritmo de Newton-Rapson[5] para atualizar as estimativas dos parâmetros até que a log-verossimilhança seja minimizada. O processo de otimização combinado com o filtro de Kalman é condicionado segundo a necessidade ou não da transformação Box-Cox para o conjunto de dados utilizado. Como a formulação sazonal adotada para o modelo estrutural trigonométrico, TSCov, requer a estimação dos $2(\dot{k}_1, \dot{k}_2, \cdots, \dot{k}_T)$ valores iniciais sazonais, nesse trabalho, nós aplicamos o método proposto por De Livera (2010) Cap.3, que se basea na regressão linear múltipla dada por

$$\sum_{i=1}^{T}\sum_{j=1}^{\dot{k}_i} a_j^{(i)}\cos(\lambda_j^{(i)}t) + b_j^{(i)}\sin(\lambda_j^{(i)}t) \tag{1.23}$$

O método consiste em fixar um único harmônico e de forma gradual adiciona-se harmônicos e testar a significância de cada um usando o *teste F*. Por exemplo, considera-se $\dot{k}_t^*$ o número de harmônicos significativos (com $p < 0.001$) para a i-ésima componente sazonal, em seguida, o modelo necessário para os dados é ajustado com $\dot{k}_i = \dot{k}_i^*$ e calcula-se o AIC. Considerando uma componente sazonal de cada vez, o modelo é ajustado repetidamente aos dados de teste,

[5] O algoritmo é aplicado usando as rotinas de otimização implementadas no R, tal como a função optim().

aumentando gradualmente o $\dot{k}_i$, mas mantendo todos os outros harmônicos constantes para cada i, até que o AIC mínimo seja alcançado. O procedimento repete-se para todos os T padrões sazonais até se obter os valores ótimos para cada $\dot{k}_i$.

No quadro computacional para o modelo TSCov, nós agregamos esse método com o filtro de kalman apresentado no fluxograma da Figura 1.1 e o método de Newton-Rapson formando assim um único processo recursivo sistemático, conforme o resumo das principais etapas:

(i) Seleciona-se os valores iniciais para os parâmetros, $\boldsymbol{\Omega}^{(0)}$. Nessa etapa, o parâmetro de transição é configurado com TRUE/FALSE para indicar se o modelo final deve ou não incluir o amortecimento na tendência. Sendo configurado como NULL, os dois casos anteriores são experimentados e a partir do AIC o melhor ajuste é selecionado;

(ii) Executa-se o filtro de Kalman usando os valores iniciais dos parâmetros, $\boldsymbol{\Omega}^{(0)}$, para se obter o conjunto de inovações e covariâncias, isto é $\{\varepsilon_t^{(0)}; t = 1, \ldots, n\}$ e $\{\boldsymbol{\Sigma}_t^{(0)}; t = 1, \ldots, n\}$;

(iii) Executa-se uma iteração do procedimento de Newton-Rapson tomando $-\ln[L(\boldsymbol{\Omega})]$ como função critério para se obter o novo conjunto de estimativas, portanto, $\boldsymbol{\Omega}^{(1)}$. Nessa etapa, o processo de seleção de harmônicas para a componente sazonal entra em jogo;

(iv) Da iteração j (com$j = 1, 2, \ldots$) repete-se a etapa (ii) usando $\boldsymbol{\Omega}^{(j)}$ no lugar de $\boldsymbol{\Omega}^{(j-1)}$ para se obter um novo conjunto de valores de inovações $\{\varepsilon_t^{(j)}\}$ e $\{\boldsymbol{\Sigma}_t^{(j)}\}$. Em seguida, a etapa (iii) é repetida para se obter novas estimativas, $\boldsymbol{\Omega}^{(j+1)}$.

(v) Enquanto se repete a etapa (iii) em (iv), o filtro de Kalman é atualizado com as novas estimativas $\boldsymbol{\Omega}^{(j+1)}$. O processo termina com a estabilização das estimativas ou a verossimilhança.

No caso do modelo estrutural básico com covariáveis, SCov, o processo é basicamente similar, com excepção a etapa (iii) que não inclui o processo de seleção de harmônicos para os termos trigonométricos.

1.5 Seleção do Modelo

Seja $\hat{\boldsymbol{x}}_0$ o vetor dos estados iniciais e $\boldsymbol{\Omega} = (\phi, \boldsymbol{\Gamma}, \boldsymbol{Q}_t, \boldsymbol{R}_t)$ o vetor dos parâmetros a estimar. A seleção do modelo que melhor se ajusta aos dados é feita através do critério de informação de Akaike que definimos como

$$AIC = \mathscr{L}^*(\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{x}}_0) + 2(K^* + \varrho) \tag{1.24}$$

onde $\mathscr{L}^*(\hat{\boldsymbol{\Omega}}, \hat{\boldsymbol{x}}_0)$ é a verossimilhança, K^* denota o número de parâmetros em $\boldsymbol{\Omega}$ e ϱ é o número de estados estimados. Assim, entre os modelos candidatos, o modelo com um AIC mínimo é selecionado.

1.6 Previsão

De acordo com Hyndman (2014), a previsão é necessária em muitas situações: decidir se construir outra usina de geração de energia nos próximos cinco anos exige previsões da procura futura; agendar funcionários em um *call center* na próxima semana exige previsões de volumes de chamadas. As previsões podem ser de longo prazo, médio prazo ou curto prazo. Quaisquer que sejam as circunstâncias ou horizontes temporais envolvidos, a previsão é uma ajuda importante para um planeamento eficaz e eficiente. Distingue-se previsão a um-passo quando se pretende prever apenas a próxima observação e previsão multi-passos quando o objetivo é obter previsões para vários momentos no futuro. Tradicionalmente, a previsão multi-passos é obtida recorrendo a uma estratégia recursiva, na qual é estimado um único modelo para a série temporal, geralmente com base na minimização do erro a um-passo. A previsão para o passo h obtém-se à custa das previsões anteriores, iterando o modelo. Um estudo importante sobre previsão multi-passos podem ser visto em Taieb et al. (2015).

Como as equações do filtro de Kalman podem lidar com observações faltantes de uma maneira natural, nós usamos como primeira estratégia de previsão a denominada estratégia de **Horizonte Crescente do Estado** (Kitagawa, 2010). Ao alargar a amostra de dados $y_1^{(\dot{a})}, \cdots, y_m^{(\dot{a})}$ como valores faltantes para $\boldsymbol{y}_t^{(\dot{a})}$ com $t = n+1, n+2, \cdots$, e aplicar o filtro de Kalman a essa amostra estendida, as previsões são produzidas. Na Secção 1.3, é assumido que $\boldsymbol{z}_t$ não contém nenhuma informação sobre $\boldsymbol{x}_t$ além daquela contida em $\mathbf{y}_{1:t-1}$, isto é:

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

Considera-se a previsão do valor de $\hat{\boldsymbol{y}}_t^{(\dot{a})}$:

$$\hat{\boldsymbol{y}}_{t|t-1}^{(\dot{a})} \equiv E(\hat{\boldsymbol{y}}_t^{(\dot{a})}|\boldsymbol{x}_t, \mathbf{y}_{1:t-1})$$

De acordo com (1.6a),

$$E(\boldsymbol{y}_t^{(\dot{a})}|\boldsymbol{x}_t, \boldsymbol{z}_t) = \boldsymbol{A}_t\boldsymbol{x}_t + \boldsymbol{\Gamma}\boldsymbol{z}_t$$

A partir da lei de projeções iteradas,

$$\hat{\boldsymbol{y}}_{t|t-1}^{(\dot{a})} = \boldsymbol{A}_t E(\boldsymbol{x}_t|\mathbf{y}_{1:t-1}, \boldsymbol{z}_t) + \boldsymbol{\Gamma}\boldsymbol{z}_t = \boldsymbol{A}_t\hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{\Gamma}\boldsymbol{z}_t \tag{1.25}$$

com MSE (Mean Squared Error) dado por

$$E\big[(\boldsymbol{y}_t^{(\dot{a})} - \hat{\boldsymbol{y}}_{t|t-1}^{(\dot{a})})(\boldsymbol{y}_t^{(\dot{a})} - \hat{\boldsymbol{y}}_{t|t-1}^{(\dot{a})})'\big] = \boldsymbol{A}_t\boldsymbol{P}_{t|t-1}\boldsymbol{A}_t' + \boldsymbol{R}_t \tag{1.26}$$

Estratégia de previsão das covariáveis

De acordo com Hyndman et al. (2008), se as covariáveis consistem em variáveis indicadoras, seus valores são conhecidos até um certo ponto futuro do tempo. Ademais, se tais variáveis indicadoras refletem o efeito de intervenções futuras conhecidas e que também ocorreram no passado, então esses valores também são conhecidos. No entanto, quando são desconhecidos, as previsões dos valores futuros das covariáveis são necessárias. Nesse trabalho, adota-se a

abordagem de média móvel exponencialmente ponderada para prever as covariáveis. Uma vez que a matriz $\boldsymbol{z}_t$ contém as covariáveis como, por exemplo, a temperatura (T_t), humidade (H_t), partículas (P_t) e vento (V_t) observadas num dado tempo, estas covariáveis são suavizadas recursivamente através do cálculo da média móvel exponencialmente ponderada. Dessa forma, a previsão de $\boldsymbol{z}_t$ no instante $t+1$ é igual a uma média ponderada entre a observação mais recente $\boldsymbol{z}_t$ e a previsão anterior $\hat{\boldsymbol{z}}_{t|t-1}$, isto é

$$\hat{\boldsymbol{z}}_{t+1|t} = \rho \boldsymbol{z}_t + (1-\rho)\hat{\boldsymbol{z}}_{t|t-1} \tag{1.27}$$

onde $0 \leq \rho \leq 1$ é o parâmetro de suavização que tipicamente é próximo de 1. Essa estratégia é similar a aplicada no trabalho de Dordonnat et al. (2008).

Devido à estrutura Markoviana na dinâmica do estado do sistema e das suposições sobre a independência condicional das observações, as distribuições preditivas podem ser calculadas recursivamente. A atualização do filtro de Kalman no instante t produz $\boldsymbol{x}_{t+1|t}$ e $\boldsymbol{P}_{t+1|t}$, usados para se obter a previsão um passo à frente de $\boldsymbol{y}_{t+1}^{(\dot{a})}$, que é dada por

$$\hat{\boldsymbol{y}}_{t+1|t}^{(\dot{a})} \equiv E(\boldsymbol{y}_{t+1}^{(\dot{a})}|\mathbf{y}_{1:t-1}, \boldsymbol{z}_t) = \boldsymbol{A}_{t+1}\hat{\boldsymbol{x}}_{t+1|t} + \boldsymbol{\Gamma}\hat{\boldsymbol{z}}_{t+1|t} \tag{1.28}$$

e o seu MSE é dado por

$$MSE(\hat{\boldsymbol{y}}_{t+1|t}^{(\dot{a})}) = \boldsymbol{A}_{t+1}\boldsymbol{P}_{t+1|t}\boldsymbol{A}_{t+1}^{'} + \boldsymbol{R}_{t+1} \tag{1.29}$$

Previsão h passos à frente

A previsão obtida em (1.28) é uma previsão de amostra finita de $\boldsymbol{y}_t^{(\dot{a})}$ com base em $\boldsymbol{z}_t$ e $\mathbf{y}_{1:t-1} \equiv (y_{t-1}, y_{t-2}, \cdots, y_1, z_{t-1}, z_{t-2}, \cdots, z_1)$. Ademais, essas previsões dependem do número de harmônicos $\dot{k}_i$ necessário para a componente sazonal i. Para calcular as previsões exatas do horizonte h da amostra finita, primeiro calcula-se a previsão, para o horizonte h, do estado do sistema,

$$\hat{\boldsymbol{x}}_{t+h|t} = \boldsymbol{\Phi}\hat{\boldsymbol{x}}_{t+h-1|t} \tag{1.30}$$

com MSE dado por

$$MSE(\hat{\boldsymbol{x}}_{t+h|t}) = \boldsymbol{P}_{t+h|t} = \boldsymbol{\Phi}\boldsymbol{P}_{t+h-1|t}\boldsymbol{\Phi}^{'} + \boldsymbol{Q}_{t+h} \tag{1.31}$$

Finalmente, a previsão h passos à frente de $\boldsymbol{y}_t^{(\dot{a})}$ e o seu MSE são dados por

$$\hat{\boldsymbol{y}}_{t+h|t}^{(\dot{a})} \equiv E(\boldsymbol{y}_{t+h}^{(\dot{a})}|\boldsymbol{y}_{1:t}^{(\dot{a})}) = \boldsymbol{A}_{t+h}\hat{\boldsymbol{x}}_{t+h|t} + \boldsymbol{\Gamma}\hat{\boldsymbol{z}}_{t+1|t} \tag{1.32a}$$

$$MSE(\hat{\boldsymbol{y}}_{t+h|t}^{(\dot{a})}) = \boldsymbol{A}_{t+h}\boldsymbol{P}_{t+h|t}\boldsymbol{A}_{t+h}^{'} + \boldsymbol{R}_{t+h} \tag{1.32b}$$

Todas as realizações de $\boldsymbol{y}_{t+h|t}^{(\dot{a})}$ passam por $\boldsymbol{y}_t^{(\dot{a})}$, tornando as previsões dependentes da série disponível. Assim, os intervalos de previsão são obtidos de forma direta. Como os erros

de previsão h passos à frente são Gaussianos, nós geramos os intervalos de previsão da taxa de cobertura nominal de 95% para $\boldsymbol{y}_{t+h}$ como

$$IP = \hat{\boldsymbol{y}}^{(\grave{a})}_{t+h|t} \pm 1.96\sqrt{MSE(\hat{\boldsymbol{y}}^{(\grave{a})}_{t+h|t})}, \tag{1.33}$$

o que significa que do modelo a estimar espera-se que o intervalo de previsão cubra todos os valores futuros com uma probabilidade de 0.95.

Capítulo 2

ANÁLISE EMPÍRICA

Índice do Capítulo

Este Capítulo ilustra a aplicação dos modelos descritos no Capítulo 1 da Parte II a casos de estudo constituídos por séries temporais do mundo real. É dada ênfase ao modelo TSCov. Ademais, o Capítulo descreve também os critérios de inicialização dos parâmetros dos modelos para o cálculo da previsão um passo e previsão multi–passos à frente para cada caso de estudo.

Considera-se $\{y_1, y_2, \cdots, y_{n-h}, y_{n-h+1}, \cdots, y_n\}$ a série temporal de dimensão n, e define-se h como o horizonte temporal de previsão. Para o teste e validação dos modelos, a primeira etapa envolve a partição das séries temporais em duas partes:

$$\underbrace{y_1, y_2, \cdots, y_{n-h}}_{\text{Série de teste}}, \overbrace{y_{n-h+1}, \cdots, y_n}^{\text{Série de validação}}$$

As séries de teste são usadas para calcular a previsão um passo a frente do vetor dos estados $\boldsymbol{x}_t$ e das observações $\boldsymbol{y}_t$, que inclui também o cálculo dos demais componentes do filtro de Kalman, tais como o erro quadrático médio, $\boldsymbol{P}_t$, as inovações $\boldsymbol{\varepsilon}_t$ e a sua covariância $\boldsymbol{\Sigma}_t$, o ganho do filtro de Kalman K_t, etc. As séries de validação são usadas para comparar as previsões com os valores observados.

2.1 Delineamentos Computacionais

Na projeção computacional dos modelos SCov e TSCov os objetos principais utilizados são o *DADO* e a *FUNÇÃO*. As funções scov() e tscov() são uma adaptação das funções bats() e tbats(), cujos seus parâmetros são configurados como:

- O vetor das observações, $\boldsymbol{y}_t$, (possivelmente com transformação Box-Cox, $\boldsymbol{y}_t^{(\grave{a})}$), é codificado como numérico e configurado como uma ts ou msts. Pode também ser codificado como matrix para permitir que o modelo lide com as séries multivariadas;
- O vetor das covariáveis, $\boldsymbol{z}_t$, é codificado como input;
- O parâmetro, use.damped.trend, é configurado com valores TRUE/FALSE para indicar se o modelo final deve ou não incluir a tendência amortecida. Se é codificado como NULL, os dois casos anteriores são experimentados e a partir do AIC o melhor ajusto é selecionado. Assim, o parâmetro de amortecimento, ϕ, é inicializado mediante um condicionamento configurado da seguinte forma: supor a priori que o modelo inclui a tendência amortecida, então, o parâmetro de amortecimento é fixado em $\phi = 0.999$, se não for o caso, o parâmetro de transição é fixado em $\phi = 1$.
- Os termos de ruído do processo para cada parâmetro no vetor de estado e do ruído de medição para cada observação no vetor de observação são estritamente positivos para garantir a controlabilidade e estabilidade do modelo.
- O parâmetro seasonal.periods é usado para especificar os períodos sazonais presentes na série temporal.
- No processo de estimação, o procedimento é executado de forma automática. O melhor modelo é selecionado de acordo com o critério de AIC descrito na secção 1.5.

Os modelos são configurados para serem flexíveis ao cálculo das previsões com ou sem a presença das covariáveis. Se não há necessidade de incluir covariáveis no modelo, precisamos somente configurar os parâmetros input = 0 e $\boldsymbol{\Gamma} = 0$ na função genérica Kf.tscov() que projeta o filtro de Kalman.

Quatro casos de estudo são apresentados nas secções seguintes. Nos primeiros três casos, o modelo TSCov é aplicado com a integração das covariáveis. No quarto caso de estudo, o modelo TSCov é aplicado sem a integração das covariáveis. Dois casos de estudo adicional são apresentados nos Apêndices B.1 e B.2. O primeiro tem a ver com a aplicação do modelo básico SCov a dados de mortalidade em Los Angels, o segundo tem a ver com o modelo TSCov aplicado a dados de níveis de concentração de NO_2 em Entre-Campos de Lisboa.

As estimativas dos parâmetros resultam minimizando o MSE dos erros de previsão um passo à frente. A otimização é realizada utilizando a função optim sob o método L-BFGS-B, uma modificação de memória limitada do algoritmo quasi-Newton, BFGS. O processo geral de estimação é realizado conforme descrito na secção 1.4. Importa sublinhar que com exceção o parâmetro de transição ϕ, a inicialização dos parâmetros, $\boldsymbol{R}_0$, $\boldsymbol{Q}_0$, $\boldsymbol{x}_0$, $\boldsymbol{P}_0$ e $\boldsymbol{\Gamma}_0$, não é um processo automático, depende das caraterísticas do conjunto de dados utilizado. Em cada caso de estudo apresentam-se os respetivos detalhes sobre a inicialização do processo, e, em termos comparativos reportamos alguns resultados dos modelos BATS e TBATS para confrontar os

resultados dos modelos SCov e TSCov. As métricas utilizadas para a avaliação do desempenho dos modelos estimados são as usuais apresentadas na Tabela 3.3, com destaque para RMSE e MAPE.

Principais Funções Implementadas no Ambiente R, ver Apêndice B.4

Função		Definição
`PhiMatrix()`	...	projeta a matriz de transição dos estados, $\boldsymbol{\Phi}$
`Amatrix()`	...	projeta a matriz do modelo de observação, $\boldsymbol{A}_t$
`Qmatrix()`	...	calcula a matriz covariância do estado $\boldsymbol{Q}_t$
`Kf.tscov()`	...	projeta o filtro de Kalman
`verossim()`	...	calcula a verossimilhança
`TSCovFit()`	...	projeta a otimização das estimativas do modelo
`tscov()`	...	projeta o melhor modelo
`TSCovFore()`	...	calcula as previsões h passos à frente

Tabela 2.1: Funções implementadas com o ambiente R

Funções Internas e Pacotes do `R` Utilizados

optim(), função que fornece algoritmos para otimizações de uso geral; minimiza uma função, variando os seus parâmetros. O primeiro argumento, dentre outros, de optim são os valores iniciais para os parâmetros a serem optimizados; o segundo argumento é a função a ser minimizada (maximizada); o terceiro argumento é o método a ser utilizado, etc.

hdr.cde(), função do pacote Hyndman (2018). Calcula as regiões de densidade alta para a densidade condicional estimada. Os principais parâmetros desta função são a densidade condicional estimada; o vetor das probabilidades de cobertura para as regiões de densidade alta.

densityplot(), função do pacote Sarkar (2008). Calcula as estimativas de densidade de Kernel. O pacote é um poderoso e elegante sistema de visualização de dados.

forecast(), função do pacote Hyndman and Khandakar (2008). Este pacote permite estimar os modelos BATS e TBATS e a função permite obter as previsões pontuais e os intervalos de previsão.

astsa, pacote que permite acessar os dados sobre a mortalidade cardiovascular em Los Angels - Estados Unidos de América (Stoffer, 2016).

2.2 Primeiro Caso de Estudo: dados com sazonalidade dupla

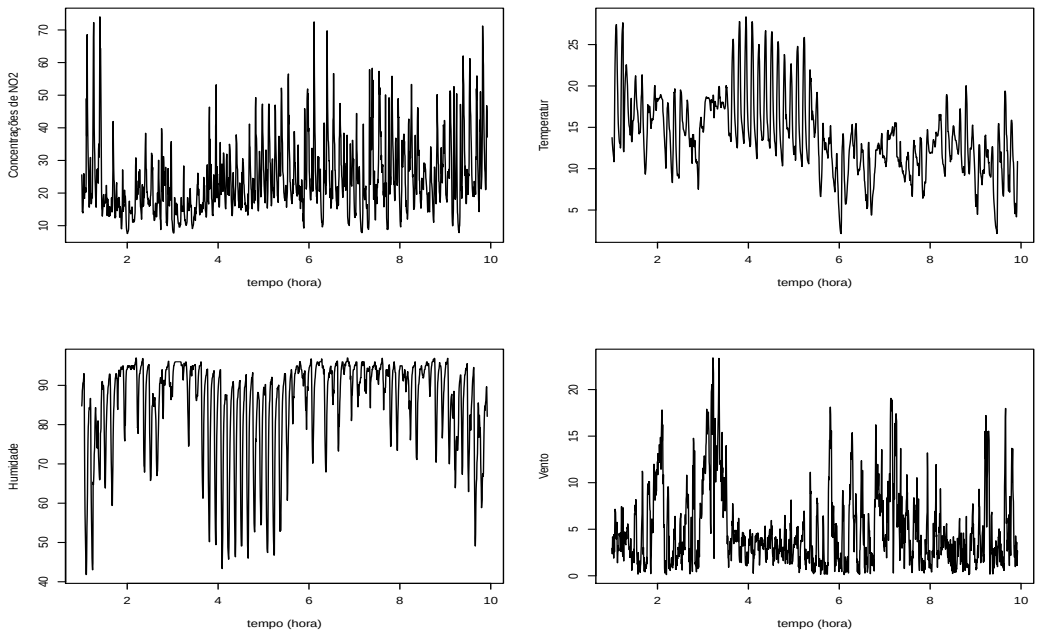


Figura 2.1: Dados horários de NO_2 observados em 2014 entre 1 de Outubro e 31 de Dezembro na estação de Paredes, Portugal. A temperatura, Humidade e Vento são as covariáveis, observadas igualmente de hora em hora (QualAr, 2015).

Para esse caso de estudo é utilizado o conjunto de dados referentes à medições horárias dos níveis de NO_2 (Dióxido de Nitrogênio), Figura 2.1. As medições foram efetivadas em 2014 entre 1 de Outubro e 31 de Dezembro a partir da estação de Paredes, Portugal. Os dados são obtidos a partir da base de dados online sobre qualidade do ar da Agência Portuguesa do Ambiente, cuja missão é propor, desenvolver e monitorizar as políticas públicas para o ambiente e o desenvolvimento sustentável. O banco de dados sobre a qualidade do ar, QualAr (2015), fornece medições por hora, resultantes de atividades de monitoramento, para vários poluentes, incluindo o NO_2.

A série temporal exibe dois padrões sazonais: um padrão diário com período 24 e um padrão semanal com período 168, Figura 2.2. As covariáveis consideradas são a temperatura, T_t, a humidade, H_t, e o vento, V_t, também observadas em intervalos de uma hora, Figura 2.1. A série das concentrações dos níveis de NO_2 é denotada por N_t.

Inicia-se uma análise prévia dos dados recorrendo a correlação cruzada. A estratégia é ajustar um modelo TBATS aos dados de NO_2 para se obter os resíduos que são utilizados para determinar a correlação cruzada com as séries de temperatura, humidade e vento. Os resultados exibidos na Figura 2.3 indicam que as correlações mais fortes ocorrem em 12h:00 de atraso com a temperatura do ar (T_{t-12}), 2h:00 de atraso com a humidade relativa (H_{t-2}) e vento corrente (V_t). O melhor modelo estimado integra as três covariáveis.

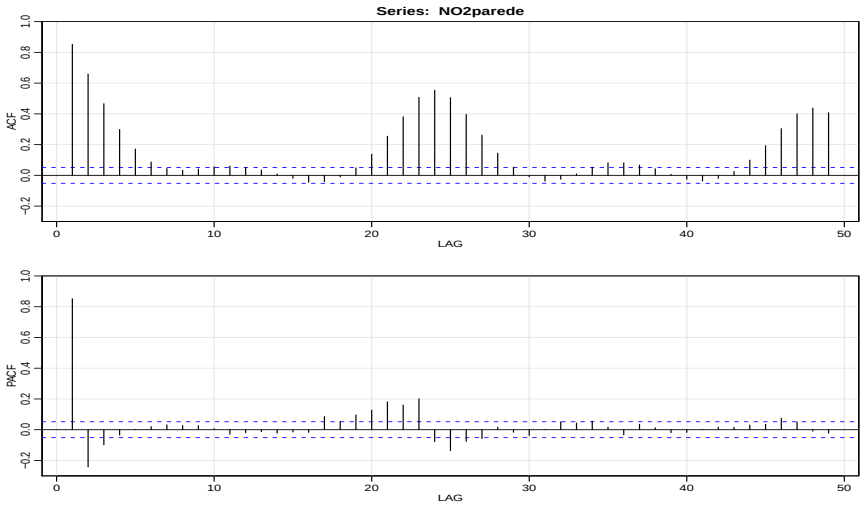


Figura 2.2: Correlograma dos níveis de concentração do NO_2 em Paredes, Portugal.

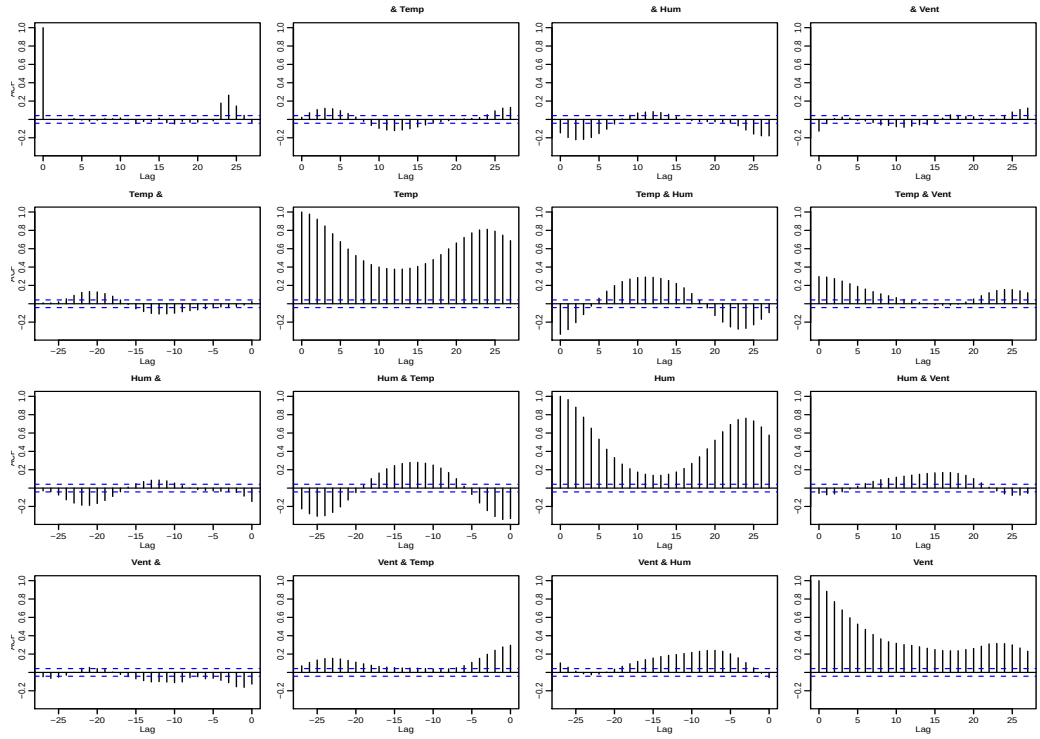


Figura 2.3: Correlação cruzada entre a série NO_2 branqueada (resíduos de **tbats**), a série de temperatura e a série de humidade relativa.

2.2.1 Estimação e previsão um passo à frente

Estimam-se três modelos: o modelo TSCov (com covariáveis reais), o modelo TSCov (com covariáveis previstas) e o modelo TBATS. A série de teste contém 1488 observações e a série de validação com 720 observações. As previsões obtêm-se considerando os valores observados das covariáveis e os valores previstos de acordo com o que está descrito secção 1.6 da Parte II.

Valores iniciais para a estimação do modelo TSCov (com covariáveis reais). A média

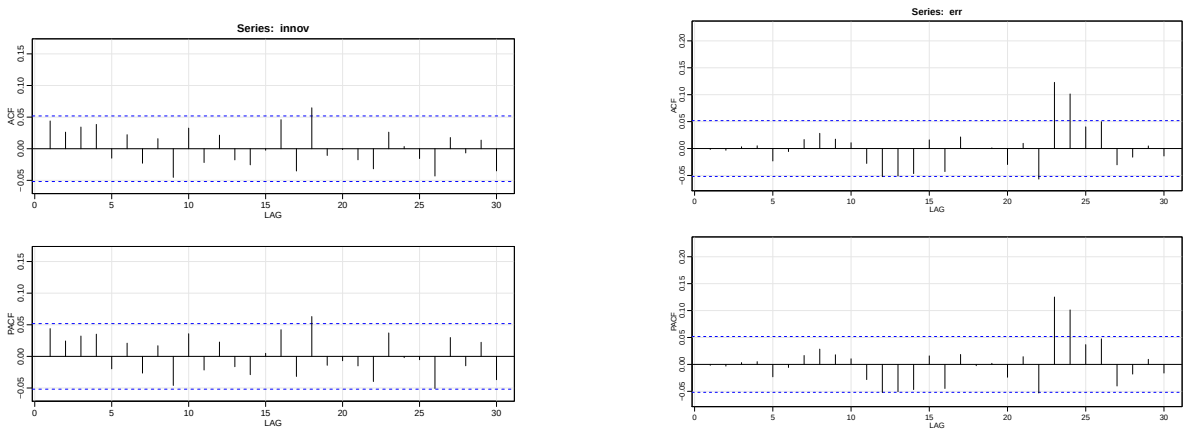
e a covariância do estado do sistema são inicializadas por $\boldsymbol{x}_0 = 0.7$ e $\boldsymbol{P}_{0i} = 4$, com $i = 16$, respetivamente. A covariância da observação é fixada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 10^{-8}$ e as variâncias do vetor de estados são inicializadas por $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^{2(i)}\} = \{0.004, 0.17, 0.0001, 0.0001\}$, com $i = 4$. Os coeficientes de regressão são inicializados por $\{\beta_1^*, \beta_2^*, \beta_3^*\} = 0.1$ e o fator de esquecimento é fixado em $\delta = 0.999$.

