

Prognóstico da série histórica do faturamento da indústria alimentícia brasileira

Forecast of the historical series of revenues of the Brazilian food industry using forecasting techniques

Matisa Andresa Maas^I, Cleber Bisognin ^{II}

Resumo

Esse trabalho tem como objetivo verificar qual é a melhor técnica de previsão, utilizando inclusive a combinação de previsões para avaliar o prognóstico do faturamento da indústria alimentícia brasileira. A série histórica do faturamento apresenta tendência determinística e sazonalidade. Visto isso, foram escolhidos para trabalhar-se os modelos: SARIMA (3,0,0)×(0,1,1)₁₂, SARIMA (4,0,0)×(2,0,0)₁₂ e Holt-Winters Multiplicativo. Analisando as medidas de acurácia, para realizar a previsão da série foi utilizada a combinação dos três modelos apresentados pelos métodos: Média Aritmética Simples, Mínimos Quadrados Ordinários e Regressão do Desvio Mínimo Absoluto. Os resultados obtidos pela previsão mostraram que o faturamento da Indústria Alimentícia Brasileira irá ter picos de crescimento e decréscimo nos próximos dois anos. Logo é necessária uma preparação no período em que ocorrerá um possível decréscimo nesse faturamento, bem como demissões dos trabalhadores, já que é o setor que mais emprega no Brasil.

Palavras-chave: Séries Temporais; Combinação de Previsão; Indústria Alimentícia.

Abstract

This paper's objective is to verify which is the best forecasting technique, including the use of the forecasts' combination to evaluate the prognosis of the Brazilian food industry's revenues. The historical series of revenues has deterministic trend and seasonality. Thereby, the models chosen to work on were: SARIMA (3,0,0)×(0,1,1)₁₂, SARIMA (4,0,0)×(2,0,0)₁₂ and Holt-Winters Multiplicative. Analyzing the accuracy measures, to perform the series' forecast it was used the combination of the three models, presented by the methods: Simple Arithmetic Mean, Ordinary Least Squares and Regression of Absolute Minimum Deviation. The results obtained by the forecast were satisfactory, showing that the Brazilian food industry's revenues will have peaks of growth and decay in the next two years. Therefore, a preparation of the sector is necessary for the period in which a possible decrease in this revenue will occur, as well as dismissal of the workers, since it is the sector that most employs in Brazil.

Keywords: Time Series; Forecast Combination; Food Industry.

^I Universidade Federal de Santa Maria, Santa Maria, Brasil. E-mail: maasmatisa@gmail.com.

^{II} Universidade Federal de Santa Maria, Santa Maria, Brasil. E-mail: cleber.bisognin@ufsm.br.



1 INTRODUÇÃO

A indústria de alimentos, segundo definição do Instituto Brasileiro de Geografia e Estatística (IBGE), compreende o processamento e transformação de produtos da agricultura, pecuária e pesca em alimentos para uso humano e animal (IBGE, 2007), sendo um dos setores mais dinâmicos e significativos na composição do Produto Interno Bruto (PIB) nacional, com sua produção voltada tanto para o mercado interno como para o mercado externo (SANTANA, 2017).

O Brasil está entre os grandes produtores de alimentos do mundo. A indústria alimentícia nacional é um importante segmento da atividade econômica do país, com grande dinamismo na produção, exportação e progresso técnico de sua cadeia produtiva. O Brasil reúne algumas características decisivas para o desempenho da indústria alimentícia: boas condições climáticas e disponibilidade de recursos naturais (que o deixam entre os maiores produtores mundiais de *commodities*, principais matérias-primas para a concepção de alimentos industrializados), além do grande mercado consumidor. Os mercados atendidos pela indústria alimentícia nacional apresentam muitas diferenças em termos econômicos e culturais, Europa e Oriente Médio, por exemplo, são os principais compradores (DA CUNHA, 2006). Em razão disso, grande parte das vendas externas do setor refere-se a preparações alimentícias, isto é, produtos com maior grau de valor adicionado e que atendem às exigências de cada segmento consumidor (DA CUNHA, 2006).

A indústria alimentícia brasileira é responsável por quase 15% do faturamento do setor industrial e por empregar mais de 1 milhão de pessoas (GOUVEIA, 2006), teve um crescimento de 2,08% em 2018, com previsão de crescimento de 2,5% a 3,0%. Sendo o setor, dentro da indústria brasileira, que mais empregou, de acordo com a Associação Brasileira da Indústria de Alimentos (ABIA, 2019), ainda com o intuito de acompanhar tendências mundiais, essa tem se preocupado com a saúde da população, trazendo temas como a redução do açúcar e nova proposta de rotulagem nacional (DUAS RODAS, 2019; AGÊNCIA BRASIL, 2019). O planejamento e a organização dos recursos são fundamentais para minimizar erros e atender às demandas dos consumidores (CECCATO, 2015). Nesse contexto, a previsão de faturamento é

importante para saber qual os resultados da indústria em períodos futuros. Com isso, esse trabalho tem como objetivo principal prever os futuros valores para a série do Faturamento da Indústria Alimentícia Brasileira, para isso foram ajustados modelos para gerar as previsões futuras e, no intuito, de “melhorar” as medidas de acurácia de previsão, utilizou-se-á métodos de combinação de previsões.

2 PROCESSOS ESTOCÁSTICOS E PREVISÃO

Para realizar a previsão utiliza-se métodos quantitativos que se baseiam em dados históricos para detectar padrões e então, prever os futuros valores das séries temporais, tais métodos empregam ferramentas matemático-estatístico. A análise de uma série temporal é composta por um conjunto de procedimentos de análise e tratamento dos dados, objetivando gerar estimativas futuras com maior precisão (ARAÚJO, 2018). Durante a análise da série temporal do Faturamento da Indústria Alimentícia Brasileira, foi constatado que a série temporal apresenta sazonalidade, e também foram aplicados testes de raiz unitária, para verificar tendência estocástica, e de Wald-Wolfowitz, para verificar tendência determinística.

