



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO**

**INSTITUTO DE CIÊNCIAS EXATAS**

**CURSO DE GRADUAÇÃO EM MATEMÁTICA**

**Elaine de Oliveira Jesuino**

**ESTUDO COMPARATIVO ENTRE MODELOS DE  
PREVISÃO PARA O CONSUMO DE ELETRICIDADE  
DO SETOR INDUSTRIAL BRASILEIRO**

**SEROPÉDICA - RJ**

**2022**



Elaine de Oliveira Jesuino

## Estudo Comparativo entre Modelos de Previsão para o Consumo de Eletricidade do Setor Industrial Brasileiro

Monografia apresentada à Banca Examinadora da Universidade Federal Rural do Rio de Janeiro, como requisito parcial para obtenção do título de Bacharel em Matemática Aplicada e Computacional, sob a orientação do Prof. Dr. Felipe Leite Coelho da Silva.

SEROPÉDICA - RJ

2022

## Resumo

O comportamento do consumo de eletricidade no setor industrial brasileiro vem sendo investigado ao longo dos últimos anos, devido principalmente ao desenvolvimento da indústria e na implementação de medidas de eficiência energética. Desta forma, as análises e projeções das séries temporais do consumo de energia elétrica podem contribuir para o planejamento do setor. A finalidade deste trabalho foi comparar o desempenho de modelos de previsão de séries temporais aplicados aos dados de consumo de energia elétrica do setor industrial brasileiro entre os anos de 2004 e 2021. Foi utilizada a técnica de validação cruzada para selecionar os hiperparâmetros para os modelos Holt-Winters, SARIMA, NNAR e Ingênuo e então foram comparados através de métricas estatísticas (MAPE, RMSE). O modelo SARIMA teve um desempenho bem próximo do modelo ingênuo e o NNAR não conseguiu modelar bem a sazonalidade da série, enquanto o modelo Holt-Winters foi o melhor em todas as comparações.

**Palavras-Chave:** energia elétrica; séries temporais; setor industrial; previsão; validação cruzada.

# Sumário

<b>Introdução</b>	<b>1</b>
<b>1 REVISÃO BIBLIOGRÁFICA</b>	<b>3</b>
<b>2 REVISÃO TEÓRICA</b>	<b>5</b>
2.1 Séries Temporais . . . . .	5
2.2 Modelos de Previsão . . . . .	8
2.2.1 Modelos Ingênuos . . . . .	8
2.2.2 Box-Jenkins . . . . .	9
2.2.3 Suavização Exponencial . . . . .	10
2.2.4 Redes Neurais Autoregressivas . . . . .	11
2.3 Métricas Estatísticas . . . . .	12
<b>3 MATERIAL E MÉTODOS</b>	<b>14</b>
3.1 Conjunto de Dados . . . . .	14
3.2 Metodologia . . . . .	14
<b>4 RESULTADOS</b>	<b>17</b>
4.1 Análise Exploratória . . . . .	17
4.2 Seleção, Ajuste e Validação . . . . .	19
4.2.1 Método Ingênuo . . . . .	19
4.2.2 Box-Jenkins . . . . .	19
4.2.3 Suavização Exponencial . . . . .	22
4.2.4 Redes Neurais Autoregressivas . . . . .	24
4.3 Previsão e Análise . . . . .	25

<b>5 CONCLUSÕES</b>	<b>27</b>
<b>Referências Bibliográficas</b>	<b>29</b>

# Introdução

O consumo de energia do setor industrial é um dos indicadores principais do desenvolvimento econômico de um país por refletir o ritmo das atividades do setor, mas seu aumento gera impactos negativos no meio ambiente e a possibilidade do esgotamento dos recursos utilizados para a produção da energia ([ANEEL, 2008](#)). Para minimizar esses impactos é preciso planejamento e, para tal, é necessário um estudo que seja capaz de projetar as tendências futuras de consumo de energia elétrica e uma das possibilidades é realizar essa predição através de ferramentas de análise de séries temporais aplicadas à série temporal do consumo de energia elétrica.

Uma série temporal é qualquer conjunto de observações ordenadas no tempo e o estudo de uma série temporal pode nos mostrar importantes características do processo de geração dos dados que pode ser usado para prever valores futuros ([MORETTIN e TOLOI, 2018](#)).

O estudo desenvolvido neste trabalho utiliza métodos de análise de séries temporais e compara seus resultados para a série histórica do consumo de energia elétrica do setor industrial brasileiro com o objetivo de encontrar o modelo que melhor se ajuste aos dados históricos e possa prever o consumo futuro.

A fim de realizar a previsão do consumo de energia elétrica do setor industrial brasileiro, foi realizado o ajuste de diferentes modelos de previsão. Primeiro foi utilizado o modelo ingênuo ([CHEN et al., 2003](#)), um modelo simples e utilizado como comparação para os modelos mais complexos. Depois foram utilizados os modelos Box-Jenkins ([BOX et al., 2008](#)) e o modelo de suavização exponencial Holt-Winters que são modelos clássicos de previsão de séries temporais. Em seguida, foi utilizado um modelo de redes neurais autoregressivas proposto por [HYNDMAN e ATHANASOPOULOS \(2021\)](#).

Este trabalho está organizado em capítulos. O capítulo 1 descreve a revisão bibli-

ográfica, as semelhanças e as diferenças deste trabalho com os trabalhos anteriores. O Capítulo 2 apresenta uma breve descrição teórica de séries temporais e dos modelos e métodos de previsão utilizados, enquanto o capítulo 3 descreve os dados utilizados e a metodologia aplicada. O Capítulo 4 mostra os resultados obtidos e as previsões realizadas pelos modelos e, por fim, no capítulo 5 são apresentadas as principais conclusões e as possibilidades para trabalhos futuros.

# Capítulo 1

## REVISÃO BIBLIOGRÁFICA

Alguns trabalhos anteriores já se propuseram a estudar a série temporal do consumo industrial de energia elétrica no Brasil utilizando somente os dados das observações passadas, sem nenhuma variável preditora. [GOTIJO et al. \(2017\)](#) utilizou os métodos ARMA e Média Móvel, que fazem parte dos modelos Box-Jenkins, além do modelo de suavização exponencial para a série do consumo industrial entre abril de 2007 e janeiro de 2017, enquanto [da SILVA et al. \(2022\)](#) utilizou o modelo de suavização exponencial Holt-Winters, o modelo Box-Jenkins SARIMA, o modelo de redes neurais autoregressivas NNAR, além de outros modelos utilizando a série entre os anos de 1979 e 2020. Já [NEVES e VIVAS \(2018\)](#) analisaram a série do consumo industrial dividida por regiões geográficas brasileiras, utilizando, entre outros métodos que utilizam variáveis exógenas, os métodos SARIMA e Holt-Winters nas observações entre 2003 e 2019.

Os modelos Box-Jenkins ([BOX et al., 2008](#)) e suas variações são os mais utilizados para realizar análises de séries temporais, seguido dos modelos de Suavização Exponencial ([HOLT, 1957](#); [WINTERS, 1960](#)) considerados de simples implementação.

Este trabalho se propõe a implementar os dois modelos mais utilizados e também complementar os trabalhos anteriores ao utilizar um recorte de observações diferente dos demais, além de utilizar a técnica de validação cruzada como metodologia de seleção de modelos.

A validação cruzada é amplamente utilizada em aprendizagem de máquina, mas geralmente considerada inadequada com dados de séries temporais pelo argumento de se tratar de valores ordenados no tempo. Porém, [HYNDMAN e ATHANASOPOULOS \(2021\)](#)



descreve uma forma de realizar validação cruzada de forma a preservar a ordem das observações, enquanto [TASHMAN \(2000\)](#), através de uma revisão bibliográfica, conclui a validade do método de validação cruzada para seleção de modelos, mesmo nos casos de modelos estatísticos como o ARIMA. A maior dificuldade para pesquisar sobre esse assunto são os vários nomes que essa técnica recebe na bibliografia, criando uma rede de citações esparsa.