Valores iniciais para a estimação do modelo TSCov (com covariáveis previstas). A média e a covariância do estado do sistema são inicializadas por $\boldsymbol{x}_0 = 0.7$ e $\boldsymbol{P}_{0i} = 3.1$, com $i = 16$, respetivamente. A covariância da observação é fixada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 10^{-8}$ e as variâncias do vetor de estados são inicializadas por $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^{2(i)}\} = \{0.004, 0.17, 0.0001, 0.0001\}$, com $i = 4$. Os coeficientes de regressão são inicializados por $\{\beta_1^*, \beta_2^*, \beta_3^*\} = 0.01$ e o fator de esquecimento é fixado em $\delta = 0.8$.

Valores iniciais para o modelo TBATS. O modelo TBATS é automático, implementado no pacote *forecast* do R, (Razbash and Hyndman, 2018). Portanto, não exige a especificação dos valores iniciais.

As estimativas dos parâmetros para os dois modelos estimados, TSCov e TBATS, estão apresentadas na Tabela 2.2. A componente irregular para o modelo TBATS é modelada com um processo ARMA(2,3). Constata-se que o coeficiente associado ao β_2^* que corresponde a humidade relativa não é significativo.

O diagnóstico dos resíduos gerados a partir dos modelos TSCov e TBATS está apresentado na Figura 2.4. Para o modelo TSCov, Figura 2.4a, o correlograma exibe uma correlação significativa positiva no *lag 18*, no entanto, o teste de Box-Ljung sobre a independência dos resíduos fornece um valor de Qui-quadrado igual a 29.154, com 19 graus de liberdade e *p-valor* = 0.164, permitindo não rejeitar a hipótese nula de que os resíduos são independentes. Para o modelo TBATS, Figura 2.4b, o teste de Ljung-Box fornece um Qui-quadrado = 19.443 com 18 graus de liberdade e *p-valor* = 0.265; também não rejeitamos a hipótese nula de que os resíduos do modelo são independentes.



(a) (b)

Figura 2.4: Correlograma dos resíduos da previsão um passo à frente dos níveis de concentração de NO_2. (a) modelo TSCov; (b) modelo TBATS.

O gráfico Q-Q normal dos resíduos do modelo estimado com TSCov, Figura 2.5, mostra a

partida da normalidade nas duas caudas, devido a possível presença de *outliers* que ocorreram nos dias 2 e 3 de Outubro, 5 de Novembro e 5, 19 e 31 de Dezembro de 2014. A Figura 2.6 mostra a previsão um passo à frente obtidas a partir dos modelos TSCov e TBATS incluindo os valores observados. Os erros de previsão um passo à frente para os dois modelos estimados (TSCov com covariáveis reais e o modelo TBATS) estão apresentados na Tabela 2.3.

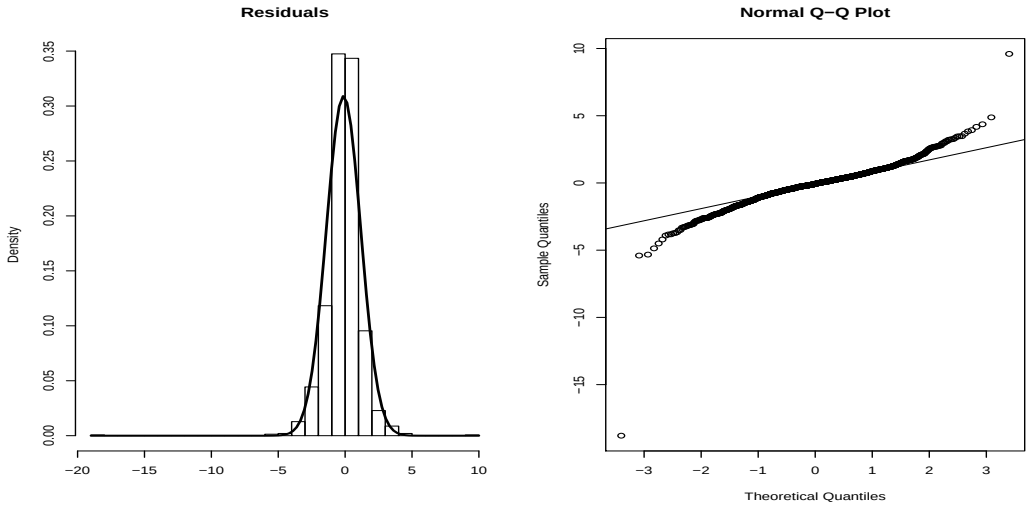


Figura 2.5: Histograma e o Q-Q normal dos resíduos do modelo TSCov estimado, referente aos níveis de concentração de NO_2 em Paredes, Portugal.

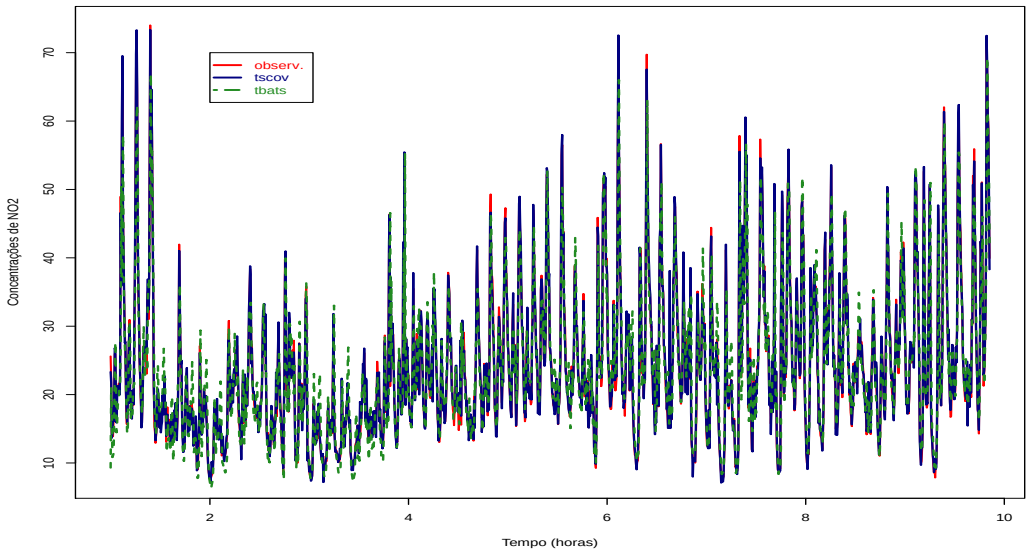


Figura 2.6: Valores observados dos níveis de concentração de NO_2 e os valores ajustados a partir dos modelos TSCov e TBATS.

Parâmetro	MLE (TSCov)	E.Padrão Ass.	MLE (TBATS)
β_1^*	-0.724	0.437	—
β_2^*	-0.007	0.013	—
β_3^*	0.257	0.035	—
$\dot{a}$	—	—	—
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_w^2	$\{0.002; 0.019\}$	$\{0.012; 0.022\}$	—
$\sigma_{w^*}^2$	$\{0.012; -0.048\}$	$\{0.164; 0.011\}$	—
γ_1	—	—	$\{-1 \times 10^{-4}; -2 \times 10^{-6}\}$
γ_2	—	—	$\{-3 \times 10^{-5}; 3 \times 10^{-5}\}$

Tabela 2.2: Estimativas dos parâmetros e os respetivos erros-padrão obtidos partir do modelo TSCov. As estimativas dos parâmetros obtidos a partir do modelo TBATS estão apresentadas na quarta coluna.

O modelo de previsão que descreve a dinâmica dos níveis de concentração de NO_2 em Paredes - Portugal, para o período em estudo, pode ser expresso como:

$$
\begin{aligned}
N_t &= \ell_{t-1} + 0.95b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} - 0.72T_t + 0.25V_t + \varepsilon_t \\
\ell_t &= \ell_{t-1} + 0.95b_{t-1} + \xi \\
b_t &= 0.95b_{t-1} + \zeta_t \\
s_t &= \sum_{j=1}^{5} s_{j,t} \quad \text{(padrão sazonal diário)} \\
s_{j,t} &= s_{j,t-1}\cos\left(\frac{2\pi jt}{24}\right) + s_{j,t-1}^{*}\sin\left(\frac{2\pi jt}{24}\right) + w_{j,t}^{(i)} \\
s_{j,t}^{*} &= -s_{j,t-1}\sin\left(\frac{2\pi jt}{24}\right) + s_{j,t-1}^{*}\cos\left(\frac{2\pi jt}{24}\right) + w_{j,t}^{*(i)} \\
s_t &= \sum_{j=1}^{3} s_{j,t} \quad \text{(padrão sazonal semanal)} \\
s_{j,t} &= s_{j,t-1}\cos\left(\frac{2\pi jt}{168}\right) + s_{j,t-1}^{*}\sin\left(\frac{2\pi jt}{168}\right) + w_{j,t}^{(i)} \\
s_{j,t}^{*} &= -s_{j,t-1}\sin\left(\frac{2\pi jt}{168}\right) + s_{j,t-1}^{*}\cos\left(\frac{2\pi jt}{168}\right) + w_{j,t}^{*(i)} \quad \text{onde} \\
\varepsilon_t &\sim \mathcal{N}(0, 2.294) \\
\xi_t &\sim \mathcal{N}(0, 1.596) \\
\zeta_t &\sim \mathcal{N}(0, 0.91) \\
w_{1,t} &\sim \mathcal{N}(0, 0.002); \quad w_{2,t} \sim \mathcal{N}(0, 0.012)
\end{aligned}
$$

$$w^*_{1,t} \sim \mathcal{N}(0, 0.019); \quad w^*_{2,t} \sim \mathcal{N}(0, 0.048)$$

Este modelo precisou de $\dot{k}^*_1 = 5$ harmônicos significativos para os termos trigonométricos do padrão sazonal diário com periodicidade 24 e $\dot{k}_2 = 3$ para o padrão sazonal semanal com periodicidade 168. O vetor dos estados estimado é de dimensão 16.

Modelo	ME	RMSE	MAE	MPE	MAPE
TSCov	0.013	3.486	3.008	-1.556	11.429
TBATS	0.195	4.973	3.324	-1.568	13.923

Tabela 2.3: Erros de previsão um passo à frente obtidos pelos modelos TSCov (com covariáveis reais) e TBATS sobre os níveis de concentração de NO_2 em Paredes, Portugal.

2.2.2 Previsão multi-passos

Calcula-se a previsão de 24 passos à frente, Figura 2.7. A Figura 2.8 mostra os intervalos de previsão de 95% gerados pelos modelos TSCov e TBATS incluindo os valores observados. Para esse caso de estudo, os intervalos gerados pelo modelo TSCov também mostram-se mais regulares ao longo do horizonte de previsão em relação os intervalos de previsão obtidos pelo modelo TBATS.

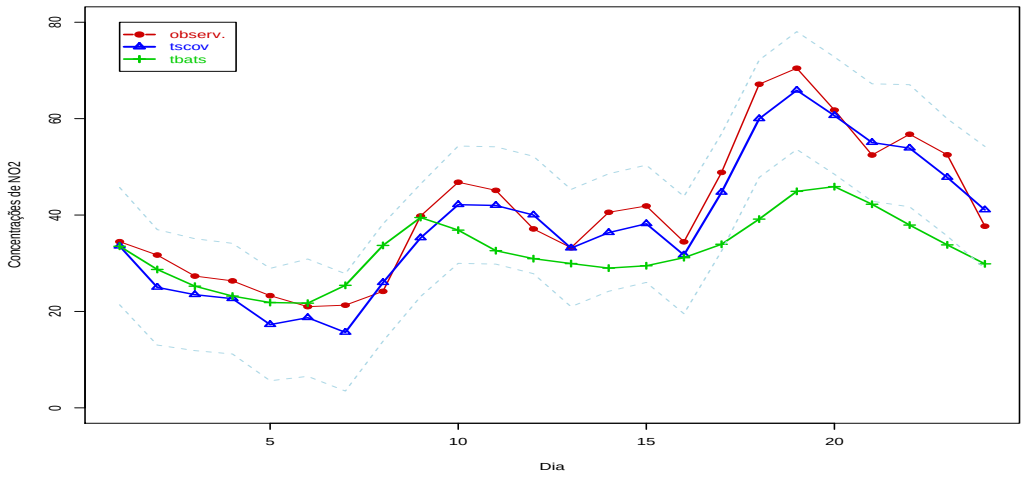


Figura 2.7: Valores observados e as previsões de 24 passos à frente dos níveis de concentração de NO_2 em Paredes, Portugal.

Conforme os resultados obtidos nesse caso de estudo, pode-se concluir que o modelo TSCov tem melhor classificação, quer em termos gráficos como em termos de medidas de precisão da previsão quando comparado com os resultados obtidos pelo modelo TBATS.

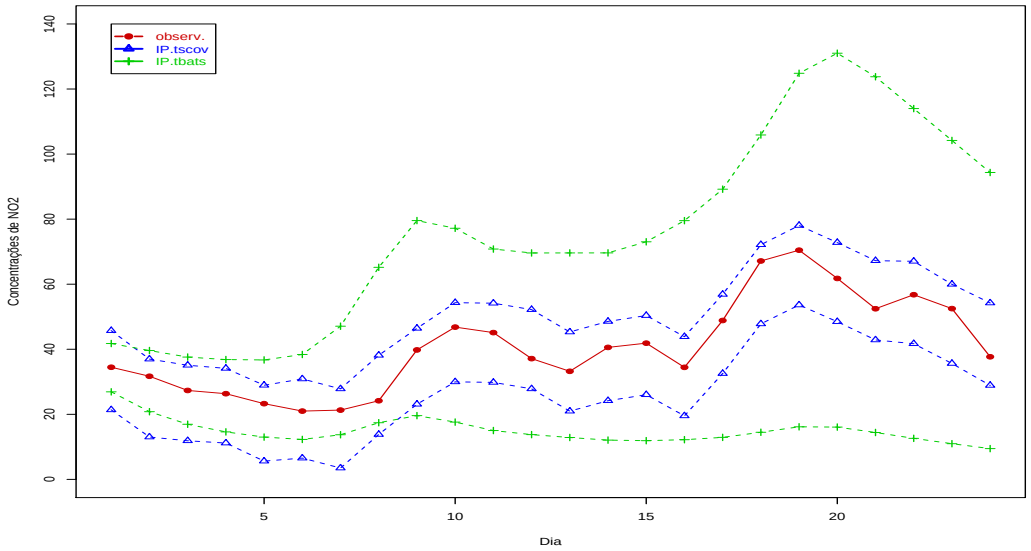


Figura 2.8: Valores observados e os intervalos de previsão de 95% gerados pelos modelos TSCov e TBATS sobre os níveis de concentração do NO_2 em Paredes, Portugal.

Horizonte	TSCov (cov. reais)		TSCov (cov. previstos)		TBATS	
	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

Tabela 2.4: Medidas de precisão de previsão até 24 passos à frente sobre os níveis de concentração de NO_2 em Paredes - Portugal, gerados pelos modelos TSCov e TBATS.

2.3 Segundo Caso de Estudo: dados com período sazonal inteiro

Trata-se de um conjunto de dados de 508 observações semanais sobre a Mortalidade[1] Cardiovascular por Poluição e Temperatura na cidade de Los Angels – Estados Unidos de América, entre 1970 e 1979, primeiro painel da Figura 2.9.

A série temporal exibe um padrão sazonal anual (aditiva) com periodicidade inteira 52 e é altamente correlacionada entre si, Figura 2.10. Ademais, exibe uma tendência aditiva descendente. Designamos a série semanal de Mortalidade Cardiovascular por $\boldsymbol{M}_t$, a série de Temperatura por $\boldsymbol{T}_t$ e a série de Partículas por $\boldsymbol{P}_t$. $\boldsymbol{T}_t$ e $\boldsymbol{P}_t$ entram no processo como covariáveis.

Uma análise prévia do conjunto de dados que inclui as covariáveis, é feita. Primeiro analisa-se a série de mortalidade cardiovascular semanal de modo individual. Pode-se obser-

[1]Este conjunto de dados está implementado no pacote de Stoffer (2016). O autor usa o objeto ts() e configura a série com periodicidade inteira, quando teria uma periodicidade não-inteira de $365.25/7 = 52.17857$, ver apêndice A.

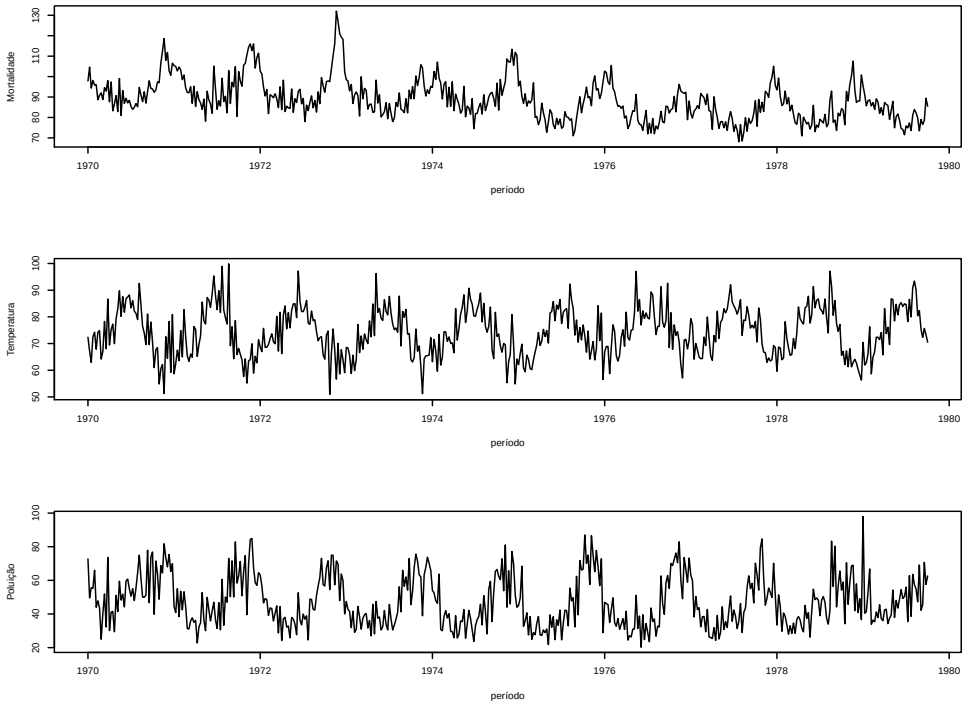


Figura 2.9: Painel I: dados sobre a mortalidade cardiovascular por temperatura e poluição em Los Angels–Estados Unidos de América entre 1970-1979. Painel II: dados sobre a temperatura. Painel III: dados sobre os níveis de poluição.

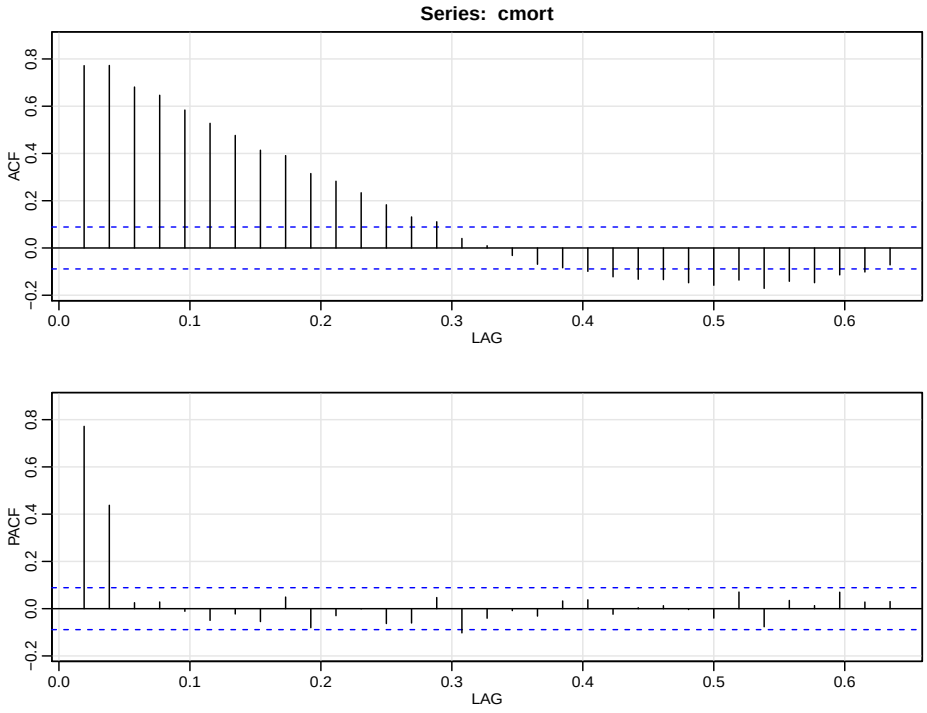


Figura 2.10: Correlograma da mortalidade cardiovascular semanal por temperatura e poluição em Los Angels–Estados Unidos de América entre 1970-1979. Painel II: dados sobre a temperatura.

var a partir das autocorrelações parciais, que apenas as autocorrelações em *lag 1* e *lag 2* são estatisticamente significantes. A estratégia é ajustar um AR(2) aos dados da mortalidade para se obter os resíduos não correlacionados que são utilizados no estudo da correlação cruzada com a série de temperatura e a série de níveis de partículas. A Figura 2.11 mostra uma forte correlação com a temperatura desfasada em uma semana $\boldsymbol{T}_{t-1}$, níveis de partículas simultâneas $\boldsymbol{P}_t$ e os níveis de partículas desfasadas em cerca de um mês $\boldsymbol{P}_{t-4}$. É estimado um modelo que integra as três covariáveis desfasados.

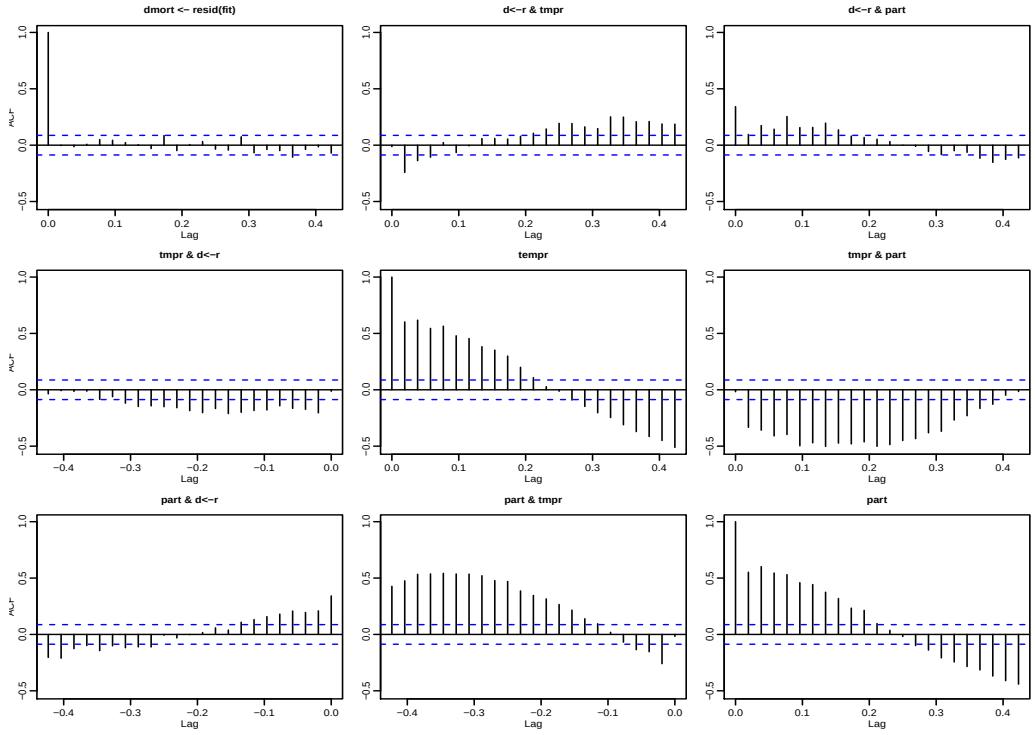


Figura 2.11: Correlação cruzada entre os resíduos de mortalidade, a série de Temperatura e a série de Partículas. Primeiro painel–série observada sobre mortalidade cardiovascular; segundo painel–série observada sobre temperatura; terceiro painel–série observada sobre poluição

2.3.1 Estimação e previsão um passo à frente

Estimam-se três modelos: o primeiro modelo que se designa por TSM (é o modelo TSCov sem o uso das covariáveis), o segundo modelo é TSCov com a integração das covariáveis, e o terceiro é o modelo TBATS. O conjunto de dados é dividido em dois segmentos: a série de teste com 450 observações e a série de validação com 58 observações. Os resultados aqui apresentados resultam da aplicação de covariáveis reais no modelo TSCov.

Valores iniciais para o modelo TSM (TSCov sem covariáveis). A média e a covariância do estado do sistema são inicializadas por $\boldsymbol{x}_0 = 1.2$ e $\boldsymbol{P}_{0i} = 80$, com $i = 8$. A covariância da observação é fixada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 0.1$ e as variâncias do vetor de estados são inicializadas por $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^{2(i)}\} = \{0.002, 0.3, 0.001, 0.001\}$, com $i = 2$. O fator de esquecimento é

fixado em $\delta = 0.75$. Para este modelo, apenas reportamos os resultados referentes a precisão da previsão, que estão apresentados na Tabela 2.7.

Valores iniciais para o modelo TSCov com covariáveis. A média e a covariância do estado do sistema são inicializadas por $\boldsymbol{x}_0 = 2.5$ e $\boldsymbol{P}_{0i} = 80$, com $i = 8$, respetivamente. A covariância da observação é fixada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 0.1$ e as variâncias do vetor de estados são inicializadas por $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^{2(i)}\} = \{0.002, 0.3, 0.001, 0.001\}$, com $i = 2$. Os coeficientes de regressão são inicializados por $\{\beta_1^*, \beta_2^*, \beta_3^*\} = 0.1$; o fator de esquecimento é fixado em $\delta = 0.99$.

As estimativas dos parâmetros para os dois modelos estimados, TSCov e TBATS, estão apresentadas na Tabela 2.5, e constata-se que o parâmetro β_1^* que corresponde a temperatura é não significativo. A componente irregular para o modelo TBATS é modelada com um processo MA(2).

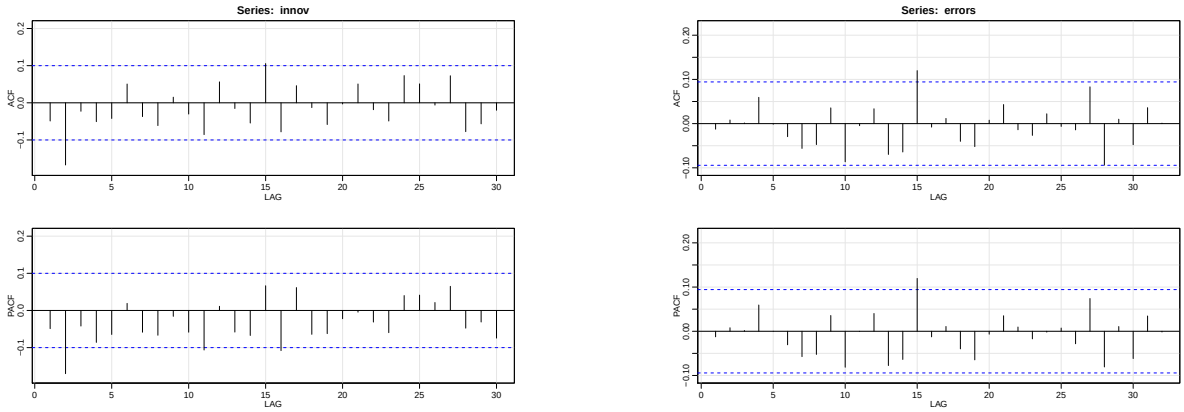
Parâmetro	MLE (TSCov)	E.Padrão Ass.	MLE (TBATS)
β_1^*	-0.019	0.021	—
β_2^*	0.134	0.056	—
β_3^*	0.140	0.034	—
α	—	—	0.316
β	—	—	-0.058
ϕ	0.802	0.331	0.803
σ_ε^2	2.536	0.012	—
σ_ξ^2	0.267	0.113	—
σ_ζ^2	0.019	0.071	—
σ_w^2	8×10^{-6}	0.002	—
$\sigma_{w^*}^2$	0.001	NA	—
γ_1	—	—	-0.0058
γ_2	—	—	-4.42×10^{-5}

Tabela 2.5: Estimativas dos parâmetros e os respetivos erros-padrão obtidos a partir do modelo TSCov (segunda e terceira colunas).

A seguir realiza-se a inspeção dos gráficos dos resíduos para os dois modelos (TSCov e TBATS), fundamentalmente, Figura 2.12. Para o modelo TSCov, o correlograma dos resíduos, Figura 2.12a, exibe uma correlação significativa negativa no *lag 2*, no entanto, o teste de Box-Ljung sobre a independência dos resíduos fornece um valor de Qui-quadrado igual a 23.131, com 21 graus de liberdade e *p-valor* = 0.3812, o que permite não rejeitar a hipótese nula de que os resíduos são independentes. Para o modelo TBATS, o teste de Ljung-Box fornece um Qui-quadrado = 23.193 com 22 graus de liberdade e *p-valor* = 0.3909; o que permite, igualmente, não rejeitar a hipótese nula de que os resíduos do modelo são independentes.

O gráfico Q-Q normal dos resíduos do modelo estimado com TSCov, Figura 2.13, mostra a partida da normalidade na cauda esquerda, fundamentalmente, devido a presença de *outliers* que ocorreram principalmente no ano 1973.

A Figura 2.14 mostra o ajuste gráfico dos dois modelos estimados, TSCov e TBATS, incluindo os valores observados, onde se pode observar que o ajuste obtido pelo modelo TSCov



(a) (b)

Figura 2.12: Correlograma dos resíduos resultantes da previsão um passo à frente da mortalidade cardiovascular por temperatura e poluição em Los Angels. (a) modelo TSCov, (b) modelo TBATS.

