

Métodos de suavização exponencial com modelagem em espaço de estado para previsão automática de séries temporais via ambiente R

Elisa Henning (UDESC) elisa.henning@udesc.br
Custodio da Cunha Alves (UNIVILLE) custodio.alves@univille.br
Andréa Cristina Konrath (UFSC) andreack@inf.ufsc.br
Olga Maria Formigoni Carvalho Walter (UFSC) olgaformigoni@gmail.com

Resumo:

Os métodos de suavização exponencial com abordagem em espaço de estado nos últimos dez anos têm sido amplamente utilizados no contexto da análise e previsão de séries temporais. Este trabalho apresenta uma aplicação das principais inovações desses métodos para estabelecer previsão automática via ambiente R. Nessa aplicação propõe-se uma análise de série temporal da venda de um produto que apresenta efeito sazonal e que tem impacto significativo nos custos e investimentos de uma empresa, localizada na região norte do Estado de Santa Catarina. Os resultados obtidos com a previsão automática estabelecida para a demanda de vendas desse produto por meio de modelos espaços de estado foram satisfatórios. Esses resultados encontrados a partir dessa análise foram fundamentais para o analista de previsão direcionar especial atenção ao grau de acurácia desses modelos preditivos, aspecto importante para tomar de decisões que contemple a busca contínua da gestão otimizada do processo produtivo dessa empresa.

Palavras chave: Análise de Séries Temporais, Modelos Espaço de Estado, Previsão Automática.

Exponential smoothing methods with state-space modeling for automatic forecasting of time series via environment R

Abstract

The exponential smoothing methods within the state-space approach in the last ten years have been widely used in the context of analyzing and forecasting time series. This paper presents an application of the main innovations of these methods to establish automatic forecasting via environment R. In this application, we propose a time series analysis of the sale of a product having a strong seasonal effect and has significant impact on costs and investments of the company located in northern Santa Catarina State. The results obtained for sales forecasting of this product by state space models were satisfactory. The results from this analysis were critical for the forecasting analyst to measure the degree of accuracy of these predictive models, and to make important decisions that contemplate the optimal management of the production process.

Key-words: Time Series Analysis, State Space Models, Automatic Forecasting

1. Introdução

A utilização de métodos estatísticos de suavização exponencial com modelagem em espaço de estado para a suavização e previsão automática de séries temporais nos últimos dez anos, tem

sido objeto de estudo de vários trabalhos científicos. O bom desempenho desses métodos aliado a grande quantidade de séries temporais que constantemente precisam ser atualizadas demandam estudos que proporcionam um grande avanço em sua modelagem. Estes estudos iniciam praticamente a partir de trabalhos de Ord *et al.* (1997), Hyndman *et al.* (2002), Hyndman *et al.* (2008) e Taylor (2003) que propõem uma nova estrutura estatística para tais modelos a partir de filtros de espaço de estado. Essa abordagem incorpora a esses métodos uma classe paramétrica, permitindo-se obter cálculo de verossimilhança, intervalos de previsão e procedimentos para a seleção de modelos.

Assim, o objetivo principal deste trabalho é a aplicação de métodos estatísticos de suavização exponencial com modelagem em espaços de estado para estabelecer previsão automática de séries temporais via ambiente R de acordo com o proposto por Hyndman *et al.* (2008). Esta aplicação envolve a previsão para o planejamento de demanda de um produto que é fabricado numa empresa do setor metal mecânico, localizada na região norte do estado de Santa Catarina. Com esse planejamento de demanda foi possível antecipar cenários futuros de vendas desse produto, tendo como propósito a redução de riscos na tomada de decisão dessa empresa.

Este artigo está estruturado em cinco seções, incluindo a presente introdução. A seção 2 traz o referencial teórico sobre métodos estatísticos de suavização exponencial incorporados à estrutura espaço de estado; na seção 3 está a metodologia aplicada; a seção 4 mostra os resultados obtidos com a aplicação. Na seção 5 estão as considerações finais.

2. Métodos Estatísticos de Previsão por Séries Temporais

A aplicação de métodos de previsão por séries temporais são adequados para prever quantitativamente uma variável e medir a qualidade dessa previsão, tem sido um diferencial cada vez mais utilizado para antecipar cenários futuros. Neste contexto, o objetivo da análise de séries temporais é encontrar modelos (estatísticos e/ou matemáticos) que permitem descrever de forma adequada o processo de geração de dados para prever o comportamento futuro dessa variável para um determinado horizonte de planejamento (VERÍSSIMO *et al.*, 2012). A essência dos métodos de previsão que utilizam análise de séries temporais consiste em identificar o padrão da série, separando-o do ruído contido nas observações individuais, e utilizá-lo para prever os valores futuros da série. Segundo Morettin e Toloí (2006) esses métodos aplicados para descrever séries temporais são processos estocásticos, isto é, processos cuja evolução no tempo é gerada e controlada por leis probabilísticas.