Nas seções a seguir, são apresentados os modelos usados para realizar a análise e previsão da série temporal do Faturamento da Indústria Alimentícia Brasileira, a saber, os modelos $SARIMA(p,d,q) \times (P,D,Q)_s$ e Suavização Exponencial, com efeitos sazonais multiplicativos, (considerando que a série temporal apresenta tendência estocástica e sazonalidade) e os modelos $SARMA(p,q) \times (P,Q)_s$ (considerando que a série temporal apresenta tendência determinística e sazonalidade), além de quatro métodos para realizar combinações de previsões: média aritmética, variância mínima e por regressão pelo método dos mínimos quadrados e pelo erro médio absoluto.

2.1 Processos Estocásticos

A seguir são definidos os processos $SARIMA(p,d,q) \times (P,D,Q)_s$ proposto por Box; Jenkins (1976).

Definição 1: Seja $\{X_t\}_{t \in \mathbb{Z}}$ um processo estocástico satisfazendo a equação (1).

$$\Phi(B^s)\phi(B)(1 - B^s)^D(1 - B)^d(X_t - \mu) = \theta(B)\theta(B^s)\varepsilon_t, \quad (1)$$

onde μ é a média do processo, $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ é o processo ruído branco, $s \in \mathbb{N}$ é a sazonalidade, B é o operador defasagem ou de retardo, isto é, $B^j(X_t) = X_{t-j}$ e $B^{sj}(X_t) = X_{t-sj}$, para $j \in \mathbb{N}$, $\phi(\cdot)$, $\theta(\cdot)$, $\Phi(\cdot)$ e $\Theta(\cdot)$ são os polinômios de ordem p , q , P e Q , respectivamente, definidos pelas equações (2) e (3).

$$\phi(z) = \sum_{l=1}^p (-\phi_l)z^l, \quad \theta(z) = \sum_{m=1}^q (-\theta_m)z^m, \tag{2}$$

$$\Phi(z) = \sum_{r=1}^P (-\Phi_r)z^r, \quad \Theta(z) = \sum_{l=1}^Q (-\theta_l)z^l, \tag{3}$$

com $\phi_l, 1 \leq l \leq p$, $\theta_m, 1 \leq m \leq q$, $\Phi_r, 1 \leq r \leq P$ e $\theta_l, 1 \leq l \leq Q$, constantes reais e $\phi_0 = \Phi_0 = -1 = \theta_0 = \Theta_0$. Então $\{X_t\}_{t \in \mathbb{Z}}$ é um processo sazonal autorregressivo integrado de média móvel de ordem $(p,d,q) \times (P,D,Q)_s$ com sazonalidade s , denotado por $SARIMA(p,d,q) \times (P,D,Q)_s$, onde d e D são, respectivamente, o grau de diferenciação e o grau de diferenciação sazonal.

A equação (4) apresenta a previsão de erro quadrático médio mínimo de um processo $SARIMA(p,d,q) \times (P,D,Q)_s$ invertível,

$$\hat{X}_t(h) = \sum_{j \geq 0} \pi_j \hat{X}_t(j-h), \tag{4}$$

onde $h \geq 1$ e $\{\pi_j\}_{j \geq 0}$, são os coeficientes da representação auto regressiva infinita do processo. Percebe-se que $\hat{X}_t(h)$ pode ser expressa com uma soma ponderada das observações presentes e passadas $\{X_t\}_{t \leq n}$, onde n é o tamanho da série temporal.

Os modelos $SARMA(p,q) \times (P,Q)_s$ são um caso particular dos modelos $SARIMA(p,d,q) \times (P,D,Q)_s$, ou seja, satisfazem a equação (1), quando $D=0=d$. Os modelos de Suavização Exponencial descrevem dados em que pode se verificar a ocorrência de tendência linear, como também de sazonalidade (PELEGRINI & FLOGIATTO, 2000). As vantagens desse modelo são o entendimento fácil, aplicação não dispendiosa e adequada para séries de comportamento mais geral (MORETIN & TOLOI, 2004). No caso de séries temporais que apresentam componente sazonal, os modelos mais tradicionais são os Holt-Winters. Estes modelos podem ser definidos para efeitos

sazonais aditivos e multiplicativos. O modelo de Holt-Winters com efeitos sazonais multiplicativo é utilizado quando a amplitude da variação sazonal aumenta com o tempo, ou seja, a diferença entre o maior e o menor ponto de demanda nos ciclos cresce com o passar do tempo (ALBUQUERQUE & SERRA, 2006). A seguir são definidos os modelos de Holt-Winters para efeitos sazonais multiplicativos. Maiores detalhes podem ser encontrados em Makridakis et al. (1998) e em Morettin; Tolo (2006).

Considere uma série temporal $\{X_t\}_{t=1}^n$ com período s .

Definição 2: Os Modelos de Holt-Winters para efeitos sazonais multiplicativos, considera-se que tal série é formada por uma tendência aditiva, por um fator sazonal multiplicativo, como apresentado pela equação (5).

$$X_t = L_t \times S_t + T_t + \varepsilon_t, \text{ para } t = s + 1, \dots, n. \quad (5)$$

As estimativas do nível da série temporal no tempo t , denotado por L_t , da tendência, denotada por T_t e do fator sazonal, denotado por S_t , são dadas, respectivamente pelas equações de (6) a (8).

$$L_t = \alpha \frac{X_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (6)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (7)$$

$$S_t = \gamma \frac{X_t}{L_t} + (1 - \gamma)S_{t-s}, \quad (8)$$

onde α é o coeficiente de ponderação exponencial do nível ($0 \leq \alpha \leq 1$), β é o coeficiente de ponderação exponencial da tendência ($0 \leq \beta \leq 1$) e γ é o coeficiente de ponderação exponencial do fator sazonal ($0 \leq \gamma \leq 1$).

As previsões h passos à frente, são dadas conforme a equação (9).

$$\hat{X}_t(h) = (L_t + hT_t) \times S_{t-s+h}. \quad (9)$$

A notação $\hat{X}_t(h)$ indica a previsão de origem t e horizonte $h \geq 1$.

