

## ESTUDO COMPARATIVO DE SÉRIES TEMPORAIS PARA PREVISÃO DE VENDAS DE UM PRODUTO

William Jacobs<sup>1</sup>

Roselaine Ruviaro Zanini<sup>2</sup>

Manfred Costa<sup>3</sup>

**RESUMO:** Este artigo apresenta um estudo comparativo de modelos de séries temporais para a previsão de vendas de determinado produto. O estudo teve como objetivo investigar a eficácia, em termos da capacidade preditiva de cada um dos modelos, utilizando uma série temporal real da demanda de determinado produto. Foram utilizados três tipos de modelos para a previsão dos valores futuros: (i) suavização exponencial (SE); (ii) autorregressivo integrado de médias móveis (ARIMA); e, (iii) redes neurais artificiais (RNAs). Após a modelagem, foram selecionados os modelos que apresentaram o melhor resultado em cada categoria supracitada e então comparado o desempenho entre cada. Os resultados demonstraram o modelo de RNAs MLP(6,10,1) como aquele mais eficaz para a série temporal utilizada (MAPE de ajustamento e previsão de 28,55% e 22,33%, respectivamente). Verificou-se que o modelo de RNAs MLP(6,10,1) apresentou um resultado 58% e 48% melhor, em termos de modelagem da série, em relação aos modelos de SE e ao modelo ARIMA, respectivamente. Em termos de capacidade preditiva, verificou-se que o modelo de RNAs MLP(6,10,1) apresentou um resultado 73% e 65% melhor em relação aos modelos de SE e ARIMA, respectivamente.

**Palavras-chave:** Previsão de venda. Modelos de suavização exponencial. Modelos ARIMA. Redes neurais artificiais.

### 1 INTRODUÇÃO

Em mercados altamente competitivos, prever os resultados futuros é uma importante informação para as organizações que nestes estão inseridas (FLORES, 2009). Quando se pode estimar, com certo grau de confiança, a quantidade a ser vendida nos próximos períodos, é possível estimar a quantidade de recursos necessários para o atendimento dessa demanda. A partir de então, desdobrar essa necessidade de recursos entre as principais funções e funções

---

<sup>1</sup> Mestre em Engenharia de Produção pela UFSM, Professor do curso de Engenharia de Produção do Centro Universitário UNIVATES, Lajeado, Rio Grande do Sul, Brasil. E-mail: [eng.williamjacobs@gmail.com](mailto:eng.williamjacobs@gmail.com).

<sup>2</sup> Doutora em Epidemiologia pela UFRGS, Departamento de Estatística, UFSM (Universidade Federal de Santa Maria), Santa Maria, Rio Grande do Sul, Brasil. E-mail: [rrzanini@terra.com.br](mailto:rrzanini@terra.com.br).

<sup>3</sup> Mestre em Engenharia de Produção pela UFRGS, Coordenador e Professor do curso de Engenharia de Produção do Centro Universitário UNIVATES, Lajeado, Rio Grande do Sul, Brasil. E-mail: [manfredcosta@yahoo.com](mailto:manfredcosta@yahoo.com).

de apoio da organização, faz com que o planejado esteja mais ajustado com a realidade, resultando em maior eficiência na utilização dos recursos e eficácia nos resultados obtidos.

Uma etapa fundamental no processo de previsão é a decisão de qual o modelo que será utilizado. Na área de análise de séries temporais, área que estuda os modelos e metodologias de modelagem e previsão de séries temporais, existem diversos modelos para a utilização. O fato, porém, é que se deve observar os diversos pressupostos em que se baseiam, fato esse que pode, muitas vezes, dificultar a escolha (FLORES, 2009). Para Zhang (2003), é necessário que o pesquisador utilize mais de um modelo para a previsão, tendo em vista a impossibilidade de conhecer, previamente, o processo gerador da série temporal. Lee e Tong (2011), acrescentam que séries temporais reais frequentemente apresentam tanto padrões lineares quanto padrões não lineares e também, que alguns modelos capturam adequadamente os padrões lineares e outros os padrões não lineares. Considerando essas recomendações e restrições, opta-se, neste estudo, de se utilizar de mais de um modelo para a modelagem e previsão da série, otimizando assim os resultados da análise.

Com isso, o objetivo do presente estudo é investigar a eficácia de previsão de três modelos: (i) SE; (ii) ARIMA; e, (iii) RNAs; para a previsão de vendas de um determinado produto de uma empresa de laticínios localizada na região do vale do Taquari, interior do Estado do Rio Grande do Sul. A série temporal analisada neste estudo é composta de 292 dados de vendas mensais de um produto produzido pela empresa. Destes 292 dados mensais de venda, 268 são utilizados para a modelagem da série temporal e 24 (últimos) são utilizados para a comparação entre os valores previstos gerados pelo modelo e os ocorridos.

Para tanto, o trabalho encontra-se dividido da seguinte forma: a Seção 2 aborda o tema previsão de vendas por intermédio de séries temporais e apresenta os três métodos de previsão abordados neste estudo; a Seção 3 apresenta a metodologia utilizada para desenvolvimento e obtenção dos resultados; a Seção 4 apresenta os resultados da modelagem da série temporal de vendas do produto e os resultados de previsão de valores futuros para o modelo; e, por último, a Seção 5 apresenta a conclusão do presente estudo e faz considerações para a realização de trabalhos futuros sobre o tema.

## **2 PREVISÃO DE VENDAS POR MEIO DE MODELOS DE SÉRIES TEMPORAIS**

A previsão de vendas é um processo sustentado por uma metodologia de trabalho clara e definida, que apoiada em modelos estatísticos, matemáticos ou econométricos, ou ainda, em

modelos subjetivos, busca determinar os valores futuros de uma série temporal de vendas (MARTINS; LAUGENI, 2003). Segundo Figueredo (2008), tal processo possibilita ao usuário obter conhecimento a respeito da provável evolução da série no futuro. Para Tubino (2007), a necessidade e importância desse processo é em função de que a quantidade vendida é a variável mais importante na definição de um sistema de produção, especialmente para as funções desenvolvidas pela área de planejamento e controle da produção.

Os métodos de previsão de vendas são classificados conforme suas abordagens, que são: (i) quantitativos; (ii) qualitativos; ou, (iii) que utilizam a combinação dos dois. Para a previsão utilizando os métodos quantitativos, requer-se a construção de modelos matemáticos baseados em dados históricos que descrevam a variação da demanda ao longo do tempo, os modelos de séries temporais (PELLEGRINI, 2000). As séries temporais são medidas de determinadas variáveis tomadas a intervalos regulares de tempo (MOORE et al., 2006).

Na previsão de vendas por meio de séries temporais, são coletados e analisados dados históricos das vendas do produto a ser analisado, com o objetivo de obter um modelo que descreva a evolução dos valores no tempo (ZHANG, 2003). Após a coleta dos dados e obtenção do modelo que melhor se ajusta, valores futuros da série temporal são previstos com base no modelo obtido (ZOU et al., 2007). Porém, há de se ressaltar, que não existe técnica de previsão perfeita, sempre haverá erros envolvidos no processo, pois são vários os fatores que influenciam o ambiente de negócios. Assim, uma estratégia importante é a de se estabelecer uma prática de revisão contínua das previsões (CHASE; JACOBS; AQUILANO, 2006).

