

ANÁLISE DO CONSUMO DE ENERGIA ELÉTRICA NO SETOR INDUSTRIAL BRASILEIRO: UMA APLICAÇÃO DOS MODELOS DE ALISAMENTO EXPONENCIAL E BOX & JENKINS

Tuany Esthefany Barcellos de Carvalho Silva (PONTIFÍCIA UNIVERSIDADE CATOLICA, RJ) tuanybarcellos@id.uff.br

Daiane Rodrigues dos Santos (UNIVERSIDADE CÂNDIDO MENDES) daianesantoseco@gmail.com

Daiane de Souza Oliveira (PONTIFÍCIA UNIVERSIDADE CATOLICA, RJ) daiane.dso@gmail.com

Resumo

O consumo de energia elétrica industrial é considerado fundamental para o sistema econômico, este consumo vem crescendo de forma significativa no Brasil nos últimos 15 anos, tornando-se um indicador de crescimento econômico. Isto posto, se torna imprescindível a aplicação de técnicas capazes de gerar boas previsões, a fim de auxiliar na análise de séries referentes ao consumo de energia elétrica industrial. Este trabalho teve como objetivo modelar a série supracitada e gerar previsões do consumo elétrico industrial, através de técnicas de modelagem como Alisamento Exponencial, Box & Jenkins e Combinação de ambos. Para verificar a capacidade preditiva dos modelos utilizou-se o critério MAPE que apontou a melhor aderência do modelo Box & Jenkins. O trabalho também abordou o impacto da covid19 no setor de consumo de energia elétrica industrial, evidenciando que a queda no consumo ocasionada por esta pandemia. Os modelos se ajustaram bem aos dados, e os resultados obtidos foram satisfatórios.

Palavras-Chaves: Amortecimento Exponencial; Box & Jenkins; Previsão

1. Introdução

Nos períodos que antecedem a revolução industrial, o uso de máquinas a vapor pelas pequenas manufaturas era algo muito comum, pois a eletricidade apresentava alto custo e suas redes eram pequenas, o que tornava o consumo de energia pouco viável para a maioria da população. Os anos seguintes foram marcados por avanços tecnológicos, fazendo com que a energia elétrica se tornasse cada vez mais acessível e confiável, as

usinas passaram a produzir geradores maiores e as novas tecnologias fizeram com que as redes crescessem. Todo este avanço tornou o consumo de energia elétrica mais acessível para todos os setores, e aos poucos a eletricidade passou a dominar e impulsionar diversos ramos como o comercial e o industrial (MACHADO, 2010). Atualmente a energia elétrica ou térmica, é considerada insumo fundamental para setor econômico, seu consumo vem crescendo de forma significativa no Brasil nos últimos 15 anos, principalmente após a grande expansão do setor industrial, que representa cerca de 34% do consumo final (SILVA, 2018).

No ano de 2020, mais precisamente a partir do mês de março o Brasil foi acometido pelo novo corona vírus (Covid19), o rápido contágio ocasionou uma pandemia, gerando impactos em diversos ramos, inclusive no setor de energia elétrica. Buscando minimizar o contágio, algumas medidas de prevenção foram aplicadas, como distanciamento social, fechamento de comércios e indústrias, objetivando manter a maior parte da população em casa, sendo estes os principais fatores responsáveis pela alteração do consumo de energia.

A busca por conseguir entender o histórico e o comportamento de uma variável, torna de extrema importância o estudo de séries temporais, o mesmo possibilita identificar padrões não aleatórios na série de uma variável de interesse, e a análise do comportamento passado permite com que sejam feitas previsões que auxiliam na tomada de decisão. Este trabalho implementou a análise de séries temporais, buscando observar o consumo de energia elétrica industrial no Brasil. Para o ajuste da série, foram utilizados modelos de amortecimento exponencial e *Box&Jenkins*.

2. Referencial teórico

2.1. Modelos de Amortecimento Exponencial

Os modelos de suavização exponencial fazem parte de uma grande classe de métodos de previsão baseados na ideia de que observações passadas contêm informações significativas sobre o padrão da série temporal estudada. O método consiste em distinguir um padrão de comportamento de qualquer outro ruído que possa estar contido nas observações da série, utilizando esse padrão para prever valores futuros.

- **Modelo Naive**

O modelo *Naive* é simples e ingênuo para a previsão de séries temporais, ele utiliza para previsão o último dado disponível. Este método é aquele em que $\hat{a}_1(T)$ é obtido para todo T como:

$$\hat{a}_1(T) = Z_t \quad (1)$$

Onde $\hat{a}_1(T)$ é o estimador do parâmetro do modelo $a_1(T)$ no instante T , e Z_T é a série temporal no instante T .

Como neste método a previsão é dada pelo último dado disponível, para a estimação de a_1 , utiliza-se a previsão τ -passos-à-frente, para todo $\tau = 1, 2, \dots$, ou seja, igual ao último valor observado (SOUZA, 2015).

$$\hat{Z}_t(\tau) = Z_t; \tau = 1, 2, \dots \quad (2)$$

- **Modelo Holt**

O modelo *Holt*, é em parte semelhante com o *Naive*, pois os dois não consideram a sazonalidade da série. Este método apresenta a série com a formulação de HOLT (1957), também conhecida por amortecimento exponencial duplo considerando a existência de tendência na série, sua equação de previsão deslocada é dada por,

$$\hat{Z}_t(\tau) = \hat{a}_1(t) + \tau \hat{a}_2(t) \quad (3)$$

O grande diferencial do método *Holt* está na atualização dos parâmetros, pois o mesmo usa diferentes constantes de amortecimento (hiperparâmetros) em cada parâmetro (nível de tendência). Para esse caso, os estimadores de $\hat{a}_1(t)$ e $\hat{a}_2(t)$, são calculados via amortecimento exponencial com constantes de amortecimento, α e β (SOUZA, 2015). Esta expressão é dada por,

$$\begin{aligned} \hat{a}_1(t) &= \alpha Z_t + (1 - \alpha)[\hat{a}_1(t - 1) + \hat{a}_2(t - 1)] \\ \hat{a}_2(t) &= \beta(\hat{a}_1(t) - \hat{a}_1(t - 1)) + (1 - \beta)\hat{a}_2(t - 1) \end{aligned} \quad (4)$$

Onde \hat{a}_1 representa a componente de nível, \hat{a}_2 a componente de tendência, $\alpha(0 < \alpha < 1)$ é a constante de suavização da componente de nível de \hat{a}_1 e $\beta(0 < \beta < 1)$ é a constante de suavização da componente de tendência \hat{a}_2 .