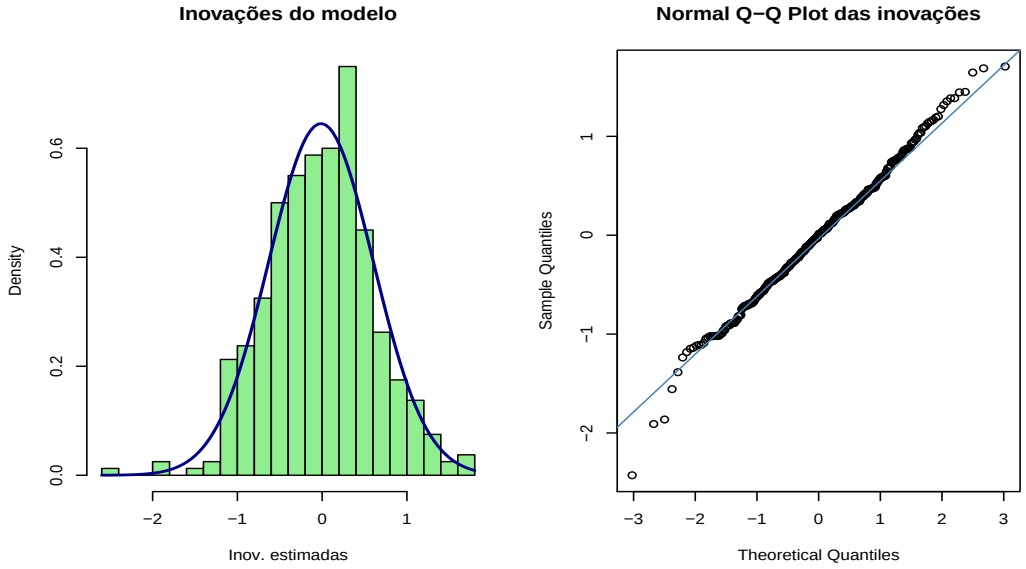


Figura 2.13: Histograma e o Q-Q normal dos resíduos do modelo TSCov estimado sobre a mortalidade cardiovascular em Los Angels - Estados Unidos de América.

coloca-se mais próximo dos valores observados em relação o ajuste obtido pelo modelo TBATS. Os erros de previsão um passo à frente estão apresentados na Tabela 2.6.

Modelo	ME	RMSE	MAE	MPE	MAPE
TSCov	-0.052	5.703	4.179	-0.289	4.77
TBATS	-0.362	7.524	5.876	-0.854	6.44

Tabela 2.6: Erros de previsão um passo à frente obtidos pelos modelos TSCov e TBATS sobre a mortalidade cardiovascular por poluição e temperatura em Los Angels.

É usual apresentar o modelo estimado na sua forma matemática. Assim, o modelo de

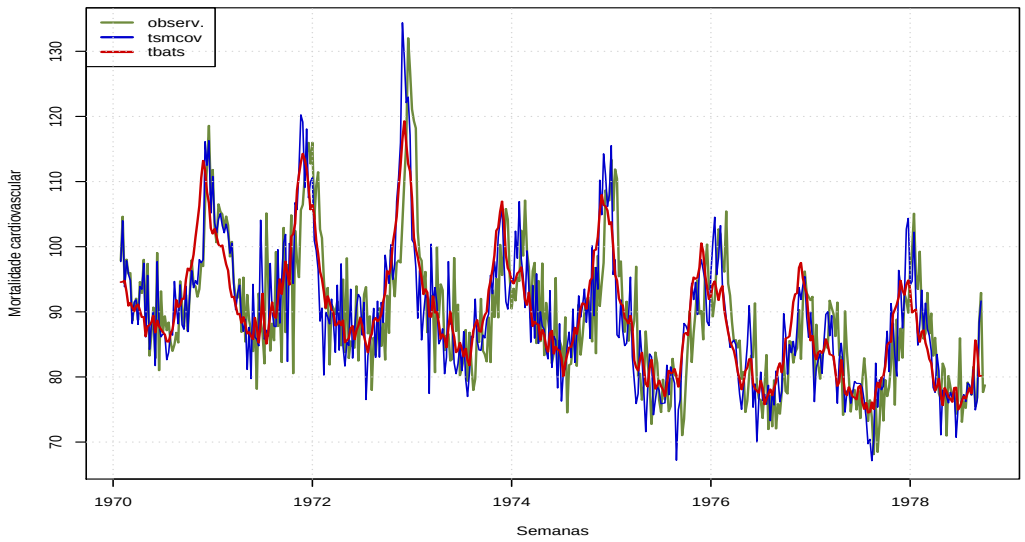


Figura 2.14: Valores observados e os ajustados a partir dos modelos TSCov e TBATS.

previsão que descreve a dinâmica da mortalidade cardiovascular por poluição e temperatura em Los Angels, para o período de estudo, pode ser expresso na forma:

$$
\begin{aligned}
M_t &= \ell_{t-1} + 0.80 b_{t-1} + \sum_{i=1}^{T} s_{t-1}^{(i)} + 0.134 P_t + 0.14 P_{t-4} + \varepsilon_t \\
\ell_t &= \ell_{t-1} + 0.80 b_{t-1} + \xi \\
b_t &= 0.80 b_{t-1} + \zeta_t \\
s_t &= \sum_{j=1}^{4} s_{j,t} \quad \text{(padrão sazonal anual)} \\
s_{j,t} &= s_{j,t-1} \cos\left(\frac{2\pi j t}{52}\right) + s_{j,t-1}^{*} \sin\left(\frac{2\pi j t}{52}\right) + w_{j,t}^{(i)} \\
s_{j,t}^{*} &= -s_{j,t-1} \sin\left(\frac{2\pi j t}{52}\right) + s_{j,t-1}^{*} \cos\left(\frac{2\pi j t}{52}\right) + w_{j,t}^{*(i)} \quad \text{onde} \\
\varepsilon_t &\sim \mathcal{N}(0, 2.536) \\
\xi_t &\sim \mathcal{N}(0, 0.267) \\
\zeta_t &\sim \mathcal{N}(0, 0.019) \\
w_t &\sim \mathcal{N}(0, 8 \times 10^{-6}) \\
w_t^{*} &\sim \mathcal{N}(0, 0.001)
\end{aligned}
$$

Para este modelo, o total de harmônicas significativos para os termos trigonométricos da componente sazonal para toda a amostra de teste de 400 valores é $\dot{k}_1^{*} = 4$. O vetor dos estados estimado é de dimensão 8.

2.3.2 Previsão multi-passos

Para esse primeiro caso de estudo, calcula-se a previsão (com covariáveis reais) de 52 passos à frente. A Figura 2.15 mostra a previsão um passo à frente juntamente com os valores observados e as previsões até 52 passos à frente (sobre a mortalidade cardiovascular em Los Angels - Estados Unidos de América) obtidas a partir dos modelos TSCov e TBATS. A Figura 2.16a é uma parte do gráfico da Figura 2.15, exibe apenas as previsões até 52 passos à frente dos modelos TSCov e TBATS, incluindo os valores observados. A área em cinza do mesmo gráfico representa os intervalos de previsão de 95% gerados pelo modelo TSCov.

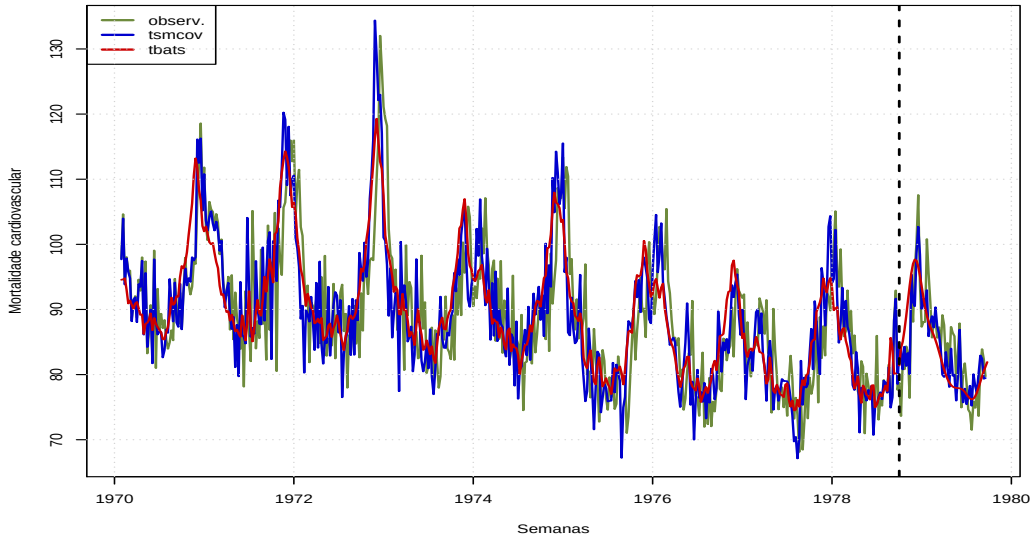
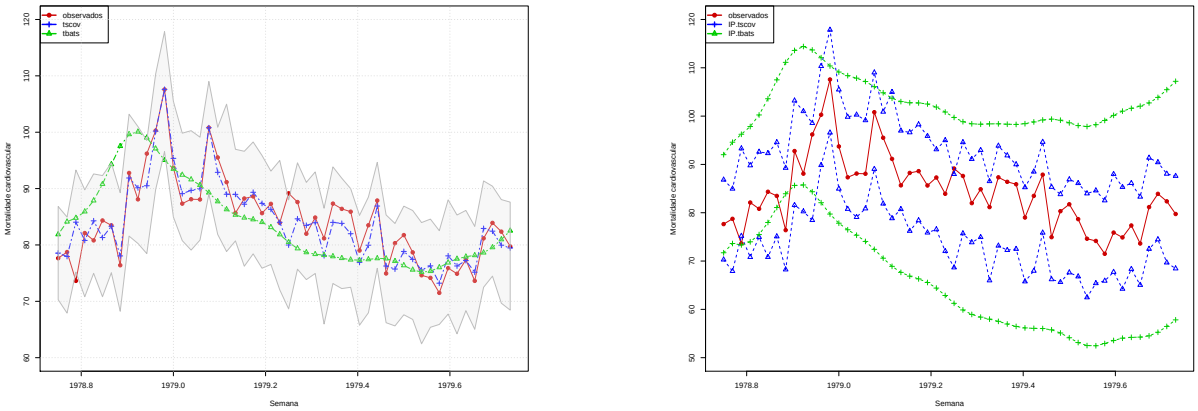


Figura 2.15: Valores observados e os ajustados incluindo as previsões até 52 passos à frente obtidos com os modelos TSCov e TBATS.

A Figura 2.16b apresenta os intervalos de previsão até 52 passos à frente gerados pelos modelos TSCov e TBATS, incluindo os valores observados. Nela se pode observar que os intervalos obtidos pelo modelo TSCov são mais regulares ao longo do horizonte de previsão quando comparados com os intervalos de previsão obtidos pelo modelo TBATS, que aumentam de amplitude conforme aumenta o horizonte de previsão.

A Tabela 2.7 mostra a avaliação da capacidade preditiva dos três modelos estimados – o modelo TSM (TSCov sem covariáveis), o modelo TSCov (com covariáveis) e o modelo TBATS.



(a) (b)

Figura 2.16: (a) Valores observados e as previsões até 52 passos à frente obtidas com os modelos TSCov e TBATS. A área em cinza, Figura 2.16a, representa os intervalos de previsão de 95% gerados pelo modelo TSCov; (b) Valores observados e os intervalos de previsão de 95% gerados pelos modelos TSCov e TBATS.

Horizonte	TSM		TSCov		TBATS	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1 – 7	5.703	4.099	4.255	2.875	4.599	4.703
1 – 14	5.411	3.051	3.403	3.014	3.504	4.768
1 – 21	5.622	3.346	3.098	3.232	3.726	4.606
1 – 28	5.534	3.386	3.104	3.289	3.901	4.588
1 – 35	5.711	3.561	3.184	3.198	3.983	4.876
1 – 42	5.813	3.565	3.331	3.401	4.054	5.214
1 – 49	5.942	3.734	3.664	3.455	4.433	5.548
1 – 52	6.446	4.653	3.836	3.657	3.769	5.677

Tabela 2.7: Medidas de precisão de previsão até 52 passos à frente sobre a mortalidade cardiovascular por poluição e temperatura em Los Angels – modelos estimados TSM, TSCov e TBATS.

Portanto, os resultados obtidos nesse primeiro caso de estudo mostram que em termos gráficos como em termos das medidas de precisão da previsão, todos concordam que o modelo TSCov tem melhor desempenho quer para a previsão um passo à frente como para a previsão multi-passos à frente, quando comparado com os resultados obtidos pelo modelo TBATS.

2.4 Terceiro Caso de Estudo: dados com sazonalidade múltipla e efeito duplo de calendário

O conjunto de dados tem a ver com a procura diária de eletricidade na Turquia, observados entre 1 de janeiro de 2000 e 31 de Dezembro de 2008, Figura 2.17. Trata-se de um conjunto de dados com efeitos sazonais duplo de calendário De Livera et al. (2011), ou seja, a série tem um padrão sazonal semanal e dois padrões sazonais anuais: um para o calendário Hijri com um período de 354.37 e o outro referente ao calendário gregoriano com um período de $365,25$. Não constitui objetivo desta secção o cálculo das previsões multi-passos. Este conjunto de dados foi utilizado no trabalho de De Livera et al. (2011) e pode ser encontrado na página web do professor *Rob J. Hyndman*: https://robjhyndman.com/publications/complex-seasonality/.

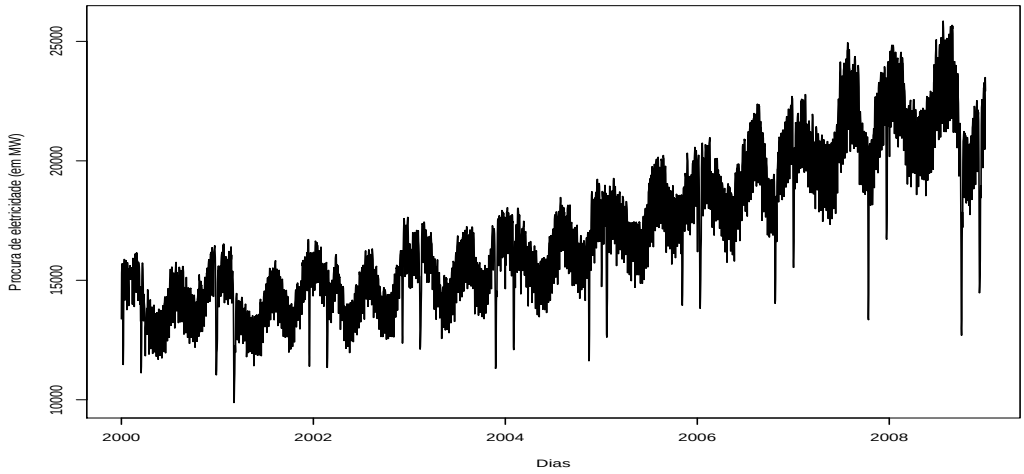


Figura 2.17: Dados de procura de eletricidade na Turquia, de 1 de janeiro de 2000 a 31 de dezembro de 2008.

De Livera (2010) constatou que a consideração de feriados religiosos e nacionais separadamente como efeitos determinísticos na regressão levou à derivação de componentes sazonais mais definidos com menos aleatoriedade usando a abordagem trigonométrica, razão convincente para investigar se um modelo trigonométrico com covariáveis seria mais apropriado para a previsão desta série temporal, uma vez que permitiria que os componentes sazonais fossem capturados sem serem contaminados pelos efeitos determinísticos do feriado. De acordo com a formulação do modelo TSCov, e indo ao encontro da sugestão do autor, duas variáveis são construídas: uma para os feriados religiosos, $\boldsymbol{F}_r$, e outra para os feriados nacionais, $\boldsymbol{F}_n$.

$$\boldsymbol{F}_r = \begin{cases} 1 & \text{se o período de tempo t ocorre quando os feriados religiosos estão em vigor} \\ 0 & \text{caso contrário} \end{cases}$$

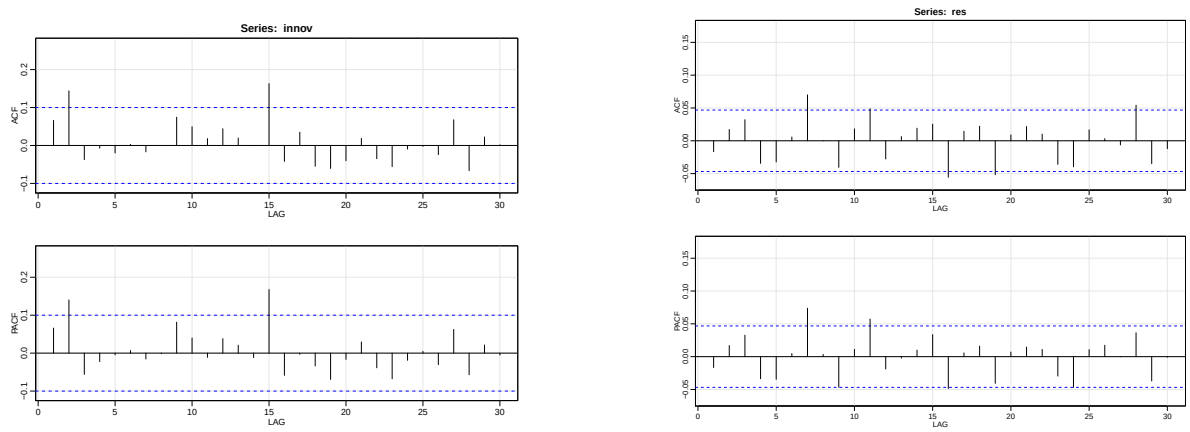
$$\boldsymbol{F}_n = \begin{cases} 1 & \text{se o período de tempo t ocorre quando os feriados nacionais estão em vigor} \\ 0 & \text{caso contrário} \end{cases}$$

Estas variáveis são construídas conforme a Tabela B.10 sobre os feriados da Turquia observados entre 1 de Janeiro de 2000 e 31 de Dezembro de 2006.

Valores iniciais para estimação do modelo TSCov. O processo é inicializado com a média e covariância do estado fixados em $\boldsymbol{x}_0 = 0$ e $\boldsymbol{P}_{0i} = 30$ com $i = 28$, respetivamente. A covariância da observação é fixada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 10^{-7}$ e as variâncias do vetor de estados são inicializadas por $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_{w_i}^2\} = \{0.005, 0.005, 0.001, 0.001, 0.001, 0.001, 0.001, 0.001\}$, com $i = 6$. Os coeficientes de regressão são inicializados por $\{\beta_1^*, \beta_2^*\} = 0.1$ e o fator de esquecimento é fixado em $\delta = 0.9$.

As estimativas dos parâmetros dos dois modelos incluindo os erros-padrão estão apresentadas na Tabela 2.8. O ajustamento gráfico dos modelos estimados e os respetivos erros de previsão um passo à frente estão mostrados na Figura 2.20 e Tabela 2.9. A componente irregular do modelo TBATS é modelada com um processo ARMA(4,2).

A inspeção dos resíduos gerados pelos modelos TSCov e TBATS estão apresentados na Figura 2.18. A Figura 2.18a, exibe duas correlações significativas nos *lags = 2 e 15*, no entanto, o teste de Box-Ljung sobre a independência dos resíduos fornece um Qui-quadrado = 23.131, com 18 graus de liberdade e *p-valor* = 0.081, o que permite não rejeitar a hipótese nula de que os resíduos são independentes. Para o modelo TBATS, o correlograma exibe correlações significativas nos lags = 7 e 16, alguns picos marginais nos lags = 19 e 28, e o teste de Ljung-Box fornece um Qui-quadrado = 26.222 com 14 graus de liberdade e *p-valor* = 0.024; o que permite rejeitar a hipótese nula de que os resíduos do modelo são independentes.



(a) (b)

Figura 2.18: Correlograma dos resíduos resultantes da previsão um passo à frente sobre a procura diária de eletricidade na Turquia. (a) correlograma do modelo TSCov, (b) correlograma do modelo TBATS.

Apesar de ter-se um histograma dos resíduos com um aspeto Gaussiano, o gráfico Q-Q normal dos resíduos do modelo estimado com TSCov, Figura 2.19, mostra, igualmente, a partida da normalidade nas duas caudas, devido possivelmente, a presença de *outliers* que ocorreram ao longo de todo período de observação.

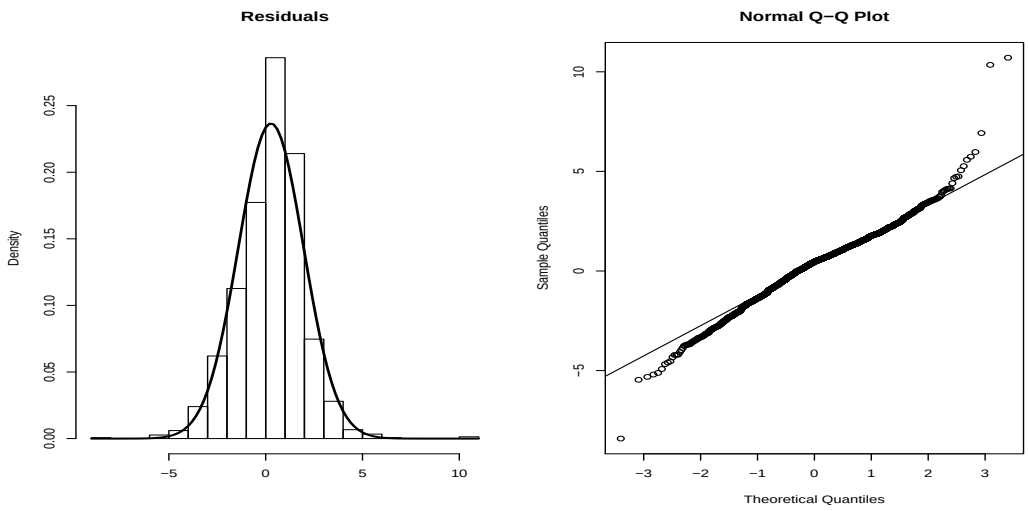


Figura 2.19: Histograma e o Q-Q normal dos resíduos do modelo estimado com TSCov sobre a procura diária de eletricidade na Turquia.

Parâmetro	MLE (TSCov)	E.Padrão Ass.	MLE (TBATS)
β_1^*	-0.655	0.023	—
β_2^*	0.384	0.035	—
ȧ	—	—	0.332
α	—	—	-0.127
β	—	—	0.034
ϕ	0.934	0.133	0.827
σ_ε^2	0.007	0.004	—
σ_ξ^2	0.089	0.007	—
σ_ζ^2	0.028	0.004	—
σ_w^2	$\{4.3-06; 2.4-03; 3.3-03\}$	$\{NA; 0.006; 0.001\}$	—
w^2	$\{2.5-06; 5.0-03; 7.7-04\}$	$\{NA; 0.005; NA\}$	—
γ_1	—	—	$\{3.3-04; 6.8-05; 7.5-05\}$
γ_2	—	—	$\{4.1-04; -5.1-04; 2.0-04\}$

Tabela 2.8: Estimativas dos parâmetros obtidas a partir dos modelos TSCov e TBATS, incluindo os erros-padrão das estimativas dos parâmetros do modelo TSCov.

Modelo	ME	RMSE	MAE	MPE	MAPE
TSCov	-0.014	292.05	249.44	-0.947	1.756
TBATS	4.408	396.78	228.53	-0.034	1.553

Tabela 2.9: Erros de previsão um passo à frente obtidos pelos modelos TSCov e TBATS sobre a procura diária de eletricidade na Turquia.

2.4.1 Previsão multi-passos

Para esse caso de estudo, calculam-se previsões até 14 passos à frente. A Figura 2.20 mostra a previsão um passo à frente obtida aplicando os modelos TSCov e TBATS; no gráfico estão presentes também os valores observados. A Figura 2.21a exibe os valores observados juntamente com as previsões até 14 passos à frente incluindo os intervalos de previsão gerados

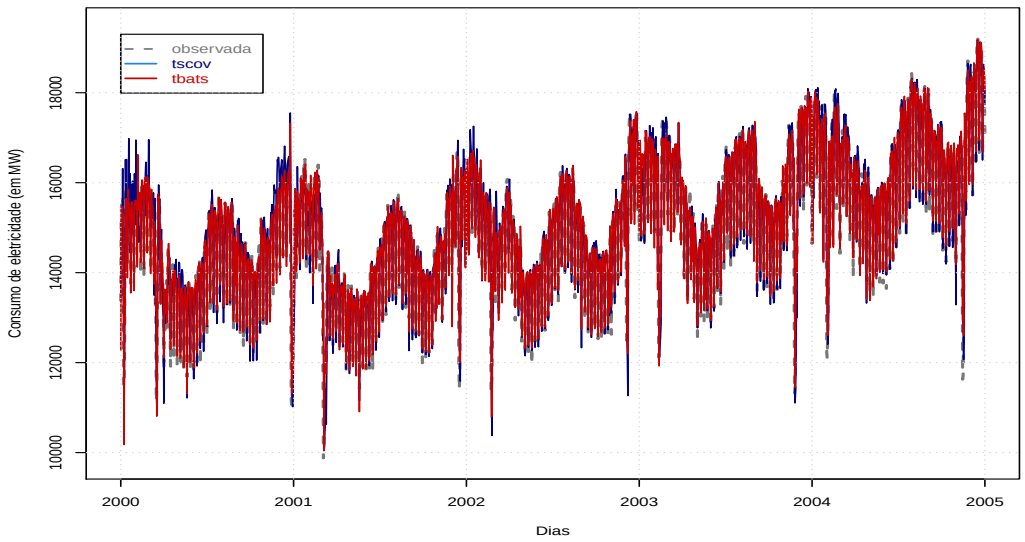
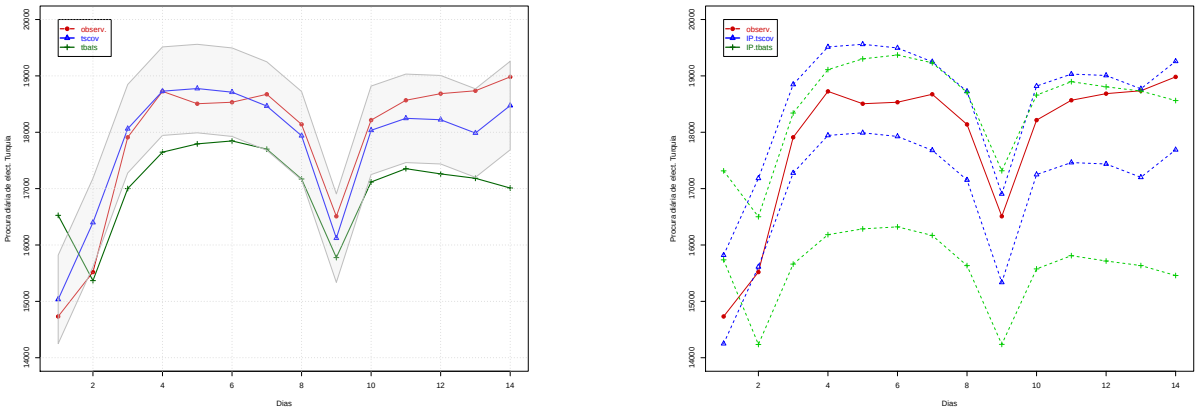


Figura 2.20: Valores observados e a previsão um passo à frente obtida a partir dos modelos TSCov e TBATS sobre a procura diária de eletricidade na Turquia.

pelo modelo TSCov. A Figura 2.21a exibe apenas os intervalos de previsão de 95% gerados pelos modelos TSCov e TBATS incluindo os valores observados. As medidas de precisão da previsão estão apresentadas na Tabela 2.10.



(a) (b)

Figura 2.21: (a) Valores observados e a previsão até 14 passos à frente obtida pelos modelos TSCov e TBATS sobre a procura diária de eletricidade na Turquia; (b) Valores observados e os intervalos de previsão gerados pelos modelos TSCov e TBATS sobre a procura diária de eletricidade na Turquia.

Os resultados obtidos nesse caso de estudo, relacionados com o ajuste gráfico e precisão de previsão, também mostram o desempenho satisfatório do modelo TSCov quando comparado com o modelo TBATS.

Horizonte	TSCov		TBATS	
	RMSE	MAPE	RMSE	MAPE
1 – 3	362.157	1.961	844.468	5.480
1 – 6	368.829	1.730	879.507	4.223
1 – 9	406.221	1.730	923.885	4.476
1 –12	413.658	1.833	993.773	4.963
1 –14	543.373	2.854	1138.068	5.724

Tabela 2.10: Precisão de previsão até 14 passos à frente dos modelos TSCov e TBATS. A Previsão refere-se a procura diária de energia elétrica na Turquia.

2.5 Quarto Caso de Estudo: dados com período sazonal não inteiro

Nessa Secção, o modelo TSCov é aplicado sem a integração das covariáveis e com o mesmo espírito de trabalho, os resultados obtidos pelo modelo é confrontado com os do modelo TBATS. O conjunto de dados utilizado é constituído por 745 observações e tem a ver com a produção semanal de gasolina nos Estados Unidos de América, entre Fevereiro de 1991 e Julho de 2005, Figura 2.22. O referido conjunto de dados também foi utilizado no trabalho de De Livera et al. (2011) e pode ser encontrado na página web do professor *Rob J. Hyndman*: https://robjhyndman.com/publications/complex-seasonality/.

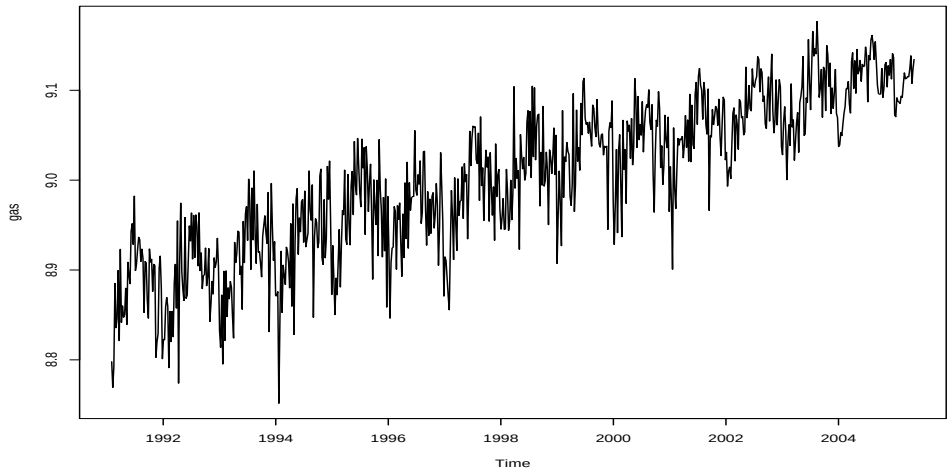


Figura 2.22: Dados sobre produção de gasolina a motor dos EUA (em milhares de barris por dia), entre Fevereiro de 1991 e Julho de 2005.