## Capítulo 2

# REVISÃO TEÓRICA

### 2.1 Séries Temporais

Uma das possíveis definições de série temporal é qualquer conjunto de observações ordenadas no tempo ([MORETTIN e TOLOI, 2018](#)). A ordem dos dados é essencial e uma característica importante dessa ordenação é o fato dos dados serem autocorrelacionados, ou seja, as observações são correlacionadas às observações passadas. Ao analisar as observações ordenadas no tempo, alguns padrões implícitos no dados podem ser visualizados como a tendência e a sazonalidade. A tendência demonstra o padrão de crescimento e/ou decrescimento das observações, enquanto a sazonalidade é a repetição de um padrão a cada período fixo de tempo. Esses padrões podem aparecer sozinhos ou ao mesmo tempo, se relacionando de forma aditiva ou multiplicativa, como pode-se ver na figura 2.1.

A definição formal, segundo [MORETTIN e TOLOI \(2018\)](#), é entender uma série temporal como uma realização particular de um processo estocástico, ou seja, cada observação obtida é uma variável aleatória com uma distribuição de probabilidade, como exemplificado na figura 2.2. Dessa forma, entende-se que uma série temporal contém um componente determinístico, composto dos padrões implícitos da série como a tendência e sazonalidade, e um componente estocástico que gera sempre um ruído probabilístico nas medições da série, às vezes também chamado de erro.



Figura 2.1: Visualização gráfica da tendência e da sazonalidade e a relação aditiva ou multiplicativa entre eles.

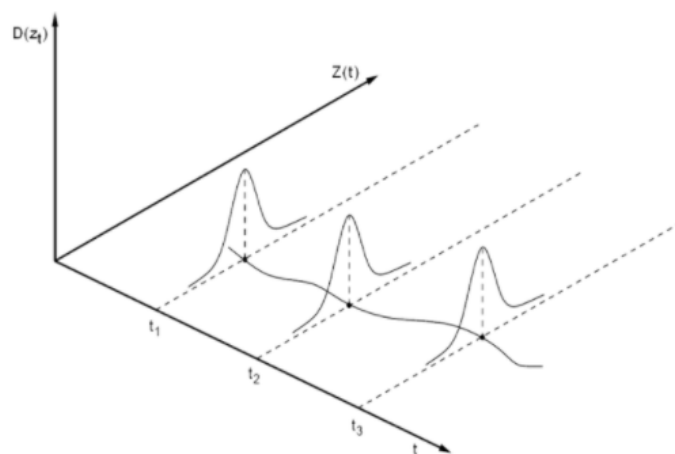


Figura 2.2: Visualização gráfica da tendência e da sazonalidade e a relação aditiva ou multiplicativa entre eles.

Ao analisar a série temporal como um processo estocástico, podemos pensar sobre o conceito de estacionariedade. Um processo estocástico estacionário tem suas propriedades estatísticas invariantes no tempo, ou seja, uma série temporal estacionária tem a propriedade que a média, a variância e a estrutura das autocorrelações das observações não mudam no decorrer do tempo. Alguns modelos de previsão, como os modelos Box-Jenkins, tem como suposição que a série é estacionária, então precisa-se testar estatisticamente a estacionariedade da série antes de aplicar modelos de previsão. Neste trabalho, foi utilizado o teste de Dickey–Fuller Aumentado (ADF) (PHILLIPS e PERRON, 1988) para aceitar ou rejeitar a hipótese nula de não estacionariedade da série.

A maioria das séries temporais medidas no mundo real não são estacionárias, então às vezes se faz necessária realizar transformações na série para se obter uma série estacionária. Existem diferentes transformações possíveis na literatura, mas neste trabalho foram utilizadas a transformação das diferenças e a transformação Box-Cox (MORETTIN e TOLOI, 2018). A transformação das diferenças, equação (2.1), consiste em obter a série das diferenças entre as observações. A primeira diferença ( $m=1$ ) significa obter a série de diferenças entre as observações consecutivas, já a diferença sazonal ( $m$  é igual ao período sazonal) significa obter a série de diferenças entre uma observação e a observação do período sazonal passado. A transformação de Box-cox, equação (2.2), tem o objetivo de estabilizar a variância da série temporal e tornar as observações mais semelhantes à distribuição normal. Como mostra a figura 2.3, além da variância, a transformação também muda a escala da série temporal.

$$Y'_t = Y_t - Y_{t-m} \quad (2.1)$$

$$Y_t^{(\lambda)} = \begin{cases} \log(Y_t), & \lambda = 0 \\ \frac{Y_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \end{cases} \quad (2.2)$$

Onde:  $Y'_t$  é a série transformada com diferenças;  $Y_t^{(\lambda)}$  é a série transformada com Box-Cox;  $Y_t$  é a série temporal observada; e  $\lambda$  é um parâmetro a ser estimado para a série, depois de transformada, se tornar estacionária.

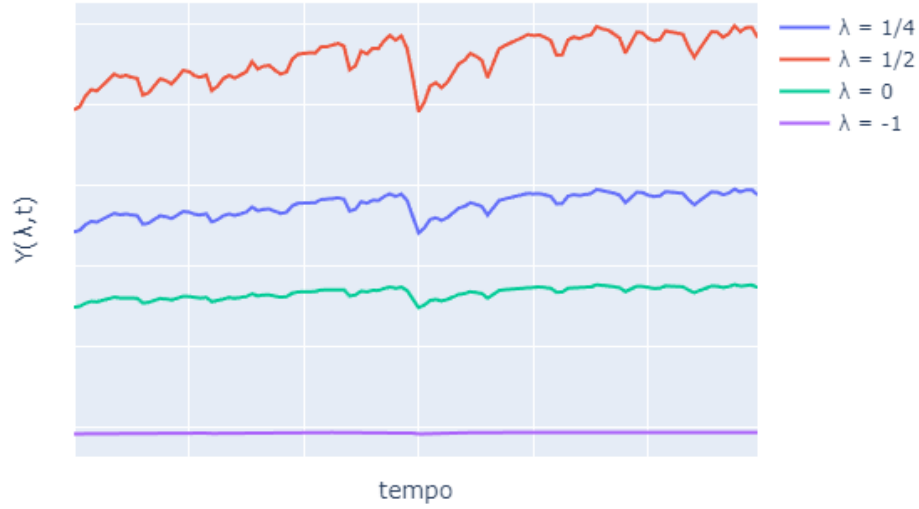


Figura 2.3: Exemplos de Transformações de Box-cox com lambdas diferentes

## 2.2 Modelos de Previsão

### 2.2.1 Modelos Ingênuos

Os modelos ingênuos realizam a previsão para o instante seguinte através do valor imediatamente anterior, sem fazer nenhuma suposição sobre o comportamento da série. Esses modelos são fáceis de usar, não precisam de nenhum ajuste e podem ser usados para previsão mesmo quando existe poucas observações passadas. Esses métodos são utilizados como referência para os outros métodos, determinando se os métodos mais complexos apresentam uma melhoria em previsão que compense o maior gasto de tempo e custo computacional utilizado para ajustar tais modelos. (CHEN et al., 2003)

O modelo de persistência repete a última observação da série para todos os passos futuros que se pretende prever. Já o método ingênuo sazonal repete as observações do último período para todos os próximos períodos. A figura 2.4 mostra um exemplo da aplicação dos métodos e as equações são mostradas em (2.3) e (2.4).

$$\textbf{Persistência} \quad \hat{Y}_{N+h} = y_N \quad (2.3)$$

$$\textbf{Sazonal} \quad \hat{Y}_{N+h} = y_{N+h-m(k+1)} \quad (2.4)$$

onde  $N$  é o número de observações antes da janela de previsão,  $h$  são os passos de tempo na previsão,  $m$  é o período sazonal e  $k$  é o número de períodos já completados na janela de previsão no passo  $h$ .