- **Modelo Holt-Winters aditivo e multiplicativo**

Holt (1957) e Winters (1960), desenvolveram modelos mais sofisticados, criando o método de amortecimento exponencial que consiste na modelagem de séries com tendência e sazonalidade.

O método Holt-Winters aborda dois tipos de modelos, o aditivo e o multiplicativo baseados em três equações alisadoras, sendo elas para o nível, tendência e sazonalidade. A suavização se dá por um método estatístico iniciado por uma equação de médias móveis ponderadas exponencialmente, com o principal objetivo de produzir ajustes nas variações aleatórias de uma determinada série temporal (MORETTIN; TOLOI, 2006).

O modelo Holt-Winters aditivo é adequado para aquelas séries em que a variância é constante. As equações de atualização dos parâmetros e de previsão deste modelo com sazonalidade aditiva se dá por,

Equação do nível,

$$a(t) = \alpha(Z_t - S_{t-12}) + (1 - \alpha).(a(t - 1) + b(t - 1)) \quad ; \quad 0 \leq \alpha \leq 1. \quad (5)$$

Equação de tendência,

$$b(t) = \beta(a(t) - a(t - 1)) + (1 - \beta)b(t - 1) \quad ; \quad 0 \leq \beta \leq 1. \quad (6)$$

Equação de Sazonalidade,

$$S_t = \gamma(Z_t - a(t)) + (1 - \gamma)S_{t-12} \quad ; \quad 0 \leq \gamma \leq 1. \quad (7)$$

Equação de previsão,

$$\hat{Z}(h) = a(t) + b(t)h + S_{t-12+h} \quad (8)$$

No modelo Holt-Winters multiplicativo a variância cresce com o nível da série observada. As equações de atualização dos parâmetros e de previsão deste modelo com sazonalidade multiplicativa se dá por,

Equação de nível,

$$a(t) = \alpha \left[\frac{Z_t}{S_{t-12}} \right] + (1 - \alpha) \cdot (a(t - 1) + b(t - 1)) \quad ; \quad 0 \leq \alpha \leq 1. \quad (9)$$

Equação de tendência,

$$b(t) = \beta(a(t) - a(t - 1)) + (1 - \beta)b(t - 1) \quad ; \quad 0 \leq \beta \leq 1. \quad (10)$$

Equação de Sazonalidade,

$$S_t = \gamma \frac{Z_t}{a(t)} (1 - \gamma) S_{t-12} \quad ; \quad 0 \leq \gamma \leq 1.$$

(11)

Equação de previsão,

$$\hat{Z}(h) = (a(t) + b(t)h) S_{T-12+h}$$

(12)

Para as equações de atualização dos parâmetros e das previsões dos modelos *Holt-Winters* aditivo e multiplicativo temos que S corresponde ao comprimento da sazonalidade, $a(t)$ nível da série, $b(t)$ tendência, S_t componente sazonal, Z_h previsão h passos à frente, Z_t valor observado, (α, β, γ) são os parâmetros de amortecimento do nível, da tendência e da sazonalidade respectivamente.

2.2 Modelos Box & Jenkins

A metodologia Box & Jenkins (1970) trata-se de uma abordagem muito utilizada na construção de modelos paramétricos para séries temporais univariadas. Este método consiste em ajustar modelos lineares estacionários ou não, a uma série de tempo observada baseando-se em um ciclo iterativo, onde a estrutura do modelo depende dos próprios dados.

Os modelos de Box & Jenkins, também conhecidos genericamente como modelos da classe ARIMA (*Auto Regressive Integrated Moving Averages*), esses modelos consistem em captar e observar o comportamento da correlação e autocorrelação seriada, baseando-

se nesse comportamento observado para assim realizar as previsões. Analisa-se as correlações e autocorrelações a fim de obter uma boa modelagem, o que resultara em uma previsão. Após observar os dados, se os mesmos não forem estacionários, pode-se torná-lo estacionário por meio de sucessivas diferenciações na série original (SOUZA,2004).

A modelagem proposta por Box & Jenkins (1970) é da forma,

$$Z_t = \mu + \sum_{k=0}^{\infty} \Psi_k u_{t-k} = \mu + \Psi(B)u_t \quad (13)$$

Sendo o filtro linear Ψ é definido por,

$$\Psi(B) = \frac{\theta(B)}{\phi(B)} \quad (14)$$

Onde, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ e $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ são polinômios de graus p e q , respectivamente. Sendo $\tilde{Z}_t = Z_t - \mu$, obtém-se $\tilde{Y}_t = \Psi(B)u_t$. Logo os modelos Box & Jenkins são dados pela equação,

$$\phi(B)\tilde{Z}_t = \theta(B)u_t \quad (15)$$

u_t é definido como ruído branco. Segundo Box & Jenkins (1970) este modelo é denominado ARMA (p, q), podendo ser reescrito como,

$$\tilde{Z}_t = \theta(B)\phi^{-1}(B)u_t \quad (16)$$

Assumindo sem perda de generalidade que $\mu = 0$.

Os modelos Box & Jenkins abrangem a grande classe ARIMA, porém no presente trabalho devido ao comportamento da série estudada, optou-se por utilizar o modelo SARIMA, o mesmo faz parte da classe dos modelos propostos por Box & Jenkins.