2.1 Métodos de Suavização Exponencial

O procedimento de suavização ou ajuste exponencial de um método estatístico de previsão parte de uma equação de médias móveis, ponderadas exponencialmente, com o objetivo de produzir ajustes nas variações aleatórias dos dados de determinada série temporal (RAGSDALE, 2009). Os métodos de previsão por suavização exponencial se utilizam apenas de observações da própria série de dados, não dependendo de nenhuma variável externa para realizar previsões (SOUZA *et al.*, 2008). Esses métodos são reconhecidos como uma das ferramentas mais empregadas na previsão à curto prazo de séries temporais (GARDNER, 1985, 2006; BERMÚDEZ *et al.*, 2008). Sua popularidade se deve principalmente a sua simples formulação, facilidade de ajustes e boa precisão (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998).

Embora os métodos de suavização exponencial tenham sido desenvolvidos em torno 1950, uma estrutura de modelagem incorporando modelos estocásticos, cálculo de probabilidade, intervalos de previsão e procedimentos para a seleção dos modelos, só foram desenvolvidos a

partir do trabalho de Hyndman *et al.* (2002) incluindo uma classificação com modelagem em espaço de estados de acordo com Tabela 1.

Por exemplo, a célula (NN) descreve o método de suavização exponencial simples, célula (NA) descreve o método linear de Holt. O método aditivo de Holt-Winters é dado pela célula (AA) e o método Holt-Winters multiplicativo pela célula (AM). As células (A_d) e (M_d) correspondem aos métodos de tendência amortecida. Por exemplo, a célula ($A_d N$) descreve o método de tendência amortecida de Holt. As demais células correspondem aos métodos de utilização menos frequentes, porém análogos. Esta classificação é decorrente da taxonomia de métodos de suavização exponencial onde cada método tem uma componente de tendência e uma componente sazonal proposta por Pegels (1969), estendida por Gardner (1985) e modificada por Hyndman *et al.* (2002) e Taylor (2003).

Componente Tendência	Componente Sazonal		
	N Nenhuma	A Aditiva	M Multiplicativa
N (Nenhuma)	NN	NA	NM
A (Aditiva)	NA	AA	AM
M (Multiplicativa)	MN	MA	MM
A_d (Amortecida)	$A_d N$	$A_d A$	$A_d M$
M_d (Amortecida)	$M_d N$	$M_d A$	$M_d M$

Fonte: Adaptado de Hyndman *et al.* (2008)

Tabela 1 – Classificação de Métodos de Suavização Exponencial com Modelagem em Espaço de Estado

2.2 Previsões Pontuais para todos os Métodos de Suavização Exponencial

Hyndman *et al.* (2008, p.18) fornecem um conjunto de fórmulas recursivas para calcular as previsões pontuais de h períodos à frente experimentando cada um dos 30 métodos de suavização exponencial com abordagem em espaço de estado para selecionar o melhor método. Em cada caso, l_t indica o nível de série no tempo t , b_t a tendência no instante t , s_t a componente sazonal da série no tempo t , e m indica o número de estações de um ano; α, β^*, γ e ϕ são constantes, $\phi_h = \phi + \phi^2 + \dots + \phi^h$ e $h_m^+ = [(h-1) \text{ mod } m] + 1$.

Alguns casos especiais relevantes podem ser obtidos ajustando os parâmetros de suavização para valores extremos. Por exemplo, se $\alpha = 0$, o nível é constante ao longo do tempo, se $\beta = 0$, a tendência é constante ao longo do tempo; e $\gamma = 0$, o padrão sazonal é constante longo do tempo. Em outro extremo, as previsões naturais (isto é, $\hat{y}_{t+h|t} = y_t$ para todo h), são obtidas utilizando o método (NN) com $\alpha = 1$. Finalmente, o método de tendência aditiva e multiplicativa são casos especiais dos seus homólogos amortecidos obtidos quando se considera $\phi = 1$.