O decaimento gradual, Figura 2.23, é típico de uma série temporal que contém uma tendência e o pico em 1 ano indica variação sazonal. Geralmente, uma tendência nos dados mostra um decaimento lento nas autocorrelações, que a priori, são grandes e positivas devido a valores similares que ocorrem próximos uns dos outros no tempo, e a Figura 2.23 mostra isso. Os picos são estatisticamente significativos. Isso significa que no período de observação, a produção de gasolina é altamente correlacionada entre si; ou seja, quando a produção de

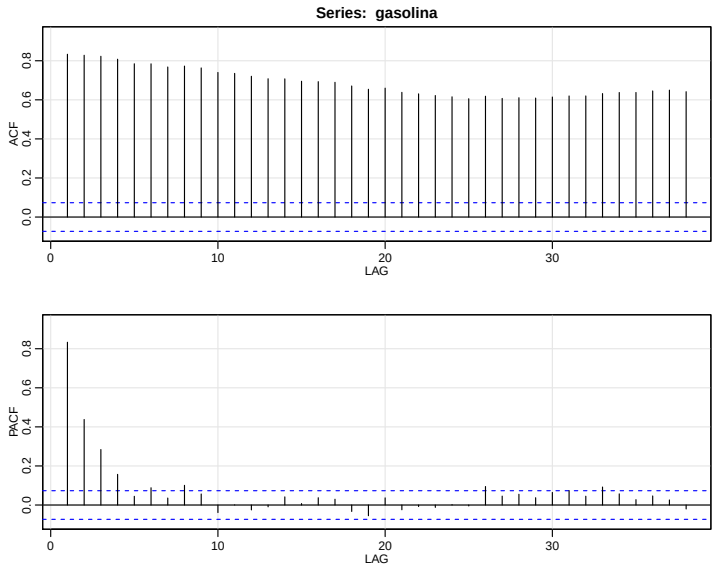
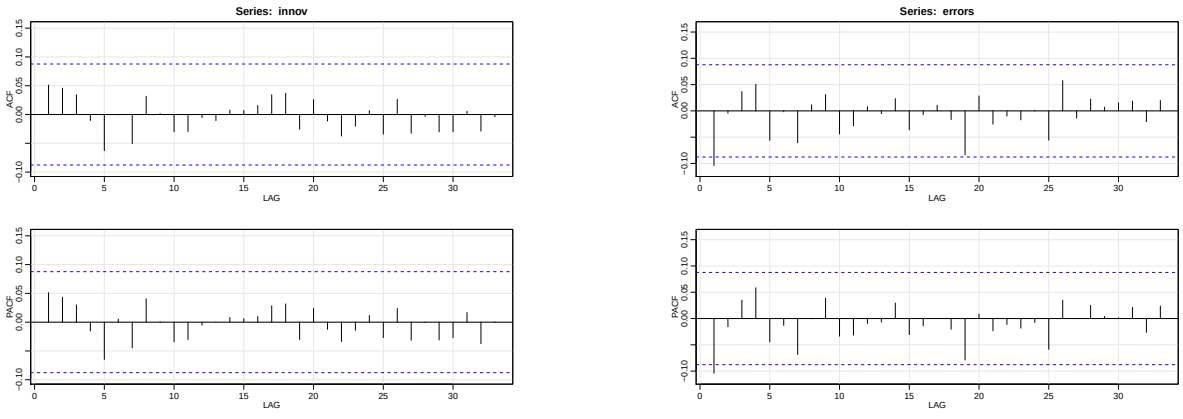


Figura 2.23: Correlograma referente a produção da gasolina nos Estados Unidos de América.

gasolina soube, ela tende a continuar subindo. Trata-se de uma tendência crescente e aditiva. A série tem um padrão sazonal anual com periodicidade não-inteira de $365.25/7 = 52.17857$.

2.5.1 Estimação e previsão um passo à frente

A série de teste envolve 520 observações e a de validação 225 observações. Para o modelo TSCov, o processo de estimação consiste em fixar a média do vetor de estados e sua covariância em $\hat{\boldsymbol{x}}_0 = 0$ e $\boldsymbol{P}_{0ii} = 6.5$, com $i = 1, \cdots, 18$; o fator de esquecimento aplicado às covariâncias é $\delta = 0.999$. A covariância da observação é inicializada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 10^{-8}$; as variâncias do estado são inicializadas em $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^2\} = \{0.0005, 0.0005, 0, 0\}$.



(a) (b)

Figura 2.24: Correlograma dos resíduos - previsão um passo à frente da produção de gasolina nos Estados Unidos de América. (a) correlograma do modelo TSCov sem a integração das covariáveis; (b) correlograma do modelo TBATS.

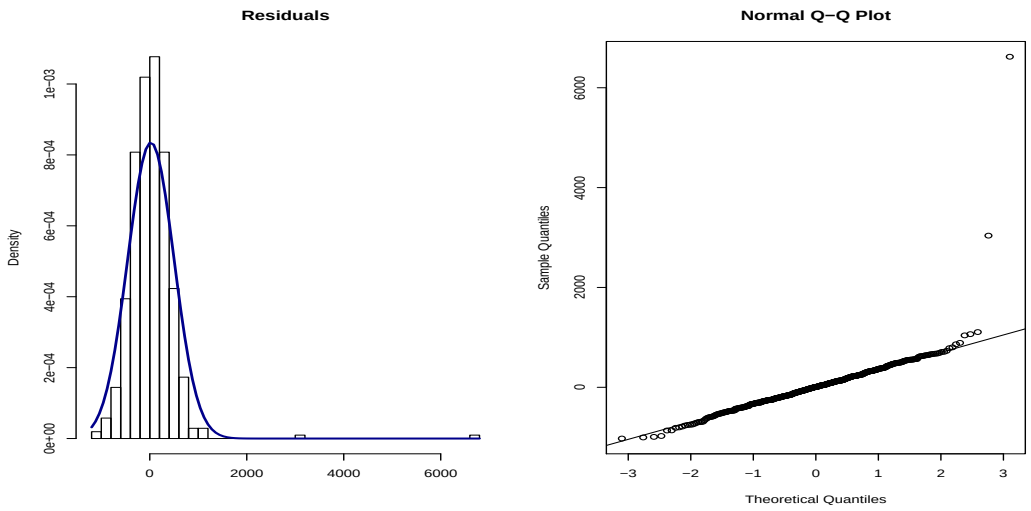


Figura 2.25: Histograma dos resíduos e Q-Q normal dos dos resíduos do modelo estimado com TSCov sobre a produção de gasolina nos Estados Unidos de América.

O modelo estimado com TBATS fornece um vetor de estados de dimensão 16 e o número de harmônicas significativo para os termos trigonométricos da componente sazonal é $\dot{k}_1 = 7$. As estimativas dos parâmetros estão apresentados na Tabela 2.11. De acordo com as caraterísticas do conjunto de dados utilizados, o modelo TSCov sem covariáveis comporta-se melhor com a matriz covariância $\boldsymbol{Q}_t$ completa. A dimensão do vetor dos estados estimado é 18 e o número de harmônicas significativo para os termos trigonométricos da componente sazonal é $\dot{k}_1^* = 7$. As estimativas dos parâmetros também constam na Tabela 2.11.

A Figura 2.24 apresenta o diagnóstico dos resíduos gerados pelos modelos TSCov e TBATS. Importante notar que mesmo com a aplicação da transformação Box-Cox, a função de auto-correlação empírica resultante dos resíduos do modelo estimado pelo TBATS, exibe uma correlação negativa no *lag 1*, Figura 2.24a, o que é improvável que isso seja devido à variação aleatória da amostragem. No entanto, o teste de Ljung-Box fornece um valor de Q-quadrado igual a 19.392, com 24 graus de liberdade e *p-valor* = 0.731; o que permite não rejeitar a hipótese nula de que os resíduos são independentes.

Para o modelo gerado por TSCov, a função de auto-correlação empírica apresentada em 2.24b não exibe correlações significativas. Estão todas dentro do intervalo de confiança de 95%, que é satisfatório. Ademais, o teste de Box-Ljung sobre a independência dos resíduos, fornece um valor de Qui-quadrado igual a 11.485, com 24 graus de liberdade e *p-valor* = 0.985; o que permite, igualmente, não rejeitar a hipótese nula de que os resíduos são independentes. Quanto a normalidade dos resíduos, o gráfico Q-Q normal dos resíduos do modelo estimado com TSCov, Figura 2.25, mostra a partida da normalidade na cauda direita, visivelmente confirmando a presença de *outliers* que ocorreram principalmente nos anos de 1994 e 2001.

A Figura 2.26 e a Tabela 2.12 mostram o ajuste gráfico e os erros de previsão um passo à frente, respetivamente, para os dois modelos, TSCov e TBATS. De salientar que o modelo TSCov comporta-se razoavelmente melhor (no que tange o RMSE), quer em termos gráficos como na precisão da previsão um passo à frente quando comparado com o modelo TBATS.

Parâmetro	MLE (TSCov)	E.Padrão Ass.	MLE (TBATS)
α	—	—	-0.06
β	—	—	0.03
ϕ	0.83	0.154	0.83
σ^2_{ε}	220.69	1.751	—
σ^2_{ξ}	1054.29	9.597	—
σ^2_{ζ}	860.39	3.819	—
σ^2_{w}	7.055	0.489	—
$\sigma^2_{w^*}$	2.892	0.216	—
γ_1	—	—	-0.003
γ_2	—	—	0.002

Tabela 2.11: Estimativas dos parâmetros e os respetivos erros-padrão obtidos partir do modelo TSCov. As estimativas dos parâmetros obtidas a partir do modelo TBATS estão apresentadas na quarta coluna.

Modelo	ME	RMSE	MAE	MPE	MAPE
TSCov	0.333	266.57	189.98	-0.144	3.015
TBATS	30.321	274.036	217.498	0.271	2.785

Tabela 2.12: Erros de previsão um passo à frente obtidos pelos modelos TSCov e TBATS sobre a produção de gasolina nos Estados Unidos de América.

2.5.2 Previsão multi-passos

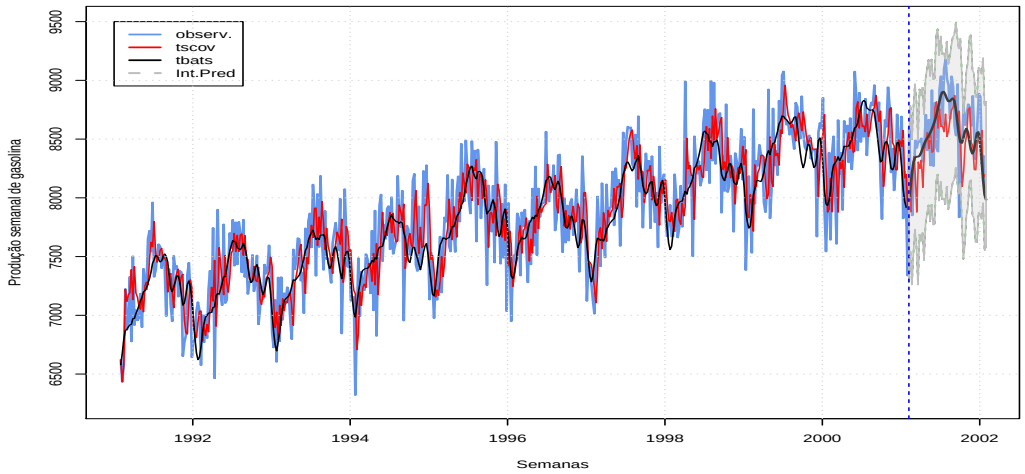
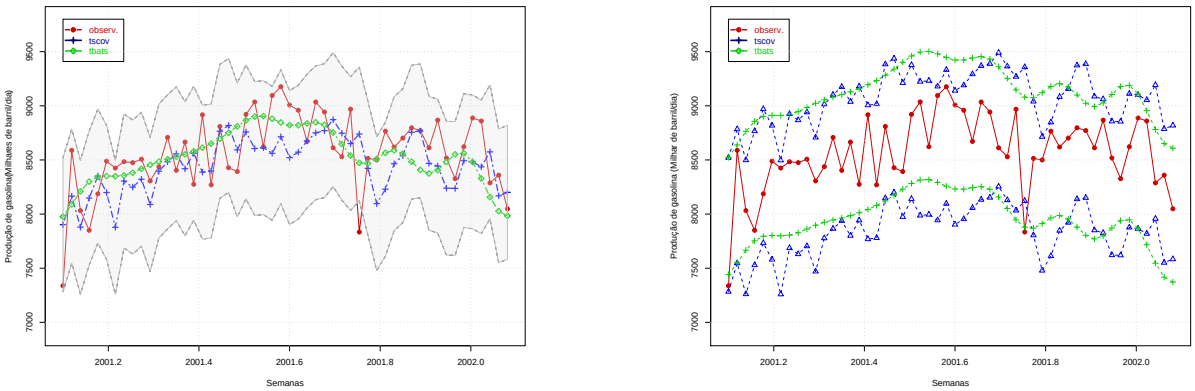


Figura 2.26: Valores observados e os ajustados incluindo as previsões até 52 passos à frente obtidos a partir dos modelos TSCov e **TBATS** sobre a produção de gasolina nos Estados Unidos de América. A área em cinza indica os intervalos de previsão superior e inferior de 95% obtidos a partir do modelo TSCov.

Calculam-se previsões até 52 passos à frente, tal como apresentadas na Figura 2.26. A série de teste é constituída de 520 observações, e a série de validação com 225 observações. A figura 2.26 apresenta os valores ajustados e a previsão de até 52 passos à frente (dos modelos

estimados com TSCov e TBATS), incluindo os intervalos de previsão obtidos a partir do modelo TSCov. A Figura 2.27a exibe apenas as previsões de até 52 passos obtidas pelo TSCov e TBATS incluindo os intervalos de previsão gerados pelo modelo TSCov. Na Figura 2.27b são apresentados, para além das previsões, os intervalos de previsão gerados pelos dois modelos. A área em cinza escura exibe os intervalos de previsão gerados pelo modelo TBATS e a área em cinza clara exibe os intervalos gerados pelo modelo TSCov. Ambos os modelos geram intervalos de previsão regulares ao longo do horizonte de previsão para os dados utilizados nesse estudo.

As medidas de precisão da previsão estão apresentadas na Tabela 2.11. Baseando-se nessas medidas, pode-se concluir que os dois modelos concorrem. No entanto, o ajustamento gráfico determinado pelo modelo TSCov mostra-se mais próximo dos valores observados quando comparado com o modelo TBATS.



(a) (b)

Figura 2.27: (a) Valores observados e as previsões até 52 passos à frente obtidas a partir dos modelos TSCov e TBATS sobre a produção de gasolina nos Estados Unidos de América. A área em cinza indica os intervalos de previsão superior e inferior de 95% obtidos a partir do modelo TSCov. Este gráfico é parte do gráfico da Figura 2.26: (b) Valores observados e os intervalos de previsão superior e inferior de 95% gerados pelos modelos TSCov e TBATS.

Horizonte	TSCov		TBATS	
	RMSE	MAPE	RMSE	MAPE
1 – 7	204.524	2.834	269.031	2.446
1 – 14	253.443	3.116	269.543	2.691
1 – 21	273.329	3.203	272.326	2.687
1 – 28	270.323	3.250	275.282	2.711
1 – 35	291.233	3.314	278.056	2.745
1 – 42	297.547	3.425	281.356	2.760
1 – 49	323.651	3.671	296.847	2.936
1 – 52	358.524	3.993	359.863	3.906

Tabela 2.13: Medidas de precisão da previsão até 52 passos à frente sobre a Produção de Gasolina nos Estados Unidos de América. Os modelos estimados são: TSCov sem a integração das covariáveis e o modelo TBATS.

2.6 Considerações do Capítulo

O objetivo desse Capítulo é testar a performance do procedimento de estimação descrito na secção 1.4 aplicando o quadro de modelos estruturais proposto no Capítulo 1 da Parte II. Os resultados obtidos no estudo empírico permitiram perceber que valores muito grandes na diagonal da matriz $\boldsymbol{Q}_t$, implica elevada incerteza na evolução do estado, e como consequência, a perda de informação durante o processo de atualização do estado a priori, $\hat{\boldsymbol{x}}_{t|t-1}$; a informação transmitida pelas observações passadas $\mathbf{y}_{1:t-1}$ acerca de $\hat{\boldsymbol{x}}_{t|t}$ torna-se pouco relevante para a previsão do estado $\hat{\boldsymbol{x}}_{t|t-1}$. Do ponto de vista da previsão, para o primeiro caso de estudo, ao comparar a precisão das previsões obtidas pelo modelo TSCov (com covariáveis reais e covariáveis previstas), verificou-se que o RMSE para a previsão de um dia (das concentrações de NO_2) para o modelo TSCov com covariáveis previstas produz um MAPE ligeiramente semelhante ao modelo TSCov com covariáveis reais. O quadro de modelos que propomos têm, em geral, melhor classificação em relação os modelos BATS e TBATS.

Embora não seja uma solução para todos os problemas, o quadro de modelos propostos é uma importante estrutura matemática que obtém bons resultados quando aplicado à dados como os apresentados no Capítulo 1 para o cálculo de previsões de curto prazo e comparado com o desempenho do modelo TBATS. É, portanto, um quadro promissor para futuros estudos nessa linha de abordagem. Como o principal objetivo do nosso quadro de modelos é a previsão de curto prazo, os resultados da estimação e previsão obtidos permitem razoavelmente considera-los de satisfatórios. No entanto, algumas melhorias podem ser feitas no quadro da previsão.

O próximo desafio para a nossa abordagem é desenvolver um procedimento de reamostragem por *bootstrap* para a previsão de curto prazo, pois acreditamos que isso é viável e pode melhorar as previsões aqui obtidas e proporcionar mais uma contribuição valiosa no quadro dos modelos estruturais com efeitos das covariáveis para previsão de séries temporais com sazonalidade complexa.

Capítulo 3

PREVISÃO *Bootstrap*: APLICAÇÃO À DADOS COM SAZONALIDADE COMPLEXA

Índice do Capítulo

O objetivo desse Capítulo é construir um procedimento *bootstrap* não paramétrico que aprimore o método de previsão descrito na secção 1.6 do Capítulo 1. O procedimento é baseado no método de estimação descrito na secção 1.4 e é aplicado para previsão de curto prazo das séries temporais com sazonalidade complexa. O modelo estrutural trigonométrico com covariáveis, TSCov, apresentado em (1.2), é utilizado para estimar o modelo *bootstrap* que se designa por Boot.TSCov. Na projeção do procedimento *bootstrap*, o espírito de abordagem dos métodos propostos por Cordeiro and Neves (2011) e Rodriguez and Ruiz (2009), são as referências. Por simplicidade, a notação utilizada nesse Capítulo omite o uso do parâmetro da transformação Box–Cox, ou seja, representar-se-á $y_t^{(\grave{a})}$ por y_t.

3.1 Bootstrap em Modelos de Espaço de Estados

O *bootstrap* é uma técnica estatística que se enquadra no título mais amplo de reamostragem. Envolve procedimentos relativamente simples, mas repetidos n vezes, facto que o torna fortemente dependente da capacidade do computador (Courtney, 2018). É uma maneira de estimar os parâmetros estatísticos de uma amostra através da reamostragem com reposição. Tal como em outras abordagens não paramétricas, o *bootstrap* não estabelece quaisquer suposições sobre a distribuição da amostra. O seu principal pressuposto é que a distribuição da amostra seja uma boa aproximação da distribuição da população, ou seja, que

a amostra seja representativa da população (Desmond, 2014). O método refere-se à inferência sobre uma distribuição amostral de uma estatística por reamostragem da própria amostra com reposição. Na medida em que a distribuição de reamostragem imite a distribuição amostral original, as inferências são precisas. A precisão melhora à medida que o tamanho da amostra original aumenta, se o teorema do limite central se aplica. É uma das várias técnicas que hoje fazem parte do amplo conjunto de estatísticas não paramétricas que são comumente chamadas de métodos de reamostragem.

Segundo Bergmeir et al. (2015), a literatura apresenta vários métodos para reamostrar por *bootstrap* séries temporais, alguns desses métodos são o *tapered block bootstrap*, o *dependent wild bootstrap* e o *extended tapered block bootstrap*. Outro tipo de *bootstrap* é o *bootstrap sieve* proposto por Buhlmann (1997) e utilizado por Cordeiro and Neves (2011). Há várias áreas da estatística onde a análise de séries temporais têm beneficiado do uso de procedimentos computacionais intensivos que ajudam na modelação e previsão nas situações analíticas mais complexas. Entre esses procedimentos, a metodologia *bootstrap* é uma das mais conhecidas, e, é aplicado com maior frequência na reamostragem residual de modelos de séries temporais (Cordeiro and Neves, 2011).

A combinação dos modelos de espaço de estados e os métodos *bootstrap* é uma tendência atual. Vários pesquisadores fazem essa combinação, (Wall and Stoffer, 2002; Menezes et al., 2006; Alonso et al., 2008; Cordeiro and Neves, 2011; Hafida and Hamdi, 2015; Shumway and Stoffer, 2017), e os resultados desses trabalhos mostram que a combinação entre o *bootstrap* e os modelos de espaço de estados é valiosa para previsão de séries temporais, uma vez que pode fornecer previsões mais precisas. Das abordagens básicas do *bootstrap*, uma delas é o *bootstrap* residual em modelos de espaço de estados. Teoricamente, se o modelo estiver corretamente ajustado, os resíduos do modelo seriam independentes e identicamente distribuídos. Então, é possível reamostrar esses resíduos com reposição para se obter uma réplica da amostra original. Em seguida, ajustar o modelo à réplica da amostra original e repetir o processo (Chernick and A.LaBudde, 2011).

Seja o modelo de espaço de estados dado por,

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{d} + \boldsymbol{\nu}_t && \text{(3.1a)} \\ \boldsymbol{x}_t &= \boldsymbol{\Phi}\boldsymbol{x}_{t-1} + \boldsymbol{c} + \boldsymbol{\Theta}\boldsymbol{w}_t \qquad t = 1, \cdots, n && \text{(3.1b)} \end{aligned}$$

Sabe-se que, o filtro de Kalman permite estimar o vetor de estado, $\boldsymbol{x}_{t+1}$ e o seu MSE, dada a informação disponível no período t, isto é

$$\begin{aligned} \boldsymbol{x}_{t+1|t} &= \boldsymbol{\Phi}\boldsymbol{x}_{t|t-1} + \boldsymbol{c} + \boldsymbol{K}_t\boldsymbol{\Sigma}_t^{-1}\boldsymbol{v}_t && \text{(3.2a)} \\ \boldsymbol{P}_{t+1|t} &= \boldsymbol{\Phi}\boldsymbol{P}_{t|t-1}\boldsymbol{\Phi}^{'} - \boldsymbol{K}_t\boldsymbol{\Sigma}_t^{-1}\boldsymbol{K}_t^{'} + \boldsymbol{\Theta}\boldsymbol{Q}\boldsymbol{\Theta}^{'} && \text{(3.2b)} \end{aligned}$$

onde, $\boldsymbol{v}_t = \boldsymbol{y}_t - \boldsymbol{d} - \boldsymbol{A}\boldsymbol{x}_{t|t-1}$, $\boldsymbol{\Sigma}_t = \boldsymbol{A}\boldsymbol{P}_{t|t-1}\boldsymbol{A}^{'} + \boldsymbol{R}$ e $\boldsymbol{K}_t = \boldsymbol{\Phi}\boldsymbol{P}_{t|t-1}\boldsymbol{A}^{'}$. A forma de inovação é dada pela equação (3.2a) juntamente com

$$\boldsymbol{y}_t = \boldsymbol{A}\boldsymbol{x}_{t|t-1} + \boldsymbol{d} + \boldsymbol{v}_t \qquad \text{(3.3)}$$

A previsão h-passos à frente de $\boldsymbol{y}_{t+h}$ e o seu MSE são dados por

$$\boldsymbol{y}_{t+h|t} = \boldsymbol{A}\boldsymbol{\Phi}^h\boldsymbol{x}_{t|t-1} + \boldsymbol{A}\sum_{j=0}^{h-1}\boldsymbol{\Phi}^j\boldsymbol{c} + \boldsymbol{d}, \quad h = 1, 2, \cdots \tag{3.4a}$$

$$\boldsymbol{\Sigma}_{t+h|t} = \boldsymbol{A}(\boldsymbol{\Phi}^h)\boldsymbol{P}_t(\Phi^h)'\boldsymbol{A}' + \boldsymbol{A}\sum_{j=0}^{h-1}[(\boldsymbol{\Phi}^j)\boldsymbol{\Theta}\boldsymbol{Q}\boldsymbol{\Theta}']\boldsymbol{A}' + \boldsymbol{R}, \quad h = 1, 2, \cdots \tag{3.4b}$$

Com mais detalhes, ver Rodriguez and Ruiz (2009).

Como, geralmente, os dados de séries temporais são auto-correlacionados, na literatura podemos encontrar várias versões de procedimentos *bootstrap* adaptadas. Exemplo disso é o método proposto por Rodriguez and Ruiz (2009), que consiste em construir os intervalos de previsão *bootstrap* diretamente das observações, sem precisar da representação *backward* do modelo, como é caso das propostas de Wall and Stoffer (2002) e Hafida and Hamdi (2015).

Em referência ao procedimento de Rodriguez and Ruiz (2009), as etapas podem ser resumidas como se segue: (i) estimar o modelo inicial para obter os resíduos; (ii) obter a sequência *bootstrap* dos resíduos padronizados com reposição; (iii) gerar a réplica *bootstrap* da série original, resolvendo a relação que resulta da equação de estado, (3.2a), e da observação, (3.3), usando os resíduos *bootstrap* e as estimativas iniciais dos parâmetros; (iv) executar o filtro de Kalman usando os parâmetros *bootstrap* e a série de observações originais para obter a réplica *bootstrap* do vetor de estados no instante t; (v) obter as previsões condicionais *bootstrap* h passos à frente. Muitos autores usam também esse tipo de abordagem com algumas adaptações, (Stoffer and Wall, 2004; Menezes et al., 2006; Rodriguez and Ruiz, 2009; Hafida and Hamdi, 2015) e Shumway and Stoffer (2017). Uma dessas adaptações, é o uso de um modelo Auto-Regressivo para filtrar a série de resíduos do modelo inicial estimado, e a réplica *bootstrap* da série original é obtida usando o ajustamento inicial e a sequência *bootstrap* dos resíduos Auto-Regressivos. Esse tipo de abordagem pode ser vista em Menezes et al. (2006) e Cordeiro and Neves (2011).

No âmbito da previsão de séries temporais com padrões sazonais complexos, o paradigma de reamostragem residual por *bootstrap* apresenta até o momento um vazio na literatura. O nosso procedimento *bootstrap* que designamos por Boot.TSCov é inspirado nas metodologias de Rodriguez and Ruiz (2009) e Cordeiro and Neves (2011).

3.2 Procedimento Geral de Boot.TSCov

Primeiro, estima-se o modelo inicial, TSCov, dado em (1.2) para se obter a sequência das inovações, $\boldsymbol{\varepsilon}_t$ (que devem ser não correlacionados) e os valores ajustados, $\hat{\boldsymbol{y}}_t$. Segundo, as inovações padronizadas (3.5) (com a garantia de que têm, pelo menos, os mesmos dois primeiros momentos) são reamostradas com reposição b vezes para se obter a amostra *bootstrap* das inovações padronizadas, $\boldsymbol{v}_t^{*s}$. Considera-se $\{\hat{y}_1, ..., \hat{y}_n\}$ a série dos valores ajustados no modelo inicial estimado. Então, a réplica *bootstrap* da série original é obtida usando (3.6) (Cordeiro and Neves, 2011).

$$\boldsymbol{v}_t^s = \boldsymbol{\Sigma}_t^{-1/2}\boldsymbol{\varepsilon}_t \tag{3.5}$$

$$\hat{\boldsymbol{y}}_t^* = \hat{\boldsymbol{y}}_t + \boldsymbol{\Sigma}_t^{1/2}\boldsymbol{v}_t^{*s} \tag{3.6}$$

Em seguida, com a réplica *bootstrap*, $\hat{\boldsymbol{y}}_t^*$, Figura 3.1, estimar o modelo Boot.TSCov para se obter as estimativas *bootstrap*, $\hat{\boldsymbol{\Omega}}^*$ e outros derivados do filtro de Kalman, tais como inovações a priori ($\hat{\boldsymbol{\varepsilon}}_t^*$) e a posteriori ($\hat{\boldsymbol{\eta}}_t^*$), o vetor de estados e outros. A previsão h-passos à frente é obtida usando as recursões do filtro de Kalman com as estimativas *bootstrap*. A seguir apresenta-se os principais passos do procedimento Boot.TSCov.

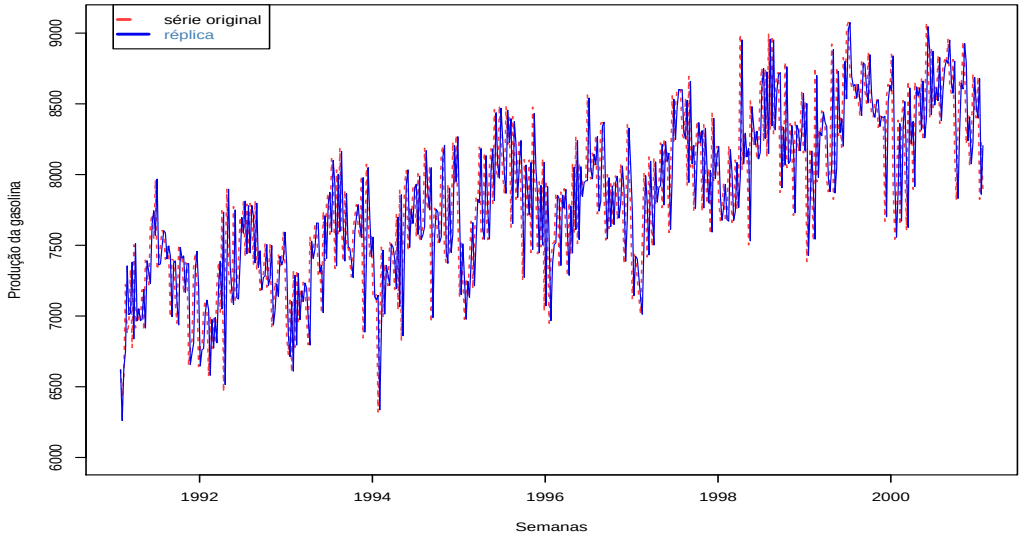


Figura 3.1: Réplica da série de produção semanal de gasolina nos Estados Unidos.

Algoritmo Boot.TSCov

1. Estimar o modelo inicial e obter as inovações $\boldsymbol{\varepsilon}$;
2. Calcular as inovações padronizadas $\boldsymbol{v}_t^s$ usando (3.5);
3. Para cada réplica B,
   - 3.1 Obter as inovações bootstrap padronizadas com reposição $\boldsymbol{v}_t^{*s}$;
   - 3.2 Calcular a réplica bootstrap $\hat{\boldsymbol{y}}_t^*$ usando a equação (3.6). Estimar os correspondentes parâmetros bootstrap, $\hat{\boldsymbol{\Omega}}^*$, incluindo os estados;
   - 3.3 Obter previsões bootstrap h-passos à frente, $\hat{\boldsymbol{y}}_{t+h|t}^*$, usando as recurssões seguintes:

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

$$
\begin{aligned}
\hat{\boldsymbol{Q}}^*_{t+h|t} &= \delta\hat{\boldsymbol{Q}}^*_{t+h-1} + (1-\delta)(\hat{\boldsymbol{K}}^*_t\hat{\boldsymbol{\varepsilon}}^*_t\hat{\boldsymbol{\varepsilon}}^{*'}_t\hat{\boldsymbol{K}}^{*'}_t) \\
\hat{\boldsymbol{R}}^*_{t+h|t} &= \delta\hat{\boldsymbol{R}}^*_{t+h-1} + (1-\delta)(\hat{\boldsymbol{\eta}}^*_t\hat{\boldsymbol{\eta}}^{*'}_t + \hat{\boldsymbol{A}}^*_{t+h}\hat{\boldsymbol{P}}^*_{t+h-1|t}\hat{\boldsymbol{A}}^*_{t+h})
\end{aligned}
$$

onde, $\hat{\boldsymbol{\eta}}^*_t = \boldsymbol{y}_t - \hat{\boldsymbol{A}}^*_t\hat{\boldsymbol{x}}^*_{t|t} - \hat{\boldsymbol{\Gamma}}^*_t\boldsymbol{z}_t$, $\hat{\boldsymbol{\varepsilon}}^*_t = \boldsymbol{y}_t - \hat{\boldsymbol{A}}^*_t\hat{\boldsymbol{x}}^*_{t|t-1} - \hat{\boldsymbol{\Gamma}}^*_t\boldsymbol{z}_t$.