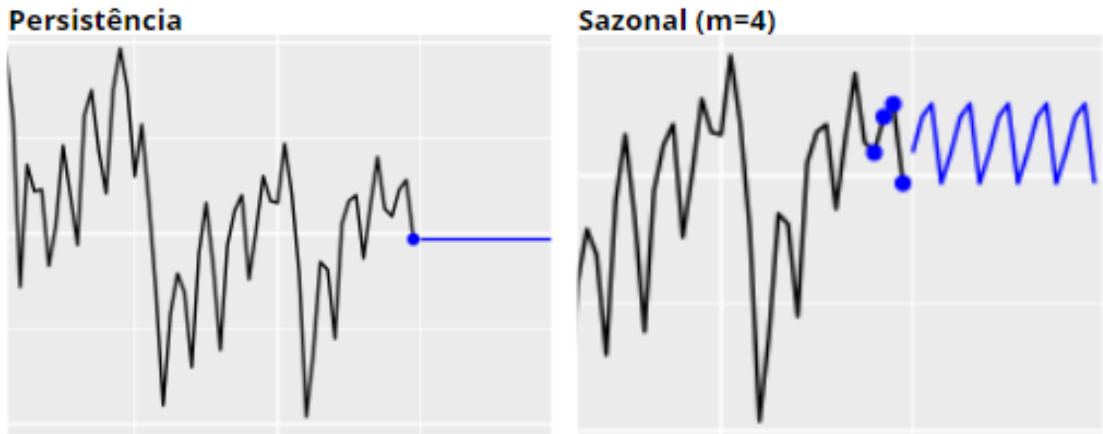


Figura 2.4: Exemplo de aplicação dos métodos ingênuos.

### 2.2.2 Box-Jenkins

Os modelos de Box-Jenkins utilizam uma combinação de diferenciações com modelos autoregressivos e média móvel para encontrar um modelo que melhor explique o comportamento de uma série histórica.

Modelos autoregressivos (AR( $p$ )) seguem da hipótese que as observações presentes podem ser explicadas como combinações lineares de  $p$  observações passadas. Já os modelos de média móvel (MA( $q$ )) explicam o presente através dos  $q$  ruídos previamente observados. Esses modelos podem ser aplicados de forma separada ou de forma conjunta, através do modelo ARMA( $p, q$ ). Tanto o modelo AR quanto o MA só podem ser aplicados em séries estacionárias, então para lidar com séries não estacionárias, Box-Jenkins desenvolveram o ARIMA( $p, d, q$ ) que diferencia a série  $d$  vezes para que os modelos possam ser utilizados em séries não estacionárias.

Para lidar com séries com sazonalidade, o SARIMA( $p, d, q$ )( $P, D, Q$ ) $s$ , detalhado na equação (2.5), é o método que modela a autoregressão e a média móvel tanto da série original quanto da série defasada no período  $s$  da sazonalidade.

$$\Phi(B^s)\phi(B)Z_t = \Theta(B)\theta(B^s)e_t \quad (2.5)$$

Onde:

$Z_t = (1 - B)^d(1 - B^s)Y_t$  é a série temporal univariada  $Y_t$  diferenciada em  $D$  vezes em relação ao período  $s$  de sazonalidade seguida de  $d$  diferenciações sem sazonalidade;

$B$  é o operador de retardo onde  $By_t = y_{t-1}$ ;

$\Phi(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps})$  é o polinômio autoregressivo AR(P) com componente sazonal de período  $s$ ;

$\phi(B) = (1 - \phi_1 B - \dots - \phi_p B^p)$  é o polinômio autoregressivo AR(p);

$\Theta(B) = (1 - \Theta_1 B - \dots - \Theta_q B^q)$  é a média móvel do processo MA(q);

$\theta(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{sQ})$  é a média móvel do processo MA(Q) com componente sazonal de período  $s$ ;

$e_t$  é um ruído branco;

$\Phi, \phi, \Theta, \theta$  são os parâmetros do modelo estimados por máxima verossimilhança.

Segundo [HYNDMAN e ATHANASOPOULOS \(2021\)](#), podemos determinar os valores de  $p$  e  $q$  através dos gráficos das funções de autocorrelação (*Autocorrelation Function - ACF*) e autocorrelação parcial (*Partial Autocorrelation Function - PACF*). O gráfico ACF mostra a correlação entre as observações e as observações defasadas da série temporal. O gráfico PACF mostra a correlação entre as observações defasadas que não pode ser explicada por defasagens anteriores. Se o ACF demonstra um decaimento amortecido e o PACF um declínio rápido para zero depois de  $p$  defasagens, isso indica um processo AR(p). Se os gráficos demonstram os mesmos comportamentos mas de forma invertida (ACF declínio e PACF decaimento) o processo indicado é o MA(q). Para o caso sazonal em vez de olhar para todas as defasagens, analisa-se somente as defasagens no período sazonal.

### 2.2.3 Suavização Exponencial

Os Modelos de Suavização Exponencial seguem a hipótese de que as observações mais recentes são mais importantes para descrever a observação presente do que as observações mais antigas. Para modelar essa hipótese, é aplicada uma média móvel ponderada com decaimento exponencial em cada um dos componentes da série temporal (nível, tendência e/ou sazonalidade). As relações aditivas ou multiplicativas entre os componentes da série também podem ser incorporadas pelos modelos. Neste trabalho utilizaremos somente a tendência aditiva, pois a modelagem multiplicativa tende a gerar previsões ruins ([HYNDMAN e ATHANASOPOULOS, 2021](#)).

As equações do modelo Holt-Winters utilizadas são:

Holt-Winters Aditivo	Holt-Winters Multiplicativo	
$\hat{y}_{t+h t} = \ell_t + hb_t + s_{t+h-m(k+1)}$	$\hat{y}_{t+h t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$	(2.6)
$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1})$	$\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(\ell_{t-1} + b_{t-1})$	
$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1}$	
$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m},$	$s_t = \gamma \frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1 - \gamma)s_{t-m}$	

Holt-Winters Aditivo Amortecido	Holt-Winters Multiplicativo Amortecido
$\hat{y}_{t+h t} = \ell_t + \Phi_h b_t + s_{t+h-m(k+1)}$	$\hat{y}_{t+h t} = [\ell_t + \Phi_h b_t] s_{t+h-m(k+1)}$
$\ell_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$	$\ell_t = \alpha(y_t / s_{t-m}) + (1 - \alpha)(\ell_{t-1} + \phi b_{t-1})$
$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$	$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)\phi b_{t-1}$
$s_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1 - \gamma)s_{t-m},$	$s_t = \gamma \frac{y_t}{(\ell_{t-1} + \phi b_{t-1})} + (1 - \gamma)s_{t-m}$

Onde:

$y_t$  são as observações da série temporal;

$\hat{y}_{t+h}$  é a previsão para o passo  $h$  no futuro,  $m$  é o período sazonal e  $k$  é o número de períodos já completados na janela de previsão no passo  $h$ ;

$\alpha$ ,  $\beta$  e  $\gamma$  são os fatores de suavização e  $\phi$  é o fator de amortecimento;

$\ell_t$ ,  $b_t$  e  $s_t$  são os valores do nível, tendência e sazonalidade, respectivamente, calculados para cada ponto  $t$ .

## 2.2.4 Redes Neurais Autoregressivas

O modelo de Redes Neurais Autoregressivas  $NNAR(p, P, k)_m$ , proposto por [HYNDMAN e ATHANASOPOULOS \(2021\)](#), é um modelo de rede neural que contém uma camada de entrada que recebe  $p + P$  entradas, sendo  $p$  a quantidade de defasagens não sazonais e  $P$  a quantidade de defasagens sazonais. Além da camada de entrada, a rede contém uma única camada escondida com  $k$  neurônios com uma função de ativação logística e uma camada de saída que retorna apenas um valor. Os valores  $p$  e  $P$  equivalem aos passos da autoregressão, equivalentes a  $SARIMA(p,0,0) \times (P,0,0)_m$  e  $m$  é o período



sazonal. Cada neurônio da camada escondida recebe os valores de entrada, realizam a combinação linear ponderada ( $z$ ) e, neste resultado, aplica a função não linear logística  $g(z)$ . O valor dessa operação é repassado para a próxima camada e novamente combinado para gerar o resultado da autoregressão. Os pesos e constantes (bias) usados para realizar as combinações ponderadas são treinados pelo algoritmo *feed-forward* e utilizados para realizar previsões um passo no futuro. A função logística adiciona não linearidade à autoregressão e a rede neural não tem as restrições de estacionariedade dos modelos autoregressivos Box-Jenkins.