Seja a série observada y_1, y_2, \dots, y_n . Uma previsão de y_{t+h} com base em todos os dados até o tempo t é indicada por $\hat{y}_{t+h|t}$. Para ilustrar todos os métodos de suavização exponencial são apresentadas neste trabalho as previsões pontuais e atualizações das equações para o método de sazonalidade multiplicativa de Holt-Winters (AM).

$$\text{Nível: } l_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1}), \quad (1)$$

$$\text{Tendência: } b_t = \beta (l_t - l_{t-1}) + (1 - \beta) b_{t-1}, \quad (2)$$

$$\text{Sazonalidade: } s_t = \frac{y_t \gamma}{l_{t-1} + b_{t-1}} + (1 - \gamma) s_{t-m}, \quad (3)$$

$$\text{Previsão: } \hat{y}_{t+h|t} = (l_t + b_t h) s_{t-m+h_m^+}, \quad (4)$$

onde m é o comprimento da sazonalidade (por exemplo, o número de meses ou trimestre num ano), l_t o nível da série, b_t a tendência, s_t a componente sazonal, $\hat{y}_{t+h|t}$ é a previsão para h períodos à frente, e $h_m^+ = [(h-1) \bmod m] + 1$. Os parâmetros (α, β^* e γ) são normalmente restringidos em situar-se no intervalo (0,1). Tal como acontece com todos os métodos de suavização exponencial, precisamos dos valores iniciais das componentes l_0, b_0 e s_{1-m}, \dots, s_0 e das estimativas dos valores dos parâmetros.

2.3 Modelos Espaço de Estado

O estudo de modelos espaço de estado teve início com Kalman (1960), alcançando êxito a partir dos trabalhos de Hyndman *et al.* (2002) e Hyndman *et al.* (2008). É importante ressaltar que há uma diferença significativa. Enquanto o método de suavização exponencial é um algoritmo que fornece apenas previsões pontuais, o modelo espaço de estado estocástico além de fornecer estas mesmas previsões pontuais fornece ainda uma estrutura para calcular os intervalos de predição e outras propriedades (HYNDMAN *et al.*, 2008).

O trio (E, T, S) refere-se as três componentes: erro, tendência e sazonalidade. Assim, o modelo ETS (A,A,N) tem erros aditivos, tendência aditiva e nenhuma sazonalidade. Em outras palavras, representa o método linear de Holt com erros aditivos. Da mesma forma ETS (M,M_d,M) refere-se a um modelo com erros multiplicativos, uma tendência multiplicativa amortecida e Sazonalidade multiplicativa. A notação ETS (. , . , .) orienta a ordem em que as componentes são especificadas. Além disso, ETS pode também ser uma abreviação de suavização exponencial.

Uma vez especificado o modelo, pode-se a partir daí estudar a distribuição de probabilidade dos valores futuros da série e, descobrir, por exemplo, a média condicional para uma observação futura tendo o conhecimento do passado. Isso é denotado como $\mu_{t+h|t} = E(y_{t+h} / x_t)$ onde x_t contém as componentes não observadas, tais como l_t, b_t e s_t . Para $h=1$ utiliza-se $\mu_t = \mu_{t+1|t}$ como uma anotação abreviada. Em muitos modelos essas médias condicionais serão idênticas às previsões pontuais conforme Hyndman *et al.* (tabela 2.1, pág.18, 2008) de modo que $\mu_{t+h|t} = \hat{y}_{t+h|t}$. No entanto, para outros modelos, a previsão média condicional e a previsão pontual são um pouco diferente para $h \geq 2$.

2.4 Equações para os Métodos de Suavização Exponencial com modelagem em Espaço de Estados

No cálculo das equações para esses métodos é ilustrado neste trabalho a ideia de compor o método de suavização exponencial linear de Holt incorporado às inovações em espaço de estado com erros aditivos e multiplicativos.

As previsões pontuais e atualizações das equações para o método suavização exponencial linear de Holt (NA) sugeridas por Hyndman *et al.* (2008, p.18).

$$\text{Nível: } l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}), \quad (5)$$

$$\text{Tendência: } b_t = \beta^* (l_t - l_{t-1}) + (1 - \beta^*) b_{t-1}, \quad (6)$$

$$\text{Previsão: } \hat{y}_{t+h|t} = l_t + b_t h. \quad (7)$$

Para todas as 30 variações de suavização exponencial existem modelos em espaço de estados. O modelo geral envolve um vetor estado $x_t = (l_t, b_t, s_t, s_{t-1}, \dots, s_{t-m+1})'$ e equações em espaço de estado da forma:

$$y_t = w(x_{t-1}) + r(x_{t-1})\varepsilon_t, \quad (8)$$

$$x_t = f(x_{t-1}) + g(x_{t-1})\varepsilon_t, \quad (9)$$

onde $\{\varepsilon_t\}$ é um processo de ruído branco gaussiano com variância σ^2 e $\mu_t = w(x_{t-1})$. O modelo com erros aditivos tem $r(x_{t-1})=1$, de modo que $y_t = \mu_t + \varepsilon_t$. O modelo com erros multiplicativos tem $r(x_{t-1})=w_t$, de modo que $y_t = \mu_t(1 + \varepsilon_t)$.