No modelo SARIMA (p, d, q)(P, D, Q) $_s$ em prática as séries observadas raramente são de processos estacionários, muita das vezes as mesmas além de não estacionárias apresentam componentes sazonais. Isto posto, Box & Jenkins (1970) formularam modelos para séries sazonais originando a classe ARIMA sazonal multiplicativo, denominado por SARIMA(p, d, q)(P, D, Q) $_s$, onde s consiste no número de fatores sazonais. O modelo é representado por,

$$\phi(B)\Phi(B_s)W_t = \theta(B)\Theta(B)u_t \quad (17)$$

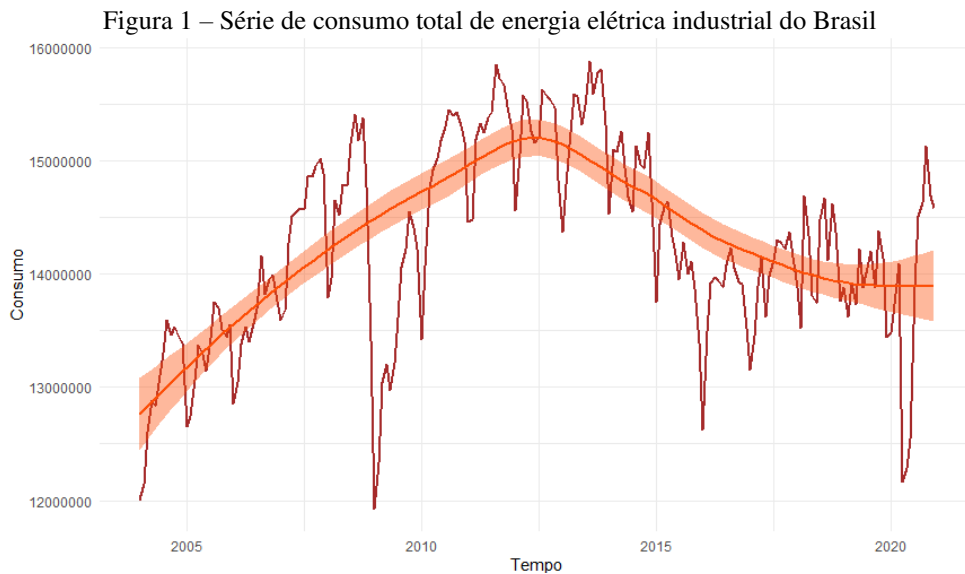
Sendo, $\phi(B) = 1 - \phi_1(B) - \phi_2(B)^2 - \dots - \phi_p B^p$ e $\Phi(B) = 1 - \Phi_1(B_s) - \Phi_2(B)^2 - \dots - \Phi_P B^P$, $W_t = \Delta^d \Delta_S^D Z_t$, $\theta(B) = 1 - \theta_1(B) - \theta_2(B)^2 - \dots - \theta_q B^q$ e $\Theta(B) = 1 - \Theta_1(B) - \Theta_2(B)^2 - \dots - \Theta_Q B^Q$. Onde $\phi(B)$ é o operador autorregressivo AR(p), $\theta(B)$ corresponde ao operador de médias moveis MA(q), $\Phi(B)$ é o operador AR da parte sazonal de ordem P , $\Theta(B)$ o operador MA da parte sazonal de ordem Q , já o Δ^d consiste no operador de diferenças, Δ_S^D é o operador de diferenças da parte sazonal e u_t representa o ruído branco.

- **Ruído Branco**

O ruído branco é uma série de tempo com algumas propriedades especiais, uma série pode ser chamada de ruído branco se for uma sequência de variáveis aleatórias independentes e identicamente distribuídas, possuir média zero, variância constante e finita. Se a série apresentar uma distribuição normal com média 0 e variância σ^2 , pode-se dizer que a mesma é um ruído branco gaussiano.

3. Análise dos resultados

Este estudo aborda a análise de uma série temporal sobre o consumo de energia industrial total do Brasil no período de janeiro de 2004 a dezembro de 2020, para a modelagem da série serão utilizados os métodos de amortecimento exponencial e Box & Jenkins respectivamente. Na figura 1 pode-se observar o comportamento da série, nota-se que a mesma atinge seu ponto máximo no ano de 2013, já seu ponto mínimo foi em janeiro de 2009, a tabela 1 apresenta as estatísticas descritivas da série para melhor visualização dos valores de mínimo, máximo e média dos dados analisados.



Fonte: Elaboração própria com base nos dados analisados

Tabela 1 – Estatística descritiva – série consumo de energia

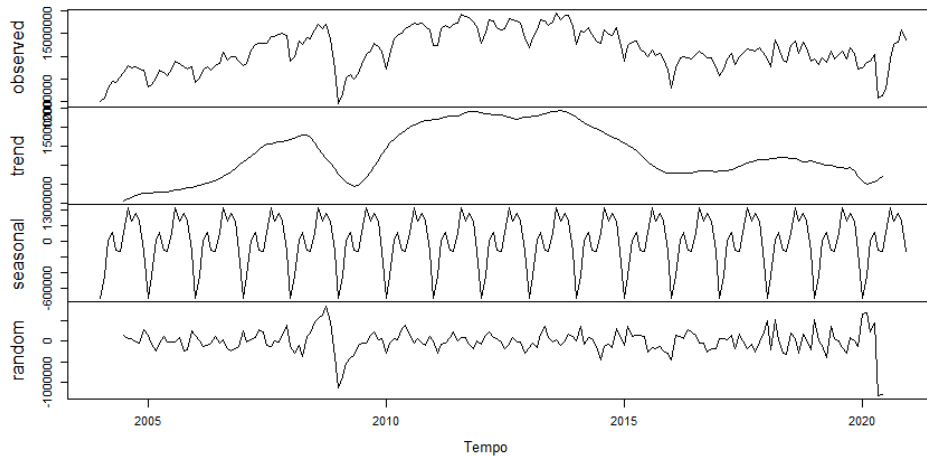
Estatística descritiva	Consumo
Mínimo	11.923.920
Média	14.284.502
Máximo	15.886.314

Fonte: Elaboração própria com base nos dados analisados.