Os intervalos de previsão são gerados tomando a distribuição empírica de $\hat{\boldsymbol{y}}^*_{t+h|t}$, sendo $\hat{\boldsymbol{y}}^*_H$ a média de $\hat{\boldsymbol{y}}^*_{t+h|t}$. Em termos computacionais, variáveis globais que agregam as estimativas *bootstrap* são construídas para permitir o seu acesso para o cálculo das previsões. O chapéu sobre as matrizes significa que são matrizes estimadas por *bootstrap* e usadas para o cálculo das previsões. Os intervalos de previsão podem, igualmente, ser obtidos de forma direta conforme (1.33).

3.3 Análise Empírica

Esta Secção apresenta os resultados da aplicação do procedimento *bootstrap* na previsão. Três conjuntos de dados reais constituem os casos de estudo, são eles: a série temporal dos níveis de concentração de NO_2 em Entre-Campos, a série temporal da procura diária de eletricidade na Turquia e a série temporal da produção de gasolina nos Estados Unidos de América. A primeira série tem dois padrões sazonais: um padrão diário e um padrão semanal, ver Apêndice B.2. A segunda série tem três padrões sazonais: um padrão semanal e dois parões anuais, ver secção 2.4. A terceira série tem um padrão sazonal anual com periodicidade não–inteira, ver secção 2.5.

Os modelos iniciais para os conjuntos de dados em causa são ajustados aplicando o procedimento descrito na secção 1.4. Os delineamentos experimentais utilizados nos casos de estudos empíricos apresentados no Capítulo 2 da Parte II e Apêndice B.2 são válidos para os estudos apresentados nessa secção, e alguns daqueles resultados são reproduzidos aqui para fins de comparação. Portanto, não são apresentados aqui os procedimentos de inicialização dos processos para estimação dos modelos iniciais cujos resíduos são utilizados para a reamostrar por *bootstrap*. As previsões e os intervalos de previsão a calcular com o modelo Boot.TSCov são obtidos tomando a média da réplica *bootstrap*, $\hat{\boldsymbol{y}}^*_{t+h|t}$. As subsecções seguintes apresentam os resultados dos casos de estudo.

3.3.1 Aplicação a dados com múltiplos padrões sazonais: níveis de concentração de NO_2 em Entre-Campos de Lisboa

Os dados referem-se aos níveis de concentração de NO_2 observados na estação de Entre-Campos de Lisboa. O modelo inicial estimado e o diagnóstico dos resíduos do modelo estão apresentados no Apêndice B.2. Para esse conjunto de dados, as estimativas dos parâmetros *bootstrap* resultam também de B = 1500 réplicas. A Tabela 3.1 mostra as estimativas dos parâmetros dos modelos TSCov, Boot.TSCov e TBATS incluindo os erros-padrão de cada estimativa calculados aplicando os modelos TSCov e Boot.TSCov.

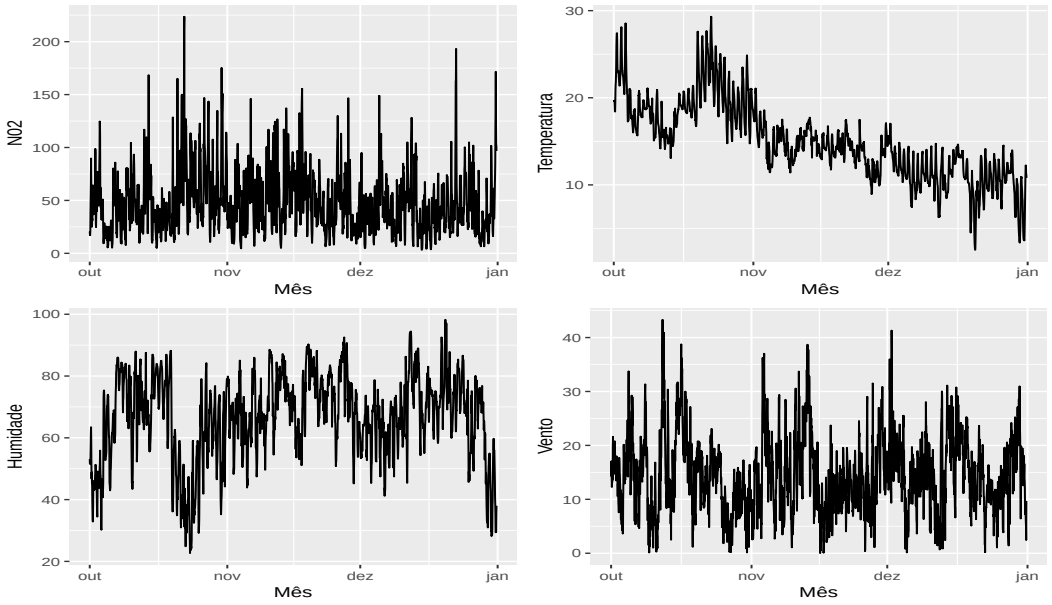


Figura 3.2: Níveis de concentração de NO_2 observados em 2014 entre 1 de Outubro e 31 de Dezembro em Entre-Campos de Lisboa. A temperatura, Humidade e Vento são as covariáveis também observadas em intervalos de uma hora.

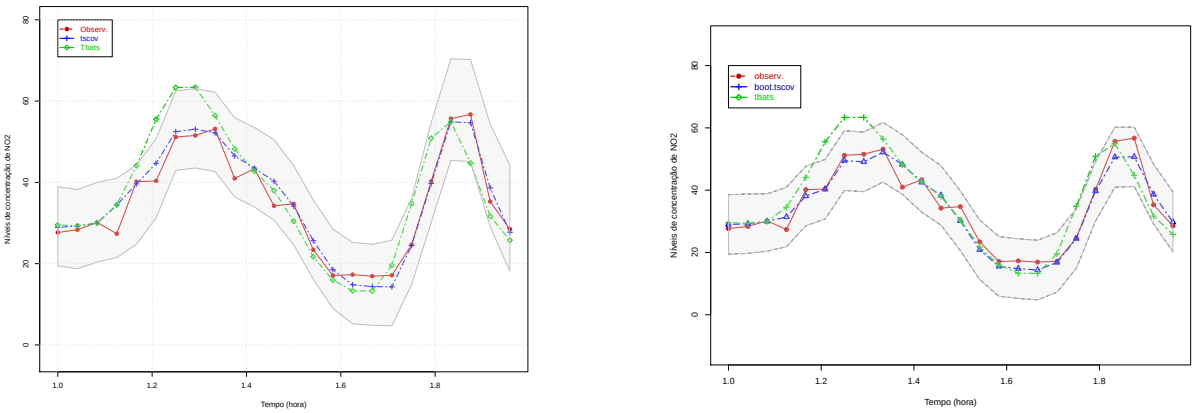
Parameter	MLE (TSCov)	St.Error	Boot.TSCov	St.Error	MLE(TBATS)
β_1^*	0.421	0.026	0.913	0.133	—
β_2^*	-0.503	0.254	-0.544	0.517	—
β_3^*	0.461	0.037	0.353	0.116	—
$\dot{a}^1$	—	—	—	—	0.437
α	—	—	—	—	1.403
β	—	—	—	—	-0.256
ϕ	0.916	0.242	0.920	0.186	0.891
σ_ε^2	3.091	0.153	3.366	0.411	—
σ_ξ^2	0.043	0.042	0.027	0.063	—
σ_ζ^2	11.453	0.374	13.313	0.624	—
σ_w^2	$\{0.032; 0.012\}$	$\{0.003; 0.014\}$	$\{0.061; 0.137\}$	$\{0.131; 0.062\}$	—
$\sigma_{w^*}^2$	$\{0.043; 0.223\}$	$\{0.237; 0.063\}$	$\{0.051; 0.302\}$	$\{0.501; 0.043\}$	—
γ_1	—	—	—	—	$\{0.0004; -0.0001\}$
γ_2	—	—	—	—	$\{0.0002; -0.0004\}$

Tabela 3.1: Estimativas dos parâmetros e os respetivos erros-padrão obtidos a partir dos modelos TSCov e Boot.TSCov. As estimativas dos parâmetros obtidos a partir do modelo TBATS estão apresentadas na última coluna.

Calculam-se as previsões até 24 passos à frente. Importante notar que as previsões *bootstrap* apresentadas são obtidas com o uso de covariáveis reais. A Figura 3.3 exibe as previsões até 24 passos obtidas com os três modelos. As medidas de precisão de previsão até 24 passos à frente estão apresentadas na Tabela 3.2.

Os resultados desse estudo também mostram que o modelo Boot.TSCov tem melhor classificação quando comparado com os resultados obtidos pelos modelos TSCov e TBATS. Ademais, os intervalos de previsão obtidos pelo modelo *Boot.TSCov* são mais precisos em relação os intervalos de previsão obtidos pelo modelo TSCov, isso permite concluir que o

procedimento *bootstrap* melhora consideravelmente as previsões obtidas usando a estratégia de horizonte crescente do estado.



(a) (b)

Figura 3.3: Valores observados e a previsão até 24 passos à frente dos níveis de concentração de NO_2 em Entre-Campos: (a) obtida pelos modelos TSCov (com covariáveis reais) e TBATS; (b) obtida pelos modelos Boot.TSCov (com covariáveis reais) e TBATS. As áreas em cinza representam os intervalos de previsão de 95% obtidos pelos modelos TSCov e Boot.TSCov.

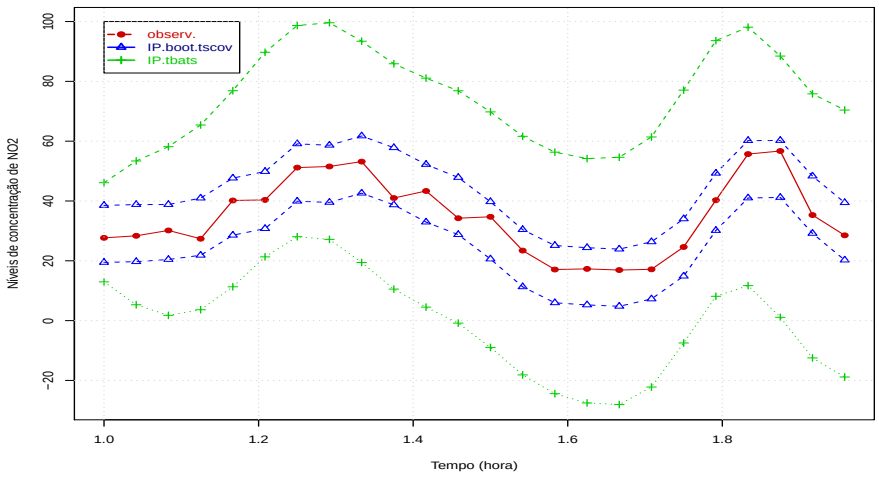


Figura 3.4: Valores observados e os intervalos de previsão de 95% obtidos pelos modelos Boot.TSCov e TBATS.

Horizonte	TSCov[1]		TSCov[2]		Boot.TSCov		TBATS	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1–3	1.536	2.422	3.607	4.147	1.613	2.113	3.071	5.416
1–6	1.945	4.536	3.368	4.864	1.827	3.753	3.422	8.354
1–9	1.865	5.505	3.319	4.917	1.874	3.941	4.217	10.171
1–12	2.551	5.867	4.754	5.521	1.735	5.344	5.135	13.434
1–15	2.734	5.713	4.938	5.176	2.430	5.811	6.116	13.278
1–18	3.046	6.487	5.177	6.733	2.913	6.232	6.435	14.622
1–21	3.372	6.923	6.326	6.857	3.462	6.547	6.581	15.065
1–24	3.611	6.964	6.461	7.714	3.661	6.833	6.803	15.275

Tabela 3.2: Precisão de previsão até 24 passos à frente dos níveis de concentração de NO_2 em Entre-Campos, obtidos com os modelos TSCov[1] (com covariáveis reais), TSCov[2] (com covariáveis previstas), Boot.TSCov (com covariáveis reais) e TBATS.

3.3.2 Aplicação a dados com múltiplos padrões sazonais: procura diária de eletricidade na Turquia

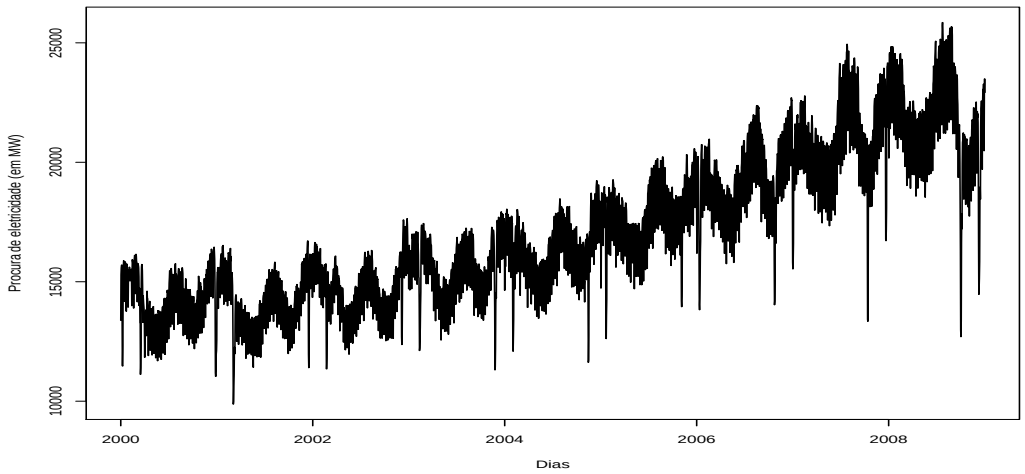
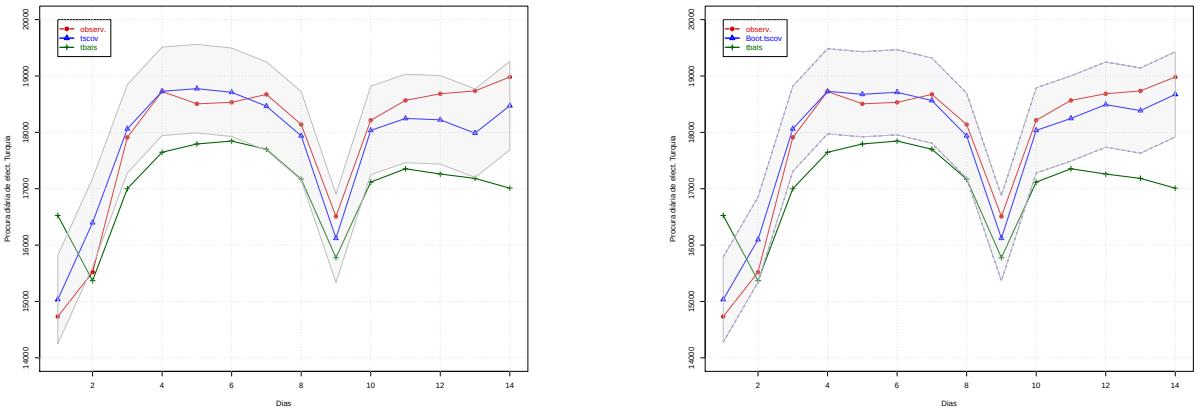


Figura 3.5: Dados de procura de eletricidade na Turquia, de 1 de janeiro de 2000 a 31 de dezembro de 2008.

Para esse caso de estudo, o conjunto de dados utilizado, Figura 3.5, é referente a procura diária de eletricidade na Turquia. A descrição desse conjunto de dados e o diagnóstico dos resíduos do modelo inicial estimado estão apresentados na secção 2.4 do Capítulo 2 - Parte II. O referido diagnóstico é feito com base no correlograma dos resíduos e o teste de Ljung-Box que fornece um Qui-quadrado = 23.131, com 18 graus de liberdade e *p-valor* = 0.081, permitindo não rejeitar a hipótese nula de que os resíduos são independentes.

O processo de estimação dos parâmetros *bootstrap* para o modelo Boot.TSCov, resulta de B = 1500 réplicas. A Tabela 3.3 exibe as estimativas dos parâmetros dos modelos TSCov, Boot.TSCov e TBATS incluindo os respetivos erros-padrão gerados pelos modelos TSCov e Boot.TSCov.

Calculam-se previsões de 14 passos à frente. A Figura 3.6 mostra as previsões obtidas com a aplicação das duas estratégias de previsão, a de horizonte crescente do estado e a estratégia bootstrap. A sobreposição dos intervalos de previsão gerados pelos dois modelos, Boot.TSCov e TBATS, estão apresentados na Figura 3.7. Em termos gráficos, o modelo Boot.TSCov produz valores futuros mais próximos dos valores observados. Este facto é também verificado na Tabela 3.4 referente a precisão de previsão, o modelo Boot.TSCov tem melhor classificação em relação os modelos TSCov e TBATS. Portanto, para esse conjunto de dados, o procedimento *bootstrap* permite razoavelmente melhorar as previsões obtidas utilizando a primeira estratégia de previsão.



(a) (b)

Figura 3.6: (a) Valores observados e a previsão de até 14 passos à frente com o modelo TSCov e TSCov; (b) valores observados e a previsão de até 14 passos à frente obtida com os modelos Boot.TSCov e TBATS. As áreas em cinza representam os intervalos de previsão de 95% gerados pelos modelos TSCov e Boot.TSCov.

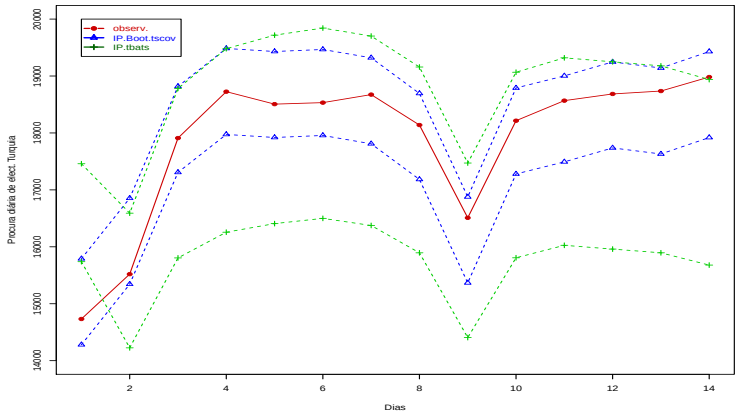


Figura 3.7: Valores observados e os intervalos de previsão de 95% obtidos pelos modelos Boot.TSCov e TBATS.

Parâmetro	MLE (TSCov)	St.Error	Boot.TSCov	St.Error	MLE(TBATS)
β_1^*	-0.655	0.023	-0.437	0.132	—
β_2^*	0.384	0.035	0.415	0.051	—
$\dot{\mathbf{a}}^1$	—	—	—	—	0.332
α	—	—	—	—	-0.127
β	—	—	—	—	0.034
ϕ	0.934	0.133	0.926	0.261	0.827
σ_ε^2	0.007	0.004	0.034	0.042	—
σ_ξ^2	0.089	0.007	0.083	0.154	—
σ_ζ^2	0.028	0.004	0.036	0.045	—
σ_w^2	$\{4.3 \times 10^{-6}; 2.4 \times 10^{-3}; 3.3 \times 10^{-3}\}$	{NA; 0.006; 0.001}	$\{4 \times 10^{-5}; 3 \times 10^{-4}; 1 \times 10^{-3}\}$	{0.031; 0.043; 0.001}	—
$\sigma_{w^*}^2$	$\{2.5 \times 10^{-6}; 5 \times 10^{-3}; 7.7 \times 10^{-4}\}$	{NA; 0.005; NA}	$\{2 \times 10^{-6}; 4 \times 10^{-5}; 4 \times 10^{-4}\}$	{0.171; 0.004; 0.014}	—
γ_1	—	—	—	—	$\{3.3 \times 10^{-4}; 6.8 \times 10^{-5}; 7.5 \times 10^{-5}\}$
γ_2	—	—	—	—	$\{4.1 \times 10^{-4}; -5.1 \times 10^{-4}; 2 \times 10^{-4}\}$

Tabela 3.3: Estimativas dos parâmetros e os respetivos erros-padrão obtidos a partir dos modelos TSCov, Boot.TSCov e TBATS

Horizonte	TSCov		Boot.TSCov		TBATS	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1 – 3	362.157	1.961	271.467	1.359	844.468	5.480
1 – 6	368.829	1.730	280.597	1.396	879.507	4.223
1 – 9	406.221	1.730	281.226	1.421	923.885	4.476
1 –12	413.658	1.833	291.365	1.415	993.773	4.963
1 –14	543.373	2.854	386.820	2.209	1138.068	5.724

Tabela 3.4: Precisão da previsão até 14 passos à frente dos modelos TSCov, Boot.TSCov e TBATS. A Previsão refere-se a procura diária de energia elétrica na Turquia.

3.3.3 Aplicação a dados de frequência não-inteira

O conjunto de dados em causa refere-se à produção semanal de gasolina nos Estados Unidos de América, Figura 3.8. A sua análise preliminar também consta na Parte II, Capítulo 2 - secção 2.5. Ambos os modelos TSCov e Boot.TSCov são aplicados sem a integração dos efeitos das covariáveis. As estimativas dos parâmetros *bootstrap* são obtidas com B = 1000 réplicas.

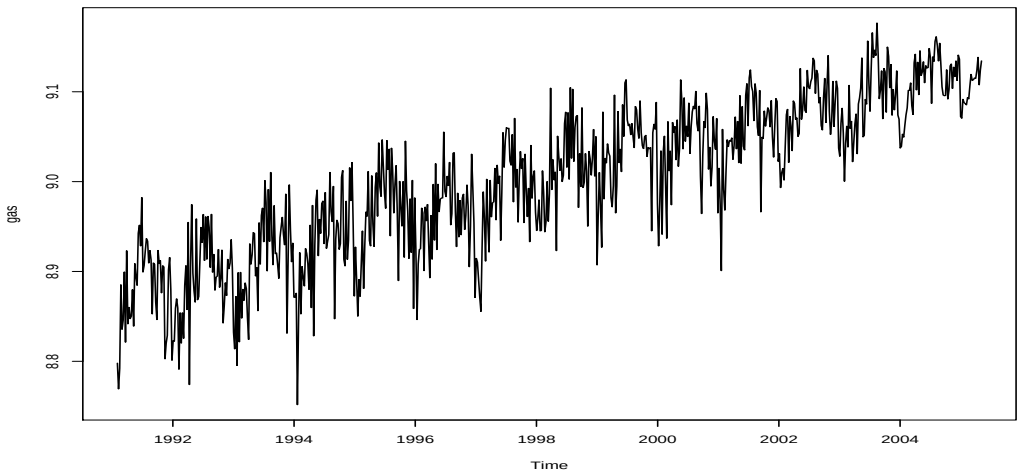


Figura 3.8: Dados sobre produção de gasolina a motor dos EUA (em milhares de barris por dia), entre Fevereiro de 1991 e Julho de 2005.

A Tabela 3.5 exibe as estimativas dos parâmetros e os erros padrão assintóticos e *bootstrap* gerados pelos modelos TSCov, Boot.TSCov e TBATS. De realçar que as estimativas dos parâmetros obtidas a partir dos dois modelos são razoavelmente semelhantes. No entanto, os erros padrão assintóticos são ligeiramente menores do que os erros padrão obtidos pelo modelo Boot.TSCov. De acordo com Shumway and Stoffer (2017), o *bootstrap* permite investigar a distribuição da amostra dos estimadores e avaliar os erros de previsão condicionais dos modelos. De facto, o nosso objetivo não é tanto da avaliação dos erros padrão das estimativas dos parâmetros, tão pouco comparar entre a aproximação assintótica das distribuições real e *bootstrap*.

Parameter	MLE(TSCov)	St.Error	Boot.TSCov	St.Error	MLE(TBATS)
$\grave{a}^1$	—	—	—	—	0.709
α	—	—	—	—	-0.063
β	—	—	—	—	0.031
ϕ	0.829	0.154	0.801	0.457	0.834
σ^2_ε	220.69	1.751	222.25	1.916	—
σ^2_ξ	1094.29	9.597	1143.19	9.947	—
σ^2_ζ	860.39	3.819	991.51	4.022	—
σ^2_w	7.055	0.489	1.249	0.604	—
$\sigma^2_{w^*}$	2.892	0.216	0.029	0.263	—
γ_1	—	—	—	—	-0.003
γ_2	—	—	—	—	0.002

Tabela 3.5: Estimativas dos parâmetros e os respetivos erros-padrão assintóticos e *bootstrap* obtidos a partir dos modelos TSCov, Boot.TSCov e TBATS.

Calculam-se previsões até 52 passos à frente, tal como apresentadas nas Figuras 3.9 e 3.10. A Figuras 3.9 exibe, não só, os valores ajustados, como também apresenta as previsões até 52 passos à frente resultantes dos modelos Boot.TSCov e TBATS incluindo os intervalos de previsão de 95% gerados pelo modelo Boot.TSCov. A Figura 3.10 apresenta partes das Figuras 2.26 e 3.9, que exibem apenas as previsões de 52 passos à frente obtidas pelos modelos TSCov, Boot.TSCov e TBATS. Finalmente, a Figura 3.11 apresenta apenas os intervalos de previsão gerados pelos modelos Boot.TSCov e TBATS.

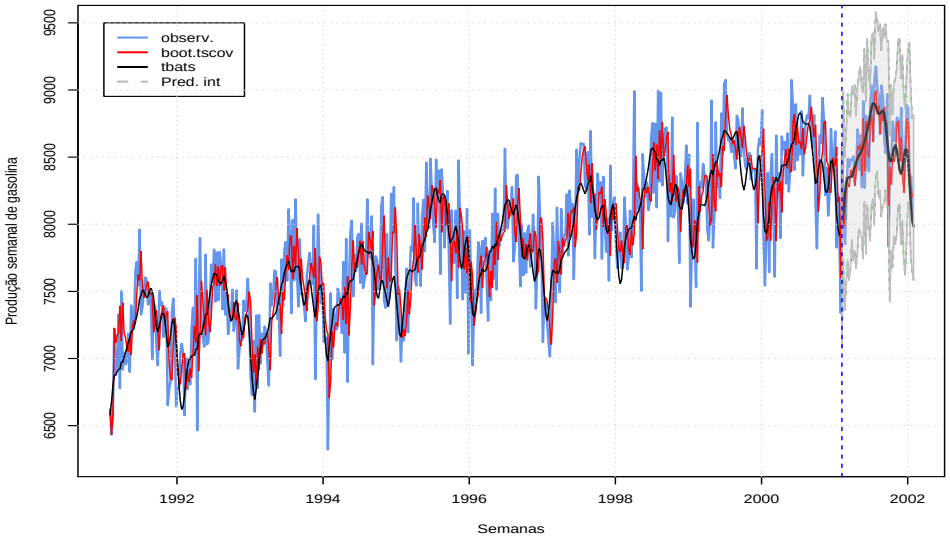
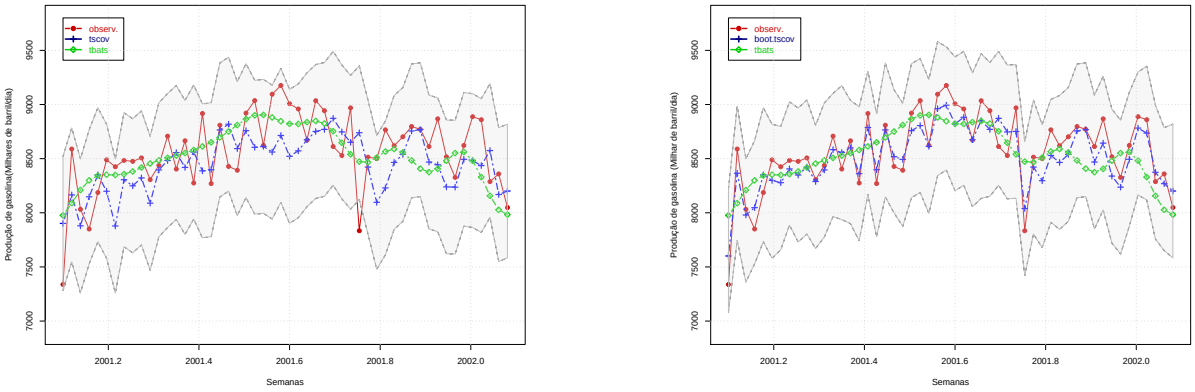


Figura 3.9: Valores observados e os ajustados incluindo as previsões até 52 passos à frente obtidos a partir dos modelos Boot.TSCov e **TBATS** sobre a produção semanal de gasolina nos Estados Unidos. As linhas pontilhadas verdes indicam os intervalos de previsão, superior e inferior, de 95% obtidos pelo modelo Boot.TSCov.



(a) (b)

Figura 3.10: (a) previsão até 52 passos à frente da produção semanal de gasolina nos Estados Unidos da América obtida: (a) pelos modelos TSCov com covariáveis reais e TBATS; (b) pelos modelos Boot.TSCov com covariáveis reais e TBATS. As áreas em cinza representam os intervalos de previsão de 95% obtidos pelos modelos TSCov e Boot.TSCov.

Figura 3.11: Valores observados e os intervalos de previsão de 95% obtidos pelos modelos Boot.TSCov e TBATS.

Conforme os resultados da Tabela 3.6 e o ajuste gráfico, pode-se concluir que todos os resultados são concordantes da melhoria substancial observada nos resultados obtidos a partir do modelo Boot.TSCov. Essa melhoria é mais expressiva quando comparado com o modelo TSCov. Quanto ao modelo TBATS, ambos os modelos continuam concorrentes.

Horizonte	TSCov		Boot.TSCov		TBATS	
	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

Tabela 3.6: Erros de previsão até 52 passos à frente sobre a Produção de Gasolina nos Estados Unidos de América obtidos a partir dos modelos TSCov, Boot.TSCov e TBATS.

3.4 Considerações do Capítulo

Nesse Capítulo um procedimento bootstrap para obter previsão de curto prazo no contexto do modelo estrutural trigonométrico formulado para séries temporais com sazonalidade complexa foi apresentado. O principal objetivo foi o aprimoramento das previsões de curto prazo obtidas nos casos anteriores estudados, cujos resultados estão apresentados no Capítulo 2. O procedimento foi satisfatoriamente aplicado para estimar o modelo e calcular as previsões. Os resultados obtidos mostram que o procedimento *bootstrap* construído fornece bons resultados quando aplicado para a estimação e cálculo de previsões para o conjunto de dados utilizados. Obtém melhores resultados quando comparado com os obtidos no Capitulo 2, particularmente para o conjunto de dados referente as concentrações de NO_2, quer para o modelo TSCov como para TBATS. Com relação ao conjunto de dados sobre a produção semanal de gasolina nos Estados Unidos de América, o modelo *bootstrap* tem melhor classificação quando comparado com os resultados do modelo TSCov, mas considera-se ainda concorrente com o modelo TBATS. Importante lembrar que para o conjunto de dados referenciado, o modelo TBATS é confrontado com o modelo *bootstrap* sem a integração dos efeitos das covariáveis.

	user	system	elapsed
T. Computacional	28532.97s	196.10	28888.19s

Tabela 3.7: Tempo computacional para estimação e previsão do modelo bootstrap usando os dados de NO_2.