Assim, $\varepsilon_t = (y_t - \mu_t) / \mu_t$ é o erro relativo para o modelo multiplicativo. Os modelos não são únicos. Obviamente, qualquer valor $r(x_{t-1})$ vai levar previsões pontuais idênticas para y_t . Todos os 30 métodos sugeridos por Hyndman *et al.*(2008) podem ser escritos na forma das Equações (8) e (9).

Algumas das combinações de tendência, sazonalidade e erro podem ocasionalmente apresentar dificuldades numéricas, especificamente, qualquer equação do modelo que requer a divisão por uma componente de estado pode envolver a divisão por zero. Isto é um problema para modelos com erros aditivos e de qualquer tendência ou sazonalidade multiplicativa, bem como para modelos com erros multiplicativos, tendência multiplicativa e sazonalidade aditiva.

Os modelos com erros multiplicativos são úteis quando os dados são estritamente positivos. No entanto, não são numericamente estáveis quando os dados apresentam zeros ou valores negativos. Assim, quando a série temporal não for estritamente positiva, apenas os seis modelos totalmente aditivos poderão ser aplicados.

As previsões pontuais são facilmente obtidas a partir desses modelos pelas equações de iteração sugeridas por Hyndman e Athanasopoulos (2013). Na maioria dos casos as previsões pontuais podem ser mostradas como sendo iguais para a equação $\mu_{t+h|t} = E(y_{t+h} / x_t)$, a esperança condicional do modelo em espaço de estado correspondente.

Os modelos também oferecem um meio de obtenção de intervalos de previsão. A abordagem mais direta que funciona para todos os métodos é simplesmente simular vários trajetos futuros da amostra condicionada à última estimativa do vetor estado, x_t .

Em seguida, os intervalos de previsão podem ser obtidos a partir de percentil dos trajetos de amostra simulada. Previsões pontuais podem ser obtidas também desta forma, tomando a média dos valores simulados em cada período de tempo futuro. Uma vantagem desta abordagem é a possibilidade de gerar uma estimativa completa da distribuição preditiva. Isso é especialmente útil em aplicações em que os custos esperados dependem de toda a distribuição.

2.6 Estimação

Para utilizar esses modelos de previsão, é necessário selecionar o tipo de modelo a ser utilizado, conhecer os valores de x_0 (valores iniciais) e os valores dos parâmetros α, β^*, γ e ϕ (estimativa). Conhecendo o modelo paramétrico desenvolvido a partir de inovações em espaço de estado a probabilidade do modelo pode ser facilmente calculada (Equações 8 e 9) e, assim obter estimativas de máxima verossimilhança. Ord *et al.* (1997) mostram que

$$L^*(\theta, \mathbf{x}_0) = n \log \left(\sum_{t=1}^n \varepsilon_t^2 \right) + 2 \sum_{t=1}^n \log |r(x_{t-1})| \quad (10)$$

é igual a duas vezes o negativo do logaritmo da função verossimilhança, eliminando os constantes, condicionados nos parâmetros $\theta = (\alpha, \beta, \gamma, \phi)'$ e os estados iniciais

$\mathbf{x}_0 = (l_0, b_0, s_0, s_{-1}, \dots, s_{-m+1})'$ onde n é o número de observações. Isto pode ser facilmente calculado, utilizando simplesmente as equações recursivas sugeridas por Hyndman *et al.* (2008).

Os parâmetros θ e o estado inicial \mathbf{x}_0 podem ser estimados através da minimização de L^* . Alternativamente, as estimativas podem ser obtidas minimizando-se o erro quadrático médio (MSE) a um passo da previsão, minimizando-se a variância residual σ^2 , ou através de algum outro critério que mensure o erro de previsão.