A fim de inferir sobre algumas componentes dos dados, foi feita a decomposição da série, na figura 2 pode-se observar as componentes separadamente, nota-se que a série apresenta pequeno indício de sazonalidade, já na componente de tendência não fica claro se a série

possui características de uma série com tendência, logo faz-se necessário a utilização de testes estatísticos que complementem a análise gráfica.

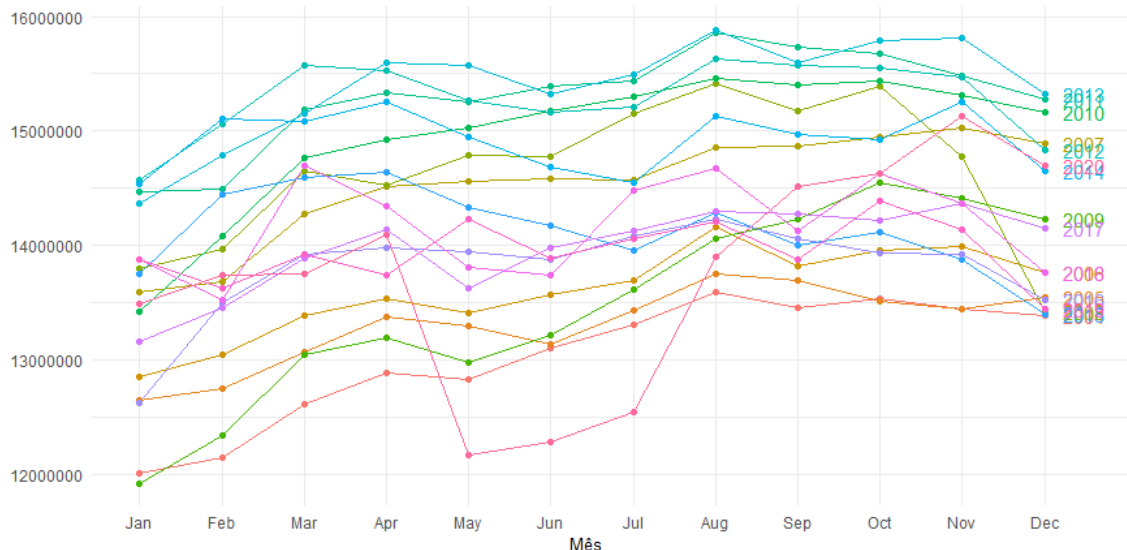
Figura 2 – Decomposição da série



Fonte: Elaboração própria com base nos dados analisados

Para verificar a componente de tendência aplicou-se o teste de *Mann-Kendall*, obtendo o p-valor de 0,04152, logo ao nível de significância de 5% rejeita-se a hipótese nula, ou seja, a série apresenta uma tendência. Na figura 3 pode-se observar a tendência da série de forma mensal, nota-se que a maioria dos anos observados exibem comportamento crescente, existe uma queda significativa em abril de 2020, apresentando comportamento crescente após o mês de julho. Um comportamento comum da série nos anos observado é a queda do consumo nos últimos meses do ano.

Figura 3- Decomposição em meses

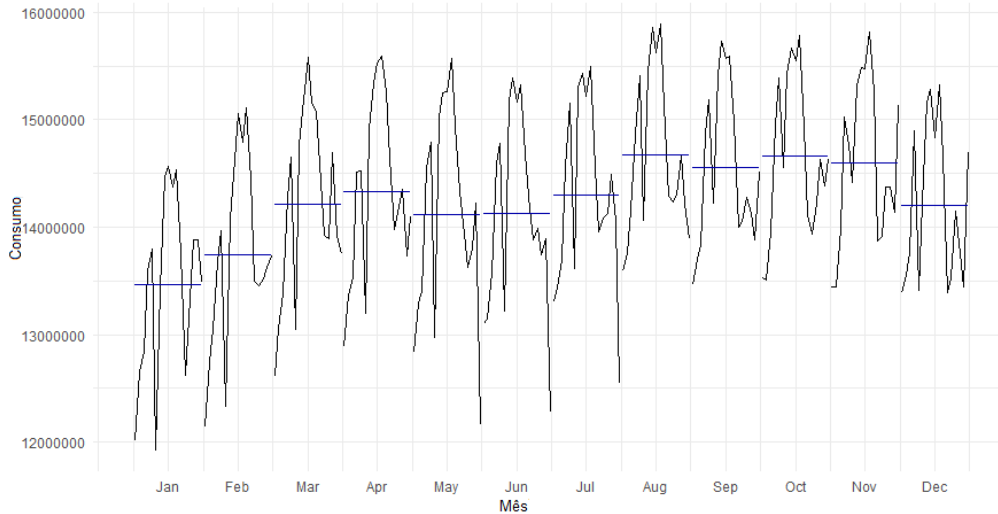


Fonte: Elaboração própria com base nos dados analisados.

Objetivando verificar a sazonalidade da série, aplicando o teste “*WO*” a função retornou um booleano *TRUE*, ou seja, segundo o teste os dados observados apresentam comportamento sazonal, a fim de testar as hipóteses de sazonalidade, obteve-se um p-valor muito pequeno de aproximadamente $6e-12$, logo ao nível de significância de 5%

rejeita-se a hipótese nula, o que indica que há indícios de sazonalidade na série. Na figura 4 pode-se notar o comportamento sazonal da série observada.

Figura 4 - Sazonalidade



Fonte: Elaboração própria com base nos dados analisados.

3.1 Implementação do método de alisamento exponencial

Na elaboração dos modelos de previsão com o método de alisamento exponencial, utilizou-se para os ajustes o período de janeiro de 2004 a dezembro de 2019, e para comparação foi utilizado 12 passos à frente, os modelos foram o *Naive*, *Holt*, *Holt-Winters* aditivo e multiplicativo. Na tabela 2 pode-se observar os valores dos hiperparâmetros de amortecimento e os coeficientes (a e b) de cada modelo ajustado.