Na seção a seguir apresenta-se as metodologias de combinações de previsão que serão utilizadas neste trabalho.

2.2 Combinações de Previsão

A combinação é uma abordagem para a realização de previsões, visto que ao invés de tentar escolher a melhor técnica, formula-se o problema perguntando que técnicas poderiam ajudar na melhoria da acurácia. Como as previsões podem ser afetadas por diversos fatores, cada técnica pode contribuir capturando algum tipo de informação que influencia esses fatores; logo, um número maior de fatores poderá ser analisado através da combinação (CLEMEN, 1989).

De acordo com Werner (2005) há muitos estudos sobre combinação de previsões propostos na literatura. O foco prioritário desses estudos, até 1989, foi o estabelecimento de métodos de combinação que de fato melhorem a acurácia e reduzissem erros de previsão (MACKAY & METCALFE, 2002).

Rausser & Oliveira (1976) comentam que, dadas duas ou três previsões do mesmo evento, a prática usual é determinar qual a melhor delas através de alguma medida de acurácia. A melhor previsão é aproveitada e as outras, descartadas. Através desse descarte de previsões classificadas como inferiores, muitas vezes, alguma informação útil é perdida. Então, ao invés de escolher uma única técnica de previsão, parece razoável considerar informações provenientes de várias técnicas e combiná-las.

Nas seções a seguir, são apresentadas as metodologias de combinações de previsões.

2.2.1 Método da Média Aritmética Simples (MAS)

O método da média aritmética simples de combinação de previsões é considerado um estimador simples e de fácil utilização. De acordo com Clemen e Winkler (1986), não é necessário o conhecimento prévio da precisão ou do grau de dependência entre as previsões para a sua utilização. Como apenas previsões acuradas devem compor as combinações, grandes diferenças nos desvios padrão dos erros das previsões individuais não são esperadas. Nesses casos, é razoável pensar que a média simples pode apresentar um desempenho melhor que o obtido com a média ponderada (MENEZES; BUNN; TAYLOR, 2000). A média aritmética é um dos métodos mais populares para combinação de diferentes previsões (FLORES; WHITE, 1989).

A combinação de duas previsões por média aritmética simples, em sua apresentação, pode ser considerada uma simplificação do método da variância mínima, diferenciando-se por apresentar pesos fixos e iguais para cada previsão individual a ser combinada. O método da média aritmética, composto por duas previsões, é apresentado na Equação (10).

$$f_t^c = \frac{f_{1t} + \dots + f_{kt}}{k} \text{ para } t = 1, \dots, n; k = 2 \text{ ou } 3, \quad (10)$$

onde f_1, \dots, f_k , para $k = 2$ ou 3 , são as previsões individuais a serem combinadas (MARTINS & WERNER, 2014).

2.2.2 Método da Variância Mínima (VM)

Proposta inicialmente por Bates e Granger (1969) onde combina-se duas previsões objetivas, que devem ser não viesadas. Em 1974, Newbold e Granger ampliaram o número de técnicas combinadas, mantendo todas as suposições anteriores, passando de duas para k técnicas combinadas, com a restrição $\sum_{j=1}^k w_j = 1$. Este método satisfaz a equação (11).

$$f_t^c = w_1 f_{1t} + \dots + w_k f_{kt}, \text{ para } t = 1, \dots, n; k = 2 \text{ ou } 3, \quad (11)$$

onde: w_1, \dots, w_k são os pesos atribuídos à previsão f_k , para $k = 2$ ou 3 e f_1, \dots, f_k são as previsões individuais a serem combinadas. Para maiores detalhes de como estimar os pesos w_1, \dots, w_k , considerando $k \geq 3$, podem ser encontrados em Newbold e Granger (1974) e Werner (2005).

Considerando a combinação de duas previsões objetivas não viesadas, é atribuído o peso w para a primeira combinação e $(1-w)$ para a segunda. Portanto o método de combinação apresentado na equação (11), para $k = 2$ pode ser observada na equação (12).

$$f_t^c = w f_{1t} + (1 - w) f_{2t} \quad (12)$$

em que f_1 e f_2 são as previsões a serem combinadas.

Os pesos a serem atribuídos a cada previsão poderiam ser iguais, porém é desejável que se dê um peso maior para a previsão que possui os menores erros. Assim

objetivando encontrar o valor dos pesos, se propôs a minimização da variância dos erros da previsão combinada, que é dada pela equação (13).

$$\sigma_c^2 = w^2\sigma_1^2 + (1-w)^2\sigma_2^2 + 2\rho w\sigma_1(1-w)\sigma_2 \quad (13)$$

em que σ_1^2 e σ_2^2 são as variâncias dos erros das previsões a serem combinadas, ρ é o coeficiente de correlação entre os erros de previsão e w é o peso dado à previsão 1.

Para minimizar a variância, diferencia-se a equação (13) em relação a w e iguala-se o resultado a zero. O resultado é dado pela equação (14), que ficou conhecida, por minimizar a variância, como método da variância mínima.

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad (14)$$

2.2.3 Método da Regressão

Na concepção de Granger; Ramanathan (1984) a combinação de previsões poderia ser uma forma estruturada de regressão, conforme consta na equação (15).

$$f_t^c = \beta_0 + \beta_1 f_{1t} + \dots + \beta_k f_{kt}, \text{ para } t = 1, \dots, n; k = 2 \text{ ou } 3, \quad (15)$$

onde: β_0 é o intercepto da equação de regressão, β_1, \dots, β_k são os pesos das previsões, $k = 2$ ou 3 e f_1, \dots, f_k são as previsões individuais (geradas pelos modelos) a serem combinadas. Cabe destacar que para a estimação dos coeficientes de regressão $\beta_0, \beta_1, \dots, \beta_k$ é considerada com variável dependente a série temporal real e as variáveis independentes são as previsões individuais de cada modelo. A proposta para obter o intercepto e os pesos de cada previsão utilizada na combinação consiste em estimar os coeficientes da regressão pelo método dos Mínimos Quadrados Ordinários (MQO) e pelo método dos Desvios Mínimos Absolutos (LAD).