Existem diversas sugestões para restringir o espaço dos parâmetros α, β e γ , a tradicional abordagem é a de assegurar que as várias equações possam ser interpretadas como uma média ponderada, exigindo, portanto $\alpha, \beta^* = \beta/\alpha, \gamma^* = \gamma/(1-\alpha)$ e ϕ onde todos se encontram no intervalo (0,1). Isto sugere que $0 < \alpha < 1, 0 < \beta < \alpha, 0 < \gamma < 1-\alpha$ e $0 < \phi < 1$. Neste trabalho, se restringe os estados iniciais \mathbf{x}_0 para que os índices sazonais sejam adicionados ao zero em sazonalidade aditiva e adicionados a m em sazonalidade multiplicativa.

2.7 Seleção do Modelo

As medidas de precisão das previsões, tais como a média do erro quadrático médio (MSE) pode ser utilizada para selecionar um modelo de um dado conjunto de dados, desde que os erros sejam calculados a partir de dados de um conjunto de retenção de saída e não a partir dos mesmos dados que foram utilizados para o desenvolvimento do modelo. No entanto, muitas vezes há muito poucos erros fora da amostra para emitir conclusões confiáveis.

O objetivo da seleção de modelos é encontrar um modelo econômico dentre as várias possibilidades, isto é, o modelo que envolva o mínimo de parâmetros possíveis a serem estimados e que explique bem o comportamento da variável resposta. Desta forma, espera-se um modelo com bom ajuste e conseqüentemente apresente um bom poder de previsão. Para selecionar um único modelo dentre as diversas possibilidades, normalmente opta-se pela utilização de critérios de informação.

Os critérios de informação geralmente penalizam a verossimilhança, que é comparada com o número de parâmetros em questão no modelo para compensar o ajuste em potencial dos dados. Um método que apresenta uma determinada abordagem que utiliza uma probabilidade penalizada é o Critério de Informação de Akaike (AIC): $AIC = L^*(\hat{\theta}, \hat{\mathbf{x}}_0) + 2q$, onde q é o número de parâmetros em θ mais o número de estados livres em \mathbf{x}_0 , e $\hat{\theta}$ e $\hat{\mathbf{x}}_0$ indicam as estimativas de θ e \mathbf{x}_0 . No cálculo do AIC, aplicado neste trabalho, exige-se que o modelo em espaço de estado não possua estados redundantes (HYNDMAN *et al.*, 2008).

2.8 Intervalos de Predição

As distribuições de previsão para a classe 1 são claramente normais, enquanto os modelos são lineares e os erros são normais. Conseqüentemente, $100(1-\alpha)$ intervalos de previsão podem ser calculados através de previsão e variações da forma usual, ou seja, assumindo que se ε_t é um processo gaussiano, então y_{n+h}/x_n também será gaussiano. Portanto, os intervalos de previsão são facilmente obtidos a partir da Equação

$$\mu_{n+h} \pm z_{\alpha/2} \sqrt{v_n + h/n} \quad (11)$$

onde z_q indica o q -ésimo quantil de uma distribuição normal padronizada. O estimador para a variância σ^2 é o estimador de máxima verossimilhança é dado por

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^n \hat{\varepsilon}_t^2, \quad (12)$$

com $\hat{\varepsilon}_t = y_t - \mu_{t|t-1}$.

As distribuições de previsão das classes 2 e 3 são não normais devido a não linearidade das equações espaço de estado. No entanto, o intervalo de previsão baseado na Equação (11) para distribuição normal geralmente fornece resultados razoavelmente precisos.

2.9 Previsão Automática

Estabelecer previsões automáticas diante dos inúmeros recursos computacionais hoje existentes é fundamental principalmente na área de gestão de negócios onde muitas vezes é grande o número de séries temporais. Mesmo onde é necessário um número menor de previsões, pode haver situações onde ninguém esteja adequadamente treinado para produzir modelos de séries temporais. Nessas circunstâncias, um algoritmo de previsão automática pode ser uma ferramenta essencial.

É importante ressaltar que um algoritmo de previsão automática deve determinar um modelo de série temporal adequado, estimar os parâmetros e calcular as previsões. Além disso, deve ser robusto para padrões de série temporal pouco comuns, e aplicável a um grande número de séries, sem intervenção do usuário.