Tabela 2 – Ajustes do modelo

Modelos	α	β	Γ	Coefficiente a	Coefficiente b
<i>Naive</i>	0.9999286	False	False	13627515	-
<i>Holt</i>	1	0.01861907	False	13627497.000	-7383.736
<i>Holt-Winters (adit)</i>	0.7977118	0	0.7612603	14191377.312	20911.915
<i>Holt-Winters (mult)</i>	0.7946649	0	0.7523775	14196837.212	20911.915

Fonte: Elaboração própria com base nos dados analisados

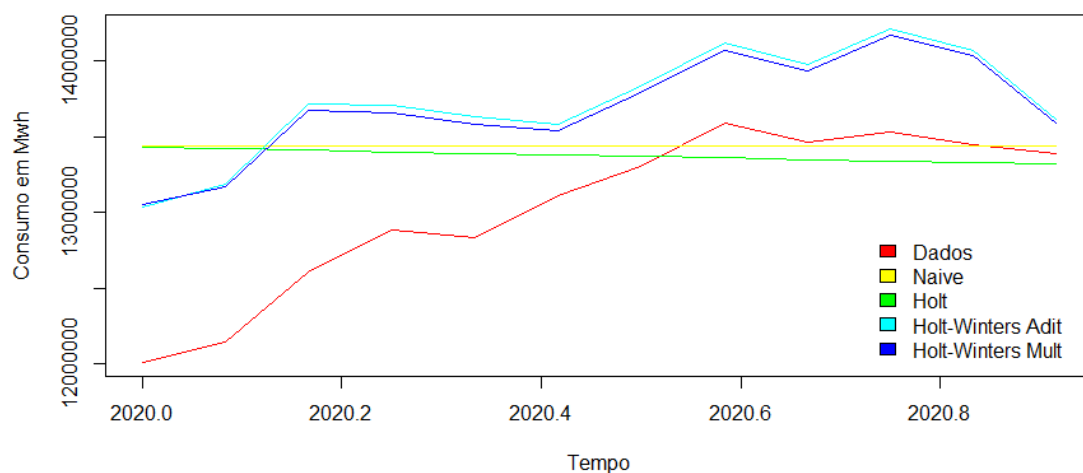
Holt-Winters aditivo e multiplicativo são modelos de amortecimento exponencial com tendência e sazonalidade, os mesmos possuem três hiperparâmetros (α, β, Γ), 1 coeficiente de nível, 1 de tendência e 12 coeficientes de sazonalidade. A tabela 3 apresenta as previsões 12 passos à frente, para melhor visualização dos modelos pode-se observar a figura 5, nota-se graficamente que os modelo *Holt* se ajustaram um pouco melhor aos dados, principalmente nos últimos meses da série.

Tabela 3 – Previsões 12 passos à frente por alisamento exponencial

Mês	Dados	Naive	Holt	Holt-Winters (ad it)	Holt-Winters (mult)
Jan/2020	13.447.777	13.7671.64	13.546.276	12.760.747	13.767.164
Fev/2020	13.391.884	13.889.504	13.538.892	13.899.012	13.889.504
Mar/2020	14.098.067	13.440.852	13.410.052	13.719.160	13.669.682
Abr/2020	12.167.393	13.440.852	13.399.796	13.708.014	13.652.872
Mai/2020	12.280.641	13.440.852	13.389.541	13.630.873	13.581.717
Jun/2020	12.544.540	13.440.852	13.379.285	13.582.989	13.541.668
Jul/2020	13.895.884	13.440.852	13.369.029	13.834.548	13.788.277
Ago/2020	14.515.814	13.440.852	13.358.774	14.120.887	14.064.951
Set/2020	14.633.124	13.440.852	13.348.518	13.977.991	13.931.824
Out/2020	15.133.359	13.440.852	13.338.262	14.213.636	14.172.972
Nov/2020	14.693.552	13.440.852	13.328.007	14.067.043	14.031.401
Dez/2020	14.578.765	13.440.852	13.317.751	13.617.794	13.591.965

Fonte: Elaboração própria com base nos dados analisados.

Figura 5 – Previsão da série



Fonte: Elaboração própria com base nos dados analisados.

Na tabela 4 são apresentados os resultados dos erros de previsão dos modelos, considerando o período de janeiro 2020 a dezembro de 2020, pode-se observar que o modelo *Holt* obteve taxas de erro mais baixa.

Tabela 4 – Erros de previsão 12 passos à frente

Modelos	MAPE
<i>Naive</i>	3,7%
<i>Holt</i>	3,6%
<i>Holt-Winters (adit)</i>	5,4%
<i>Holt-Winters (Mult)</i>	5,1%

Fonte: Elaboração própria com base nos dados analisados.

Na série de ruído branco todos os ACFs são zero em todas as defasagens. Para verificar se os resíduos dos modelos ajustados possuem essas características inicialmente aplicou-se teste de *Ljung-Box* para testar se a série é i.i.d, na tabela 5 pode-se observar os p valores dos modelos ajustados, ao nível de significância de 5% não rejeita-se a hipótese nula, logo os resíduos são independentes e identicamente distribuídos.

Tabela 5 – Resultado do teste de Ljung-Box

Modelos	p-valor
<i>Naive</i>	0,2954
<i>Holt</i>	0,3225
<i>Holt-Winters (adit)</i>	0,6911
<i>Holt-Winters (Mult)</i>	0,6702

Fonte: Elaboração própria com base nos dados analisados

Para testar a normalidade dos resíduos aplicou-se o teste de *Kolmogorov-Smirnov*, na tabela 6 são apresentados os resultados em p-valor para cada modelo, nota-se que ao nível de significância de 5%, não rejeita-se a hipótese nula, ou seja, os resíduos são normalmente distribuídos.