A figura 2.5 mostra o exemplo da arquitetura de uma NNAR. Neste exemplo, temos 3 neurônios de entrada, sendo 2 entradas não sazonais ( $p=2$ ) e 1 entrada sazonal ( $P=1$ ) e a camada oculta contém somente 1 neurônio ( $k=1$ ), logo temos o modelo  $NNAR(2, 1, 1)_m$ .

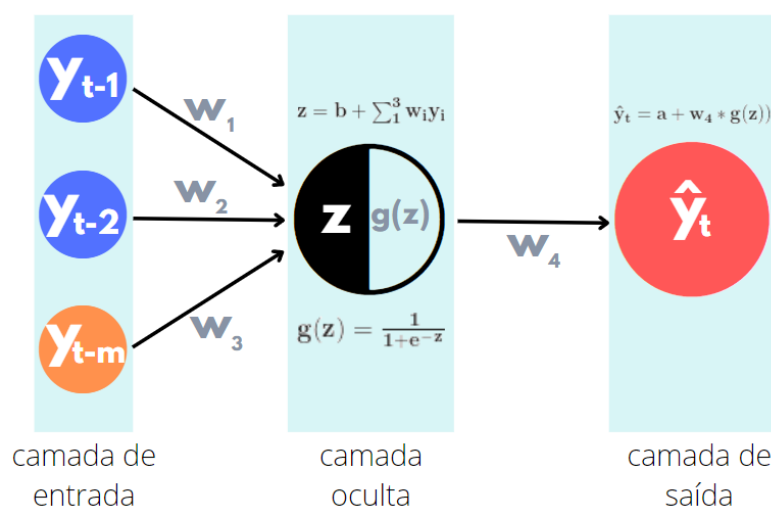


Figura 2.5: Exemplo da arquitetura de uma  $NNAR(2, 1, 1)_m$ .

## 2.3 Métricas Estatísticas

Para comparar a qualidade da previsão e ajuste dos modelos, foram utilizadas duas métricas neste trabalho:

- **Erro Médio Absoluto Percentual (MAPE - Mean Absolute Percent Error):** Esta métrica calcula o erro relativo de previsão e, ao multiplicar por 100, indica a porcentagem de erro variando entre 0% e 100%. Por ser invariante de escala, pode ser

usada para comparar diferentes modelos e séries.

$$MAPE(\%) = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| * 100 \quad (2.7)$$

onde  $N$  é o número de passos no tempo previsto,  $y_i$  são os valores observados e  $\hat{y}_i$  são os valores previstos pelo modelo.

- **Raíz do Erro Quadrático Médio (RMSE - Root Mean Squared Error):** Esta métrica é a raiz do erro médio quadrático, assim retornando o erro calculado à unidade de medida dos dados do modelo.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (2.8)$$

onde  $N$  é o número de passos no tempo previsto,  $y_i$  são os valores observados e  $\hat{y}_i$  são os valores previstos pelo modelo.

Ao aplicar os modelos Box-Jenkins, foi utilizada uma métrica que mensura a qualidade de um modelo estatístico visando também a sua simplicidade, fornecendo uma medida de comparação e seleção de modelos.

- **Crítério de Informação Bayesiana (BIC - Bayesian Information Criterion):** Esta métrica é baseada na teoria da informação e estima a quantidade de informações perdidas pelo modelo enquanto penaliza a complexidade.

$$BIC = k \ln(n) - 2 \ln(\hat{L}) \quad (2.9)$$

onde:  $k$  é o número de parâmetros estimados;  $L$  é a função de verossimilhança; e  $n$  é o número de observações.

## Capítulo 3

# MATERIAL E MÉTODOS

### 3.1 Conjunto de Dados

A Empresa Brasileira de Pesquisa Energética disponibiliza, em seu sítio eletrônico, dados mensais do consumo de energia elétrica no Brasil ([EPE, 2022](#)) a partir do ano de 2004, atualizados mensalmente. Para este estudo, foram extraídos os dados de consumo de energia elétrica do setor industrial de todo Brasil dos 216 meses entre Janeiro de 2004 até Dezembro de 2021.

Para testar o poder de ajuste à série histórica e também a generalização das previsões, os dados foram divididos em dois conjuntos distintos: o conjunto de treino, que abrange os dados entre 2004 e 2018, e o conjunto de teste, que abrange o período de janeiro de 2019 a dezembro de 2021. O conjunto de treino foi utilizado em todo o processo de exploração dos dados, ajustes e validação dos modelos, sendo usado para avaliar o desempenho de previsão e capacidade dos modelos de explicar os dados observados. O conjunto de teste foi utilizado para avaliar a capacidade preditiva dos modelos escolhidos em dados desconhecidos.

### 3.2 Metodologia

A primeira etapa do estudo consiste em uma análise exploratória na série temporal para entender melhor seu comportamento e suas características, como estacionariedade, tendência e sazonalidade, informações importantes para identificar os modelos a serem

ajustados. A segunda etapa consiste em aplicar, para cada um dos modelos utilizados, as fases de seleção, ajuste e diagnóstico dos modelos.

Por vezes, antes de iniciar a aplicação dos modelos é necessário realizar transformações no dados para que eles possam satisfazer as hipóteses dos modelos. A figura 3.1 trás um diagrama das fases da metodologia aplicada para cada um dos quatro modelos de previsão testados: Ingênuo, Box-Jenkins, Holt-Winters e NNAR.

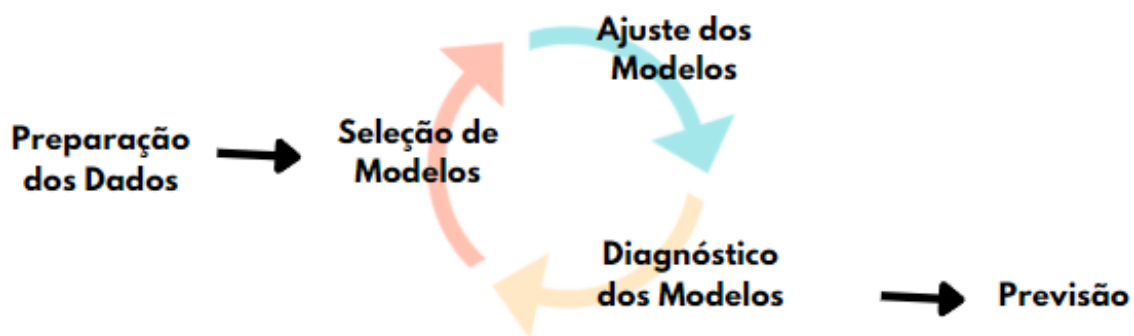


Figura 3.1: Fases da metodologia

A fase de seleção de modelos é onde se escolhe os hiperparâmetros a serem utilizados na fase de ajuste. Os hiperparâmetros são os parâmetros que são definidos antes do início do aplicação dos modelos baseados em escolhas e/ou conhecimentos prévios sobre a série temporal.

A fase de ajuste é a etapa de aplicação do modelo, quando se faz a estimação e/ou treinamento dos parâmetros utilizados para modelar a série temporal. Nesta fase aplica-se a técnica de validação cruzada, também conhecida como *rolling origin evaluation* (TASHMAN, 2000), que tem por objetivo avaliar a capacidade de generalização dos modelos, deixando os erros menos sensíveis às observações atípicas, permitindo selecionar um modelo mais robusto. A técnica consiste em utilizar diferentes partes do conjunto de teste para treino e previsão em cada iteração e o modelo com o menor erro médio é selecionado e, se adequado, utilizado na etapa de previsão. A figura 3.2 demonstra as quatro iterações de treino e validação utilizadas neste trabalho. A adequação ou não dos modelos é verificada na fase de diagnóstico.