Um algoritmo de previsão automática foi desenvolvido por Hyndman *et al.* (2013) no pacote *Forecast*, envolvendo um conjunto de etapas que combinam todo o referencial teórico até aqui apresentado. As etapas envolvidas estão resumidas abaixo:

1. Para cada série, aplicam-se todos os modelos que são apropriados, otimizando os parâmetros (ambos os parâmetros de suavização e a variável estado inicial) do modelo, em cada caso.
2. Selecionar o melhor dos modelos de acordo com o AIC.
3. Produzir previsões pontuais utilizando o melhor modelo (com parâmetros otimizados) para o maior número de passos à frente, conforme necessário.
4. Obter intervalos de predição para o melhor modelo utilizando resultados analíticos, ou simulando caminhos futuros para a amostra $\{y_{n+1}, \dots, y_{n+h}\}$ e buscando $\alpha/2$ e $(1-\alpha)/2$ percentil dos dados simulados em cada horizonte de previsão.

3. Metodologia

A sistemática da modelagem para determinar a previsão automática de demanda envolve a aplicação do algoritmo proposto na seção 2.8 via ambiente R. O algoritmo e a estrutura de modelagem para determinar a previsão automática de séries temporais são implementados no pacote *Forecast* (HYNDMAN *et al.*, 2013) em R que é aplicável para dados de séries temporais com ou sem efeito de sazonalidade.

O pacote *Forecast* para o sistema de computação estatística R utilizado no desenvolvimento da metodologia apresenta algumas funcionalidades específicas para séries temporais univariadas tais como funções de previsão para estabelecer previsão univariada. Neste trabalho a previsão de demanda é determinada para um horizonte de quatro trimestres de 2012, e para tal aplicação são utilizados dados reais referentes ao histórico de vendas de uma indústria do setor metal mecânico do norte do estado de Santa Catarina, cujos dados revelam tanto características de tendência quanto de sazonalidade.

A série temporal referente às vendas trimestrais de um produto dessa indústria envolve os

dados relevantes e relativos às características técnicas de montagem de equipamentos, encontrados no contexto real durante o período de 20 trimestres de 2007 a 2011. Durante esses cinco anos de operação essa indústria obteve um determinado crescimento na venda desse produto, conforme Tabela 2, que mostra os dados históricos referentes às vendas trimestrais no período dos 20 trimestres.

Trimestre	1	2	3	4	5	6	7	8	9	10
Vendas	139	103	301	101	168	124	363	122	146	282
Trimestre	11	12	13	14	15	16	17	18	19	20
Vendas	462	165	416	380	514	170	331	381	261	267

Tabela 2 – Dados históricos de vendas trimestrais (em milhares)

4. Discussões e Resultados

Os resultados obtidos com a aplicação dessa metodologia para os dados reais da série temporal das vendas da empresa mencionada num período de cinco anos revelam que tais modelos podem ser eficientemente aplicados via ambiente R para selecionar o modelo de série temporal mais apropriado e, sobretudo otimizar em tempo real a previsão de demanda.

As inovações da estrutura de modelagem espaço de estado descritas na seção 2 são implementadas pelo algoritmo de suavização exponencial automático em ambiente R via função ETS ().

O modelo escolhido através do algoritmo automático para o conjunto de dados históricos das vendas da empresa mencionada foi ETS (M,M_d,M) trimestral, considerando o erro multiplicativo (M), com tendência amortecida (M_d) e com sazonalidade multiplicativa (M). São sete parâmetros no total, sendo que três correspondem às constantes de suavização $\alpha = 0,1e-04$ (nível), $\beta = 1e-04$ (tendência) e $\gamma = 0,948$ (sazonalidade); as componentes sazonais relativas aos trimestres do ano, conforme Tabela 3.

Trimestre	Quantidade
1°	0,5618
2°	1,5060
3°	0,9482
4°	0,9839

Tabela 3 – Componentes Sazonais trimestrais

A previsão determinada a partir desse modelo escolhido inclui um conjunto de informações tais como os parâmetros do modelo, o valor de vetor estado x_t para todo t , as componentes sazonais trimestrais. As expressões resultantes para o modelo estão sumarizadas abaixo.

$$l_t = 0,00001 \frac{y_t}{s_{t-m}} + (1 - 0,00001)(l_{t-1} + b_{t-1}), \quad (13)$$

$$b_t = 0,00001 (l_t - l_{t-1}) + (1 - 0,00001) b_{t-1}, \quad (14)$$

$$s_t = \frac{y_t 0,948}{l_{t-1} + b_{t-1}} + (1 - 0,948) s_{t-m}, \quad (15)$$

$$\hat{y}_{t+h|t} = (l_t + b_t h) s_{t-m+h_m^+}. \quad (16)$$

A Tabela 4 apresenta os valores de alguns erros de previsão dentro da amostra conforme medidas de desempenho e otimização estatística.