Tabela 6 – Resultado do teste de Kolmogorov-Smirnov

Modelos	p-valor
<i>Naive</i>	0,4116
<i>Holt</i>	0,1896
<i>Holt-Winters (adit)</i>	0,4399
<i>Holt-Winters (Mult)</i>	0,4505

Fonte: Elaboração própria com base nos dados analisados

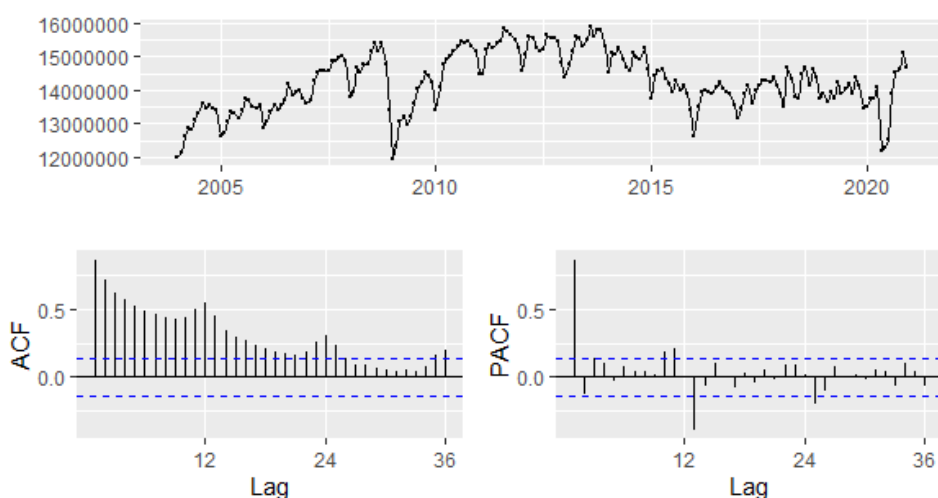
Com base nas análises gráficas e nas métricas escolhidas, foi possível evidenciar que dentre os modelos de alisamento exponencial o *Holt* foi o modelo que melhor se ajustou aos dados no período analisado.

4.2 Implementação do método de Box & Jenkins

A fim de comparar qual metodologia apresenta o modelo que melhor se ajusta aos dados, será implementado o método de previsão de Box & Jenkins para a mesma série de consumo de energia industrial total no Brasil, no período de janeiro de 2004 a dezembro de 2020 deixando os 12 últimos meses para análise da capacidade preditiva do modelo.

Para verificar a estacionariedade, pode-se observar na figura 6 o gráfico do comportamento da série de dados e suas FAC e FACP respectivamente. Nota-se que a série não possui característica estacionária, pois a mesma não é constante em torno da média e sua variância também não é constante, apresentando uma sazonalidade.

Figura 6 – Comportamento da série FAC e FACP



Fonte: Elaboração própria com base nos dados analisados.

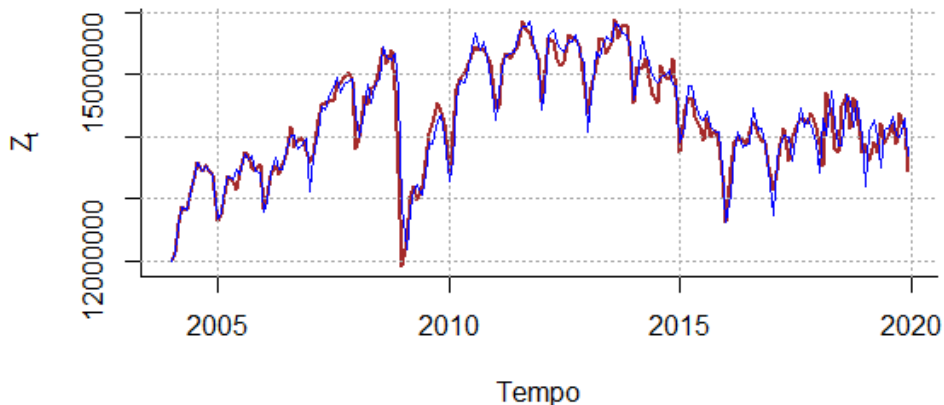
Para que a série se aproxime da distribuição normal estabilizando a variância, utilizou-se a função Box-Cox e o lambda obtido foi de 0.2269183. A fim de complementar a análise gráfica que evidencia que a série é não estacionária, aplicou-se o teste de Dickey-Fuller aumentado o teste consisti na regressão da primeira diferença da série contra sua defasagem. O p-valor obtido no teste foi de 0,8324, logo ao nível de significância de 5% não rejeita-se a hipótese nula, ou seja, a série é não estacionária. Para que a série se torne estacionária foi feita uma diferenciação. O teste de Dickey-Fuller aumentado foi aplicado novamente e o p-valor obtido foi muito menor que 0,05, logo ao nível de significância de 5% rejeita-se a hipótese nula, ou seja, a série se tornou estacionária.

Para elaboração do modelo Box & Jenkins a série foi dividida em duas partes, o período de janeiro de 2004 a dezembro de 2019 foi reservado para treino, e janeiro de 2020 a dezembro de 2020 foi reservado para o teste do modelo. Após o teste de sobrefixação (SOUZA,2004), o modelo escolhido para gerar as previsões foi o SARIMA(1,1,2)(0,1,1)₁₂. Após analisar e verificar o resultado apresentado por alguns métodos de acurácia, este modelo foi considerado o melhor dentre os testados e seus coeficientes foram significativos.

A figura 7 apresenta os dados no período de 2004 a dezembro de 2019, onde a cor marrom representa os dados reais observados e a cor azul corresponde aos valores ajustados, nota-

se que os valores se ajustam bem aos dados. Utilizando as métricas de acurácia obteve-se um MAPE de 1,27% para o período de ajuste, o mesmo é significativamente pequeno, ou seja, existem indícios de que este seja um bom ajuste.

Figura 7 – Ajuste dos dados



Fonte: Elaboração própria com base nos dados analisados

Após a análise do ajuste, foram geradas as previsões 12 passos à frente. Na tabela 7 pode-se observar os valores reais observados e as previsões obtidos pelo modelo SARIMA(1,1,2)(0,1,1)₁₂.