O diagnóstico dos modelos se dá através da análise dos resíduos entre os valores observados da série original e os valores modelados pelo modelo. O objetivo é obter



Figura 3.2: Iterações da Validação Cruzada

resíduos que são considerados ruídos brancos, ou seja, resíduos com média zero, variância constante, e sem autocorrelação. Na prática, nem sempre é possível obter resíduos com todas essas características, então considera-se o modelo válido quando os resíduos têm média zero e não possuem autocorrelação, hipóteses testadas através de testes estatísticos.

Por último, é realizada a etapa de previsão do consumo de eletricidade para o período de teste utilizando os modelos escolhidos. Os modelos são comparados utilizando métricas estatísticas (MAPE, RMSE) com o objetivo de encontrar aquele com a melhor capacidade de previsão para série temporal do consumo de energia elétrica do setor industrial brasileiro.

Para a realização de todo processamento computacional, foi utilizado o ambiente Jupyter Notebook, onde foi possível utilizar as linguagens Python e R e várias de suas bibliotecas concomitantemente, de forma a utilizar os melhores algoritmos e processamentos para cada modelo proposto.

## Capítulo 4

# RESULTADOS

### 4.1 Análise Exploratória

O primeiro passo na análise de uma série histórica é observar o gráfico temporal para se ter intuições sobre seu comportamento. Ao analisar o gráfico 4.1, a série parece conter uma tendência variável durante os anos, além de uma clara sazonalidade anual, ou seja, a série temporal observada contém um padrão de repetição com um período de 12 meses. A série também demonstra ser não estacionária, pois mudanças na tendência e na sazonalidade podem ser visualizadas através dos *boxplots* na figura 4.2.



Figura 4.1: Observações da série temporal do consumo.

Para confirmar a intuição anterior foi realizada a decomposição STL (*Seasonal and Trend Decomposition using Loess*), mostrada na figura 4.3. Esta decomposição divide os valores observados da série temporal nos componentes de tendência, sazonalidade e resíduo, confirmando a presença de uma tendência e uma sazonalidade com período de

12 meses que mudam durante os anos.

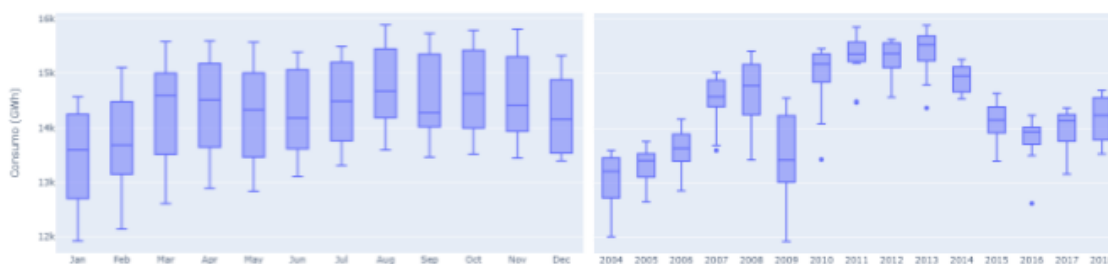


Figura 4.2: Observações da série temporal do consumo.

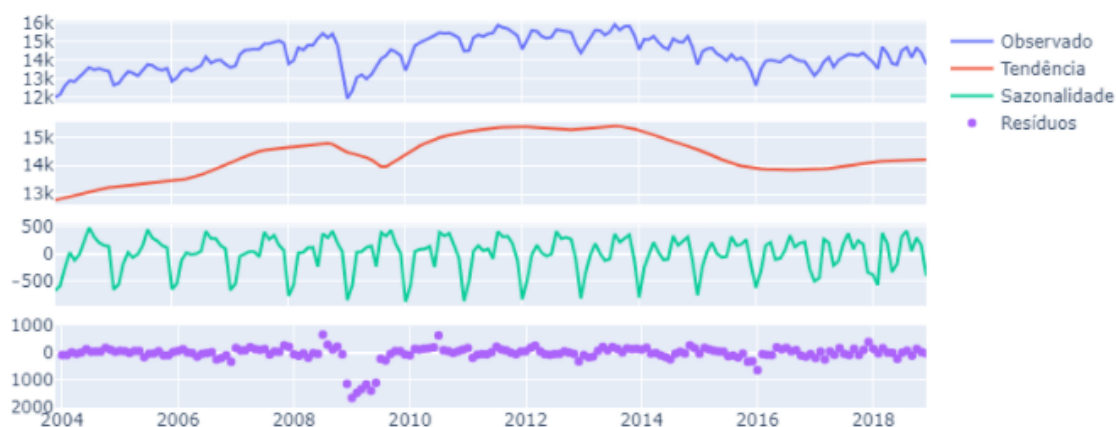


Figura 4.3: Decomposição STL

Pode-se perceber, através dos resíduos, a dificuldade de descrever a série entre os anos de 2008 e 2010, o que é explicado por serem observações atípicas influenciadas pela crise financeira mundial ocorrida entre os anos 2008 e 2009, que teve por consequência uma forte retração de atividade industrial (EPE, 2008). Algumas mudanças na série, como a tendência de queda a partir de 2014, podem ser explicadas pela aumento de autoprodução de energia de alguns setores industriais, gerando uma redução da demanda à rede brasileira.

Para testar a estacionariedade da série foi realizado o teste de Dickey-Fuller Aumentado (ADF) ao nível de significância de 5%. O teste encontrou um p-valor maior que o nível de significância, aceitando a hipótese de não estacionariedade da série, confirmando a intuição de que a média e a variância da série mudam ao decorrer dos anos.

Através das estatísticas descritivas, descritas na tabela 4.1, percebemos que os dados

têm uma assimetria positiva e uma distribuição mais achatada que a normal, enquanto o coeficiente de variação demonstra uma baixa dispersão dos valores em relação à média.

<b>Média</b>	<b>Mediana</b>	<b>Desvio Padrão</b>	<b>Coef. Variação</b>
14301.97	14293.91	878.34	6.14%
<b>Valor Mínimo</b>	<b>Valor Máximo</b>	<b>Assimetria</b>	<b>Curtose</b>
11923.91	15886.31	-0.26	-0.54

Tabela 4.1: Tabela de Estatísticas Descritivas

## 4.2 Seleção, Ajuste e Validação

### 4.2.1 Método Ingênuo

Por ser uma série sazonal, foi selecionado o modelo ingênuo sazonal e, ao realizar a validação cruzada, foi encontrado um MAPE médio de 6.01% e um RMSE médio de 988.89 GWh. Esses valores devem ser usados com cautela, pois o modelo ingênuo não modela a tendência e nem as mudanças na sazonalidade, mas servirá como um referencial durante a validação dos outros modelos.

A figura 4.4 demonstra os resultados obtidos para cada iteração da validação cruzada.

### 4.2.2 Box-Jenkins

Uma das suposições deste modelo é que a série deve ser estacionária, então devemos transformar a série do consumo até que ela se torne estacionária antes de continuar com a seleção dos modelos. Como a série contém sazonalidade com período de 12 meses, devemos primeiro realizar a diferença sazonal e verificar sua estacionariedade. Feito isso, a série continua não estacionária, o que foi verificado pelo teste ADF (p-valor de 0.10), então uma nova diferença foi realizada e só então obtivemos uma série estacionária. A figura 4.5 mostra a série observada e a série obtida depois das transformações realizadas.