Medida de desempenho	Valor
ME (Erro Médio)	-6,5353
RMSE (Raiz do Erro Quadrático Médio)	86,4383
MAE (Erro Absoluto Médio)	53,1267
MAPE (Erro Percentual Absoluto Médio)	22,6828

Tabela 4 – Valores de alguns erros de previsão dentro da amostra

O conjunto de previsões pontuais mostrando o intervalo de predição (80% e 95%) desse método para dois períodos conforme Tabela 5.

Item	2012/1	2012/2
Previsão Pontual	403.3640	401.0348
LI (80 %)	282.8519	278.8717
LS (80 %)	522.3418	520.7906
LI (95 %)	219.6032	220.8057
LS (95 %)	584.8550	581.7753

Tabela 5 – Previsões pontuais e intervalos de predição para dois trimestres

Os resíduos podem ser aproximados por uma distribuição normal (p -valor = 0,5387) e não apresentaram autocorrelação. Não há motivos para a não aceitação do modelo.

A Figura 1 traz a série original (em preto) as previsões pontuais (em azul) e o intervalo de predição (em cinza), obtidos utilizando método de suavização exponencial com modelagem em espaço de estado.

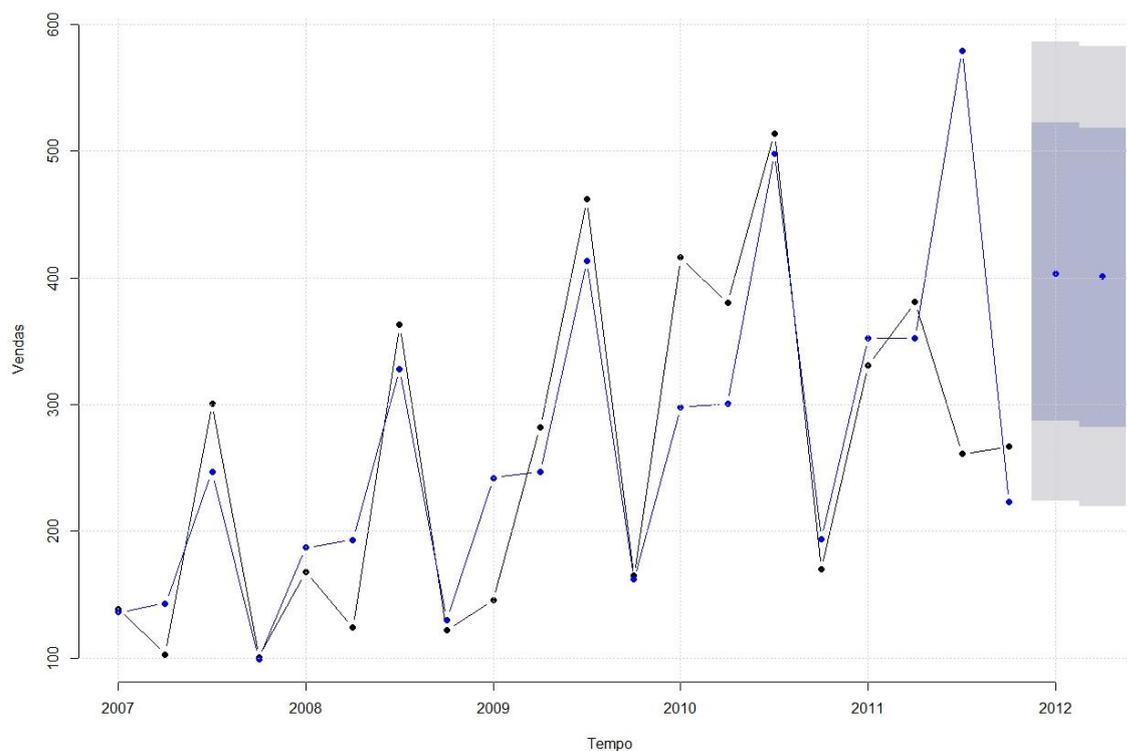


Figura 1 – Previsões utilizando método de suavização exponencial com modelagem em espaço de estado

Embora exista uma grande quantidade de computação envolvida, o algoritmo de previsão proposto neste trabalho para execução em ambiente R pode ser tratado com maior rapidez em computadores modernos. As previsões para o conjunto de dados históricos de vendas determinadas com o pacote de previsão em ambiente R levou não mais do que alguns segundos em um PC padrão.