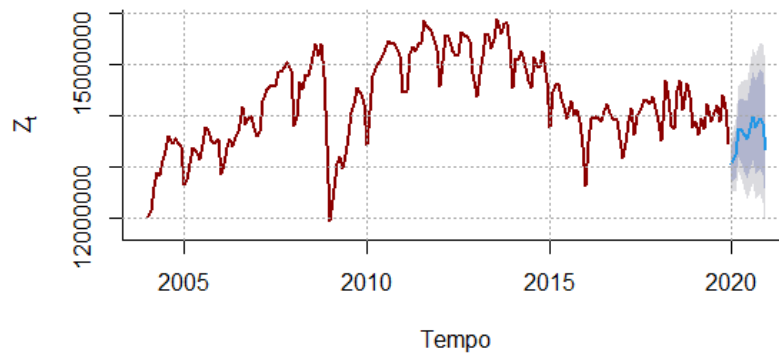
Tabela 7– Previsões 12 passos à frente modelo Box & Jenkins

Mês	Dados	Previsão
Jan/2020	13.447.777	13.044.411
Fev/2020	13.391.884	13.215.468
Mar/2020	14.098.067	13.734.880
Abr/2020	12.167.303	13.715.793
Mai/2020	12.280.641	13.614.191
Jun/2020	12.544.540	13.527.502
Jul/2020	13.895.884	13.751.942
Ago/2020	14.515.814	13.990.251
Set/2020	14.633.124	13.760.586
Out/2020	15.133.359	13.942.166
Nov/2020	14.693.552	13.855.627
Dez/2020	14.578.765	13.352.270

Fonte: Elaboração própria com base nos dados analisados.

A figura 8 possibilita a melhor visualização das previsões 12 passos à frente, onde a linha azul no gráfico corresponde a estimativa pontual, o azul mais escuro corresponde ao intervalo de confiança de 80%, já o azul mais claro representa o intervalo de confiança de 95%.

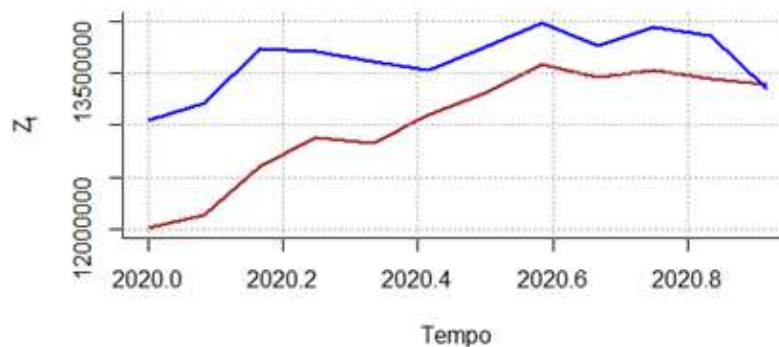
Figura 8 – Gráfico das previsões 12 passos à frente (Método Box & Jenkins)



Fonte: Elaboração própria com base nos dados analisados

Na figura 9 é possível visualizar o comportamento da série de dados observados no período de janeiro de 2020 a dezembro de 2020 (representada em marrom) e as previsões desses 12 meses (representada em azul). Utilizando a métrica de acurácia MAPE, obteve-se um resultado de 4,4%, ou seja, para este período de 12 meses a previsão obtida pelo modelo SARIMA(1,1,2)(0,1,1)₁₂ errou em média 4,4%.

Figura 9 – Série das previsões 12 passos à frente (Método Box & Jenkins)



Fonte: Elaboração própria com base nos dados analisados.

Objetivando verificar se os resíduos do modelo ajustado pelo método Box & Jenkins possui característica de um ruído branco, foi feito o teste de *Ljung-Box* para testar se a série é i.i.d, o p-valor obtido foi de 0,9839, logo ao nível de significância de 5% não rejeita-se a hipótese nula, ou seja, os resíduos do modelo são independentes e identicamente distribuídos. Em seguida aplicou-se o teste de normalidade *Kolmogorov-Smirnov*, para verificar se os resíduos correspondem a uma distribuição normal, o p-valor obtido no teste foi de 0,3025, logo ao nível de significância de 5% não rejeita-se a hipótese nula, ou seja, os resíduos são normalmente distribuídos.

4.3 Combinação dos modelos

Objetivando alcançar uma previsão mais assertiva, aplicou-se a técnica de combinação de modelos. Neste trabalho os dois modelos combinados serão aqueles que apresentaram o melhor resultado, ou seja, o modelo *Holt* pelo método de alisamento exponencial, e o modelo SARIMA (2,1,1)(0,1,1)₁₂ com transformação de Box-Cox, obtido pelo método de Box & Jenkins.

Para encontrar o valor de K que representa o peso do modelo SARIMA, foi aplicada a seguinte equação,

$$K = \frac{\text{Var do resíduo modelo SARIMA}}{\text{Var do resíduo modelo SARIMA} + \text{Var do resíduo modelo Holt}} \quad (18)$$

Para o peso do modelo *Holt*, foi feito o complementar, ou seja, $1 - K$. Isto posto, a previsão dos dados será realizada a partir da combinação dos modelos, representada pela seguinte equação,

$$\text{previsão combinada} = K * \text{SARIMA} + (1 - k) * \text{Holt} \quad (19)$$

Na tabela 8 é possível observar os dados reais e a previsão combinada dos modelos. A fim de verificar quanto assertiva foi a previsão, utilizou-se como método de acurácia o MAPE, e o valor obtido foi de 3,56%, ou seja, em média o modelo combinado está errando 3,56%.