Existem várias versões *bootstrap* que a literatura apresenta. Os resultados obtidos com o método aqui proposto demonstram que é uma possibilidade valiosa para previsão de curto prazo de séries temporais com sazonalidade complexa e particularmente para séries temporais com período não-inteiro. Do ponto de vista estrutural, é um procedimento bastante simples. No entanto, em termos computacionais acarreta custos elevados, pois, exige do computador maior capacidade de processamento. A Tabela 3.7 mostra o tempo computacional para

estimação e previsão do modelo *bootstrap* usando os dados referentes às concentrações do NO_2 com 1500 observações, cerca de 8 horas de tempo de execução do código *R*.

Capítulo 4

Modelo de Previsão da Procura de Energia Elétrica em Cabinda

Índice do Capítulo

4.1 Introdução

Nesse Capítulo, dois conceitos base da economia serão amplamente utilizados. Trata-se da *Procura* e *Consumo*. Do ponto de vista econômica, entende-se por *Procura* à quantidade de um produto que um consumidor ou comprador estaria disposto a comprar a qualquer preço. O *Consumo* consiste na aquisição e utilização de um bem ou serviço. De acordo com essas definições, doravante nos referiremos da *Procura* na ótica da intenção de obter o produto, no caso, a energia elétrica. Usaremos o termo *Consumo* na vertente do benefício efetivo.

O estudo da procura de energia elétrica é uma das preocupações mais importantes para os gestores de empresas de produção e distribuição de energia elétrica, sobre tudo quando se pretende estimar a quantidade de capacidade adicional necessária para garantir a oferta suficiente de energia. Devido ao seu papel fundamental para o funcionamento eficaz e econômico dos sistemas de geração da energia elétrica, é fundamental que os gestores primem por um planeamento baseado em previsões da procura de energia elétrica para garantir a oferta.

O objetivo nesse Capítulo é estimar um modelo para gerar previsões pontuais e probabilísticas sob a forma de densidades preditivas[1] de curto prazo da procura diária total de energia elétrica na cidade de Cabinda. Aplica-se o modelo TSCov à dados referentes ao consumo total diário de energia elétrica em Cabinda.

[1]Essas previsões são calculadas usando o método HDR proposto por Hyndman (1996). O objetivo é gerar as regiões de densidade relativamente alta.

A vantagem de se fazer previsões probabilísticas está no fato das empresas precisarem da previsão de energia probabilística presente em todo o planeamento e operações da cadeia de valores de energia elétrica; e as previsões probabilísticas sob a forma de densidades preditivas são necessárias para resolver muitos dos desafios enfrentados pelos gestores de empresas. Por exemplo, a energia elétrica não pode ser armazenada, a geração instantânea deve coincidir com a procura a partir do sistema. Para assegurar esse equilíbrio entre procura e oferta, bem como a segurança e qualidade no fornecimento de energia elétrica, a previsão do consumo elétrico de curto prazo é necessária. Tais previsões fornecem a base para a geração e manutenção da programação, e podem ser utilizadas para estimar os fluxos de carga elétrica de forma mais eficiente, impedindo que o sistema sofra perturbações graves (Rebennack et al., 2010). Mas, a questão chave, do ponto de vista operacional, é de fato, saber se haverá problemas em atender a procura máxima; e a incapacidade de atendê-la, obviamente, pode resultar em apagões inesperados. Um cenário possível para esse problema é fazer previsão e estimar as regiões de maior densidade de consumo, o que permite expor as características mais marcantes da procura.

As previsões probabilísticas sob a forma de densidades preditivas fornecem visualização da distribuição e transmitem consideravelmente mais informações do que pode ser obtida com base na função de distribuição empírica. A percepção da forma das densidades preditivas estimadas a partir das recursividades do filtro de kalman, resulta da aplicação do método proposto por Hyndman (1996), que visa estimar as regiões de previsão. O método consiste em resumir a distribuição de probabilidade por uma região do espaço amostral com o fim de identificar um conjunto relativamente pequeno que contém a maior parte da probabilidade, que Hyndman (1996) denomina por HDR (*Highest Density Regions*).

4.2 Panorama do Setor de Energia Elétrica em Cabinda

Angola, República situada entre os paralelos $4°22'$ e $18°02'$ e os meridianos $4°05'$ e $11°45'$ a Este de Greenwich, no Hemisfério Sul, na parte Ocidental da África Austral e ocupa uma área de $1.246.700km^2$. É limitada a Norte pela República do Congo e por uma parte da República Democrática do Congo (ex-Zaíre); a Leste pela República da Zâmbia e por uma parte da República Democrática do Congo; a Sul, pela República da Namíbia e a Oeste, pelo Oceano Atlântico, Figura 4.1. Está dividido entre uma faixa costeira árida, que se estende desde a Namíbia até Luanda, um planalto interior húmido, uma savana seca no interior Sul e Sudoeste, e floresta tropical no Norte. É um país marcado por apenas duas estações distintas: o verão e o inverno.

O setor de energia elétrica, em geral, apresenta duas componentes distintas: uma voltada para o exterior, obedecendo à lógica de funcionamento do mercado internacional e outra, obedecendo um modelo puramente interno. Devido à guerra civil que teve início em 1975 e continuou, com alguns intervalos, até 2002, a componente eletricidade do sistema energético nacional foi abalada; que só depois de 2002 foram encontradas as condições para a sua estruturação, organização e funcionamento, Figura 4.2.

Figura 4.1: Situação geográfica de Angola no contexto Africano.

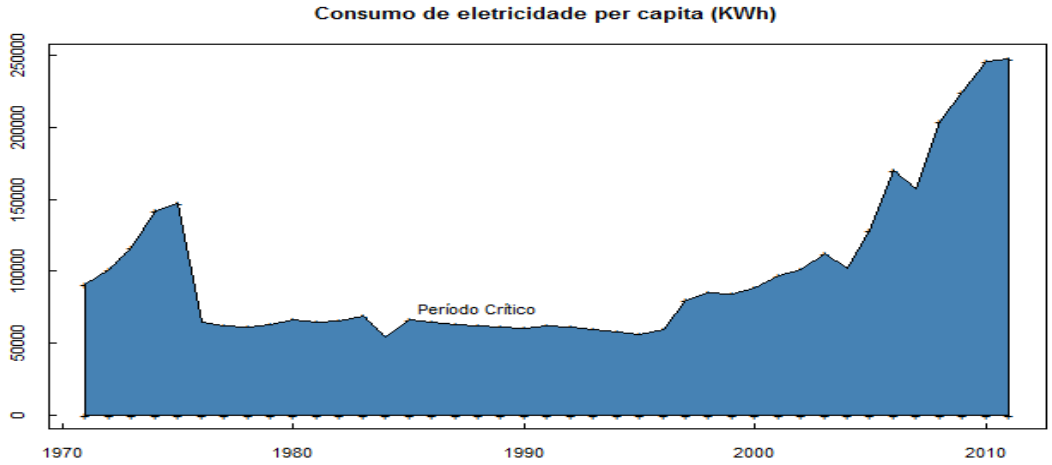


Figura 4.2: Consumo per capita de energia elétrica (em KWh) entre 1971 e 2010 em Angola.

A produção e distribuição de eletricidade recaem na Rede Nacional de Transporte de Energia Elétrica. A EPE (Empresa de Produção de Eletricidade) e ENDE (Empresa Nacional de Distribuição de Eletricidade) são as empresas estatais com a responsabilidade fundamental sobre o setor elétrico. Estas empresas têm levado a cabo vários projetos para a melhoria do setor elétrico de Angola, orientadas pelo Ministério da Energia e Águas.

Cabinda é uma das 18 províncias que constituem a República de Angola, sendo um enclave limitado ao Norte pela República do Congo, a Leste e ao Sul pela República Democrática do Congo, e a Oeste pelo Oceano Atlântico, Figura 4.1. Com uma superfície de $7.283 km^2$ e cerca de 300.000 habitantes, o clima é tropical quente e húmido, com precipitações anuais em torno de $800 mm$. A temperatura média anual varia entre os $25°$ e os $30°$ Celsius. O verão - húmido e quente, decorre de setembro/outubro a maio/junho; e o inverno - fresco e seco, que vai de

Figura 4.3: Uma das três centrais térmicas da cidade de Cabinda: central térmica de Malembo.

maio/junho a setembro/outubro. O sistema elétrico está baseado em três centrais térmicas com funcionamento em paralelo, Figura 4.3. Todas sincronizadas com mesma frequência e voltagem para produção de um total de 77.5 MW de energia elétrica. Todavia, a província continua registar défices no fornecimento de energia elétrica, provocados pela falta de capacidade de resposta à crescente procura da eletricidade.

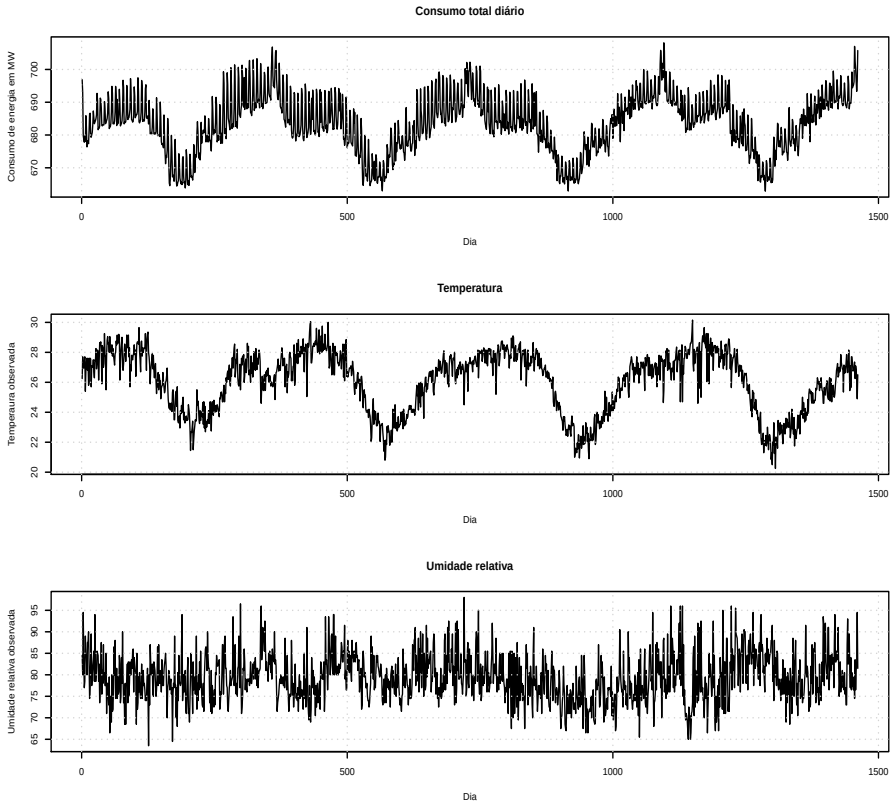


Figura 4.4: Procura diária de energia elétrica na cidade de Cabinda. I painel: consumo total diário de energia elétrica; II painel: variação diária da temperatura; III painel: variação diária da humidade relativa. Todas as variáveis são observada no período entre 01 de Janeiro de 2011 e 31 de Dezembro de 2014.

Os dados utilizados para esse estudo, Figura 4.4, são fornecidos pelas seguintes Instituições de Angola: a ENDE (Empresa Nacional de Distribuição de Energia Elétrica) em Cabinda forneceu os dados sobre o consumo horário; o INAMET (Instituto Nacional de Meteorologia) em Cabinda forneceu o conjunto de dados referentes a variação diária da temperatura e umidade relativa. Importante notar que a leitura do consumo de energia elétrica é feita em intervalos de um hora, enquanto a temperatura e a humidade relativa as leituras são feitas em intervalos de um dia. Por esse motivo, nós transformamos os dados horários de energia elétrica em total diário.

De acordo com a Figura 4.6, o efeito de calendário na procura de energia elétrica é visível entre o verão, concretamente para o período de altas temperaturas (Dezembro, Janeiro, Fevereiro, Março e Abril), e o inverno. O máximo de consumo total diário é de 708.1MW cujo pico desse consumo é registado no período entre $20h$ e $22h$, justificado fundamentalmente pelo uso da iluminação pública e aos diferentes níveis no comportamento das famílias no que tange o uso de eletro-domésticos como aparelhos de ar condicionado, lâmpadas residenciais, etc. O mínimo do consumo total diário é de 662.9MW e o pico de mínimo decorre às $9h$ matinal e vai crescendo lentamente até atingir o pico. Em média o consumo total diário é de 683.7MW. O histograma do consumo total diário de energia elétrica na cidade de Cabinda, apresentado na Figura 4.5, revela uma distribuição bimodal.

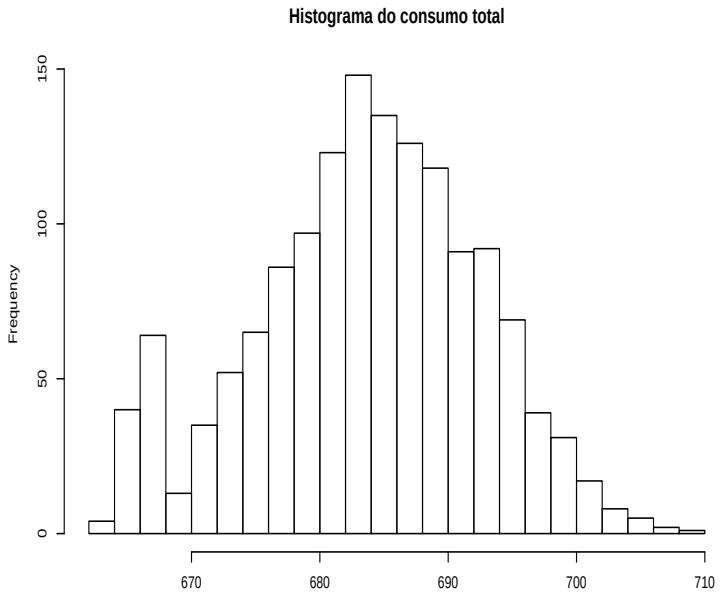
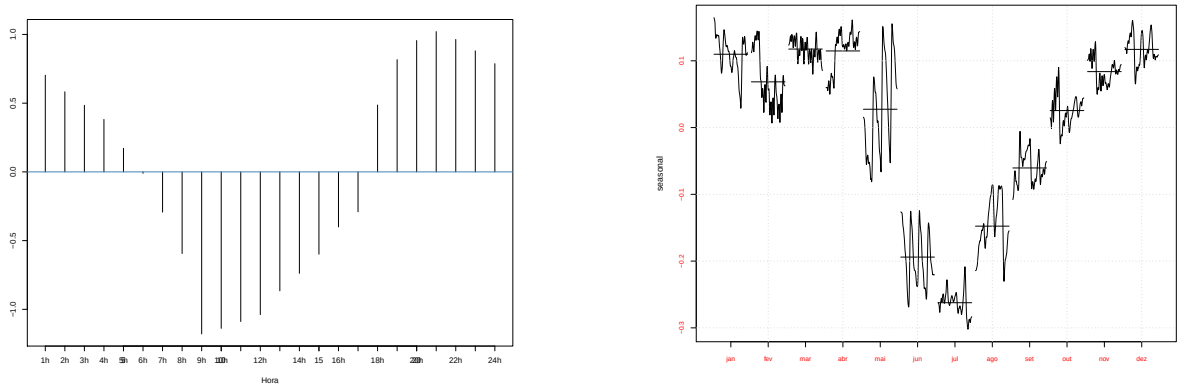


Figura 4.5: Histograma do consumo total diário de energia elétrica na cidade de Cabinda.



(a) (b)

Figura 4.6: Consumo horário de energia elétrica na cidade de Cabinda. (a) dinâmica da procura horária por dia, (b) dinâmica da procura horária por mês.

4.3 Análise Prévia do Conjunto de Dados

Primeiro analisa-se a série em si, Figura 4.7, que revela ser altamente correlacionada entre si. Ademais, a série tem dois padrões sazonais: um padrão semanal com periodicidade 7 e um padrão anual com periodicidade 365.25.

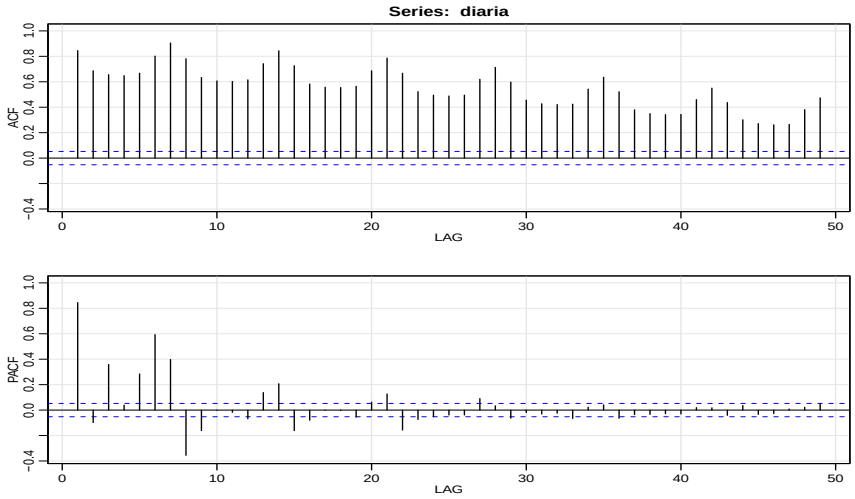


Figura 4.7: Correlograma do consumo total diário de energia elétrica na cidade de Cabinda.

Em seguida, uma análise preliminar é feita para determinar os possíveis desfasamentos nas covariáveis. A estratégia é aplicar o modelo TBATS à série de consumo total diário de energia elétrica para se obter os resíduos e determinar a correlação cruzada entre os resíduos do consumo, a série de temperatura e a série de humidade relativa. A Figura 4.8 apresenta o resultado que a priori revela a inexistência de correlação com a temperatura e a humidade relativa. Não obstante, o contexto do uso das covariáveis não exige tanto o rastreio do grau de similaridade (ou relação) entre as séries temporais envolvidas no processo. Importante lembrar que a relação de causalidade entre o consumo de eletricidade e as covariáveis para esse contexto

não é tida em conta. As covariáveis apenas representam séries temporais disponíveis em tempo oportuno para melhorar o processo de previsão (Hyndman et al., 2008).

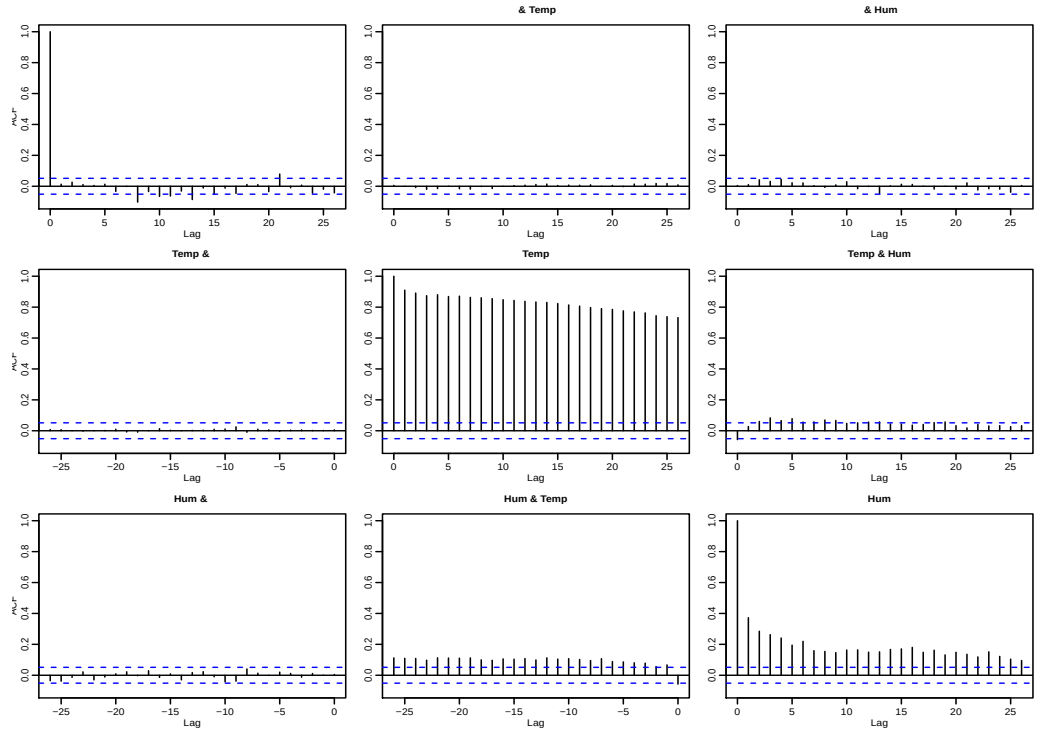


Figura 4.8: Correlação cruzada entre os resíduos do consumo total diário de energia elétrica, a série de temperatura e a série de humidade relativa.

Designa-se a série de consumo total diário por $\boldsymbol{C}_t$, a temperatura por $\boldsymbol{T}_t$ e a humidade relativa por $\boldsymbol{H}_t$. Estas duas últimas entram no processo como covariáveis. De acordo com a Figura 4.9, que apresenta o consumo de energia elétrica em relação à temperatura, revela uma relação linear entre as duas variáveis. Um aumento no consumo de energia elétrica relacionado aos fatores descritos acima durante o verão é visível. Definimos essa relação como linear por partes com um corte na linha de regressão em torno de $25°C$ para indicar o efeito dos níveis da procura total diária de energia elétrica Gob et al. (2013).

$$\boldsymbol{I}_t \quad = \quad \begin{cases} 1 & \text{se } T_t > 25°C \\ 0 & \text{se } T_t \leq 25°C \end{cases} \tag{4.1}$$

onde $25°C$ é o limiar abaixo do qual o consumo de energia elétrica é considerado não afetado pela temperatura.

4.4 Estimação do Modelo e Previsão Pontual

A série de teste é composta de 1166 observações que corresponde o período entre 01 de Janeiro de 2011 até 11 de Março de 2014. A série de validação tem 295 observações. A média e a covariância do estado do sistema são inicializadas em $\boldsymbol{x}_0 = 0$ e $\boldsymbol{P}_{0i} = 30$, com $i = 14$, respetivamente. A covariância da observação é inicializada em $\boldsymbol{R}_0 = \sigma_\varepsilon^2 = 10^{-8}$ e as variâncias

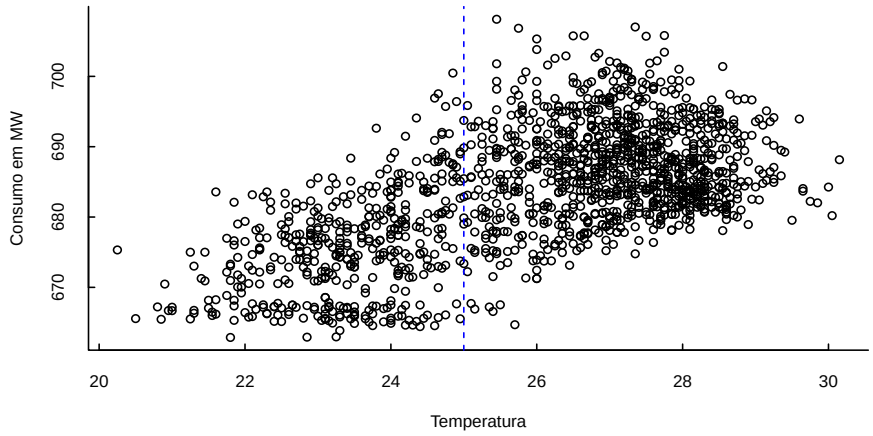


Figura 4.9: Consumo total diário (em MW) em função da temperatura (em $^{\circ}C$).

do vetor de estados são inicializadas por $\boldsymbol{Q}_0 = diag\{\sigma_\xi^2, \sigma_\zeta^2, \sigma_w^{2(i)}\} = \{4, 0.9, 0, 0\}$, com $i = 4$. Os componentes da matriz coeficiente de regressão são inicializados por $\{\beta_1, \beta_2, \beta_3\} = 0.01$; o fator de esquecimento é fixado em $\delta = 0.6$.

As estimativas dos parâmetros estão apresentadas na Tabela 4.1. A avaliação da adequação do modelo ajustado aos dados é verificada pela análise do correlograma dos resíduos, Figura (4.10), que não apresenta picos significativos em todos os *lags*. Ademais, o teste de Box-Ljung sobre a independência dos resíduos fornece um valor de Qui-quadrado igual a 22.436, com 29 graus de liberdade e *p-valor* = 0.586; o que permite não rejeitar a hipótese nula de que os resíduos são independentes. Os erros de previsão um passo à frente estão apresentados na Tabela 4.2. As estimativas dos parâmetros e os respetivos erros-padrão estão exibidos na Tabela 4.1.

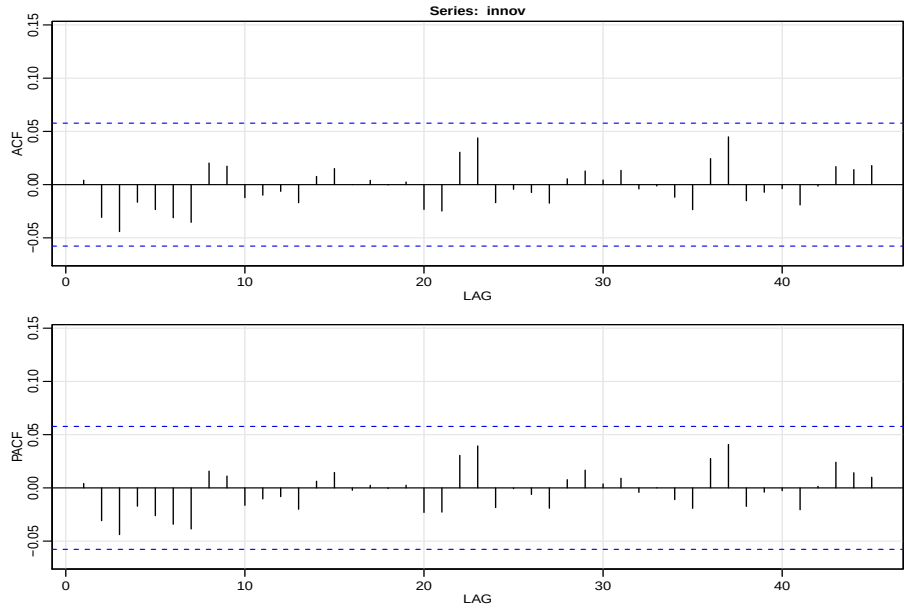


Figura 4.10: Correlograma dos resíduos da previsão um passo à frente da procura total diária de energia elétrica na cidade de Cabinda.

Parâmetro	MLE (TSMCov)	E.Padrão
σ^2_ε	2.179	0.623
β^*_1	0.909	0.113
β^*_2	0.088	0.025
β^*_3	3.299	0.402
σ^2_ξ	3.460	0.313
σ^2_ζ	0.026	0.116
ϕ	0.800	0.071
σ^2_w	$\{0.336; 0.258\}$	$\{0.322; 0.213\}$
$\sigma^2_{w^*}$	$\{0.062; 0.226\}$	$\{0.222; 0.012\}$

Tabela 4.1: Estimativas dos parâmetros e os respetivos erros-padrão do modelo de previsão da procura total diária de energia elétrica na cidade de Cabinda.

Modelo	ME	RMSE	MAE	MPE	MAPE
TSMCov	0.026	2.728	1.984	0.003	0.289

Tabela 4.2: Erros de previsão um passo à frente do modelo estimado para a previsão da procura total diária de energia elétrica na cidade de Cabinda.

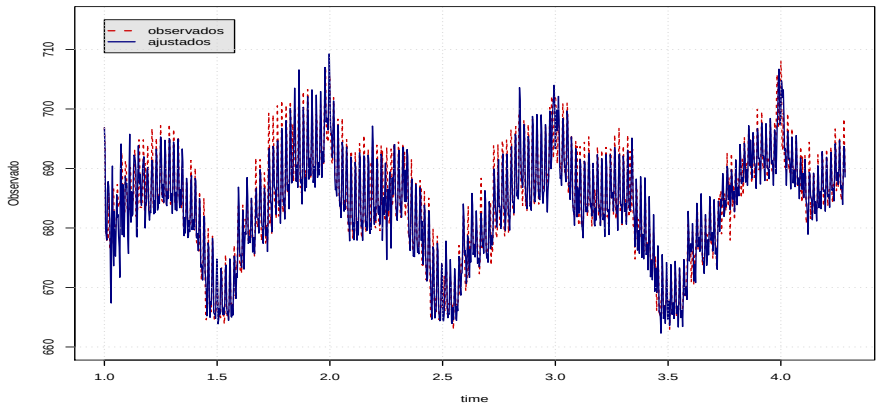


Figura 4.11: Valores observados e os ajustados do modelo estimado sobre a previsão da procura total diária de energia elétrica na cidade de Cabinda.

Para esse caso de estudo, as previsões calculadas resultam do uso de covariáveis reais, no caso a temperatura e a humidade relativa. Os resultados da previsão um passo à frente e a previsão até 7 passos à frente, assim como as medidas de precisão da previsão, estão apresentados nas Figuras 4.11, 4.12 e Tabela 4.3, respetivamente. A avaliação das distribuições de previsão estão exibidas na Figura 4.13. O modelo de previsão que descreve a dinâmica da procura total diária na cidade de Cabinda, para o período entre 01 de Janeiro de 2011 e 31 de Dezembro de 2014, é expresso por

$$\begin{aligned} C_t &= \ell_{t-1} + 0.80 b_{t-1} + \sum_{i=1}^{T} s^{(i)}_{t-1} + 0.91 T_t + 0.09 H_t + 3.29 I_t + \varepsilon_t \\ \ell_t &= \ell_{t-1} + 0.80 b_{t-1} + \xi \\ b_t &= 0.80 b_{t-1} - \zeta_t \end{aligned}$$

$$
\begin{aligned}
s_t &= \sum_{j=1}^{2} s_{j,t} \quad \text{(padrão sazonal semanal)} \\
s_{j,t} &= s_{j,t-1}\cos\left(\frac{2\pi jt}{7}\right) + s^*_{j,t-1}\sin\left(\frac{2\pi jt}{7}\right) + w^{(i)}_{j,t} \\
s^*_{j,t} &= -s_{j,t-1}\sin\left(\frac{2\pi jt}{7}\right) + s^*_{j,t-1}\cos\left(\frac{2\pi jt}{7}\right) + w^{*(i)}_{j,t} \\
s_t &= \sum_{j=1}^{5} s_{j,t} \quad \text{(padrão sazonal anual)} \\
s_{j,t} &= s_{j,t-1}\cos\left(\frac{2\pi jt}{365.25}\right) + s^*_{j,t-1}\sin\left(\frac{2\pi jt}{365.25}\right) + w^{(i)}_{j,t} \\
s^*_{j,t} &= -s_{j,t-1}\sin\left(\frac{2\pi jt}{365.25}\right) + s^*_{j,t-1}\cos\left(\frac{2\pi jt}{365.25}\right) + w^{*(i)}_{j,t} \quad \text{onde} \\
\varepsilon_t &\sim \mathcal{N}(0, 2.179) \\
\xi_t &\sim \mathcal{N}(0, 43.460) \\
\zeta_t &\sim \mathcal{N}(0, 0.026) \\
w_{1,t} &\sim \mathcal{N}(0, 336); \quad w_{2,t} \sim \mathcal{N}(0, 0.258) \\
w^*_{1,t} &\sim \mathcal{N}(0, 0.062); \quad w^*_{2,t} \sim \mathcal{N}(0, 0.226)
\end{aligned}
$$

Para a estimação deste modelo foram necessários $\dot{k}^*_1 = 2$ harmônicas significativos para os termos trigonométricos do padrão sazonal semanal com periodicidade 7 e $\dot{k}_2 = 5$ para o padrão sazonal anual com periodicidade 365.25. O vetor dos estados estimado é de dimensão 14.