Seguindo o método de seleção de modelos de [HYNDMAN e ATHANASOPOULOS \(2021\)](#) utilizando o gráfico de autocorrelação (ACF) e o gráfico de autocorrelação par-



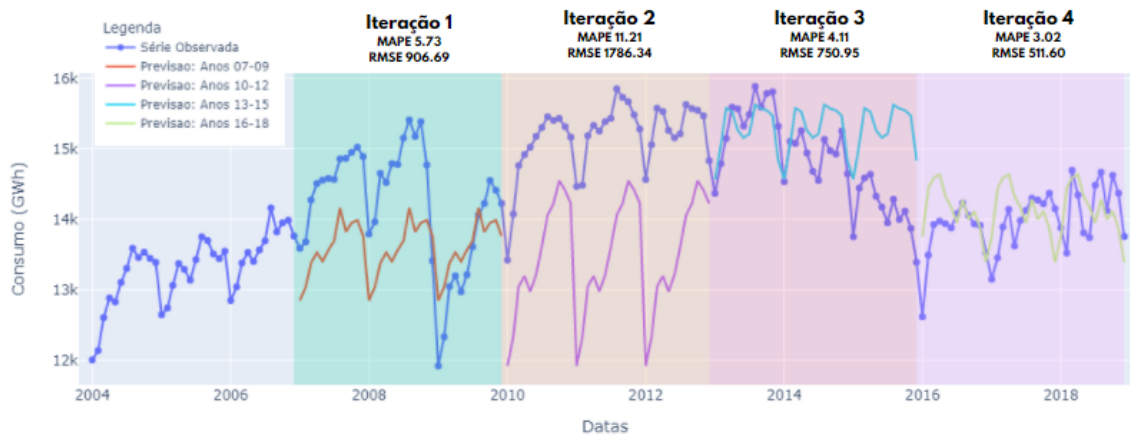


Figura 4.4: Validação Cruzada do Modelo Ingênuo Sazonal. Para cada iteração (área colorida) está expresso o MAPE, o RMSE, os valores observados (linha azul) e os valores previstos (linha colorida).

cial (PACF), mostrados na figura 4.6, selecionamos dois modelos a serem ajustados:  $SARIMA(0, 1, 0)(2, 1, 0)_{12}$  e  $SARIMA(0, 1, 0)(0, 1, 1)_{12}$ .

Além desses, foi utilizada a seleção automática de modelos da biblioteca pmdarima, que implementa em Python a função *auto.arima* do R, capaz de fazer uma busca extensiva em grade por todos os modelos válidos e selecionar aquele com menor critério de informação Bayesiano (BIC). A seleção de modelos foi realizada na série em suas três transformações (série original, série diferenciada sazonal e série diferenciada sazonal com primeira diferença) e foram encontrado dois outros modelos possíveis:

$SARIMA(0, 1, 0)(1, 0, 1)_{12}$  e  $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ .

Os quatro modelos selecionados foram então ajustados para cada uma das iterações da validação cruzada e a média dos erros percentuais foi calculada. A tabela 4.2 mostra as médias obtidas e o melhor resultado em destaque (verde).

Modelo SARIMA	$(0, 1, 0)(2, 1, 0)_{12}$	$(0, 1, 0)(0, 1, 1)_{12}$	$(0, 1, 0)(1, 0, 1)_{12}$	$(1, 0, 0)(0, 1, 1)_{12}$
MAPE Médio	9.21%	8.19%	6.04%	<b>5.93%</b>

Tabela 4.2: Resultado da validação cruzada.

Interessante perceber que, mesmo que os modelos escolhidos pelo *auto.arima* não

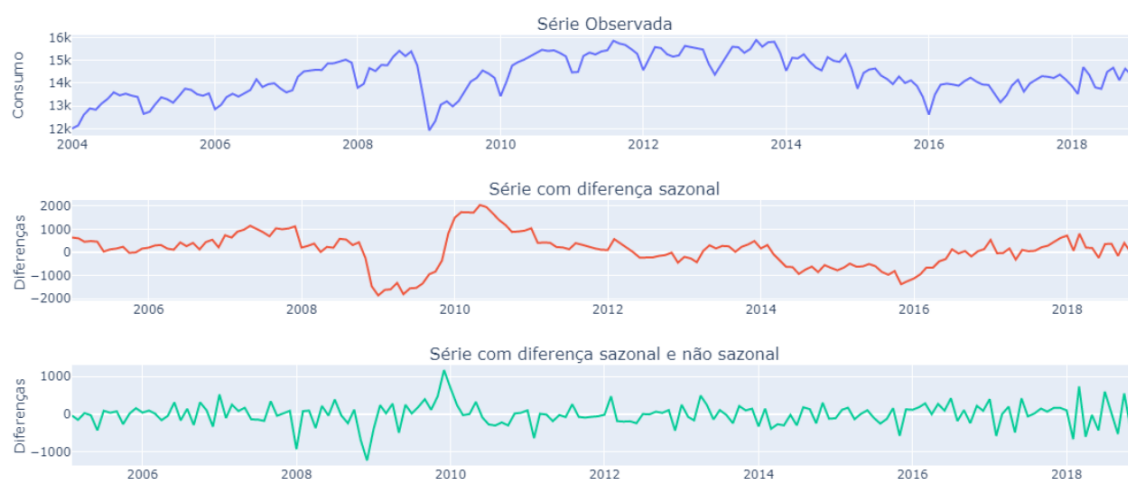


Figura 4.5: Gráfico da série observada e das séries obtidas com as transformações de diferenças realizadas.

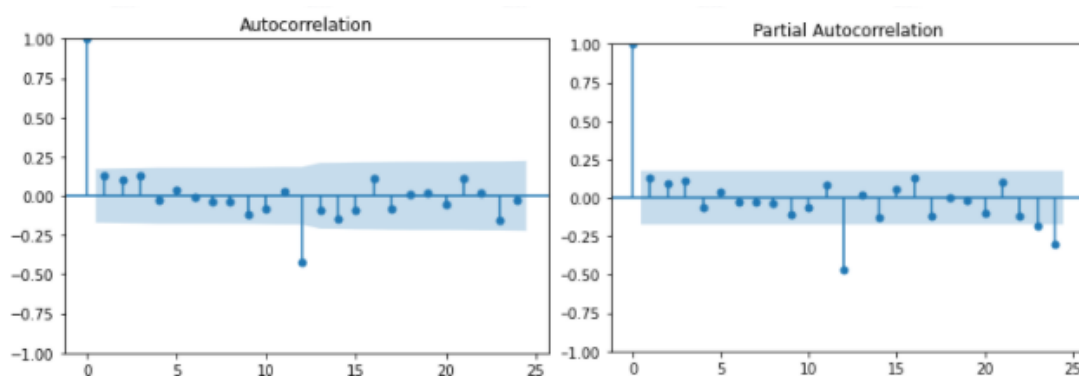


Figura 4.6: Série Transformada com diferença sazonal e primeira diferença.

utilizam as duas diferenças necessárias para transformar a série temporal em uma série estacionária, os dois modelos tiveram uma capacidade de generalização melhor do que os modelos escolhidos pelos gráficos ACF/PACF e são válidos pelos testes de hipótese necessários de acordo com [BOX et al. \(2008\)](#).

O modelo  $SARIMA(1, 0, 0)(0, 1, 1)_{12}$  obteve o melhor resultado na validação cruzada e ao ajustar o modelo em todas as observações do conjunto de treinamento e realizar o diagnóstico encontramos que todos os parâmetros ajustados são significativos e os resíduos são considerados ruído branco, o que demonstra que o modelo é adequado para ser utilizado na etapa de previsão. A figura 4.7 demonstra os resultados obtidos para cada iteração da validação cruzada para o modelo SARIMA selecionado.