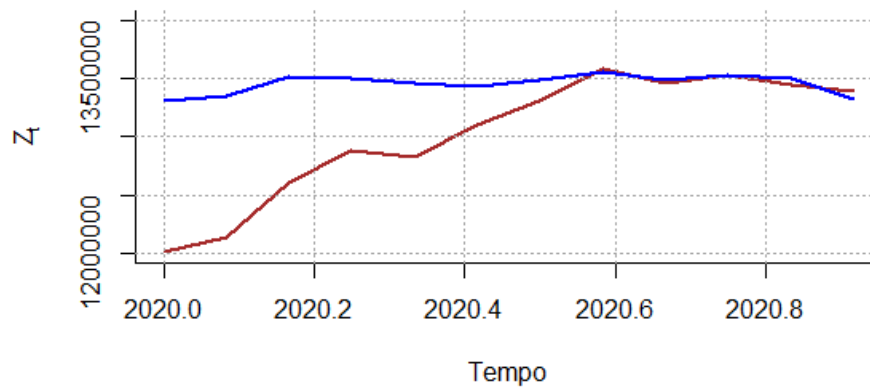
Tabela 8 – Previsões 12 passos à frente modelo Combinado

Mês	Dados	Previsão
Jan/2020	13.447.777	13.304.858
Fev/2020	13.391.884	13.353.626
Mar/2020	14.098.067	13.515.794
Abr/2020	12.167.393	13.502.664
Mai/2020	12.280.641	13.462.672
Jun/2020	12.544.540	13.427.535
Jul/2020	13.895.884	13.493.680
Ago/2020	14.515.814	13.564.340
Set/2020	14.633.124	13.482.660
Out/2020	15.133.359	13.534.853
Nov/2020	14.693.552	13.499.765
Dez/2020	14.578.765	13.328.988

Fonte: Elaboração própria com base nos dados analisados.

A figura 10 apresenta os dados observados (representado em marrom) no período de janeiro de 2020 a dezembro de 2020, e a previsão desses 12 meses (representada em azul), pode-se notar que a partir do mês 5 de 2020 os dados previstos se aproximam cada vez mais dos observados.

Figura 10 – Previsão combinada 12 passos à frente



Fonte: Elaboração própria com base nos dados analisados

5. Conclusão

No ano de 2020 o mundo foi acometido por um novo vírus, para minimizar o contágio algumas medidas de segurança foram tomadas, as mesmas impactaram diretamente alguns setores, como o consumo de energia elétrica. O presente trabalho objetivou aplicar modelos de séries temporais, buscando verificar brevemente se houve impacto significativo e se de fato eles afetariam as previsões feitas para o período de janeiro de 2020 a dezembro de 2020. Isto posto, foi possível evidenciar que existiram impactos negativos no consumo de energia elétrica pelo setor industrial brasileiro nos meses de abril a julho de 2020, período que as medidas restritivas estavam mais severas, porém nas previsões do mês de julho em diante, foi possível identificar a eficácia dos modelos, pois os mesmos apresentaram resultados significativos com uma porcentagem relativamente baixa de erro.

Com as análises dos dados, pode-se concluir graficamente e através do teste de *Mann-Kendall* que os dados no período analisado apresentam tendência. Para a componente de sazonalidade observou-se os resultados dos testes, que apontaram indícios de sazonalidade na série. Após a aplicação dos modelos de previsão *Naive*, *Holt*, *Holt-Winters* aditivo e *Holt-Winters* multiplicativo, os resultados gráficos e das métricas de erro mostram que o modelo que melhor se ajustou aos dados foi o *Holt*, ou seja, o modelo que considera apenas a tendência da série e não a sazonalidade.

O método Box & Jenkins foi implementado, e após aplicar o teste sobrefixação o modelo que melhor se ajustou aos dados foi o SARIMA(1,1,2)(0,1,1)₁₂, com a transformação Box-Cox. Após a análise de previsão dos dois melhores modelos, aplicou-se a técnica de combinação, gerando uma previsão combinada do Holt com o SARIMA (1,1,2)(0,1,1)₁₂. Para verificar qual modelo implementado foi o mais assertivo, ou seja, utilizou-se a métrica de acurácia MAPE. Para o modelo Holt o MAPE foi de 3,6%, já para o SARIMA o MAPE de ajuste foi de 1,27%, e o modelo combinado apresentou um MAPE de 3,56%, logo o modelo SARIMA é o que melhor se ajusta aos dados.

Com a análise dos resíduos, observou-se através dos testes estatísticos que todos apresentam características de ruído branco, este resultado também pode ser observado graficamente. Com isso pode-se concluir, que os modelos abordados apresentaram boas previsões, sendo o Box – Jenkins o mais assertivo.

Referências

BARBOSA, Eduardo Campana; SÁFADI, Thelma; NASCIMENTO, Moysés; NASCIMENTO, Ana Carolina Campana; SILVA, Carlos Henrique Osório; MANULI, Rômulo César. **Metodologia Box & Jenkins para previsão de temperatura média mensal da cidade de Bauru (SP)**, 2015.

BBC, News. **Corona vírus: Brasil é um dos mais afetados entre 75 países onde epidemia ainda cresce**. 2020. Disponível em: <
<https://www.bbc.com/portuguese/brasil-53047836>>

BOX G. E. P. & JENKINS, G.M. **Time Séries Analysis, Forecasting and Control**, San Francisco; Holden Day, Inc. [and (1976) revised edition].

SOUZA, R. C.; CAMARGO, M. E.; **Análise e Previsão de Séries Temporais: os modelos ARIMA**. 2ª Edição. Rio de Janeiro: Gráfica e Editora Regional, 2004.

SOUZA, Reinaldo Castro; OLIVEIRA, Fernando Luiza Cyrino. **Análise de Séries Temporais: Os Modelos de Amortecimento Exponencial**, 2015.

SILVA, Edson da. **Consumo eficiente de energia elétrica em indústrias do ramo de bebidas no Brasil**. Centro Universitário UNIFACVEST, 2018.

MACHADO, Alexandre Regueiro de Almeida. **Gerenciamento de Energia em Planta Industrial Visando a Redução dos Custos de Operação**. 2010.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. 2ª Edição. São Paulo: Blucher, 2006.