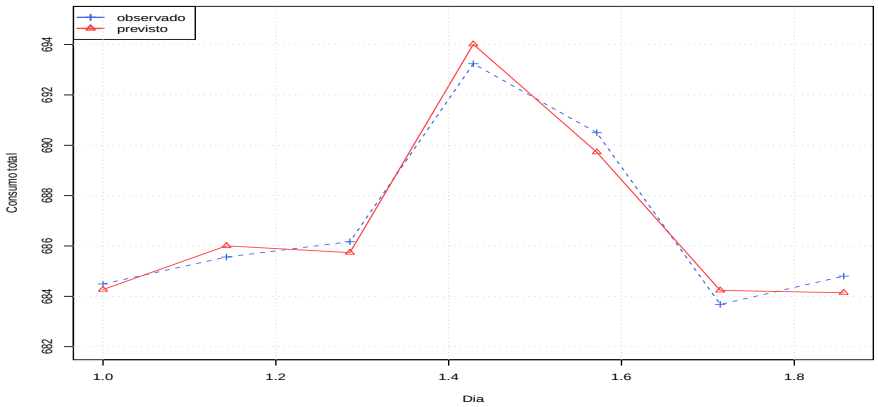
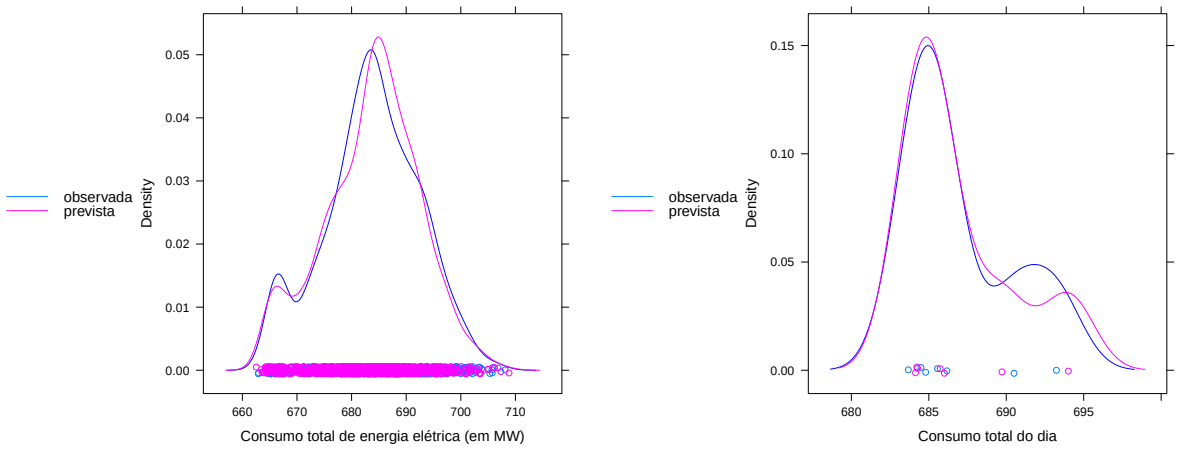


Figura 4.12: Valores observados e a previsão até uma semana à frente que corresponde o período entre 12 e 18 de Março de 2014.

4.5 Previsão Probabilística Sob a Forma de Densidades Preditivas

Existem vários métodos estatísticos que permitem resumir uma distribuição de probabilidade por uma região do espaço amostral cobrindo uma probabilidade especificada. Uma das estratégias de seleção de tal região é exigir que ela (a região) contenha pontos de densidade



(a) (b)

Figura 4.13: Avaliação das distribuições de previsão. (a) densidade do consumo total de energia elétrica observado e a dos valores ajustados; (b) densidade dos valores do consumo total de energia observado e a previsão até uma semana à frente.

relativamente alta. Quais as regiões da função HDR? Do ponto de vista matemático, Hyndman (1996) define:

- O subconjunto do espaço amostral (as regiões) $R(f_\alpha)$ com densidade maior a um valor f_α tal que a chance de encontrar uma variável nessa região seja igual ou maior que $1-\alpha$. O subconjunto dado por HDR é a menor dessas regiões,

$$R(f_\alpha) = \{x : f(x) \geq f_\alpha\}$$

onde f_α é a maior constante, tal que

$$P[X \in R(f_\alpha)] \geq 1-\alpha$$

sendo $f(x)$ a função densidade da variável aleatória X.

O propósito por trás de resumir uma distribuição de probabilidade por uma região do espaço amostral é identificar um conjunto relativamente pequeno que contém a maior parte da probabilidade, embora a densidade possa ser diferente de zero sobre regiões infinitas do espaço amostral. Do ponto de vista prático, as regiões de maior densidade são um resumo mais efetivo da distribuição de previsão.

Uma das questões fundamentais para os pesquisadores tem a ver com a forma de visualizar as previsões probabilísticas. Comunicar incertezas mais profundas resultantes de conhecimento incompleto sobre o futuro é um desafio. A previsão probabilística geralmente é representada como um histograma. A escolha mais adequada de visualização para ilustrar a incerteza futura depende muito dos objetivos do pesquisador, do contexto da comunicação e do público alvo (Spiegelhalter et al., 2011). O facto de existirem diferentes tipos de usuários, o uso de previsões probabilísticas sugere a importância da interação entre o pesquisador de previsões e os usuários, considerando os seus objetivos. Pois, é possível a compreensão das previsões probabilísticas, mesmo que não tenham formação avançada em estatística. Nesse

contexto, o método HDR é uma alternativa adequada, uma vez que permite visualizar, não só, a incerteza futura expondo as características mais impressionantes da densidade prevista, como também permite estimar o máximo local num conjunto de valores.

Os resultados aqui apresentados são obtidos com a aplicação das funções *cde()* e *hdr()* do pacote Hyndman (2018), que permitiram estimar a densidade preditiva e as regiões de maior densidade exibidas nas Figuras 4.14 e 4.15. Dadas as previsões pontuais obtidas pelo modelo TSCov, Figura 4.12, as densidades preditivas condicionais estimadas, usando a função *cde()* para o período entre 12 e 18 de Março de 2014 estão apresentadas na Figura 4.14. Essas densidades relacionam-se com a informação ilustrada na Figura 4.15.

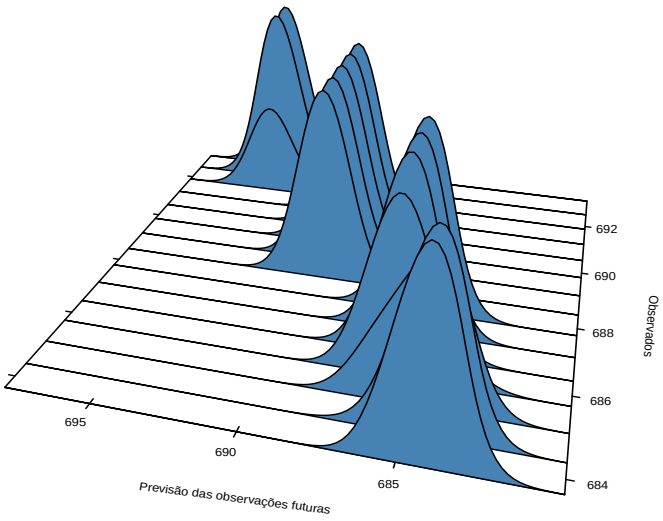


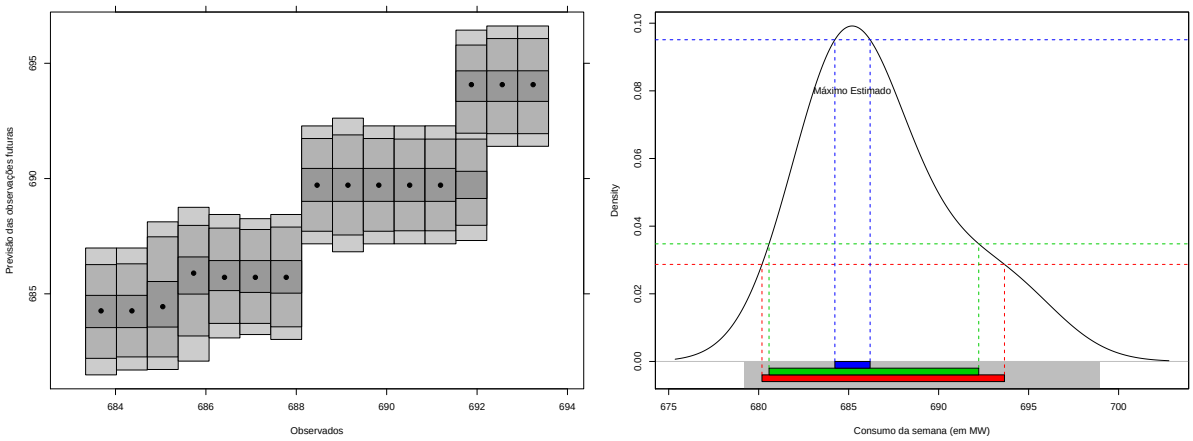
Figura 4.14: Estimativa das densidades condicionais para o período entre 12 e 18 de Março de 2014, relacionadas com as regiões de maior densidade.

As regiões onde o consumo de energia elétrica difere em densidade condicional, que chamamos de "regiões de excesso no consumo", estão apresentadas no gráfico 4.15a. Os níveis de cobertura probabilística aplicados para as regiões estimadas são fixadas em $prob = \{50, 95, 99\}$ e os intervalos estimados estão apresentados abaixo.

- Cobertura de 50%, $[684.23; 686.19]$;
- Cobertura de 95%, $[680.58; 692.23]$ e
- Cobertura de 99%, $[680.19; 693.66]$.

A moda estimada é $685.1979MW$, conforme ilustrada na Figura 4.15b.

Sabe-se que a moda populacional de uma distribuição de probabilidade contínua é o valor em que a função densidade de probabilidade atinge o valor máximo, ou seja, o valor que está no pico. Assim, os máximos globais são modas. Ademais, quando uma função densidade de probabilidade tem vários máximos locais, é comum referir-se a todos os máximos



(a) (b)

Figura 4.15: Estimativa das regiões de maior densidade para o período entre 12 e 18 de Março de 2014.

locais como modas da distribuição. Dentro de cada região estimada, Figura 4.15a, observam-se pontos – são os pontos com densidade de probabilidade relativamente alta em relação todos os pontos fora das regiões. Observa-se também que a primeira região contém a maior quantidade de pontos. Do ponto de vista prático, isso significa que na semana de 12 a 18 de Março de 2014, o fluxo de carga elétrica consumida com maior frequência fixou-se entre 684 e 688MW, com um máximo local em torno de 685MW. Este valor coincide com a moda estimada e apresentada na Figura 4.15b. Portanto, pode-se razoavelmente concluir que o Fluxo Máximo de Carga Elétrica Consumida durante a semana de 12 a 18 de Março de 2014 é aproximada à 685.2MW. Na Figura 4.15b este pico está ilustrado pelo retângulo limitado com linhas descontínuas azuis.

Horizonte	ME	RMSE	MAE	MPE	MAPE
1 – 2	-0.111	0.352	0.334	-0.016	0.048
1 – 3	0.071	0.382	0.368	0.010	0.054
1 – 4	-0.139	0.508	0.469	-0.020	0.068
1 – 5	0.043	0.572	0.530	0.006	0.077
1 – 6	-0.056	0.569	0.534	-0.008	0.078
1 – 7	0.046	0.583	0.552	0.007	0.080

Tabela 4.3: Erros de previsão até uma semana à frente.

4.6 Considerações do Capítulo

O propósito principal desse Capítulo foi o de estimar um modelo específico capaz de calcular previsões pontuais e probabilísticas sob a forma de densidades preditivas de curto prazo. Em torno desse estudo procurou-se responder as três perguntas formuladas no Capítulo introdutório dessa tese, que constituem a motivação de âmbito social desse projeto de pesquisa.

- Através da análise descritiva preliminar foi constatado que o efeito de calendário na

procura de energia elétrica na cidade de Cabinda é bastante visível nas duas principais estações do ano, verão e inverno, e de fácil caraterizar. Ou seja, a procura de eletricidade é maior no verão do que no inverno, como é obvio. O pico do consumo diário é registado no período entre $20h$ e $22h$ e o mínimo decorre às $9h$ matinal.

- Um modelo de previsão da procura diária de energia elétrica na cidade de Cabinda foi estimado. Os resultados obtidos permitem assegurar que o modelo estimado é capaz de captar as principais características presentes na série observada e é adequado para descrever a dinâmica da procura diária de energia elétrica na cidade de Cabinda e calcular previsões.
- Com o modelo estimado, foi possível calcular as regiões onde o consumo de eletricidade difere em densidade. Permitiu, igualmente, estimar o fluxo máximo de carga elétrica consumida durante o período previsto. Do ponto de vista do planeamento e operação dos sistemas de geração de energia elétrica, os resultados obtidos pela aplicação do método HDR permitem que o Gestor da ENDE tome decisões seguras sobre:
  - A produção de energia elétrica que deve coincidir com a procura para assegurar o equilíbrio entre a procura e a oferta;
  - O fluxo de carga elétrica a estimar de forma a impedir que os sistemas sofram perturbações graves e evitar desperdícios.

Parte III

Resultados da Tese, Conclusão e Trabalho Futuro

Capítulo 1

CONTRIBUIÇÕES, RESULTADOS DA TESE, CONCLUSÃO E TRABALHO FUTURO

Índice do Capítulo

1.1 Contribuições

Esta tese tem como principal objetivo fornecer contribuições inerentes à construção de modelos de previsão para séries temporais com padrões sazonais complexos. A esse respeito, as principais contribuições são as seguintes.

- É proposto um modelo estrutural básico, SCov, com efeitos das bivariáveis para previsão de séries temporais de sazonalidade não complexa. A formulação do modelo integra as três principais componentes não observáveis: o nível, a tendência e a sazonalidade.
- No contexto da previsão de séries temporais com padrões sazonais complexos, é proposto o modelo estrutural trigonométrico com efeitos das covariáveis. A sua formulação integra, igualmente, as três componentes não observáveis. A componente sazonal é modelada mediante séries de Fourier e o seu ruído é projetado no intuito de: (i) ser a fonte de aleatoriedade para a componente sazonal em si e; (ii) propagar o efeito dessa aleatoriedade nos coeficientes dos termos trigonométricos estocasticamente variantes ao longo do tempo.
- Do ponto de vista da extração do sinal de uma série temporal, um filtro de Kalman com as matrizes de covariância do sistema calculadas recursivamente é construído. O cálculo recursivo dessas matrizes é baseado nas inovações, a priori e a posteriori, do modelo. As inovações são incorporadas nas estruturas das matrizes de covariâncias de modo a influenciarem o ajusto das mesmas num processo recursivo até melhorar a precisão da estimativa do estado.

- No domínio da estimação, um procedimento computacional de estimação é construído. É um procedimento único, recursivo e sistemático, baseado na estimativa de máxima verossimilhança e congrega no mesmo processo o filtro de Kalman e o método de regressão múltipla para a seleção do número de harmônicas para os termos trigonométricos na componente sazonal.
- No âmbito da previsão, o filtro de Kalman construído permite calcular não só as previsões pontuais e os intervalos de previsão, como também permite calcular os erros padrão de cada estimativa de parâmetro. A previsão das covariáveis é baseada na abordagem de média móvel exponencialmente ponderada. Ainda no âmbito da previsão, um procedimento *bootstrap* não paramétrico para previsão de séries temporais com padrões sazonais complexos é proposto.
- Os dois modelos propostos são aplicados à séries temporais reais, com realce para o estudo feito com base à dados de procura diária de energia elétrica na cidade de Cabinda aplicando o modelo TSCov, cujo objetivo principal é calcular previsões probabilísticas sob a forma de densidades preditivas de curto prazo e estimar as regiões onde o consumo de eletricidade difere em densidade e o máximo da carga elétrica consumida para o período previsto.
- As formulações dos modelos propostos podem admitir a transformação Box-Cox.

1.2 Resultados da Tese

A presente Tese originou apresentação de resultados à comunidade científica internacional, na forma de artigo e apresentação em Congresso internacional que a seguir apresentamos:

- **Apresentação em Congresso Nacional**
  Puindi, A.C., F. Geslie, M.Eduarda Silva (2015). *Previsão multi-passos: comparação de três abordagens com aplicação ao consumo de energia elétrica em Cabinda*. Estatística: Progressos e Aplicações. Atas do XXII Congresso da Sociedade Portuguesa de Estatística, Olhão, pp. 187–198, ISBN: 978-972-8890-39-1
- **Publicação**
  Puindi, António and Silva, M., *Dynamic structural models with covariates for short-term forecasting for time series with complex seasonal patterns*. Submitted to Journal of Applied Statistics, Manuscript ID: CJAS-2018-1028.

  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 in 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.

1.3 Conclusão

Os resultados do nosso estudo empírico demonstram o potencial do quadro de modelos que propomos incluindo o processo de estimação como uma metodologia promissora para a previsão de séries temporais com padrões sazonais complexos, especialmente quando o objetivo é incluir os efeitos de variáveis de influência externas na previsão. As covariáveis utilizadas tiveram um impacto significativo na previsão e, tal como esperado, as previsões obtidas foram mais precisas sob o uso de covariáveis. O procedimento *bootstrap* formulado teve como principal objetivo melhorar as previsões de curto prazo obtidas no procedimento usual com recursões do filtro de Kalman. O procedimento foi aplicado satisfatoriamente para estimar os modelos e calcular previsões. Os resultados mostram que o procedimento *bootstrap* fornece bons resultados para o conjunto de dados usados. Os intervalos de previsão obtidos pelo modelo TSCov são menos precisos quando comparados com os obtidos pelo modelo Boot.TSCov.

Finalmente, com os resultados obtidos nos estudos empíricos realizados, uma pergunta pode ser feita aqui: qual dos modelos, TSCov e TBATS, usar? Os resultados mostram que há pouco para diferenciar os dois modelos, então a resposta é imediata, a abordagem TSCov é preferível se houver covariáveis que são preditores úteis, pois podem ser adicionados como regressores e melhorar as previsões.

A nossa formulação responde a dois problemas interessantes: (i) no campo da previsão, que relaciona a capacidade de acomodar covariáveis quando a intenção é melhorar as previsões; (ii) no campo da projeção de modelos de previsão, é uma nova ferramenta para modelos estruturais. Em suma o trabalho desenvolvido permite obter as respostas das três questões formuladas na motivação.

Não obstante, o nosso estudo proposto pode ser melhorado de várias maneiras e podemos destacar aqui algumas linhas de trabalho futuro que estão listadas na seção seguinte:

1.4 Trabalho futuro

- **O estudo do viés**
  Estudar a propagação e o impacto do viés em ambos os estimadores do filtro de Kalman, $\boldsymbol{x}_{t|t-1}$ e $\boldsymbol{x}_{t|t}$ para o modelo estrutural proposto, com referência à Costa and Monteiro (2016).

- **Inclusão da incerteza das previsões ex-ante**
  Como incluir na previsão da variável de interesse a incerteza associada à previsão das covariáveis.
- **Seleção automática de covariáveis**
  Algum estudo pode ser feito nesta linha para a seleção automática de covariáveis candidatas ao modelo final estimado, podendo ser resolvido usando métodos de penalização como o LASSO.
- **Projeção do estimador da matriz coeficiente $\boldsymbol{\Gamma}$**
  O modelo de espaço de estados dado em (1.3) envolve covariáveis na equação da medição. Uma extensão interessante será considerar que o efeito dessas covariáveis varia no tempo, isto é considerar $\boldsymbol{\Gamma}_t$. Para tal poderá ser considerado o algorítmo EM (Expectation Maximization) (Arlene, 2007).
- **Análise multivariada**
  Quando as dependências significativas entre séries temporais individuais não podem ser ignoradas, as séries multivariadas precisam de ser introduzidas e usadas. Assim, uma projeção do modelo estrutural que acomode séries multivariadas é necessária. Das extensões possíveis, é partir do pressuposto qualitativo de que todas as séries seguem o mesmo tipo de dinâmica, o que implica que os componentes dos vetores de estado têm interpretações semelhantes, mas podem assumir valores diferentes para cada série temporal a denotar por $\boldsymbol{y}_{k,t} = \{y_{1,t}, \cdots, y_{p,t}\}$, com $k = 1, \cdots, p$ séries temporais. As matrizes do sistema são as mesmas, e as matrizes de variância individuais podem não ser completamente conhecidas, dependendo, por isso, de alguns parâmetros desconhecidos. Seguindo a mesma notação, o vetor de estados do sistema seria descrito por $\boldsymbol{x}_t^{(k)}$. As matrizes do modelo de observação e do modelo de transição dos estados, respetivamente, são igualmente representadas por $\boldsymbol{A}_t$ e $\boldsymbol{\Phi}$. A matriz $\boldsymbol{\Gamma}$ representa a transição de entrada e $\boldsymbol{z}_{k,t}$ o vetor de controlo de entradas. Dessa forma, o modelo estrutural multivariado pode ser especificado pelas equações (1.1), (Petris et al., 2009).

$$\boldsymbol{y}_{k,t} = \boldsymbol{A}_t \boldsymbol{x}_t^{(k)} + \boldsymbol{\Gamma} \boldsymbol{z}_{k,t} + \nu_{k,t} \qquad t = 1, 2, \cdots, n \tag{1.1a}$$

$$\boldsymbol{x}_t^{(k)} = \boldsymbol{\Phi} \boldsymbol{x}_{t-1}^{(k)} + w_{i,t-1}^{(k)} \qquad t = 1, 2 \cdots, n \tag{1.1b}$$

onde, $\nu_{k,t} \sim N(0, \boldsymbol{R}_t)$ e $w_{i,t-1}^{(k)} \sim N(0, \boldsymbol{Q}_{t-1})$ são ruídos não correlacionados e independentes do vetor de estado. A matriz de covariância das observações, para esse contexto, é dada por

$$\boldsymbol{R}_t = \begin{bmatrix} \sigma^2_{11,\varepsilon} & \cdots & \sigma^2_{1k,\varepsilon} \\ \vdots & \ddots & \vdots \\ \sigma^2_{kk,\varepsilon} & \cdots & \sigma^2_{kk,\varepsilon} \end{bmatrix}$$

O filtro de Kalman apresentado na subsecção (1.3.1) é válido para o caso multivariado, com ou sem a integração das covariáveis.

Parte IV

Apêndice e Bibliografia

Apêndice A

DEFINIÇÃO DOS PERÍODOS SAZONAIS

A *frequência* é o número de observações por *ciclo* (normalmente um ano, mas às vezes uma semana, um dia ou uma hora). Isso é o oposto da definição de *frequência* na física, ou na análise de Fourier, em que *período* é a duração do ciclo e *frequência* é o inverso do *período*. Ao usar a função ts() em R, as opções da tabela devem ser usadas.

Usando função ts()	
Dados	Frequência
Anual	1
Trimestral	4
Mensal	12
Semanal	52

Tabela A.1: Frequências em séries temporais usando o objeto ts().

Como um ano não tem exatamente 52 semanas, mas 365,25/7 = 52,18 em média, permitindo um ano bissexto a cada quatro anos, a maioria das funções que usam objeto ts() requerem uma frequência inteira. Uma alternativa para incorporar frequências não-inteiras é usar o objeto msts() definido no pacote de Hyndman and Khandakar (2008) que lida com séries temporais de múltipla sazonalidade. Então você pode especificar todas as frequências que podem ser relevantes. Também é flexível o suficiente para lidar com frequências não inteiras. Por exemplo, um conjunto de dados observados de hora em hora, usando o objeto msts() teria

Frequências					
Dados	Minuto	Hora	Dia	Semana	Ano
Diários				7	365.25
Horários			24	168	8766
Meia hora			48	336	17532
Minutos		60	1440	10080	525960
Segundos	60	3600	86400	604800	31557600

Tabela A.2: Frequências em séries temporais usando o objeto msts().

```
1 dados <- msts(x, seasonal.periods=c(24, 168, 8766)
```

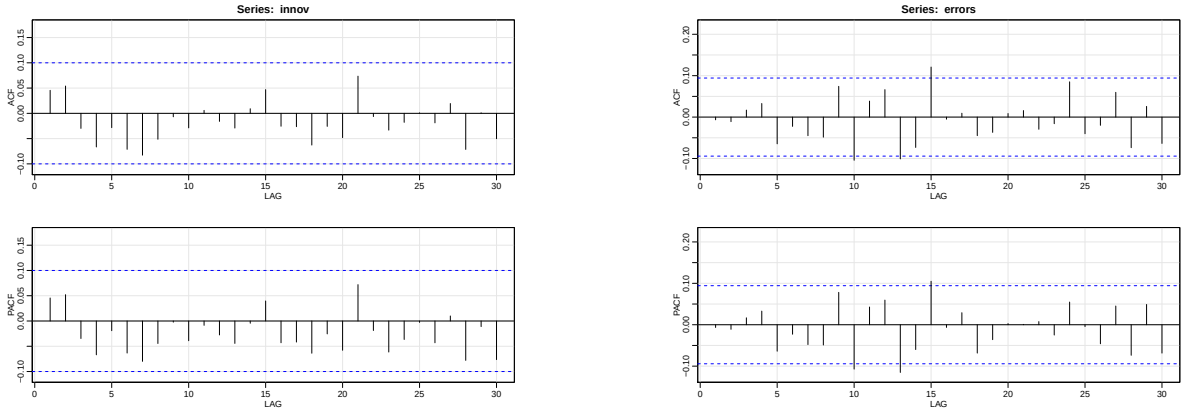
Apêndice $\mathcal{B}$

RESULTADOS RELACIONADOS COM O CAPÍTULO 2 DA PARTE II

B.1 Aplicação do modelo SCov a dados de Mortalidade Cardiovascular

Conforme a secção 2.3, aplica-se o modelo SCov a dados de Mortalidade cardiovascular (M_t); Temperatura (T_{t-1}) e Poluição (P_t e P_{t-4}) em Los Angels - Estados Unidos de América. O objetivo é avaliar o desempenho do modelo SCov, comparando com o modelo BATS.

O diagnóstico dos resíduos dos modelos estimados está apresentado na Figura B.1. O Teste de Box-Ljung sobre a independência dos resíduos do modelo estimado com SCov fornece um valor de Qui-quadrado igual a 25.498, com 21 graus de liberdade e *p-valor* = 0.418, o que permite não rejeitar a hipótese nula de que os resíduos são independentes. Para o modelo estimado com BATS, o teste de Ljung-Box fornece um valor de Qui-quadrado igual a 31.567, com 23 graus de liberdade e *p-valor* = 0.109, o que permite, igualmente, não rejeitar a hipótese nula de que os resíduos são independentes. As estimativas dos parâmetros estão apresentadas na Tabela B.1.



(a) (b)

Figura B.1: Correlograma dos resíduos resultantes da previsão um passo à frente da mortalidade cardiovascular por temperatura e poluição em Los Angels. (a) modelo SCov, (b) modelo BATS.

Parâmetro	MLE (SCov)	E.Padrão Ass.	MLE (BATS)
β_1^*	-0.108	0.182	—
β_2^*	0.140	0.032	—
β_3^*	0.090	0.054	—
α	—	—	0.029
β	—	—	0.001
ϕ	0.8	0.462	0.895
σ_ε^2	3.545	0.527	—
σ_ξ^2	6.049	0.533	—
σ_ζ^2	11.012	0.714	—
σ_w^2	0.253	0.142	—
γ	—	—	-0.170
ψ	—	—	$\{0.252; 0.354; 0.008\}$

Tabela B.1: Estimativas dos parâmetros obtidas a partir dos modelos SCov e BATS, incluindo os erros-padrão das estimativas dos parâmetros do modelo SCov.

B.1.1 Previsão

Calcula-se a previsão de 52 passos à frente. A Figura B.2 exibe a previsão um passo à frente e os respetivos erros de previsão um passo à frente estão apresentados Tabela B.2. A previsão de 52 passos à frente obtida pelos modelos Scov e BATS está apresentada na Figura B.3.

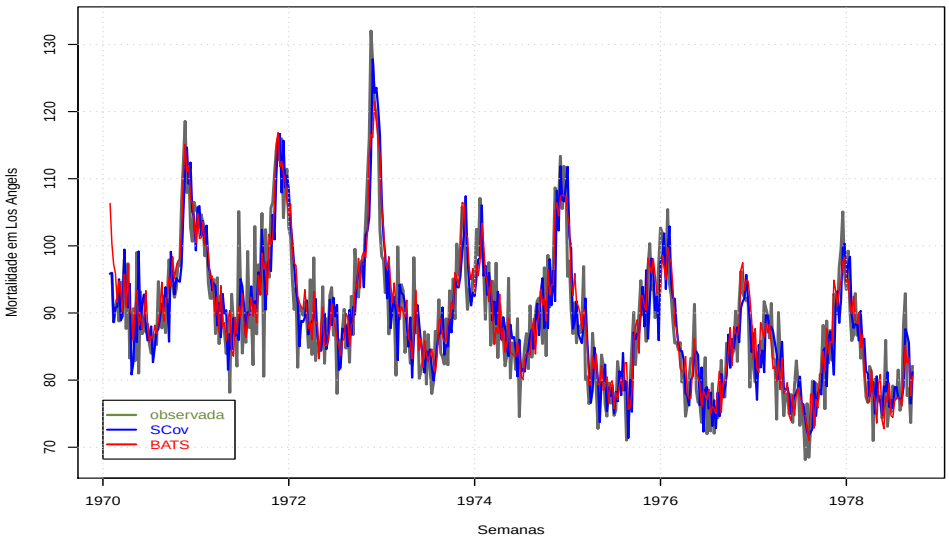


Figura B.2: Valores observados e os ajustados obtidos a partir dos modelos SCov e BATS.

Quer em termos gráficos, Figuras B.3 e B.2, como para a precisão da previsão, Tabelas B.2 e B.3, os resultados ilustram que o modelo SCov comporta-se melhor em relação o modelo BATS para esse conjunto de dados utilizado.

Modelo	ME	RMSE	MAE	MPE	MAPE
SCov	-0.535	4.119	4.393	-0.947	3.934
BATS	-0.419	4.517	3.602	-0.696	4.067

Tabela B.2: Erros de previsão um passo à frente obtidos pelos modelos SCov e BATS sobre a mortalidade cardiovascular por poluição e temperatura.