A Seção 2 foi subdividida em outros três assuntos relativos ao tema, que são os modelos utilizados no presente estudo para a modelagem e previsão da série temporal de vendas. Dessa forma, a Seção 2.1 apresenta os modelos de suavização exponencial; a Seção 2.2 introduz os modelos ARIMA; e, a Seção 2.3 apresenta os modelos de redes neurais artificiais.

## **2.1 Modelos de suavização exponencial (SE)**

Os modelos de SE são métodos conhecidos com esta denominação em virtude destes aplicarem um conjunto de pesos desiguais aos valores passados da série temporal, sendo que tais pesos decaem de forma exponencial da mais recente a mais distante observação (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998; MORETTIN; TOLOI, 2006). A larga utilização dos modelos de SE são em virtude de sua simplicidade, baixo custo e razoável precisão (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998). A principal desvantagem

é a dificuldade em determinar os valores mais apropriados das constantes de suavização (MORETTIN; TOLOI, 2006).

Os modelos de SE utilizados neste estudo foram: (i) suavização exponencial de Holt (SEH); e, (ii) suavização exponencial de Holt-Winters (SEHW). Os modelos de SEH são utilizados satisfatoriamente em séries temporais que apresentam tendência de crescimento ou de decrescimento linear. Os modelos de SEHW, ou modelos de Winters como são conhecidos, são utilizados satisfatoriamente em séries temporais que, além de tendência de crescimento ou decrescimento linear, apresentam também sazonalidade. As Equações 1, 2 e 3, apresentam os componentes nível e tendência do modelo no período  $t$  e a previsão para  $t+k$  períodos, respectivamente, utilizando os modelos de SEH (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998).

$$L_t = \alpha y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2)$$

$$\hat{y}_{t+k} = L_t + kT_t \quad (3)$$

Os modelos de SEHW, por sua vez, dividem-se em aditivos e multiplicativos. Quanto aos modelos de suavização exponencial de Holt-Winters aditivo (SEHWA), parte-se do pressuposto que a amplitude da variação sazonal permanece constante ao longo do tempo (PELLEGRINI, 2000). As Equações 4, 5, 6 e 7, apresentam os componentes nível, tendência e sazonalidade do modelo no período  $t$  e a previsão para  $t+k$  períodos, respectivamente, para os modelos de SEHWA (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998).

$$L_t = \alpha(y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (4)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (5)$$

$$S_t = \gamma(y_t - L_t) + (1 - \gamma)S_{t-s} \quad (6)$$

$$\hat{y}_{t+k} = L_t + kT_t + S_{t-s+k} \quad (7)$$

Quanto aos modelos de suavização exponencial de Holt-Winters multiplicativo (SEHWM), parte-se do pressuposto que a amplitude da variação sazonal varia ao longo do tempo, ou seja, com o passar deste (PELLEGRINI, 2000; ALLEMÃO, 2004). As Equações 8, 9, 10 e 11, apresentam os componentes nível, tendência e sazonalidade do modelo no período  $t$  e a previsão para  $t+k$  períodos, respectivamente, para os modelos de SEHWM

(MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998).

$$L_t = \alpha \frac{y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (8)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (9)$$

$$S_t = \gamma \frac{y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (10)$$

$$\hat{y}_{t+k} = (L_t + kT_t)S_{t-s+k} \quad (11)$$

Esta seção apresentou os modelos de suavização exponencial de Holt e de Holt-Winters, que devido a simplicidade e capacidade de modelagem de séries temporais com tendência linear e/ou sazonalidade, serão utilizados neste estudo para a previsão dos valores futuros das vendas de um produto.

## 2.2 Modelos autorregressivos integrados de médias móveis (ARIMA)

Os modelos ARIMA foram descritos por Box e Jenkins (1976) e caracterizam-se como uma das mais difundidas metodologias para análise de séries temporais (MADDALA, 2003; ZHANG, 2003; MORETTIN; TOLOI, 2006). Por mais de meio século tais modelos têm sido utilizados em diversas áreas que realizam previsões por intermédio de séries temporais (HO; XIE; GOH, 2002). Há diversas aplicações gerais e específicas realizadas por diversos autores ao longo dos anos, como: aplicações na área médica, ambiental, financeira, engenharia, previsão da qualidade do ar, entre outras (WERNER; RIBEIRO, 2003).

Os modelos ARIMA são modelos classificados como univariados, ou seja, utilizam apenas uma variável para a sua execução e consistem em explicar determinada variável por meio de valores passados dela mesma, dos valores passados de choques e dos valores passados dos erros ou resíduos (SARTORIS, 2003). A principal desvantagem desses modelos, segundo Zhang (2003) e Aburto e Weber (2007), é que assumem uma relação linear entre a variável dependente e independente, o que nem sempre acontece. Quanto aos benefícios, são relativamente flexíveis e podem representar diversas situações diferentes de séries temporais, podendo-se utilizar apenas um filtro, como: o autorregressivo (AR), ou o de médias móveis (MA), ou o autorregressivo de médias móveis (ARMA) (ZHANG, 2003).

A metodologia proposta por Box e Jenkins (1976) é desenvolvida com base em três passos: (i) identificação do modelo a ser utilizado; (ii) estimação dos parâmetros do modelo

identificado; e, (iii) verificação do modelo obtido. Wang (2011) e Zhang (2003), ressaltam a importância de se verificar a condição de estacionariedade e sazonalidade da série antes de iniciar os passos descritos. As etapas descritas são repetidas várias vezes testando diversos modelos, até a obtenção de um modelo satisfatório (ZHANG, 2003). Fava (2000) ressalta que após verificadas essas condições, o modelo poderá ser utilizado para previsão.

A condição de estacionariedade da série pode ser verificada de diversas maneiras, entre elas, através do uso de testes de raiz unitária. Os testes de raiz unitária objetivam diagnosticar se há a presença ou não de raiz unitária na série temporal. Caso a série seja não estacionária (que possui raiz unitária), utilizam-se procedimentos de diferenciação para que esta transforme-se em estacionária (MATOS, 2000; KUMAR; JAIN, 1999). Um teste comumente utilizado para verificação de raiz unitária é o *Augmented Dickey-Fuller* (ADF), que pode ser verificado com mais detalhe na obra de Enders (2004).

Na etapa de identificação, são identificados os filtros que irão compor o modelo ARIMA, ou seja, a presença e o número de componentes autorregressivos e de médias móveis, respectivamente. Para tanto, a técnica comumente utilizada é a análise da função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP). A decisão é tomada com base nos padrões identificados para cada uma das funções, FAC e FACP, com base na respectiva descrição apresentada no Quadro 1 (FAVA, 2000).

Quadro 1 – padrões teóricos da FAC e da FACP para identificação do modelo

Tipo de modelo	Padrão típico de FAC	Padrão típico de FACP
AR( $p$ )	Declina exponencialmente ou com padrão de onda senoidal amortecida, ou ambos.	Picos significativos através das defasagens $p$ . Truncada em $k=p$ .
MA( $q$ )	Picos significativos através das defasagens $q$ . Truncada em $k=p$ .	Declina exponencialmente.
ARMA( $p,q$ )	Declínio exponencial. Declinante a partir de $k=p$ .	Declínio exponencial.