5. Considerações Finais

O modelo estatístico de suavização exponencial com abordagem em espaço de estado proposto neste trabalho permite uma considerável flexibilidade na especificação de estruturas paramétricas tanto lineares quanto não lineares. Além disso, é apropriado para qualquer tipo de dados e ajusta-se bem à abordagem de suavização exponencial, uma vez que as equações de nível, tendência e sazonalidade possuem seus estados explícitos nesse modelo.

A análise de séries temporais, mediante a utilização de recursos computacionais tais como o pacote de previsão em R, significa uma excelente alternativa para incentivar a utilização de uma ferramenta aparentemente complexa, uma vez que permite um ambiente de contínua simulação de dados, que possibilita o teste de vários modelos e a escolha daquele que apresenta os parâmetros mais adequados.

Com a execução correta de um algoritmo de previsão automática um modelo de previsão pode ser escolhido em tempo real sem o auxílio de um analista de série temporal, ou seja, a partir da análise da amostra com um critério de seleção do modelo. Evidentemente, a interpretação dos resultados obtidos da análise de séries temporais requer experiência do analista de previsão, de forma que os cenários construídos tornem-se mais robustos e úteis na identificação da melhor alternativa de decisão.

Referências

BERMÚDEZ, J.D.; SEGURA, J.V.; VERCHER, E. *SIOPRED: a prediction and optimisation integrated system for demand*, TOP, v. 16, p. 258–271, 2008.

GARDNER, E.S. Jr. *Exponential Smoothing: The State Of The Art*. *Journal Of Forecasting*, n.4, p. 1–28, 1985.

GARDNER, E.S. *Exponential Smoothing: The State Of The Art - Part I e II*. *International Journal of Forecasting*, v. 22, p. 637–666, 2006.

HYNDMAN, R. J.; KOEHLER, A. B.; ORD, J. K.; SNYDER, R. D.; GROSE, S. *A State Space Framework For Automatic Forecasting Using Exponential Smoothing Methods*. *International Journal Of Forecasting*, n.18, p. 439–454, 2002.

HYNDMAN, R. J., KOEHLER, A. B.; ORD, J. K.; SNYDER, R. D. *Forecasting With Exponential Smoothing: The State Space Approach*, Springer-Verlag, Berlin, 2008.

HYNDMAN, R.J.; ATHANASOPOULOS, G. *Ateloprosopos Forecasting: principles and practice*. Disponível em: <http://otexts.com/fpp/>. Acesso em: 19 Jun. 2013.

HYNDMAN, R.J.; ATHANASOPOULOS, G.; SLAVA, R.; SCHMIDT, D.; ZHOU, Z.; YOUSAF, K. (2013). *Forecast: Forecasting functions for time series and linear models*. R package version 4.04. Disponível em: <http://CRAN.R-Project.org/package=forecast>. Acesso em: 08 Mai. 2013.

KALMAN, R.E. *A new approach to linear filtering and prediction problem*, *Journal of Basic Engineering*, v. 82, n. 1, p. 35–45, 1960.

MAKRIDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R.J. *Forecasting Methods And Applications*. 3. Ed. New York: John Wiley & Sons, 1998.

MORETTIN, P. A.; TOLOI, C. M. *Análise de Séries Temporais*. 2ª ed. Edgard Blucher, 2006.

ORD, J.K.; KOEHLER, A. B.; SNYDER, R. D. *Estimation and prediction for a class of dynamic nonlinear statistical models*, *Journal of the American Statistical Association*, v. 92, p. 1621–1629, 1997.

PEGELS, C.C. *Exponential forecasting: Some new variations*, Management Science, v. 15, n. 5, p. 311–315, 1969.

RAGSDALE, C.T. *Modelagem E Análise De Decisão*. Cengage Learning, 2009.

SOUZA, G.P.; SAMOHYL, R.W.; MIRANDA, R.G. *Métodos Simplificados de Previsão Empresarial*. Ciência Moderna, 2008.

TAYLOR, J.W. *Exponential smoothing with a damped multiplicative trend*, International Journal of Forecasting, v. 19, p. 715–725, 2003.

VERÍSSIMO, A.J.; ALVES, C. C. ; HENNING, E.; AMARAL, C.E.; CRUZ, A.C. *Métodos Estatísticos de Suavização Exponencial Holt-Winters para Previsão de Demanda em uma Empresa do Setor Metal Mecânico*. Gestão Industrial, v. 8, n. 4, p. 154-171, 2012.