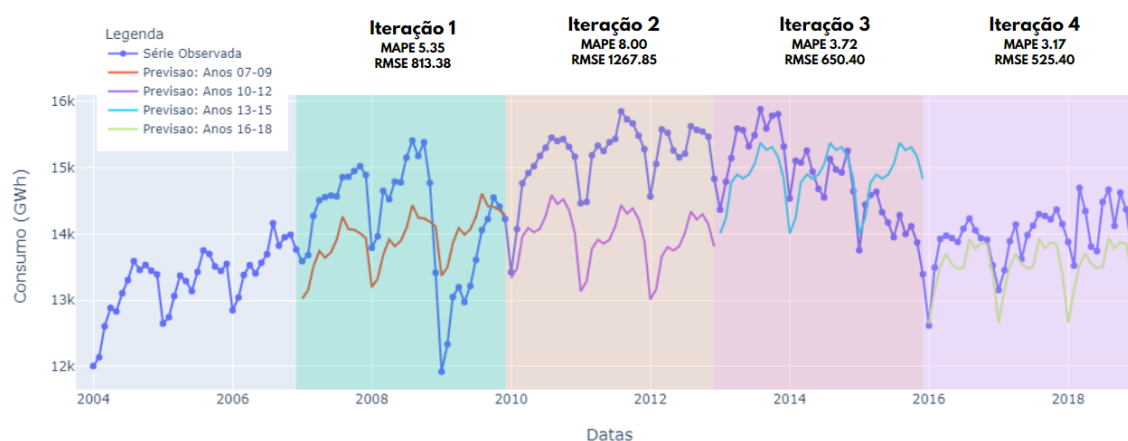


Figura 4.7: Validação Cruzada do modelo  $SARIMA(1, 0, 0)(0, 1, 1)_{12}$ . Para cada iteração (área colorida) está expresso o MAPE, o RMSE, os valores observados (linha azul) e os valores previstos (linha colorida).

### 4.2.3 Suavização Exponencial

Como vimos na fase de exploração, a série do consumo de energia contém tendência e sazonalidade, então usaremos o modelo Holt-Winters e modelaremos a tendência de forma aditiva e a sazonalidade podendo ser aditiva ou multiplicativa. Outro hiperparâmetro testado foi o amortecimento da tendência, podendo ter os valores *True* (tendência com amortecimento) ou *False* (tendência sem amortecimento). Foi testada também a influência de se utilizar ou não a transformação de Box-Cox durante o ajuste dos modelos, podendo ter os valores *True* (transformação utilizada) ou *False* (transformação não utilizada). A tabela 4.3 mostra o resultado encontrado pela validação de todos os modelos testados e os melhores valores da coluna estão em destaque. Os hiperparâmetros mostrados na primeira coluna indicam as escolhas feitas para cada modelo na seguinte ordem:

- Tendência: aditiva (*add*)
- Amortecimento da tendência: utilizado (*True*) ou não utilizado (*False*)
- Sazonalidade: Aditiva (*add*) ou Multiplicativa (*mul*)
- Transformação Box-cox: utilizada (*True*) ou não utilizada (*False*)

O resultado obtido mostra dois modelos com MAPE iguais mas um dos modelos tendo um RMSE um pouco menor. Para ajudar a decidir, os dois modelos foram ajustados em

Modelo	RMSE	MAPE (%)
<i>(add, False, mul, True)</i>	<b>643.92</b>	<b>3.85</b>
<i>(add, False, add, True)</i>	651.44	<b>3.85</b>
<i>(add, True, add, False)</i>	676.95	4.20
<i>(add, True, add, True)</i>	684.64	4.11
<i>(add, False, mul, False)</i>	697.21	4.15
<i>(add, True, mul, False)</i>	709.44	4.36
<i>(add, False, add, False)</i>	709.53	4.42

Tabela 4.3: Resultado da Validação Cruzada. Os hiperparâmetros indicam as escolhas na seguinte ordem: Tendência, Amortecimento, Sazonalidade e Transformação Box-cox.

todo o conjunto de treino e apresentaram resíduos não autocorrelacionados e com média zero. Como os dois modelos são válidos, o modelo *(add, False, mul, True)* foi escolhido, ou seja, o modelo com tendência aditiva e sem amortecimento, sazonalidade multiplicativa e transformação Box-cox, pois foi o modelo com menor RMSE na validação cruzada. A figura 4.8 mostra os resultados da validação cruzada do modelo selecionado.

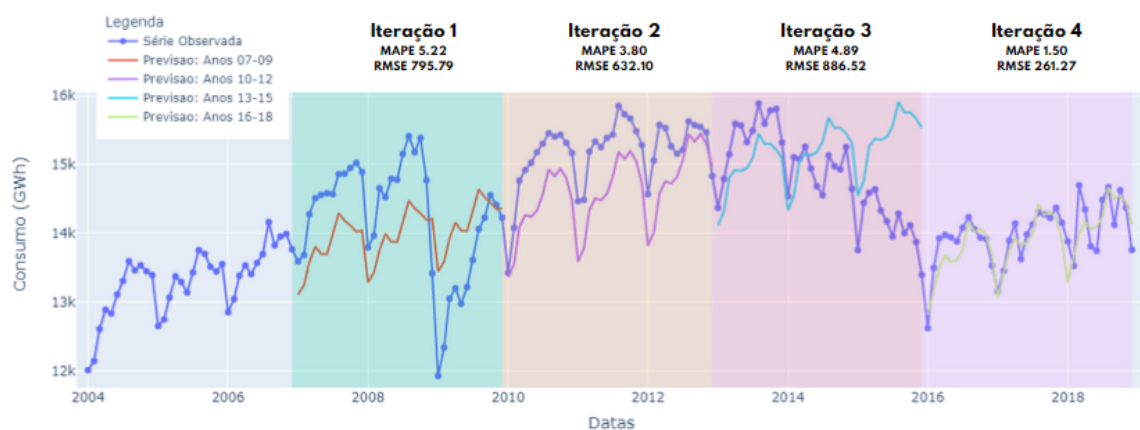


Figura 4.8: Validação Cruzada do modelo com tendência aditiva e sem amortecimento, sazonalidade multiplicativa e transformação Box-cox. Para cada iteração (área colorida) está expresso o MAPE, o RMSE, os valores observados (linha azul) e os valores previstos (linha colorida).

#### 4.2.4 Redes Neurais Autoregressivas

Para escolher os melhores valores para os hiperparâmetros de  $NNAR(p, P, k)_{12}$ , foi realizada uma validação cruzada para cada valor de  $p$  e  $P$  entre 0 e 6. O valor de  $k$  utilizado foi  $(p + P + 1)/2$ , proposto por [HYNDMAN e ATHANASOPOULOS \(2021\)](#). Para cada combinação de valores  $(p, P)$  foram treinadas 100 redes neurais e a média entre elas foi utilizada para a previsão. A figura 4.9 apresenta o resultado obtido, quanto mais claro o ponto menor o MAPE médio, então o modelo escolhido é o  $NNAR(3, 2, 3)_{12}$  com um MAPE médio de 4.56%. Pela natureza estocástica das redes neurais, esse processo foi repetido algumas vezes e um resultado parecido foi encontrado, sempre favorecendo o ponto  $(3, 2)$ .

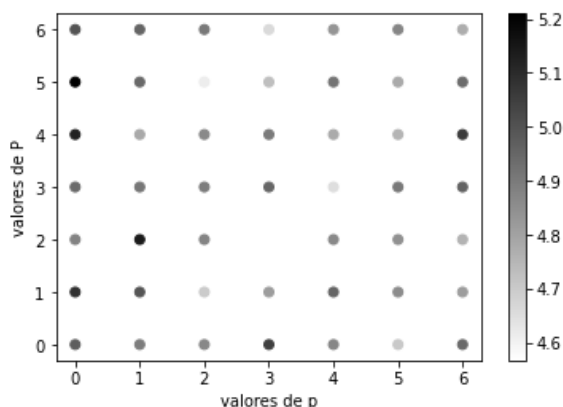


Figura 4.9: MAPE médio encontrado para cada combinação de  $p$  e  $P$ .

Mesmo não tendo nenhuma suposição sobre os resíduos como os outros modelos estatísticos, ao aplicarmos o NNAR nos dados de treino também obtemos resíduos não correlacionados e com média zero. Na figura 4.10 percebe-se que mesmo tendo métricas melhores que o modelo ingênuo, o NNAR tem previsões que não conseguem modelar a variação sazonal da série temporal.



Figura 4.10: Validação Cruzada do NNAR  $(3, 2, 3)_{12}$ . Para cada iteração (área colorida) está expresso o MAPE, o RMSE, os valores observados (linha azul) e os valores previstos (linha colorida).