Fonte: Adaptado de Gujarati (2000), Fava (2000) e Morettin (2008)

A FAC é o conjunto de valores de  $r_k$  e a representação gráfica deste conjunto de valores em relação ao seus *lags* (defasagens) é denominado correlograma. A Equação 12 apresenta o modelo para cálculo de  $r_k$  (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998).

$$r_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}, \text{ para } k=0,1,2,\dots,n. \quad (12)$$

A FACP, por sua vez, mede a correlação entre o valor de determinada série no tempo  $t$

e seu valor defasado em  $t-k$  períodos depois que a influência de  $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$  sobre  $y_t$  tenha sido descontada (FAVA, 2000). Existem diversos meios para obtenção dos valores da FACP, dentre eles o algoritmo de Box e Jenkins. Para obter maiores informações sobre a obtenção dos valores da FACP, o leitor poderá verificar em Enders (2004).

Após analisado quais são os filtros ou componentes que serão utilizados no modelo ARIMA, o próximo passo é a determinação do número de parâmetros  $p$  e  $q$  a serem inclusos no modelo. Os valores de tais parâmetros podem ser tanto obtidos através da FAC e da FACP (KUMAR; JAIN, 1999), quanto através dos critérios de Akaike ou critério informacional de Akaike (AIC) e do critério informacional Bayesiano (BIC), escolhendo-se, dentre os modelos propostos, aquele com o menor resultado dentre esses critérios. O AIC e o BIC são obtidos, respectivamente, através das Equações 13 e 14 (ENDERS, 2004).

$$AIC = n \log \left[ \frac{SQR}{(n-p)} \right] + 2p \quad (13)$$

$$BIC = n \log \left[ \frac{SQR}{(n-p)} \right] + p \log n \quad (14)$$

Obtido o modelo ARIMA(p,d,q), realiza-se então a estimação de seus parâmetros. Fava (2000) e Maddala (2003) ressaltam que quando o filtro MA estiver presente, deve-se utilizar o método dos mínimos quadrados não lineares. Caso contrário, a estimação dos parâmetros pode ser realizada pelo método dos mínimos quadrados ou pelo método da máxima verossimilhança. Segundo Morettin (2008), os pacotes estatísticos utilizados para séries temporais trazem o algoritmo adequado para a estimação dos parâmetros, motivo esse e em função da complexidade de tais métodos, que estes não serão abordados neste estudo. Para maiores detalhes, leia Morettin (2008).

Na etapa da verificação do modelo obtido, avaliam-se se os resíduos gerados pelo ajustamento comportam-se como sendo ruído branco, ou seja, sua média é zero e sua variância permanece constante ao longo do tempo. Outra consideração importante nesta etapa é a de que os resíduos do modelo não sejam autocorrelacionados. Caso não seja verificada tal situação, faz-se necessária a utilização de outro modelo (FAVA, 2000; MORETTIN, 2008).

A Seção 2.2 apresentou os principais conceitos relativos a utilização dos modelos ARIMA, aos leitores interessados em aprofundar o entendimento sobre tais modelos sugere-se o trabalho de Box e Jenkins (1976) e o de Makridakis, Wheelwright e Hyndman (1998). Para



finalizar a presente seção, a Equação 15 apresenta o modelo genérico ARIMA(p,d,q) e a Equação 16 apresenta o modelo genérico SARIMA(p,d,q)(P,D,Q), ou como conhecido, modelo ARIMA com sazonalidade (MAKRIDAKIS, WHEELWRIGHT; HYNDMAN, 1998).

$$(1 - \phi_1 B - \dots - \phi_p B^p) y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t \quad (15)$$

$$(1 - \phi_1 B - \dots - \phi_p B^p) (1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}) (1 - B)^d (1 - B^s)^D y_t = (1 - \theta_1 B - \dots - \theta_q B^q) (1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}) \varepsilon_t \quad (16)$$

Esta seção apresentou os modelos ARIMA e o processo de obtenção destes para a previsão de séries temporais, denominado metodologia de *Box-Jenkins*.

### 2.3 Redes neurais artificiais (RNAs)

Uma RNA trata-se de um conjunto de entrelaçamentos, sendo que cada conexão ou sinapse apresenta um peso associado para o ajuste dos dados de entrada aos dados de saída (SOBREIRO; ARAÚJO; NAGANO, 2009). Conforme Haykin (1999), as RNAs possuem a propensão natural de armazenar conhecimento experimental e torná-lo disponível para o uso. Isso se dá em função de as RNAs serem treinadas, por intermédio de algumas regras de treinamento, que ajustam os pesos das conexões conforme os dados disponíveis são minimizados por uma função de erro apropriada (ABURTO; WEBER, 2007).

A utilização de RNAs para a previsão de séries temporais, segundo Zhang, Patuwo e Hu (1998), ocorreu em 1968 com o trabalho de Hu, que realizava previsões referentes ao clima. Segundo Allemão (2004), em sua grande maioria, as RNAs são aplicadas a problemas relacionados à clusterização, ao reconhecimento e à previsão de séries temporais. Allemão (2004) complementa que para a prática de previsão, utilizam-se as redes conhecidas como aproximadoras de funções, ou seja, a topologia *multilayer perceptron* (MLP) e redes Elman. A maior vantagem dos modelos de RNAs sob os modelos ARIMA é em virtude de sua capacidade de capturar os padrões não lineares normalmente existentes nas séries temporais (KHASHERI; BIJARI, 2010; ZHANG, 2003), situação essa não observada nos modelos ARIMA (LEE; TONG, 2011).

A topologia de RNA mais difundida em pesquisas que envolvem modelagem e previsão por intermédio de séries temporais é a *feedforward* de uma única camada escondida. Tal modelo é descrito por três camadas (a camada de entrada, a camada escondida e a camada de



saída) e, a relação entre as saídas da RNA e suas entradas pode ser descrita matematicamente pela Equação 17 (KHASHERI; BIJARI, 2010; ZHANG, 2003).

$$y_t = w_0 + \sum_{j=1}^q w_j g\left(w_{0j} + \sum_{i=1}^p w_{ij} y_{t-i}\right) + \varepsilon_t \quad (17)$$

Onde  $w_j$  ( $j = 0, 1, 2, \dots, q$ ) e  $w_{ij}$  ( $i = 0, 1, 2, \dots, p; j = 1, 2, 3, \dots, q$ ) são os parâmetros do modelo geralmente chamados de *connection weights*;  $p$  é o número de neurônios na camada de entrada e o  $q$  é o número de neurônios na camada escondida (KHASHERI; BIJARI, 2010; ZHANG, 2003). A função de transferência geralmente utilizada na camada escondida é a função sigmóide (Equação 18) (ZHANG, 2003). Na camada de saída, a função mais utilizada é a linear, pois funções não lineares na camada de saída podem trazer distorções à capacidade preditiva da RNA (KHASHERI; BIJARI, 2010).