Definimos como resíduo do modelo a equação (16).

$$r_t = X_t - [\beta_0 + \beta_1 f_{1t} + \dots + \beta_k f_{kt}]. \quad (16)$$

O estimador de $(\beta_0, \beta_1, \dots, \beta_k)'$ pelo método dos Mínimos Quadrados Ordinários, denotado neste trabalho por MQO, é o valor de $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)' \in R^{k+1}$ tal que minimiza a função perda

$$l_1(n) = \sum_{t=1}^n r_t^2, \quad (17)$$

onde $\{r_t\}_{t=1}^n$ são dados pela equação (16).

Da mesma forma, o estimador de $(\beta_0, \beta_1, \dots, \beta_k)'$ pelo método dos Desvios Absolutos Mínimos, denotado neste trabalho por LAD, é o valor de $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)' \in R^{k+1}$ tal que minimiza a função perda

$$l_2(n) = \sum_{t=1}^n |r_t|, \quad (18)$$

onde $\{r_t\}_{t=1}^n$ são dados pela equação (16).

2.3 Medidas de Acurácia

De acordo com Souto et al. (2007) como objetivamos gerar previsões de valores futuros da série, é importante a utilização de alguns critérios de eficiência do modelo que considerem o grau de precisão das previsões futuras. Os critérios utilizados neste trabalho serão o Média dos Erros Percentuais Absolutos (MAPE), que é a medida de acurácia mais utilizada devido a vantagem de interpretação e independência na escala, Erro Médio Absoluto (MAE), que mede o valor médio de erro entre as séries observadas e ajustadas, Raiz Quadrada da Média dos Erros Quadráticos (RMSE) que representa as diferenças individuais quadráticas entre as séries temporais observadas e ajustadas e o coeficiente U de Theil que analisa a qualidade das previsões, sendo que, quanto mais próximo de zero significa que o erro de previsão, gerado por determinado modelo, é menor que da previsão ingênua.

3 METODOLOGIA

Os dados do Faturamento da Indústria Alimentícia Brasileira são oriundos do Portal da Indústria, mantido pela Confederação Nacional da Indústria (CNI) que pode

ser obtido pelo site: <http://www.portaldaindustria.com.br>. O faturamento real da indústria é definido como a soma dos ganhos obtidos pelas empresas com a venda de seus produtos medida em bilhões de Reais, já sendo descontado a inflação. O período a ser analisado é de janeiro de 2003 a maio de 2019. O uso desse período se deve ao fato de ser o mais completo disponível até a data de acesso, que foi em julho de 2019.

Na análise preliminar, os dados históricos são apresentados graficamente para poder verificar o comportamento da série. Foram aplicados os testes de raiz unitária de Dickey-Fuller e de Phillip-Perron, por meio das rotinas *adf.test* e *pp.test* do pacote *forecast*, no caso de utilizar o software R Core Team (2018) e o teste de Wald-Wolfowitz, para verificar tendência determinística, por meio da rotina *ww.test* do pacote *trend*. Nos testes de raiz unitária (Dickey-Fuller e de Phillip-Perron) a Hipótese Nula é que a série temporal apresenta raiz unitária e as Hipóteses Alternativas podem ser série temporal estacionária ou séries temporal explosiva.

Para o ajuste dos modelos foi utilizado o pacote *forecast* do software R Core Team (2018). A estimação dos parâmetros dos processos SARIMA $(p,d,q) \times (P,D,Q)_s$, foi realizada utilizando a rotina *auto.arima* que calcula a verossimilhança exata via representação de Estado de Espaço do modelo enquanto as inovações são encontradas via Filtro de Kalman. A estimação dos coeficientes dos polinômios dos processos, sob hipótese de estacionariedade, é baseada em Gardner *et al.* (1980). A seleção do modelo é feita pelos critérios AIC, AICc e BIC. Para estimar os parâmetros do processo Holt-Winters, utilizou-se a rotina *HoltWinters* com efeitos sazonais multiplicativos.

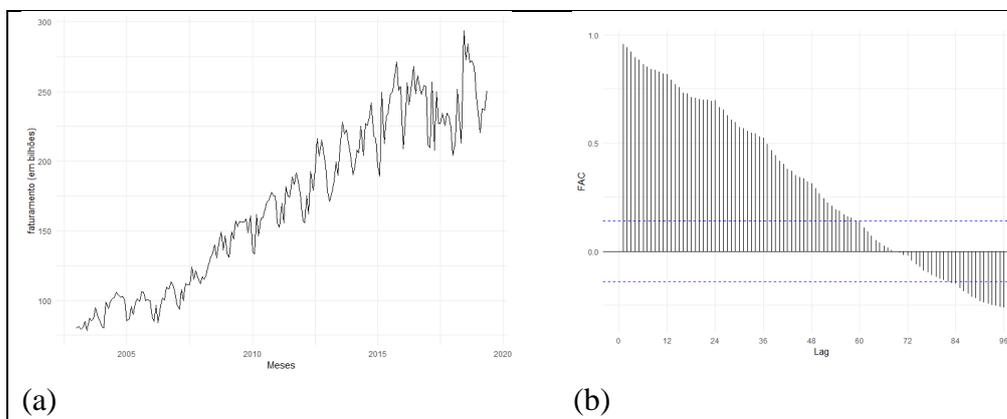
Uma vez que os modelos a serem utilizados foram definidos e os seus parâmetros foram estimados, utilizou-se as técnicas de combinação de previsão por variância mínima, média aritmética simples e regressão MQO, utilizando a função *Forecast_comb* do pacote *ForecastCombinations*. Já a combinação de previsão por regressão LAD foi utilizada a rotina *comb_LAD* do pacote *GeomComb*. A verificação da acurácia da técnica ou modelo a ser escolhido para realizar as predições, se dará por meio do RMSE, MAPE, MAE e pelo U de Theil. As medidas foram calculadas utilizando-se a rotina *accuracy* do pacote *forecast*. A elaboração dos gráficos desta análise, foi realizada utilizando os pacotes *ggplot2* e *ggfortify* do software R Core Team (2018).