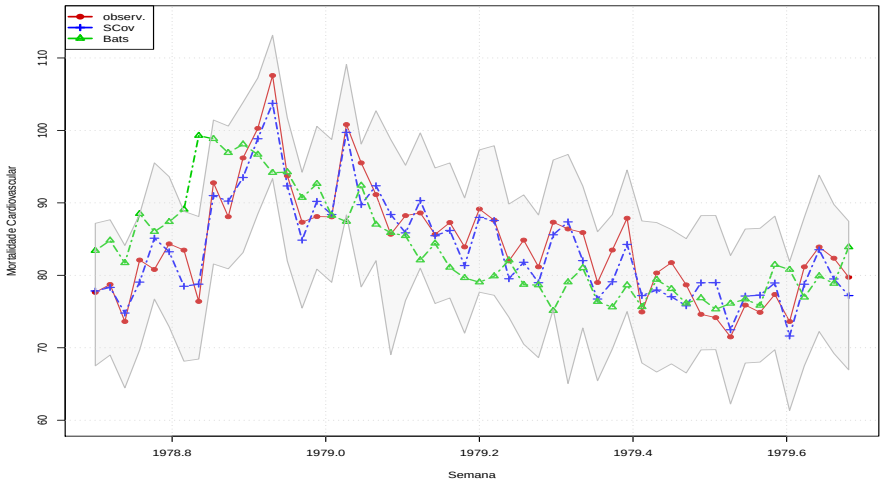


Figura B.3: Valores observados e a previsão (com covariáveis reais) até 52 passos à frente obtidas a partir dos modelos SCov e BATS. A área em cinza representa os intervalos de previsão de 95% obtidos pelo modelo SCov.

Horizonte	SCov		BATS	
	RMSE	MAPE	RMSE	MAPE
1 – 7	2.517	2.666	5.901	3.233
1 – 14	2.587	2.529	6.6477	4.264
1 – 21	2.577	2.506	6.418	4.662
1 – 28	2.579	2.555	6.398	4.406
1 – 35	2.593	2.680	6.502	4.645
1 – 42	2.632	2.703	6.579	5.316
1 – 49	2.665	2.794	6.624	6.537
1 – 52	2.698	2.983	6.897	6.763

Tabela B.3: Medidas de precisão de previsão até 52 passos à frente sobre a mortalidade cardiovascular por poluição e temperatura em Los Angels – modelos estimados SCov e BATS.

B.2 Aplicação do modelo TSCov a dados de níveis de concentração de NO_2 – Estação de Entre-Campos em Lisboa

Os dados referente aos níveis de concentração de NO_2, Figura 3.2, são obtidas a partir da base de dados online sobre qualidade do ar da Agência Portuguesa do Ambiente, cuja missão é propor, desenvolver e monitorizar as políticas públicas para o ambiente e o desenvolvimento sustentável. O banco de dados sobre a qualidade do ar, (QualAr, 2015), fornece medições por

hora, resultantes de atividades de monitoramento, para vários poluentes, incluindo o NO_2. As medições foram efetivadas em 2014 entre 1 de Outubro e 31 de Dezembro em Lisboa, Portugal. A análise detalhada desse conjunto de dados pode ser vista em Andreia et al. (2017). A análise prévia do conjunto de dados baseada no correlograma, Figura B.4, revela a existência de dois padrões sazonais na série: um padrão diário com periodicidade 24 e um padrão semanal com periodicidade 168. A correlação cruzada com as covariáveis é mostrada na Figura 2.3, que não exibe uma correlação significativa com as covariáveis.

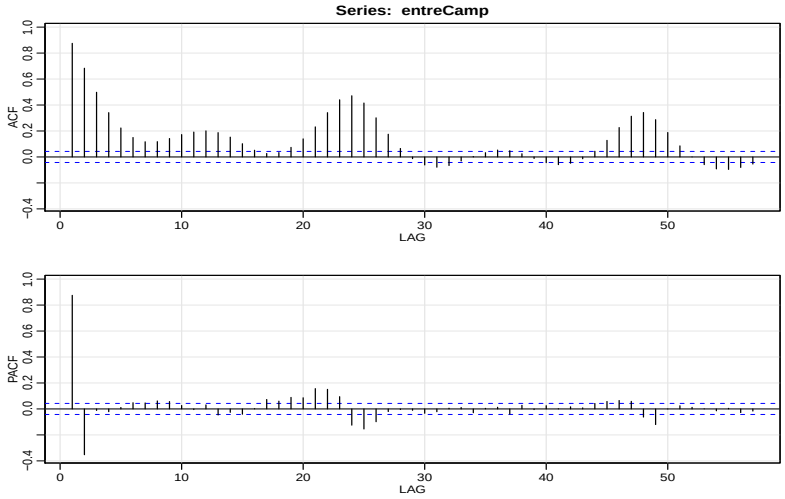


Figura B.4: Correlograma da série dos níveis de concentração de NO_2 em Entre-Campos, Lisboa.

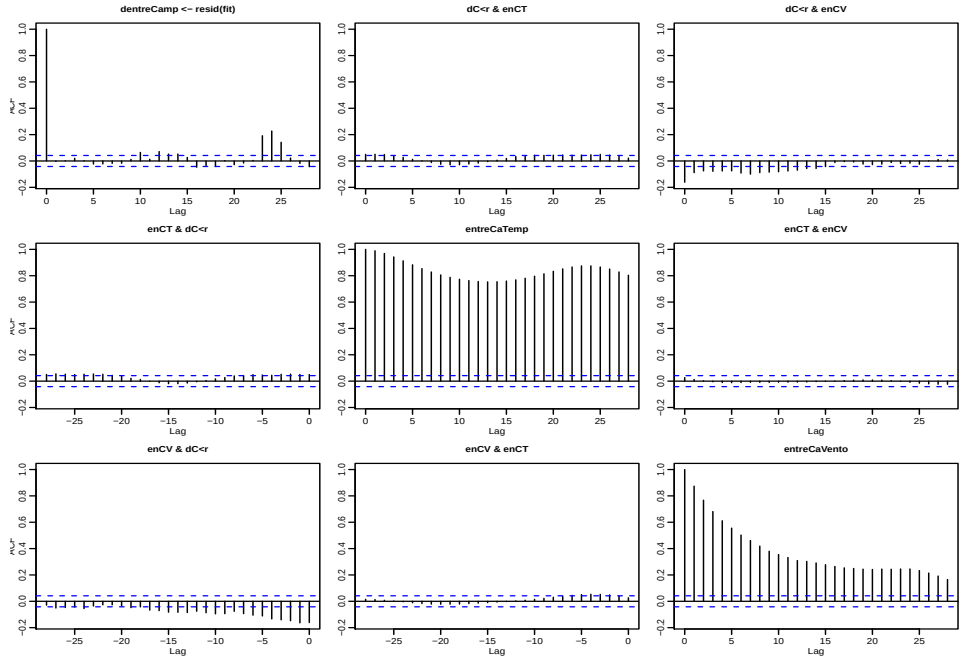
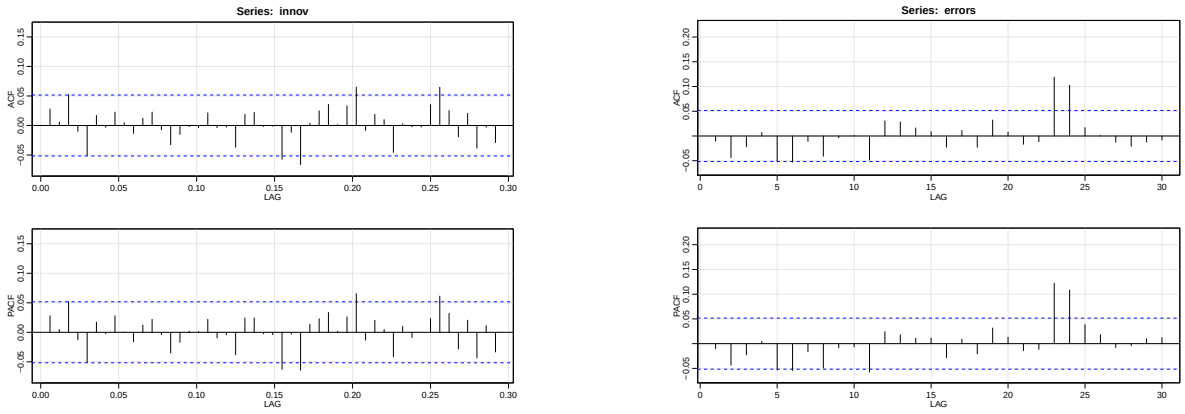


Figura B.5: Correlação cruzada entre os resíduos de NO_2 e as séries de temperatura e vento.

Para estimação do modelo, a série de teste é constituída de 1500 observações e a série de validação com 704 observações. Estimam-se dois modelos: um modelo com o uso de covariáveis reais e outro com o uso de covariáveis previstas. As covariáveis utilizadas são a Temperatura, T_t, a humidade, H_t, e o Vento, V_t. O diagnóstico dos resíduos dos modelos estimados está apresentado na Figura B.6. O teste de Ljung-Box sobre a independência dos resíduos do modelo TSCov fornece um valor de Qui-quadrado igual a 31.921, com 19 graus de liberdade e *p-valor* = 0.285, o que permite não rejeitar a hipótese nula de que os resíduos são independentes. O gráfico da normalidade dos resíduos, Figura B.7, exibe partidas nas duas caudas, devido a possível presença de *outliers*. Para o modelo estimado com TBATS, o teste de Ljung-Box fornece um Qui-quadrado = 25.401, com 22 graus de liberdade e p-valor = 0.198, permitindo igualmente não rejeitar a hipótese nula de que os resíduos são independentes. A Tabela B.5 apresenta os erros de previsão um passo à frente.



(a) (b)

Figura B.6: Correlograma dos resíduos resultantes da previsão um passo à frente dos níveis de concentração de NO_2 em Entre-Campos, Lisboa. (a) modelo SCov, (b) modelo BATS.

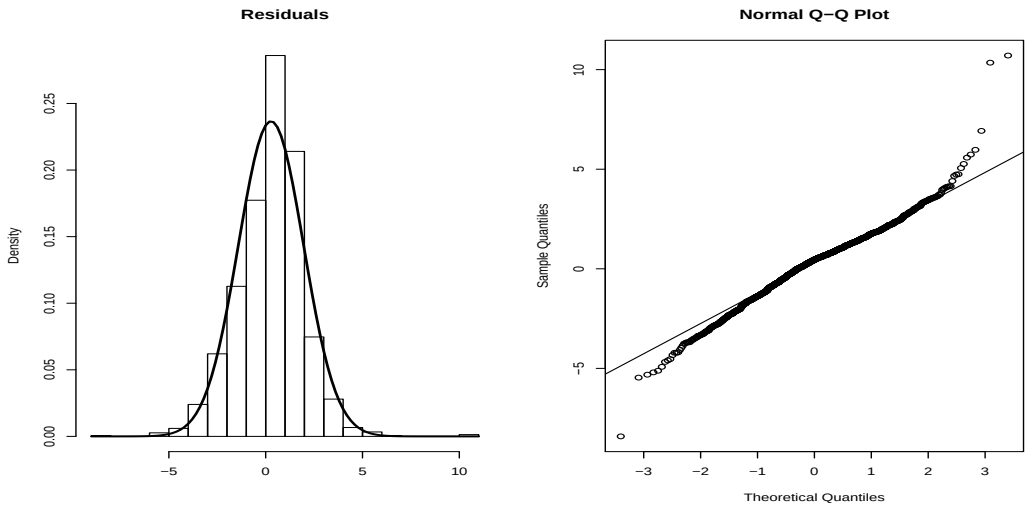


Figura B.7: Histograma dos resíduos do modelo TSCov estimado sobre os níveis de concentração de NO_2 em Entre-Campos, Lisboa.

Parâmetro	MLE (TSCov)	E.Padrão Ass.	MLE (TBATS)
β_1^*	0.421	0.026	—
β_2^*	-0.503	0.254	—
β_3^*	0.461	0.037	—
$\grave{a}$	—	—	0.437
α	—	—	1.403
β	—	—	-0.256
ϕ	0.916	0.242	0.891
σ_ε^2	3.091	0.153	—
σ_ξ^2	0.043	0.042	—
σ_ζ^2	11.453	0.374	—
σ_w^2	$\{0.032; 0.012\}$	$\{0.003; 0.014\}$	—
$\sigma_{w^*}^2$	$\{0.043; 0.223\}$	$\{0.237; 0.063\}$	—
γ_1	—	—	$\{0.0004; -0.0001\}$
γ_2	—	—	$\{0.0002; -0.0004\}$

Tabela B.4: Estimativas dos parâmetros e os respetivos erros-padrão obtidos partir do modelo TSCov. As estimativas dos parâmetros obtidos a partir do modelo TBATS estão apresentadas na quarta coluna.

Modelo	ME	RMSE	MAE	MPE	MAPE
TSCov	-0.019	10.185	7.496	-4.612	16.644
TBATS	0.585	12.414	8.649	-3.929	19.649

Tabela B.5: Erros de previsão um passo à frente obtidos pelos modelos TSCov e TBATS sobre as concentrações de NO_2 em Entre-Campos, Lisboa.

B.2.1 Previsão

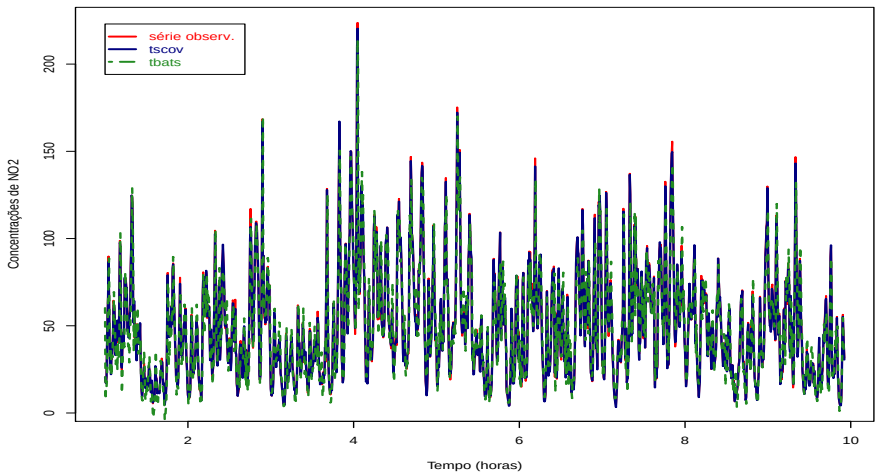


Figura B.8: Valores observados e os ajustados a partir dos modelos TSCov e TBATS sobre os níveis de concentração de NO_2 em Entre-Campos, Lisboa.

As previsões calculadas correspondem 24 passos à frente. A Figura B.9 exibe as previsões calculadas pelos modelos TSCov e TBATS incluindo os intervalos de previsão obtidos pelo modelo TSCov. Na Tabela B.6 estão apresentados os erros de previsão até 24 passos à frente.

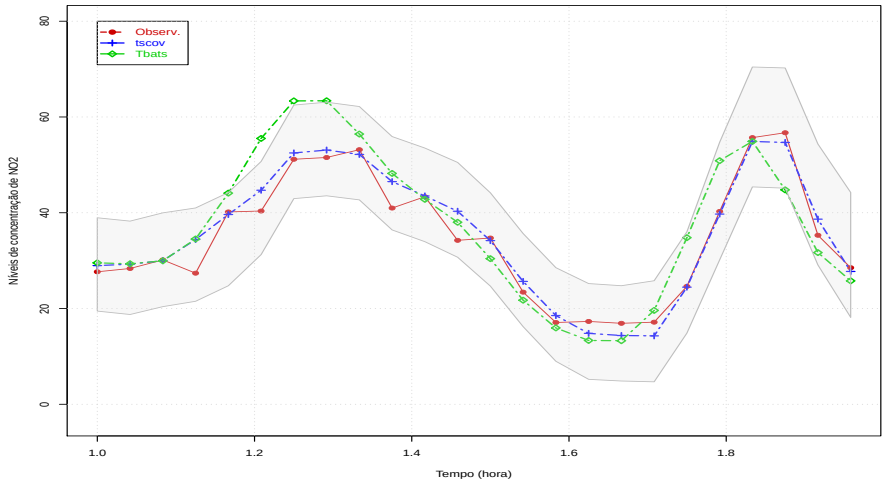


Figura B.9: Previsão de 24 passos à frente obtida pelos modelos TSCov e TBATS, incluindo os intervalos de previsão de 95% gerados pelo modelo TSCov.

Horizonte	TSCov (cov. reais)		TSCov (cov. previstos)		TBATS	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
1 – 3	1.536	2.422	3.607	4.147	3.071	5.416
1 – 6	1.945	4.536	3.368	4.864	3.422	8.354
1 – 9	1.865	5.505	3.319	4.917	4.217	10.171
1 – 12	2.551	5.867	4.754	5.521	5.135	13.434
1 – 15	2.734	5.713	4.938	5.176	6.116	13.278
1 – 18	3.046	6.487	5.177	6.733	6.435	14.622
1 – 21	3.372	6.923	6.326	6.857	6.581	15.065
1 – 24	3.611	6.964	6.461	7.714	6.803	15.275

Tabela B.6: Precisão de previsão até 24 passos à frente dos níveis de concentração de NO_2 em Entre-Campos, Lisboa. Os modelos aplicados: TSCov e TBATS.

B.3 Feriados religiosos e nacionais

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

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

Figura B.10: Dados de feriados da Turquia entre 1 de Janeiro de 2000 e 31 de Dezembro de 2006, (De Livera et al., 2011).

B.4 Principais funções implementadas no ambiente R

Foi usado a versão 3.4.1 (2017-06-30) para (64-bit).

B.4.1 Implementação das matrizes do sistema

```
PhiMatrix <- function(smallphi=NULL, seasonal.periods=NULL, kvector=NULL) {
  # Esta função calcula a matriz de transição "Phi".
  return(Phi)
}
```

Listing B.1: Matriz do modelo de transição

```
Amatrix <- function(smallphi=NULL, kvector=NULL, tau) {
  # Esta função calcula a matriz do modelo de observação "At".
  return(At)
}
```

Listing B.2: Matriz do modelo de observação

```
Qmatrix <- function(sig1, sig2, sig3, sig4, kvector=NULL,
                    seas.variance=NULL){
  # Esta função calcula a matriz de covariância de estado "Qt".
  return(Qt)
}
```

Listing B.3: Matriz de covariância de estado

B.4.2 Implementação do filtro de Kalman

```
Kf.tscov<-function(y, delta, At, mu, Sigma, Phi, Gam, Qt, Rt, input){
  # Esta função calcula os componentes do filtro de Kalman.
  list(xt1, xt2, Pt1, Pt2, yadj, like, innov, Kn = K)
}
```

Listing B.4: Filtro de Kalman

```
verossim <- function(param.vector, ydata, input, smallphi, seasonal.periods,
                     kvector, param.control, tau) {
  # Esta função calcula o valor da log-verossimilhança.
  return(like)
}
```

Listing B.5: Função da log-verossimilhança

B.4.3 Projeção do modelo

```
TESCovFit <- function(ydata, input, damping, seasonal.periods=NULL,
                      kvector=NULL, start.params=NULL){
  # Esta função calcula a log-verossimilhança.
  return(modelSaida)
}

#-------------------------
tscov <- function(ydata, input, damped=NULL, seasonal.periods=NULL){
  # Esta função projeta o melhor modelo
  return(MelhorModel)
}
```

Listing B.6: Funções para otimização das estimativas do modelo e projetar o melhor modelo

Referências Bibliográficas

Ahmad, F. and Maxwell, L. (2015). Exponential smoothing with regressors: Estimation and initialization. *Model Assisted Statistics and Applications*, (10):253–263.

Akhlaghi, S., Zhou, N., and Huang, Z. (2017). Adaptive adjustment of noise covariance in kalman filter for dynamic state estimation. *CoRR*, abs/1702.00884:1–5.

Alonso, A. M., Garcia-Martos, C., Rodriguez, J., and Sanchez, M. J. (2008). Seasonal dynamic factor analysis and bootstrap inference: Application to electricity market forecasting. *Working paper*, pages 1–44.

Andreia, M., Menezes, R., and Silva, M. E. (2017). Modelling spatio-temporal data with multiple seasonalities: the no2 portuguese case. *Spatial Statistics*, pages 1–25.

Anthanasopoulos, G., Hyndman, R. J., and University, M. (2006). Modelling and forecasting australian domestic tourism. *Tourism Management 29*, pages 19–31.

Arlene, H. N. (2007). State space models with exogenous variables and missing data. *University of Florida, PhD Thesis*, pages 12–30.

Bergmeir, C., Hyndman, R. J., and Benitez, J. M. (2015). Bagging exponential smoothing methods using stl decomposition and box-cox transformation. *Preprint submitted to International Journal of Forecasting*, pages 2–18.

Billah, B., Hyndman, R. J., and Koehler, A. B. (2005). Empirical information criteria for time series forecasting model selection. *Journal of Statistical Computation and Simulation 75*, pages 831–840.

Box, G. and Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society; Ser. B, 26(2)*, pages 211–252.

Brockwell, P. and Davis, R. (2002). Introduction to time series and forecasting. *Springer-Verlang*, Second Edition:259–261.

Buhlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli (3)*, pages 123–148.

Chernick, M. R. and A.LaBudde, R. (2011). An introduction to bootstrap methods with applications to r. *Wiley*, pages 3–129.

Cordeiro, C. and Neves, M. (2011). Forecasting with exponential smoothing methods and bootstrap. *REVSTAT–Statistical Journal*, pages 135–149.

Costa, M. and Monteiro, M. (2016). Bias–correction of kalman filter estimators associated to a linear state space model with estimated parameters. *Preprint submitted to Journal of Statistical Planning and Inference*, pages 1–33.

Cottet, R. and M, M. (2003). Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association 98:464*, pages 839–849.

Courtney, T. (2018). What is bootstrapping in statistics? (accessed february 24, 2018).

Cowpertwait, P. S. and Metcalfe, A. V. (2009). Introductory time series with r. *Springer*, Second Edition:229–243.

De Livera, A. (2010). Modeling time series with complex seasonal patterns using exponential smoothing. *PhD thesis, Monash University*, pages 91–111.

De Livera, A., Hyndman, R., and R.D.Snyder (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Jornal of the American Statistical Association 106(496)*, pages 1513–1527.

Desmond, C. O. (2014). A primer to bootstrapping; and an overview of dobootstrap. *Department of Psychology, Stanford University*, pages 1–6.

Dordonnat, V., S.J.Koopman, M.Ooms, A.Dessertaine, and J.Collet (2008). An hourly periodic state space model for modelling french national electricity load. *International Journal of Forecasting*, (24):566–587.

Douc, R., Moulines, E., and Stoffer, D. (2014). Nonlinear time series theory, methods, and applications with r examples. *Chapman & Hall/CRC*, pages 3–56.

Durbin, J. and S.J.Koopman (2011). Time series analysis by state space methods. *Oxford University Press*, pages 170–176.

Gardner, E. and McKenzie, E. (1985). Forecasting trends in time series. *Management Science 31*, pages 1237–1246.

Gob, R., K.Lurz, and A.Pievatolo (2013). Electrical load forecasting by exponential smoothing with covariates. *Applied models business and industry*, 29:629–645.

Gould, P. G., Koehler, A. B., Vahid-Araghi, F., Snyder, R. D., Ord, J. K., and Hyndman, R. J. (2008). Forecasting time-series with multiple seasonal patterns. *European Journal of Operational Research (191)*, pages 207–222.

Hafida, G. and Hamdi, F. (2015). Bootstrapping periodic state-space models. *Communications in Statistics - Simulation and Computation*, pages 374–401.

Hamilton, J. D. (1994). Time series analysis. *Princeton University Press, 41 William St.Princeton, New Jersey 08540*, pages 372–407.

Harvey, A. and Koopman, S. (1993). Forecasting hourly electricity demand using timevarying splines. *Journal of the American Statistical Association 88*, pages 1228–1236.

Harvey, A. C. (1989). Forecasting, strutural time series models and the kalman filter. *Cambridge University Press*, pages 100–127.

Helmut, L. (2005). New introduction to multiple time series analysis. *Springer Berlin Heidelberg New York*, pages 611–631.

Herman, J. (1994). Topics in advanced econometrics. estimation, testing, and specification of cross-section and time series models. *Cambridge University Press*, pages 154–167.

Hyndman, R. (1996). Computing and graphing highest density regions. *American Statistical Association, Vol.50, No.2*, pages 120–126.

Hyndman, R. (2014). Forecasting: Principles and practice. *University of Western Australia: online version*, pages 34–104.

Hyndman, R., A.B.Koehler, J.K.Ord, and R.D.Snyder (2008). Forecasting with exponential smoothing: the state space approach. *Springer-Verlang*, pages 137–143.

Hyndman, R., Koehler, A. B., Snyder, R. D., and Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting 18*, pages 439–454.

Hyndman, R. J. (2018). *hdrcde: Highest Density Regions and Conditional Density Estimation*. R package version 3.2.

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3):1–22.

Jan, G. D. G. and Hyndman, R. (2005). 25 years of iif – international institute of forecasters – time series forecasting: A selective review (working paper). *Department of Econometrics and Business Statistics, Australia*, pages 4–41.

Jerez, M., Casals, J., and Sotoca, S. (2015). Single and multiple error state-space models for signal extraction. *Journal of Statistical Computation and Simulation*, 85:2–18.

Jonathan, D. and Chan, K.-S. (2008). Time series analysis with applications in r. *Springer*, Second Edition:11–97.

Kitagawa, G. (2010). Introduction to time series modeling. *CRC Press*, pages 135–147.

Kitagawa, G. and Gersch, W. (1996). Smoothness priors analysis of time series. *New York, Springer-Verlag*, pages 55–88.

Koehler, A., R.D.Snyder, J.K.Ord, and A.Beaumont (2012). A study of outliers in the exponential smoothing approach to forecasting. *International Journal of Forecasting, 28(2)*, pages 477–484.

Makridakis, S. G., S.C.Wheelwright, and R.J.Hyndman (1998). Forecasting: methods and applications. *John Wiley and Sons*, Third Edition New York:158–159.

Menezes, J. C., V.Lopes, V., and C.Pinheiro, C. (2006). Determination of state-space model uncertainty using bootstrap techniques. *Journal of Process Control*, 16:685–692.

Mohamed, A. H. (1999). Optimizing the estimation procedure in ins/gps integration for kinematic applications. *PhD Thesis, Department of Geomatics Engineering, University of Calgary*, pages 73–142.

Ord, J., A.B.Koehler, and R.D.Snyder (1997). Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association*, 92(440):1621–1629.

Ord, J., R.D.Snyder, A.B.Koehler, R.J.Hyndman, and M.Leeds (2005). Time series forecasting: the case for the single source of error state space approach. *Working paper*, pages 2–33.

Pedregal, D. and Young, P. (2006). Modulated cycles, an approach to modelling periodic components from rapidly sampled data. *International Journal of Forecasting 22*, pages 181–194.

Petris, G., Petrone, S., and atrizia Campagnoli, P. (2009). Dynamic linear models with r. *Springer*, pages 31–83.

QualAr (2015). Online database on air quality, url: https://qualar.apambiente.pt/qualar/index.php.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramanathan, R., Engle, R., C.W.J.Granger, F.Vahid-Araghi, and C.Brace (1997). Short-run forecasts of electricity loads and peaks. *International Journal of Forecasting 13(2)*, pages 161–174.

Razbash, S. and Hyndman, R. (2018). Forecasting functions for time series and linear models. *cran.r-project.org, Package forecast*, pages 1–131.

Rebennack, S., P.M.Pardalos, V.F.Pereira, M., and N.A.Iliadis (2010). Handbook of power systems ii. *Springer*, (1):127–155.

Rodriguez, A. and Ruiz, E. (2009). Bootstrap prediction intervals in state-space models. *Journal of the time series analysis*, pages 167–178.

Sarkar, D. (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York. ISBN 978-0-387-75968-5.

Schwarz, G. (1978). Estimating the dimensional of a model. *Annals of Statistics, Hayward, v.6, n.2*, pages 461–464.

Shumway, R. H. and Stoffer, D. S. (2006). Time series analysis and its applications with r examples. *Springer*, Second Edition:11–394.

Shumway, R. H. and Stoffer, D. S. (2011). Time series analysis and its applications: With r examples. *New York: Springer*, Third Edition:319–359.

Shumway, R. H. and Stoffer, D. S. (2017). Time series analysis and its applications: With r examples. *New York: Springer*, Four Edition:287–331.

Soares, L. and Medeiros, M. (2008). Modeling and forecasting short-term electricity load: a comparison of methods with a application to brazilian data. *International Journal of Forecasting 24*, pages 630–644.

Spiegelhalter, D., Pearson, M., and Short, I. (2011). Visualizing uncertainty about the future. *American Association for the Advancement of Science, Washington, DC 20005*.

Stoffer, D. (2016). *astsa: Applied Statistical Time Series Analysis*. R package version 1.7.

Stoffer, D. S. and Wall, K. D. (2004). Resampling in state space models. *Cambridge University Press*, pages 2–26.

Taieb, S. B., Huser, R., Hyndman, R. J., and Genton, M. G. (2015). Probabilistic time series forecasting with boosted additive models: an application to smart meter data. *Working Paper*, pages 2–30.

Taylor, J. (2010). Triple seasonal methods for short–term electricity demand forecasting. *European Journal of Operational Research: 204*, pages 139–152.

Taylor, J. and P.McSharry (2008). Short-term load forecasting methods: An evaluation based on european data. *IEEE Transactions on Power Systems*, (22):213–2219.

Taylor, J. and R.Buizza (2003). Using weather ensemble predictions in electricity demand forecasting. *IEEE Transactions on Power Systems*, (19):57–70.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of Operational Research Society*, pages 799–805.

Taylor, J. W. and Snyder, R. D. (2012). Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *Omega*, 40:748–757.

Tommaso, P. (1991). Forecasting with structural time series models. *Dipartimento di Scienze Statistiche, Università di Udine. Blackwell Publishers, Oxford*, pages 1–31.

Tommaso, P. and L.Alessandra (2012). Maximum likelihood estimation of time series models: the kalman filter and beyond. *MPRA Munich Personal RePEc Archive*, (39600):1–30.

Tsay, R. S. (2005). Analysis of financial time series. *Wiley*, Second Edition:490–533.

Velasco, M. G. and Garcia, I. M. D. P. (2009). Series temporales. *Caceres: Universidad de Extremadura*, Ed.IV. Serie:37–46.

Wall, K. D. and Stoffer, D. S. (2002). A state space approach to bootstrapping conditional forecasts in arma models. *Journal of Time Series Analysis 23*, pages 733–751.

Wang, J. (2000). Stochastic modeling for real-time kinematic gps/glonass positioning. *Navigation (46); No.4*, pages 297–305.

Wang, S. (2006). Exponential smoothing for forecasting and bayesian validation of computer models. *Georgia Institute of Technology, PhD thesis*, (1):96–126.

Welch, G. and Bishop, G. (2001). An introduction to the kalman filter. *Chapel Hill, NC 27599–3175*, pages 18–24.

Weron, R. (2006). Modeling and forecasting electricity loads and prices: a statistical aproach. *Wiley*, pages 67–74.

West, M. and Harrison, J. (1997). Bayesian forecasting and dynamic models, 2nd ed. *New York, Springer-Verlag*.

Zarchan, P. and Musoff, H. (2009). Fundamentals of kalman filtering: A practical approach. *American Institute of Aeronautics and Astronautics, Inc.*, Third Edition:129–140.