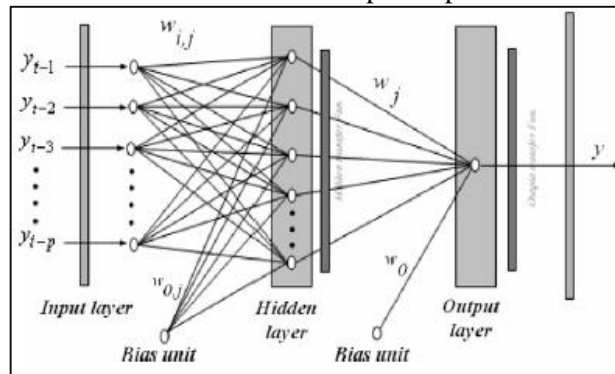
$$g(x) = \frac{1}{1 + \exp(-x)} \quad (18)$$

Da mesma forma que nos modelos apresentados nas Seções 2.1 e 2.2, os modelos de RNAs utilizam de valores passados da série temporal para a previsão dos valores futuros (KHASHERI; BIJARI, 2010, ZOU et al., 2007; ZHANG, 2003). Assim, a Equação 19 apresenta o modelo de série temporal para a modelagem utilizando RNAs.

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}, w) + \varepsilon_t \quad (19)$$

Onde  $w$  é o vetor de todos os parâmetros,  $f$  é a função determinada pela estrutura e conexões sinápticas da RNA e  $\varepsilon_t$  é o erro aleatório do período  $t$ . Desta forma, o presente modelo torna-se similar a um modelo não linear autorregressivo (ZHANG, 2003; KHASHERI; BIJARI, 2010). A topologia de uma rede MLP é apresentada em Khashei e Bijari (2010), conforme pode ser visualizado na Figura 1.

Figura 1 – Modelo estrutural de uma RNA para a previsão de séries temporais



Fonte: Khashei e Bijari (2010)

A Figura 1 apresenta o modelo estrutural de RNAs para a previsão de séries temporais, destacando os componentes presentes nesta estrutura, que são: os valores defasados da camada de entrada ( $y_{t-p}$ ), os pesos das conexões ( $w_{i,j}$ ) entre a camada de entrada (*input layer*) e a intermediária (*hidden layer*), os pesos das conexões entre a camada intermediária (ou escondida) e a de saída (*output layer*), o *bias* de cada uma das camadas ( $w_0$ ) e a saída do modelo ( $y$ ) (KHASHEI; BIJARI, 2010).

Uma vez definida a estrutura da rede, esta estará pronta para o processo de treinamento, que ajustará os pesos internos da rede conforme uma função de custo vai sendo minimizada. A função de custo é um critério que busca a minimização dos erros quadrados da rede (Equação 20), sendo que a minimização é realizada através de um algoritmo de treinamento não linear, o *back-propagation* (Equação 21) e, para acelerar o processo de treinamento, evitando também a instabilidade do algoritmo, introduz-se um termo chamado *momentum*  $\delta$  (Equação 22) (KHASHERI; BIJARI, 2010; ABURTO; WEBER, 2007; KOVÁCS, 2006).

$$E = \frac{1}{N} \sum_{n=1}^N (e_i)^2 = \frac{1}{N} \sum_{n=1}^N \left( y_t - \left( w_0 + \sum_{j=1}^Q w_j g \left( w_{0j} + \sum_{i=1}^P w_{ij} y_{t-i} \right) \right) \right)^2 \quad (20)$$

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \quad (21)$$

$$\Delta w_{ij}(t+1) = -\eta \frac{\partial E}{\partial w_{ij}} + \delta \Delta w_{ij}(t) \quad (22)$$

O *back-propagation* é considerado um algoritmo de treinamento do tipo supervisionado, pois se utiliza das informações dos padrões de entrada fornecidos à rede e a sua respectiva

saída desejada e que, por meio de um mecanismo de correção de erros (gradiente descendente), ajusta os pesos da rede aos padrões de entrada fornecidos na etapa de treinamento. Assim, o treinamento é realizado em duas etapas: (i) a fase *forward* e (ii) a fase *backward*. Na primeira fase, são apresentados conjuntos de dados à rede e esta propaga o sinal até a sua camada de saída. A segunda fase realiza a correção dos erros, alterando os pesos internos da rede, conforme a saída desejada (BRAGA; CARVALHO; LUDERMIR, 2011).

Esta seção apresentou as RNAs, especificamente a topologia MLP, para a previsão de séries temporais. A principal característica destes modelos em comparação aos de Holt-Winters e Box-Jenkins, é a capacidade de modelar tanto padrões lineares quanto não lineares que podem estar presentes na série temporal.

### 3 METODOLOGIA

O estudo foi concebido e realizado conforme os seguintes passos: (i) tabulou-se os dados para a obtenção da série temporal da demanda do produto em análise; (ii) modelou-se a série temporal segundo os modelos apresentados na Seção 2; (iii) comparou-se os resultados gerados por cada modelo; (iv) apresentou-se os gráficos de ajustamento e de previsão do modelo que obteve o melhor resultado para a série temporal (Seção 4).

Na modelagem da série temporal pelo modelo de SE, utilizou-se as suas variantes: SEH, SEHWA e SEHWM. Então, estimaram-se os parâmetros de cada modelo, por meio do vetor que minimizou a soma quadrada dos erros (ZHANG, 2003). Após, modelou-se cada um dos componentes da série temporal (nível, tendência e sazonalidade) e obteve-se as medidas de erro para a modelagem e previsão.

Para os modelos ARIMA, na primeira etapa verificou-se a condição de estacionariedade da série temporal e, para isso, utilizou-se o teste ADF – apresentado na Seção 2. Em seguida, com base na FAC e FACP, identificou-se e estimou-se cinco modelos como candidatos para a modelagem da série temporal, e o escolhido foi aquele que minimizou os critérios AIC e BIC. Então, foi verificado se os resíduos gerados pelo modelo são ruído branco – ou seja, média zero, variância constante e não autocorrelacionados.

Na modelagem das RNAs utilizou-se a topologia MLP com o *back-propagation* como algoritmo de treinamento. Implementou-se as RNAs com 1 camada intermediária para a modelagem da série temporal, e a configuração das RNAs seguem o detalhamento disposto no Quadro 2, totalizando a modelagem de 18 RNAs para a série temporal estudada.

Quadro 2 – Configuração das RNAs modeladas no presente estudo

Número de neurônio nas camadas	Rede MLP com 1 camada escondida
Camada de entrada	3, 6 e 12
Camada escondida	5, 10, 15, 20, 25 e 30
Camada de saída	1
Número de modelos	18

Fonte: Elaborado pelos autores

A modelagem das RNAs foi realizada conforme o seguinte procedimento: (i) separou-se os vinte e quatro últimos valores da série temporal para a verificação da capacidade de predição e generalização da rede; (ii) dos dados restantes da série, utilizou-se 70% para o treinamento e 30% para a validação; (iii) modelou-se a série temporal da demanda de conforme as configurações dadas no Quadro 2; (iv) avaliou-se empiricamente a possibilidade de superajustamento, comparando as medidas de erro do conjunto de treinamento e do conjunto de validação; (v) realizou-se a previsão de vinte e quatro valores futuros; (vi) obteve-se e analisou-se os resultados do ajustamento e da previsão da série temporal.