4 RESULTADOS E DISCUSSÃO

4.1 Análise da Série Temporal

A Figura 1, apresenta o gráfico da série histórica do faturamento da Indústria Alimentícia Brasileira no período de janeiro de 2003 a maio de 2019 e, sua respectiva, função de autocorrelação amostral. Percebe-se que ao longo do tempo houve um crescimento muito significativo. No final do ano de 2017 até maio de 2018, a indústria alimentícia apresentou uma pequena queda no seu faturamento. Ainda, em fevereiro de 2019 teve um decréscimo, mas apresenta um crescimento nos meses que seguem. Pode-se perceber que há um comportamento ascendente, indicativo de tendência determinística, levando a necessidade de corrigir esse efeito para prosseguir com a previsão da mesma, ainda podemos verificar que a série parece apresentar sazonalidade.

Figura 1 – Faturamento da Indústria Alimentícia Brasileira no período de janeiro de 2003 a maio de 2019: (a) gráfico da série temporal; (b) função de autocorrelação amostral



Fonte: Os Autores.

Foram aplicados os testes de raiz unitária de Dickey-Fuller e de Phillip-Perron onde ambos os testes, apresentaram um p-valor menor do que o nível de significância de 5% [$<0,01$], o que indica que a série temporal não apresentava raiz unitária. Aplicou-se o teste de Wald-Wolfowitz para verificar se existe tendência determinística na série temporal, este tem como hipótese alternativa que existe tendência determinística na série, o p-valor [$< 0,00001$] do teste foi menor do que o nível de significância de 5%, o que indica que existe tendência determinística.

Por tratar-se de uma série temporal não estacionária e com sazonalidade e, como o objetivo deste trabalho é prever os seus futuros valores, iremos tratar a tendência desta série temporal como estocástica, utilizando os modelos SARIMA, e como determinística, a tendência é tratada através de um modelo linear e após é ajustado um modelo SARMA. Será ajustado ainda o modelo Holt Winters onde a componente sazonal será tratada como multiplicativa. Conforme Werner et al. (2017), foi ajustado o modelo de Holt-Winters com efeitos sazonais aditivos e multiplicativos. Analisando os resíduos de ambos os modelos, foi obtido que a soma dos quadrados das autocorrelações amostrais para 100 *lags* foi de 119,77, para o modelo Holt-Winters com efeitos sazonais aditivos, e 108,49, para o modelo com efeitos sazonais multiplicativos. Desta forma percebemos uma pequena vantagem em utilizarmos o modelo multiplicativo.

A seguir serão calculadas as previsões utilizando os modelos e suas respectivas medidas de acurácia. Observa-se então que a nossa série é composta por: sazonalidade e tendência.

A seguir são apresentados os modelos que foram ajustadas à Série Temporal.

Modelo 1 - Modelo SARIMA $(p,d,q) \times (P,D,Q)_s$. Considerando tendência estocástica, foi ajustado, como apresentado na Tabela 1, o modelo SARIMA $(3,0,0) \times (0,1,1)_{12}$

Tabela 1 – Resultado do ajuste do Modelo SARIMA $(3,0,0) \times (0,1,1)_{12}$

Parâmetros	Estimativa	Erro Padrão	Estat. Z	$P(> z)$
ϕ_1	0,294317	0,070625	4,1673	0,00003
ϕ_2	0,372610	0,068226	5,4614	0,00001
ϕ_3	0,308914	0,071027	4,3492	<0,00001
θ_1	-0,691843	0,070994	-9,7450	<0,00001

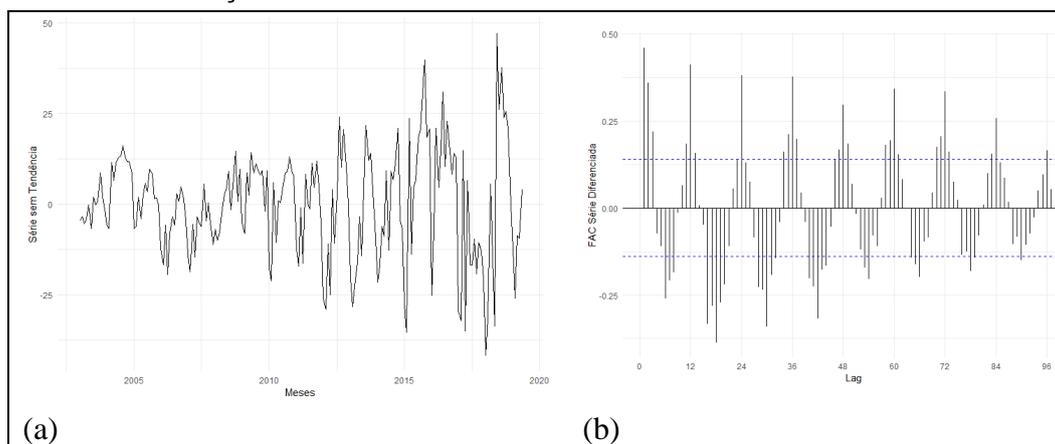
Log Verossimilhança: -698,41; AIC = 1406,81

Teste Resíduos (Box-Pierce): $df=15$ - p-valor=0,6424

Fonte: Os Autores.

Modelo 2 - Considerando tendência determinística, foi ajustado o modelo linear, onde $t=1, \dots, n$, que obteve um $R^2 = 0,9368$. As estimativas dos coeficientes e suas respectivas estatísticas são apresentadas na Tabela 2.

Figura 2 - Faturamento da Indústria Alimentícia Brasileira, entre janeiro/2003 a maio/2019 após retirada da tendência determinística: (a) gráfico da série temporal; (b) função de autocorrelação amostral



Fonte: Os Autores.