Para a verificação de superajustamento das RNAs aos dados da série, a análise empírica consistiu na comparação da variação das medidas de erro do processo de treinamento e de validação das RNAs. Caso constatada divergências acentuadas entre tais medidas, haveria fortes suposições de que determinada RNA teria memorizado os dados fornecidos para o seu treinamento, perdendo a capacidade de generalização, situação esta que não gera resultados práticos a presente pesquisa. Segundo Braga, Carvalho e Ludermir (2011), para a RNA não apresentar o problema de superajustamento, tais medidas deveriam, teoricamente, ser próximas uma da outra.

Utilizaram-se os critérios de avaliação conforme pesquisa realizada por Mentzer e Khan nos anos 1995 e 2006. Os autores, com o intuito de verificar a importância de vários critérios de avaliação de métodos de previsão, realizaram pesquisas com profissionais e acadêmicos da área no ano de 1995 e no ano de 2006, com 186 e com 86 respondentes, respectivamente. Segundo os pesquisadores, os critérios mais utilizados são o *Mean Absolute Percentage Error* (MAPE) e o *Mean Absolute Error* (MAE) que obtiveram na primeira pesquisa 75% da preferência e na segunda 65% (FELICIANO, 2009). As equações Equações 23 e 24 apresentam, respectivamente, tais critérios.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right| 100 \quad (23)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |e_i| \quad (24)$$

Após a comparação entre as modelagens realizadas no estudo, a última etapa consistiu na apresentação do gráfico com o ajustamento do modelo à série temporal e com a previsão de seus vinte e quatro valores futuros. Além do gráfico de ajustamento, realizou-se os devidos comentários a respeito dos modelos obtidos.

#### 4 RESULTADOS DO ESTUDO

Nesta seção, como forma de apresentar os resultados do estudo, são apresentadas as etapas de modelagem e da previsão da série temporal do produto em análise. Na Seção 4.1, é apresentada a modelagem da série utilizando os modelos de SE, ARIMA(p,d,q) e de RNAs. Na Seção 4.2, é apresentada a previsão de vinte e quatro valores futuros por meio do modelo que melhor descreveu o processo gerador da série temporal.

##### 4.1 Modelagem da série temporal das vendas do produto

Para a modelagem utilizando os modelos de SE, tanto a estimação dos parâmetros, quanto a modelagem da série, foram realizadas utilizando o pacote estatístico R versão 2.13.1. A Tabela 1 apresenta os respectivos modelos de SE utilizados, os valores estimados de seus parâmetros, o MAE de ajustamento (MAE<sub>a</sub>), o MAPE de ajustamento (MAPE<sub>a</sub>), o MAE de previsão (MAE<sub>p</sub>) e o MAPE de previsão da série temporal (MAPE<sub>p</sub>).

Tabela 1 – Modelagem da série temporal do produto utilizando os modelos de SE

Modelo	Parâmetros	MAE <sub>a</sub>	MAPE <sub>a</sub>	MAE <sub>p</sub>	MAPE <sub>p</sub>
SEH	$\alpha : 0,00178$ $\beta : 0,00039$	18.365	0,7708	19.977	0,9578
SEHWA	$\alpha : 0,002386$ $\beta : 0,001009$ $\gamma : 0,003937$	17.632	0,7568	18.552	0,7799
SEHWM	$\alpha : 0,002386$ $\beta : 0,001009$ $\gamma : 0,003937$	17.632	0,7568	18.552	0,7799

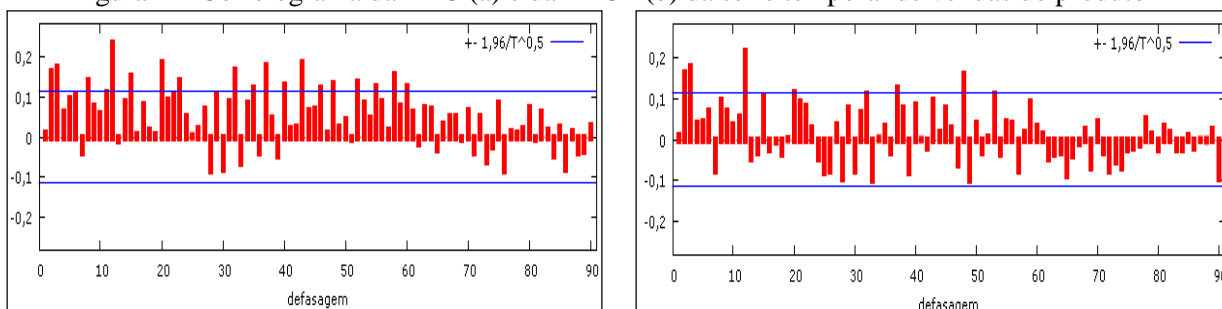
Fonte: Elaborada pelos autores

Os modelos da Tabela 1, em termos de MAPE<sub>a</sub> e MAPE<sub>p</sub>, não apresentaram resultados satisfatórios. Além disso, não houve diferenças entre a estimação dos modelos de SEHWA e de SEHWM. O modelo de SEHWA é mais adequado para séries temporais onde a variância

permanece constante ao longo do tempo e, do contrário, o modelo de SEHWM é mais adequado para séries temporais onde a variância não permanece constante.

A próxima etapa consistiu na modelagem da série temporal de vendas do produto utilizando os modelos  $ARIMA(p,d,q)$ . O processo de modelagem foi realizado utilizando o *software Gretl*, versão 1.9.5cvs, exceto o teste de raiz unitária ADF, que foi realizado no *software Eviews*, versão 4. Para o início do processo de modelagem da série, a primeira etapa consistiu na análise do correlograma de sua FAC e no correlograma de sua FACP, conforme é apresentado na Figura 2a e na Figura 2b, respectivamente.

Figura 2 – Correlograma da FAC (a) e da FACP (b) da série temporal de vendas do produto



Fonte: Elaborado pelos autores

Posteriormente, foi realizada a verificação da condição de estacionariedade. Verificou-se nas Figuras 2a e 2b, que alguns dos coeficientes de autocorrelação estão fora dos limites do intervalo de confiança em que o coeficiente é considerado como sendo estatisticamente igual a zero e que há um padrão de persistência na magnitude destes coeficientes. Para tanto, conforme apresentado na Tabela 2, foi realizado o teste de raiz unitária ADF para testar a hipótese de estacionariedade.

Tabela 2 – Teste ADF para a verificação de estacionariedade da série temporal

d	Valor do teste	p-valores dos parâmetros	Valor crítico do teste			Decisão do teste
			1%	5%	10%	
1	$\tau_\tau = -17,49$	$\alpha < 0,01$ $\beta t < 0,01$ $\gamma < 0,01$	1%	$-3,99 / \tau_\tau < \tau_c$	1%	$H_1 : \gamma \neq 0$
			5%	$-3,43 / \tau_\tau < \tau_c$	5%	$H_1 : \gamma \neq 0$
			10%	$-3,13 / \tau_\tau < \tau_c$	10%	$H_1 : \gamma \neq 0$

Fonte: Elaborada pelos autores

O teste ADF utilizado foi aquele que apresenta os componentes intercepto e tendência, pois a regressão realizada para o teste ADF apontou como significativa a utilização de tais parâmetros. O resultado do teste ADF com intercepto e tendência indica que a série é

estacionária em nível (para os níveis de significância de 1%, 5% e 10%), não sendo necessária, portanto, a realização de diferenciação na série para que esta se torne estacionária.

Quanto à sazonalidade, o modelo apresenta alguns picos significativos em determinados *lags*. Porém, não há um padrão definido possível de ser identificado no processo gerador da série temporal. Desta forma, a Tabela 3 apresenta alguns modelos identificados como possíveis para descrever satisfatoriamente o processo gerador. Além dos modelos identificados, são apresentados os *p-valores* dos coeficientes do modelo, os critérios AIC e BIC, o MAE e o MAPE do ajustamento e previsão do modelo.

Tabela 3 – Modelos ARIMA(p,d,q) identificados e estimados para a série temporal do produto

Modelo	<i>p-valores dos coeficientes</i>	AIC	BIC	MAE <sub>a</sub>	MAPE <sub>a</sub>	MAE <sub>p</sub>	MAPE <sub>p</sub>
ARIMA (3,0,0)	c: 1,56x10E-73 $\phi_1$ : 0,6949 $\phi_2$ : 0,0044 $\phi_3$ : 0,0018 $\phi_1$ : 0,0031	6.748	6.766	14.395	0,6523	13.563	0,5874
ARIMA (3,0,0)*	$\phi_2$ : 1,18x10E-13 $\phi_3$ : 1,63x10E-11	6.687	6.704	14.351	0,6263	14.159	0,5812
ARIMA (2,0,0)	$\phi_1$ : 5,08x10E-12 $\phi_2$ : 5,18x10E-23 $\phi_1$ : 0,0061	6.730	6.741	16.242	0,6285	15.928	0,6079
SARIMA (3,0,0)(1,0,0)12	$\phi_2$ : 3,06x10E-13 $\phi_3$ : 6,40x10E-11 $\Phi_1$ : 0,0283	6.693	6.707	15.401	0,6374	15.944	0,6636
SARIMA (3,0,0)(1,0,0)12	c: 1,45x10E-62 $\phi_1$ : 0,6079 $\phi_2$ : 0,0055 $\phi_3$ : 0,0029 $\Phi_1$ : 0,0296	6.745	6.767	14.383	0,6656	14.087	0,6481
SARIMA (2,0,0)(1,0,0)12	$\phi_1$ : 5,82x10E-10 $\phi_2$ : 2,41x10E-20 $\Phi_1$ : 0,0077	6.725	6.739	16.128	0,6293	15.210	0,5988

Fonte: Elaborada pelos autores

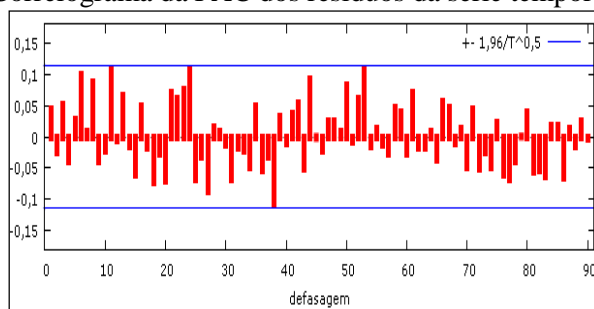
O modelo que melhor ajustou-se à série temporal do produto foi o modelo ARIMA(3,0,0), ou seja, um modelo autorregressivo descrito por três parâmetros e sem intercepto. Este modelo foi escolhido em função da significância estatística de seus



parâmetros (conforme seus *p-valores*) e em função deste ter sido aquele que minimizou os critérios AIC e BIC.

Para o atendimento aos pressupostos da metodologia, procedeu-se ao teste para a verificação da condição de autocorrelação nos resíduos gerados pelo modelo ARIMA(3,0,0). Para tanto, a Figura 3 apresenta o correlograma da FAC dos resíduos do modelo.

Figura 3 – Correlograma da FAC dos resíduos da série temporal do produto



Fonte: Elaborado pelos autores

Percebe-se pela Figura 3, que não há coeficientes de autocorrelação fora do intervalo de aceitação. Há sim, determinados coeficientes próximos ao limite máximo de aceitação (tanto no inferior quanto no superior). A existência de coeficientes fora do intervalo de aceitação é considerada uma situação indesejada na modelagem de séries temporais utilizando os modelos ARIMA(p,d,q), pois determinados padrões existentes na série não teriam sido capturados pelo modelo. Este teste comprovou que o modelo desenvolvido poderá ser utilizado.

A próxima etapa do estudo foi a modelagem da série temporal das vendas do produto utilizando os modelos de RNAs. O processo de modelagem e previsão foi realizado utilizando o *software* EasyNN, versão 14.0.g. A Tabela 4 apresenta os modelos de RNAs utilizados para a modelagem da série e o MAE<sub>a</sub>, o MAPE<sub>a</sub>, o MAE<sub>p</sub> e o MAPE<sub>p</sub>, que auxiliaram na escolha do modelo que melhor descreveu o processo gerador da série.

Tabela 4 – Modelagem da série temporal das vendas do produto utilizando RNAs

MLP	MAE <sub>a</sub>	MAPE <sub>a</sub>	MAE <sub>p</sub>	MAPE <sub>p</sub>
MLP (3,5,1)	13.831	0,4832	7.184	0,3146
MLP (3,10,1)	14.318	0,4692	6.178	0,2474
MLP (3,15,1)	14.884	0,5607	7.612	0,3312
MLP (3,20,1)	18.610	0,6950	8.630	0,3166
MLP (3,25,1)	15.787	0,5548	7.075	0,2787
MLP (3,30,1)	16.724	0,6571	9.048	0,4090
MLP (6,5,1)	15.505	0,6679	8.907	0,3732
MLP (6,10,1)*	7.523	0,2855	5.023	0,2233
MLP (6,15,1)	15.395	0,5409	9.602	0,4059
MLP (6,20,1)	9.131	0,3219	7.446	0,2986
MLP (6,25,1)	9.805	0,3518	5.492	0,2413
MLP (6,30,1)	8.240	0,3104	5.673	0,2770

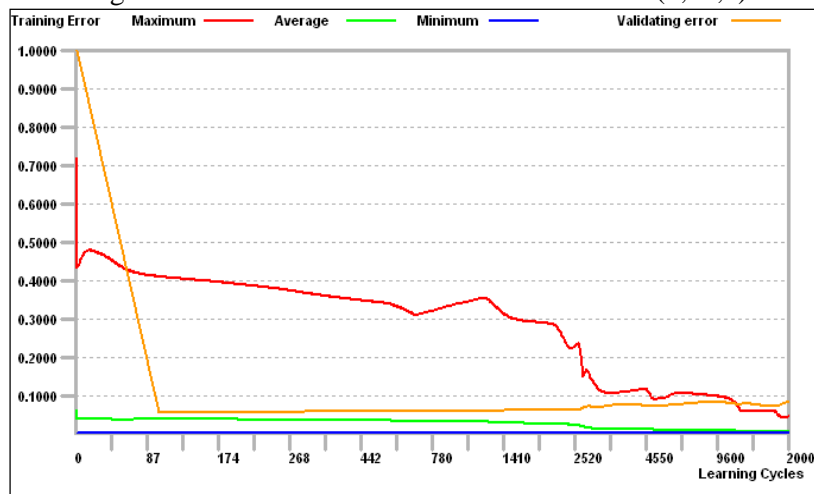
MLP	MAE <sub>a</sub>	MAPE <sub>a</sub>	MAE <sub>p</sub>	MAPE <sub>p</sub>
MLP (12,5,1)	10.140	0,4227	6.735	0,3053
MLP (12,10,1)	6.423	0,2377	5.383	0,2396
MLP (12,15,1)	9.562	0,3971	6.186	0,2804
MLP (12,20,1)	9.425	0,3488	6.965	0,2951
MLP (12,25,1)	8.287	0,3067	6.547	0,2889
MLP (12,30,1)	7.725	0,3328	7.121	0,3150

Fonte: Elaborada pelos autores

Verificou-se que o modelo de RNAs que melhor descreveu o processo gerador da série temporal foi o MLP (6,10,1), ou seja, o modelo com seis neurônios na camada de entrada, dez neurônios na camada intermediária e um neurônio na camada de saída. Observou-se que tal modelo não apresentou superajustamento aos dados de treinamento, pois a modelagem *out of sample* apresentou erros, em termos de MAPE, semelhantes.