Tabela 2 – Resultado do ajuste do modelo para tendência determinística

Parâmetros	Estimativa	Erro Padrão	Estat. Z	P(> z)
β_0	8,468e+01	2,139e+00	39,58	<0,00001
β_1	1,348e-02	5,507e-04	24,48	<0,00001
β_2	-4,732e-05	2,931e-06	-16,14	<0,00001

Fonte: Os Autores.

A Figura 2 apresenta a série temporal sem tendência e sua respectiva função de autocorrelação amostral. Foram aplicados os testes de raiz unitária de Dickey-Fuller [p-valor<0,01] e de Phillips-Perron [p-valor<0,01], ou seja, a série temporal tornou-se estacionária, após retirada a tendência determinística.

A seguir, foi ajustado um modelo SARIMA $(4,0,0) \times (2,0,0)_{12}$, cujos coeficientes encontram-se na Tabela 3.

Tabela 3 – Resultado do ajuste do Modelo SARIMA $(4,0,0) \times (2,0,0)_{12}$

Parâmetros	Estimativa	Erro Padrão	Estat. Z	P(> z)
ϕ_1	0,310928	0,071150	4,3700	0,0000001
ϕ_2	0,287617	0,074269	3,8726	0,0001077
ϕ_3	0,284448	0,078647	3,6168	0,0002983
ϕ_4	-0,191458	0,072507	-2,6406	0,0082771
Φ_1	0,275709	0,081251	3,3933	0,0006905
Φ_2	0,451667	0,087651	5,1530	0,0000001

Log Verossimilhança: -736,75; AIC = 1487,51

Teste Resíduos (Box-Pierce): df=15 - p-valor=0,8902

Fonte: Os Autores

Modelo 3 - Modelo de Suavização Exponencial Holt-Winters (Multiplicativo): as estimativas para os parâmetros do modelo encontram-se na Tabela 4.

Tabela 4 – Resultado do ajuste do Modelo de Suavização Exponencial Holt-Winters (Multiplicativo)

Parâmetro	α	β	Γ
Estimativa	0,3611	0,0032	0,1942
Teste Resíduos (Box-Pierce): df=15 - p-valor=0,0995			

Fonte: Os Autores.

4.2 Análise das Previsões

Após estabelecer os modelos e estimar os seus parâmetros, passa-se a obter os pesos w_1, \dots, w_k pelo método de variância mínima e os coeficientes $(\beta_0, \beta_1, \dots, \beta_k)'$ das combinações pelos métodos de regressão. Nas Tabelas 5 e 6, encontram-se os pesos para a combinação dos modelos, dois a dois e a combinação de todos os três modelos.

Tabela 5 – Combinação de previsão por regressão: $\hat{\beta}_0, \dots, \hat{\beta}_k$, para $k = 2$ ou 3 , são os coeficientes pelo método de regressão, utilizando o método dos Mínimos Quadrados Ordinários (MQO) e Desvios Absolutos Mínimos (LAD)

Combinação	Modelos	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
LAD	1 e 2	1,8808	0,5787	0,4088	
	1 e 3	1,6411	0,4839	0,5050	
	2 e 3	0,7962	0,4904	0,5059	
	Todos	1,0988	0,0523	0,4837	0,4575
MQO	1 e 2	1,6809	0,5042	0,4892	
	1 e 3	2,2483	0,4748	0,5126	
	2 e 3	0,0560	0,4879	0,5096	
	Todos	-	-	0,5568	0,5557
		0,3411	0,1137		

Fonte: Os Autores.

Tabela 6 – Combinação de previsão: $\hat{w}_1, \dots, \hat{w}_k$, para $k = 2$ ou 3 , são os pesos atribuídos as previsões pelo método de variância mínima

Modelos	\hat{w}_1	\hat{w}_2	\hat{w}_3
1 e 2	0,4965	0,5035	
1 e 3	0,4965	0,5035	
2 e 3	0,5000	0,5000	
Todos	0,3302	0,3349	0,3349

Fonte: Os Autores.

Tabela 7 – Medidas de acurácia das previsões de cada modelo e combinação de previsão dos modelos ajustados a série temporal Faturamento da Indústria Alimentícia Brasileira no período de janeiro/2003 a maio/2019

Combinação	Modelos	RMSE	MAE	MAPE	U de Theil
Modelo	1	10,3009	7,2141	3,9670	0,0281
	2	10,2288	7,3956	4,1652	0,0279
	3	10,2290	7,0916	4,0739	0,0278
LAD	1 e 2	10,0845	7,0948	3,9304	0,0275
	1 e 3	9,9184	6,9078	3,8610	0,0270
	2 e 3	9,7039	6,8374	3,8592	0,0264
	Todos	9,7028	6,8367	3,8456	0,0264
MAS	1 e 2	10,0588	7,1078	3,9445	0,0274
	1 e 3	9,9389	6,9207	3,8692	0,0271
	2 e 3	9,6976	6,8417	3,8571	0,0264
MQO	Todos	9,7789	6,8710	3,8239	0,0266
	1 e 2	10,0387	7,1047	3,9661	0,0274
	1 e 3	9,9118	6,9169	3,8799	0,0270
VM	2 e 3	9,6887	6,8492	3,8552	0,0264
	Todos	9,6832	6,8666	3,8746	0,0264
	1 e 2	10,0585	7,1082	3,9451	0,0274
	1 e 3	9,9387	6,9208	3,8700	0,0271
	2 e 3	9,6976	6,8417	3,8571	0,0264
	Todos	9,7776	6,8705	3,8241	0,0266

Fonte: Os Autores.