A Figura 4 apresenta o treinamento da rede MLP (6,10,1). Para todos os modelos de RNAs utilizados neste estudo, foi estipulado o número máximo de 20.000 ciclos de treinamento, tendo como objetivo evitar o superajustamento aos dados de treinamento.

Figura 4 – Gráfico do treinamento da rede MLP (6,10,1)



Fonte: Elaborado pelos autores

Observou-se que o erro médio do processo de treinamento (0,0067) convergiu ao erro meta estipulado (0,01). O erro do processo de validação foi de 0,094 e o erro máximo do processo de treinamento foi de 0,067. Embora o erro máximo e o erro de validação não tenham atingido o critério de convergência (0,01) para os 20.000 ciclos de treinamento, verificou-se que o processo de treinamento da RNA atingiu valores aceitáveis de erro (em termos do erro médio).

## 4.2 Previsão das vendas do produto

A Tabela 5 apresenta os resultados de ajustamento ( $MAE_a$  e  $MAPE_a$ ) e previsão ( $MAE_p$  e  $MAPE_p$ ) para cada uma das categorias de modelos utilizados no estudo, considerando nestas apenas aqueles que obtiveram o melhor resultado em termos de ajustamento e previsão.

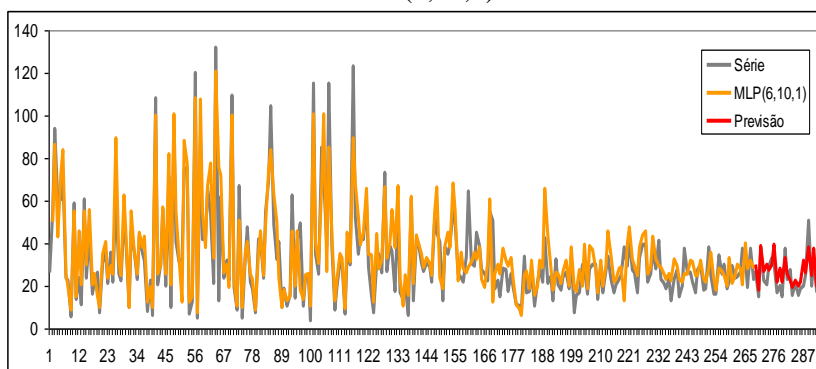
Tabela 5 – Comparação entre os modelos de previsão da série temporal de vendas

Tipos/Modelos		$MAE_a$	$MAPE_a$	$MAE_p$	$MAPE_p$
SEH	SEHWA e SEHWM	17.632	0,7568	18.552	0,7799
ARIMA	ARIMA (3,0,0)	14.351	0,6263	14.159	0,5812
RNAs*	MLP (6,10,1)	7.523	0,2855	5.023	0,2233

Fonte: Elaborada pelos autores

Observou-se que o modelo genérico de RNAs foi o que apresentou os melhores resultados em termos de ajustamento e de capacidade preditiva, sendo que o modelo MLP (6,10,1) foi o aquele que apresentou os melhores resultados dentre os demais modelos que compoem este tipo. A Figura 5 apresenta o gráfico do ajustamento do modelo à série, bem como a previsão realizada vinte e quatro períodos adiante.

Figura 5 – Gráfico do ajustamento e previsão da série temporal de vendas utilizando o modelo MLP(6,10,1)



Fonte: Elaborado pelos autores

Observou-se que o modelo MLP (6,10,1), de forma geral, descreveu adequadamente o processo gerador da série temporal. Essa constatação se dá, principalmente, devido à capacidade do modelo em capturar a heterocedasticidade presente na série temporal.

## 5 CONCLUSÕES

No presente estudo, foi modelada uma série temporal das vendas de um produto de uma empresa de laticínios, utilizando três modelos genéricos existentes na literatura que trata sobre a previsão de séries temporais. A justificativa da utilização destes modelos foi em virtude das

suas utilizações clássicas em modelagem e previsão de séries temporais, bem como devido ao comportamento linear e não linear geralmente presente nas séries temporais, onde se teve como objetivo investigar a eficácia em termos de previsão para cada um dos modelos.

Verificou-se que em termos de ajustamento à série, comparando os modelos de cada categoria que apresentaram os melhores resultados segundo os critérios adotados nas respectivas, os modelos de RNAs tiveram um desempenho 58% melhor em relação aos modelos de SE e aproximadamente 48% melhor em relação aos modelos ARIMA(p,d,q). Em termos de capacidade preditiva, os modelos de RNAs apresentaram um desempenho de aproximadamente 73% melhor em relação aos modelos de SE e aproximadamente 65% em relação aos modelos ARIMA(p,d,q). Dessa forma, constata-se a superioridade dos modelos de RNAs, para a série temporal das vendas modelada no presente estudo, em relação aos modelos de SE e ARIMA(p,d,q), tanto para a modelagem quanto para a previsão da série.

Uma possível forma de se melhorar a capacidade de previsão da série temporal estudada é por meio da combinação das previsões obtidas pelos modelos SE, ARIMA e MLP. A combinação de previsões tem como característica o aproveitamento das informações capturadas por cada modelo. Desta forma, seria possível aproveitar os resultados obtidos com a modelagem de cada modelo aqui utilizado, não descartando aquele que simplesmente apresenta um resultado global inferior ao melhor modelo.

Como sugestão para pesquisas futuras, sugere-se a utilização de técnicas de combinação de previsões para a previsão da série temporal. Poderia aqui ser analisado qual das técnicas de combinação disponíveis na literatura apresentaria o melhor desempenho.

## **COMPARATIVE STUDY OF TIME SERIES FORECASTING MODELS TO FORECAST ONE PRODUCT'S SALES**

**ABSTRACT:** This paper presents a comparative study among three generic time series models to forecast a determined product. The objective of this study was analyzing the efficacy throughout the forecasting capacity of each model, using a real sales time series. The time series models used in this research were: (i) exponential smoothing (ES); (ii) autoregressive integrated moving average (ARIMA); and, (iii) artificial neural network models. After fitted the series, the best model of each model category were selected and compared to each others. The results showed that artificial neural network multilayer perceptron (MLP) – MLP (6,10,1) is the best predictor to the time series used in relation to the adjustment and the forecasting mean absolute percentage error of 28,55% and 22,33%,

respectively. In terms of process modeling, the MLP (6,10,1) showed an improvement of 58% and 48%, comparing to the exponential smoothing and autoregressive integrated moving average models, respectively. In terms of forecasting capacity, the MLP (6,10,1) showed a result 73% and 65% better than the ES and ARIMA models, respectively.

**Keywords:** Sales forecasting. Exponential smoothing models. ARIMA models. Artificial neural network models.

## REFERÊNCIAS

ABURTO, L.; WEBER, R. Improved supply chain management based on hybrid. **Applied Soft Computing**, n. 7, p. 136-144, 2007.

ALLEMÃO, M.A.F. 2004. **Redes Neurais Aplicadas à Previsão de Demanda de Numerário em Agências Bancárias**. Dissertação (Mestrado em Ciência da Computação).. Programa de Pós Graduação em Ciência da Computação, Universidade Federal do Rio de Janeiro, 2004.

BOX, P.; JENKINS, G.M. **Time series analysis: forecasting and control**. San Francisco, CA: Holden-day Inc, 1976.

BRAGA, A.P.; CARVALHO, A.P.L.F.; LUDERMIR, T.B. **Redes Neurais Artificiais: teoria e aplicações**. 2ª ed. Rio de Janeiro: Ed. LTC – Livros Técnicos e Científicos, 2011, 226 p.

CHASE, R.B.; JACOBS, F.R.; AQUILANO, N.J. **Administração da produção para a vantagem competitiva**. 10 ed. Porto Alegre/RS: Ed. Bookman, 2006, 724 p.

ENDERS, W. **Applied Econometric Time Series**. 2 ed. United States of America: Ed. Wiley Series In Probability And Statistics, 2004, 466 p.

FARUK, D.O. A hybrid neural network and ARIMA model for water quality time series prediction. **Engineering Applications of Artificial Intelligence**, Turquia, n. 23, p.586-594, 2010.

FAVA, V.L. **Análise de Séries de Tempo**. In: Vasconcellos, M.A.S.; Alves, D. Manual de econometria: nível intermediário. São Paulo: Ed. Atlas, 2000.

FELICIANO, R.A. **Uma proposta de gerenciamento integrado da demanda e distribuição, utilizando sistemas de apoio à decisão (SAD) com business intelligence (BI)**. Dissertação (Mestrado em Engenharia). Universidade de São Paulo, 2009.

FIGUEREDO, C.J. **Previsão de séries temporais utilizando a metodologia Box & Jenkins e redes neurais para inicialização de planejamento e controle da produção**. Dissertação (Mestrado em Engenharia). Programa de Pós Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, 2008.

FLORES, J.H.F. **Comparação de modelos MLP/RNA e modelos Box-Jenkins em séries temporais não lineares**. Dissertação. Mestrado em Engenharia de

Produção, Escola de Engenharia, Universidade Federal do Rio Grande do Sul, 2009.

GUJARATI, D.M. **Econometria básica**. 3 Ed. São Paulo/SP: Ed. Makron books, 846p, 2000.

HAYKIN, S. **Redes Neurais: Princípios e práticas**. 2 ed. Porto Alegre: Ed. Artmed S. A., 1999, 903 p.

HO, S.L.; XIE, M.; GOH, T.N. A comparative study of neural network and Box-Jenkins ARIMA modeling in time series prediction. **Computers & Industrial Engineering**, n. 42, p. 371-375, 2002.

KHASHEI, M.; BIJARI, M. An artificial neural network ( $p,d,q$ ) model for timeseries forecasting. **Expert Systems with applications**, n. 37, p. 479-489, 2010.

KOVÁCS, Z.L. **Redes Neurais Artificiais: Fundamentos e aplicações**. 4 ed. São Paulo/SP: Ed. Livraria da Física, 2006, 177 p.

KUMAR, K.; JAIN, V.K. Autoregressive integrated moving averages (ARIMA) modelling of a traffic noise time series. **Applied Acoustics**, n. 58, p. 283-294, 1999.

LEE, Y.S.; TONG, L.I. Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. **Knowledge-Based Systems**, n. 24, p. 66-72, 2011.

MADDALA, G.S. **Introdução à Econometria**. 3 ed. Rio de Janeiro: Ed: LTC – Livros Técnicos e Científicos, 2003, 345 p.

MAKRIDAKIS, S.; WHEELWRIGHT, S.C.; HYNDMAN, R.J. **Forecasting: methods and applications**. 3ª ed. United States of America: Ed. Wiley & Sons, 1988, 642 p.

MARTINS, P.G., LAUGENI, F.P. **Administração da Produção**. 1. ed. São Paulo/SP: Ed. Saraiva, 2003.

MATOS, O.C.M. **Econometria básica: teoria e aplicações**. 3ª ed. São Paulo/SP: Ed. Atlas, 2000, 300 p.

MOORE, D.S.; McCABE, G.P.; DUCKWORTH, W.M.; SCLOVE, S.L. A prática da estatística empresarial: como usar dados para tomar decisões. 1ª Ed. Rio de Janeiro/RJ: Ed. LTC, 2006, 952 p.

MORETTIN, P.A. 2008. Econometria Financeira: Um curso em séries temporais financeiras. 1ª ed. São Paulo/SP: Ed: Blucher, 319 p.

MORETTIN, P.A.; TOLOI, C.M.C. Análise de séries temporais. 2. ed. São Paulo/SP: Ed. Edgard Blücher, 2006, 531 p.

PELLEGRINI, F.R. **Metodologia para implementação de sistemas de previsão de demanda**. Dissertação. Mestrado em Engenharia de Produção, Escola de Engenharia, Universidade Federal do Rio Grande do Sul, 2000.

SARTORIS, A. **Estatística e introdução à econometria**. 1ª ed. São Paulo/SP: Ed. Saraiva, 2003, 413 p.

SOBREIRO, V.A.; ARAÚJO, P.H.S.L.; NAGANO, M.S. Precificação do etanol utilizando técnicas de redes neurais artificiais. **Revista de Administração** da USP. v. 44, n. 1, p. 46-58, 2009.

TSENG, F.M.; YU, H.C.; TZENG, G.H. Combining neural network model with seasonal time series ARIMA model. **Technological Forecasting and Social Change**, n. 69, p.71-87 2002.

TUBINO, D.F. **Planejamento e controle da produção: teoria e prática**. 1ª ed. São Paulo/SP: Ed. Atlas, 2007, 190 p.

WANG, C.C. A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. **Expert Systems with Applications**, n. 38, p. 9296-9304, 2011.

WERNER, L.; RIBEIRO, J.L.D. Previsão de demanda: uma aplicação dos modelos Box-Jenkins na área de assistência técnica de computadores pessoais. **Revista Gestão & Produção**, v. 10, n. 1, p.47-67. 2003.

ZHANG, G.; PATUWO, B.E.; HU, M.Y. Forecasting with artificial neural networks: The state of the art. **International Journal of Forecasting**, n. 14, p. 35-62, 1998.

ZHANG, G.P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, n. 50, p. 159-175, 2003.

ZOU, H.F., XIA, G.P., YANG, F.T., WANG, H. Y. An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. **Neurocomputing**, n.70, p.2913-2926, 2007.

Originais recebidos em: 08/10/2013

Aceito para publicação em: 01/03/2014