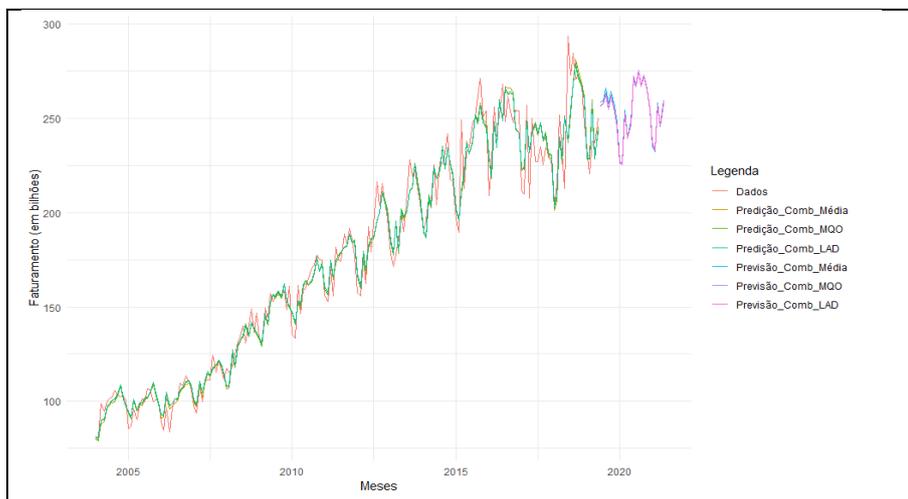
O método utilizado para a escolha do melhor modelo para realizar a previsão futura da nossa série foi a verificação das medidas de acurácia das previsões. Quanto menor for o valor da medida de acurácia, melhor é o modelo. Na Tabela 7 são apresentadas as medidas de acurácia das previsões calculadas utilizando-se os modelos e as quatro combinações de previsões.

Na Tabela 7, são apresentados, em *itálico e sublinhado*, os modelos e as combinações (dentro de cada metodologia) que apresentam menores medidas de acurácia. Assim, o Modelo 1 apresenta menor MAPE, o Modelo 2 apresenta menor RMSE e o Modelo 3 apresenta menores MAE e U de Theil. Dentre as combinações de previsão, pelo método de regressão LAD, a combinação das três previsões apresenta menores medidas de acurácia. Pelo método de regressão MQO, a combinação dos modelos 2 e 3 apresenta menores MAE e MAPE, enquanto a combinação dos três modelos apresenta menores RMSE e U de Theil. Pela metodologia de variância mínima e média aritmética simples, a combinação dos modelos 2 e 3 apresentam menores RMSE, MAE e U de Theil, enquanto a combinação dos três modelos apresenta menor MAPE.

Quando verificados as menores medidas de acurácia considerando todas as metodologias de previsão (modelos e combinações de previsão), exibidas em negrito, temos que: a combinação de todos os modelos pelo método MQO, Média e LAD devem ser considerados para a obtenção do resultado da previsão.

Após a conclusão de que os melhores resultados de previsão seriam pelas combinações de todos os modelos pelos métodos do MQO, Média e LAD, obteve-se os resultados da previsão para os próximos dois anos, evidenciados na Tabela 8. A Figura 3 apresenta as predições (dentro da amostra) e as previsões (fora da amostra), obtidas com base no último período observado, da série temporal Faturamento da Indústria Alimentícia Brasileira utilizando a combinação de previsões que apresentaram menores medidas de acurácia. Pode-se observar que estas combinações captam bem o comportamento dos dados da série. A previsão para os próximos dois anos terá seu valor mínimo em 225,42 bilhões de reais e seu valor máximo em 274,53 bilhões de reais. Terá um desvio padrão de 13,81 bilhões de reais no faturamento da indústria brasileira.

Figura 3 – Predição e Previsão do Faturamento da Indústria Alimentícia Brasileira, entre janeiro/2003 a maio/2021



Fonte: Os Autores.

Tabela 8 – Valores Previstos da série temporal do Faturamento da Indústria Alimentícia Brasileira no período de junho/2019 a maio/2020

Previsão			
Data	Média	MQO	LAD
jun/19	258,663	256,401	256,975
jul/19	259,653	257,809	258,231
ago/19	265,740	262,422	263,260
set/19	257,812	254,600	255,384
out/19	263,999	261,530	261,976
nov/19	256,520	253,509	254,256
dez/19	247,843	245,052	245,741
jan/20	227,225	226,090	226,434
fev/20	226,105	225,137	225,429
mar/20	254,247	252,473	252,747
abr/20	239,608	239,317	239,205
mai/20	248,148	245,891	246,141
jun/20	270,898	272,400	271,523
jul/20	266,930	268,056	267,212
ago/20	274,784	275,214	274,533
set/20	267,265	267,626	266,972
out/20	272,045	272,702	271,832
nov/20	265,028	265,308	264,704
dez/20	255,257	255,027	254,614
jan/21	235,380	237,405	236,483
fev/21	232,253	233,086	232,560
mar/21	258,278	256,918	256,784
abr/21	245,505	246,915	245,997
mai/21	257,631	259,498	258,124

Fonte: Os Autores.

5 CONCLUSÕES

Indiscutivelmente, as previsões fazem parte do planejamento do negócio das empresas e setores industriais. É no planejamento que se projeta os cursos de ações futuras com base em estimativas das necessidades de recursos. Estimar a demanda futura de bens e serviços é condição essencial para a elaboração de um plano de trabalho que inclui o dimensionamento das capacidades envolvidas com a definição de equipamentos, dos recursos financeiros, da disponibilidade de mão-de-obra e da quantidade de materiais necessários para a produção de bens e serviços (GONÇALVES, 2016).

Conclui-se com esse trabalho que, as técnicas escolhidas para prever a série temporal do Faturamento da Indústria Alimentícia foram, a combinação dos três modelos $SARIMA (3,0,0) \times (0,1,1)_{12}$, $SARIMA (4,0,0) \times (2,0,0)_{12}$ e *Holt-Winters* Multiplicativo, pelos métodos de Média, MQO e LAD. Pelos resultados apresentados pode-se perceber

que o faturamento para os próximos anos tende a ter picos de crescimentos e decréscimo, indicando sazonalidade, o que já acontecia nos anos anteriores. A sazonalidade é uma característica frequente na demanda por alimentos, sendo causada por variações climáticas, datas comemorativas, entre outros fatores. Com isso, é conveniente utilizar procedimentos para avaliar a sazonalidade e métodos de previsão que considerem o efeito das flutuações sazonais sobre a demanda, pois quanto mais informações se tiverem sobre o comportamento da demanda de um produto, mais acurada será a previsão e, conseqüentemente, as decisões baseadas nesta previsão (QUEIROZ, 2003).

Considerando que a indústria alimentícia brasileira é o setor industrial que mais emprega, é importante esse trabalho para que se tome atitudes referentes ao período em que irá ocorrer decréscimo no faturamento, evitando assim a demissão de trabalhadores. Ainda, para um próximo trabalho seria relevante verificar se quando ocorre esses decréscimos no faturamento, os indicadores econômicos do país modificam.

Referências

ABIA - ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE ALIMENTOS (Brasil). **Indústria De Alimentos Fecha 2018 com aumento de 2,08% em faturamento** [Internet]. Available from: <https://bit.ly/2DYutZI>. Acesso em: 9 dez. 2019.

AGÊNCIA BRASIL **Faturamento do setor de alimentos aumentou 2,08% no ano passado** [Internet]. Available from: <https://bit.ly/2N8KjqE>. Acesso em: 29 jul. 2019.

ALBUQUERQUE, Jean Carlos da Silva; SERRA, CMV. Utilização de modelos de Holt-Winters para a previsão de séries temporais de consumo de refrigerantes no Brasil. **XXVI ENEGEP, Fortaleza, Anais...** Rio de Janeiro: ABEPRO, 2006.

ARAÚJO, Gustavo Castro et al. Previsão de demanda e análise simplificada da gestão de estoque aplicadas a uma empresa do setor alimentício. **Brazilian Journal of Production Engineering-BJPE**, v. 4, n. 2, p. 48-64, 2018.

BATES, John M.; GRANGER, Clive WJ. **The combination of forecasts. Journal of the Operational Research Society**, v. 20, n. 4, p. 451-468, 1969.

CECATTO, Cristiano; BELFIORE, Patrícia. O uso de métodos de previsão de demanda nas indústrias alimentícias brasileiras. **Gestão & Produção**, v. 22, n. 2, p. 404-418, 2015.

CLEMEN, Robert T.; WINKLER, Robert L. Combining economic forecasts. **Journal of Business & Economic Statistics**, v. 4, n. 1, p. 39-46, 1986.

DA CUNHA, Denis Antonio; DIAS, Roberto Serpa; GOMES, Adriano Provezano. **Uma análise sistêmica da indústria alimentícia brasileira**. 2006.

DE MENEZES, Lilian M.; BUNN, Derek W.; TAYLOR, James W. Review of guidelines for the use of combined forecasts. **European Journal of Operational Research**, v. 120, n. 1, p. 190-204, 2000.

DUAS RODAS. **ABIA indica 4 prioridades para a indústria alimentícia em 2019** [Internet]. Available from: <https://bit.ly/2SYMzR1>. Acesso em: 29 jul. 2019.

FLORES, Benito E.; WHITE, Edna M. Subjective versus objective combining of forecasts: an experiment. **Journal of Forecasting**, v. 8, n. 3, p. 331-341, 1989.

GRANGER, Clive WJ; RAMANATHAN, Ramu. **Improved methods of combining forecasts. Journal of forecasting**, v. 3, n. 2, p. 197-204, 1984.

GONÇALVES, Paulo. **Administração de materiais**. Elsevier Brasil, 2016.

GOUVEIA, Flávia. Indústria de alimentos: no caminho da inovação e de novos produtos. **Inovação Uniemp**, v. 2, n. 5, p. 32-37, 2006.

IBGE. Instituto Brasileiro de Geografia e Estatística. **Classificação Nacional de Atividades Econômicas - versão 2.0**. Rio de Janeiro: IBGE, 2007.

MACKAY, Marisa Maio; METCALFE, Mike. Multiple method forecasts for discontinuous innovations. **Technological Forecasting and Social Change**, v. 69, n. 3, p. 221-232, 2002.

MORETINN, Pedro Alberto; TOLOI, Célia M. de Castro. **Análise se Séries Temporais**. São Paulo, ABE, 2004.

NEWBOLD, Paul; GRANGER, Clive WJ. Experience with forecasting univariate time series and the combination of forecasts. **Journal of the Royal Statistical Society: Series A (General)**, v. 137, n. 2, p. 131-146, 1974.

PELLEGRINI, Fernando R.; FOGLIATTO, F. Estudo comparativo entre modelos de Winters e de Box-Jenkins para a previsão de demanda sazonal. **Revista Produto & Produção**, v. 4, p. 72-85, 2000.

QUEIROZ, Abelardo A.; CAVALHEIRO, Darlene. Método de previsão de demanda e detecção de sazonalidade para o planejamento da produção de indústrias de alimentos. **Anais do Encontro Nacional de Engenharia de Produção**, v. 23, 2003.

R CODE TEAM (2018). R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. URL <https://www.R-project.org/>.

RAUSSER, Gordon C.; OLIVEIRA, Ronald A. An econometric analysis of wilderness area use. **Journal of the American Statistical Association**, v. 71, n. 354, p. 276-285, 1976.

SANTANA, Daniel Filipe da et al. **Indústria de alimentos e bebidas no Estado da Paraíba: contribuição socioeconômica e comportamento ambiental sob a perspectiva empresarial**. 2017.

SOUTO, Eduardo et al. Uma Abordagem Simples para Previsão do Consumo de Energia em Redes de Sensores sem Fio. **23º. Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, Fortaleza**, v. 2, p. 1139-1142, 2005.

WERNER, Liane. **Um modelo composto para realizar previsão de demanda através da integração da combinação de previsões e do ajuste baseado na opinião**. 2005.

WERNER, Liane; BISOGNIN, Cleber; Menegotto, Letícia. Previsão da umidade Relativa do Ar de Porto Alegre. **Cadernos do IME – Série Estatística**. Rio de Janeiro, v. 43, p. 39-59, 2017.