### 4.3 Previsão e Análise

Depois de selecionar os melhores modelos utilizando o conjunto de treino e a técnica de validação cruzada, os modelos são usados para realizar a previsão dos anos de 2019 a 2021 e elas são comparadas com os dados observados no conjunto de teste. A tabela 4.4 mostra a comparação das métricas entre os modelos, tanto na fase de validação quanto na previsão, mostrando em vermelho o pior desempenho da coluna, enquanto em verde mostra-se o melhor. Já a figura 4.11 mostra as previsões realizadas por cada modelo e também os dados observados do conjunto de teste.

Modelo	Hiperparâmetros	Validação Cruzada		Previsão	
		MAPE(%)	RMSE	MAPE(%)	RMSE
Ingênuo	Sazonal	6.01	988.89	4.76	819.75
SARIMA	$(1, 0, 0) \times (0, 1, 1)_{12}$	5.93	1089.28	4.62	842.73
Holt-Winters	(add,False,mul,True)	3.85	643.92	3.57	698.06
NNAR	$(3, 2, 3)_{12}$	4.92	806.13	3.79	758.57

Tabela 4.4: Comparação da previsão dos modelos

O modelo SARIMA selecionado não conseguiu se ajustar bem aos dados, sendo pior em RMSE que o modelo ingênuo, mesmo tendo um MAPE um pouco melhor. O modelo NNAR teve um desempenho superior ao ingênuo e ao SARIMA nas duas métricas, mas percebe-se pela figura 4.11 que ele não consegue descrever bem a sazonalidade da série. Entre os modelos comparados, o Holt-Winters foi o que melhor se ajustou à série, conseguindo captar um pouco da tendência e da sazonalidade da série e foi o melhor nas duas métricas utilizadas, tanto na validação quanto na previsão.

O resultado obtido se compara ao resultado obtido por [da SILVA et al. \(2022\)](#), mostrando que o modelo Holt-Winters supera tanto o SARIMA quanto o NNAR para a série de consumo industrial, mesmo que neste trabalho tenha se utilizado menos observações.

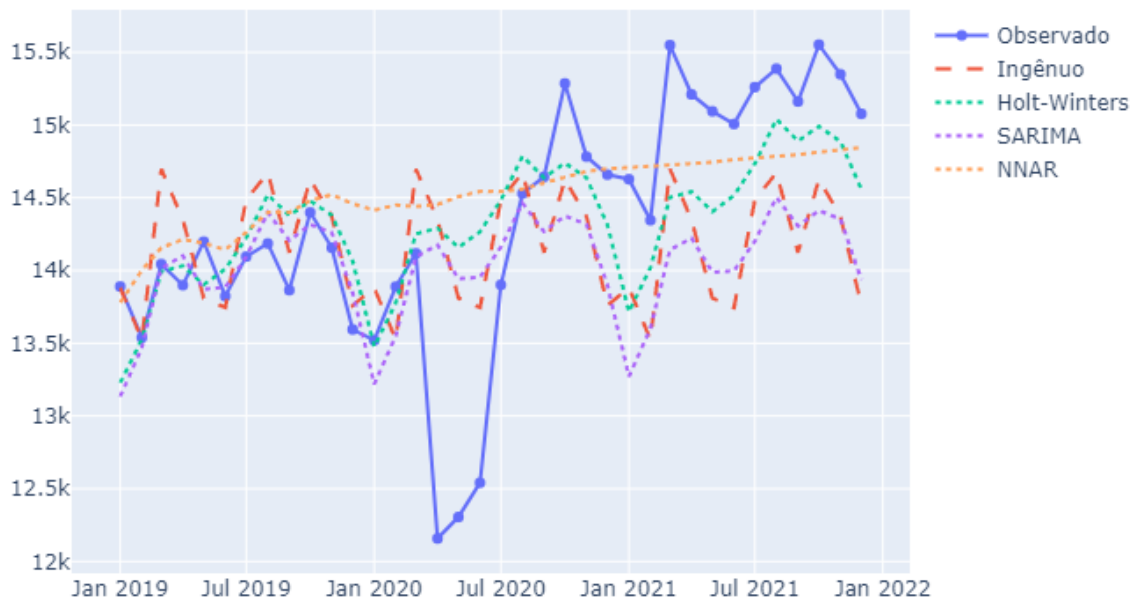


Figura 4.11: Gráfico das previsões de cada um dos modelos utilizados e dos valores observados do conjunto de teste.

## Capítulo 5

# CONCLUSÕES

Este trabalho buscou comparar a qualidade das previsões de quatro métodos de previsão distintos ajustados ao consumo de energia elétrica do setor industrial brasileiro. Analisando os resultados, percebemos que a série temporal utilizada é complexa, com observações consideradas atípicas e com tendência e sazonalidade variáveis, o que é explicado pela influência de fatores econômicos e/ou ambientais externos.

O método ingênuo foi utilizado como um referencial, pois modelos mais complexos devem ser utilizados para previsão somente se conseguem descrever melhor a série do que o método mais simples. O modelo Box-Jenkins não teve êxito em explicar a série melhor que o método ingênuo, obtendo métricas muito parecidas e às vezes piores que o método ingênuo, tanto na validação quanto na previsão, mostrando que essa metodologia não conseguiu capturar a estrutura da série temporal modelada. Já o modelo Holt-Winters teve a melhor performance entre todos os modelos utilizados, conseguindo captar a estrutura da série e foi capaz de realizar uma previsão adequada.

O modelo de redes neurais autoregressivas obteve um resultado satisfatório nas fases de validação e previsão da série. Mesmo que o modelo tenha a desvantagem de ser estocástico, ou seja, cada implementação da rede gera pesos e resultados um pouco diferentes, o modelo ainda pode ser otimizado além dos hiperparâmetros otimizados neste trabalho. Uma possibilidade para trabalhos futuros seria otimizar o número de neurônios na camada oculta, o número de redes utilizadas ao mesmo tempo e a forma que as redes são combinadas para gerar a previsão.

Além das opções acima, também seria interessante, em um trabalho futuro, implemen-



tar outras arquiteturas de redes neurais com mais camadas ocultas (como o perceptron multicamada) e/ou com outros tipos de camadas (como LSTM). Esses novos modelos poderão ser usados sozinhos para previsão ou agregados com os modelos Holt-Winters e NNAR, já testados neste trabalho, para gerar previsões combinadas.

## Referências Bibliográficas

ANEEL: *Atlas de Energia Elétrica no Brasil*. 2008. – Disponível em: [http://www2.aneel.gov.br/arquivos/pdf/atlas\\_par1\\_cap2.pdf](http://www2.aneel.gov.br/arquivos/pdf/atlas_par1_cap2.pdf). – Acesso em: 16 jan. 2022

BOX, George E. P. ; JENKINS, Gwilym M. ; REINSEL, Gregory C.: *Time Series Analysis: Forecasting and Control*. Hoboken, New Jersey : 4ª edição. John Wiley & Sons, Inc., 2008

CHEN, Rachel J. C. ; BLOOMFIELD, Peter ; FU, Joshua S.: An Evaluation of Alternative Forecasting Methods to Recreation Visitation. Em: *Journal of Leisure Research* - Vol. 35, No. 4, pp. 441-454 (2003)

EPE: *Resenha Mensal do Mercado de Energia Elétrica - Dezembro*. 2008. – Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes>. – Acesso em: 07 abr. 2022.

EPE: *Consumo Mensal de Energia Elétrica por Classe (regiões e sub-sistemas)*. 2022. – Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes>. – Acesso em: 16 jan. 2022.

GOTIJO, Tiago S. ; MORO RODRIGUES, Fabiano de ; CÁSSIO RODRIGUES, Alexandre de ; SILVA, Stela A. da ; AZEVEDO, Andressa A. de: Consumo Industrial de Energia Elétrica: um estudo comparativo entre métodos preditivos. Em: *Brazilian Journal of Production Engineering*, Vol. 3, N.º 3, p. 31-45 (2017). – Disponível em: [https://periodicos.ufes.br/index.php/bjpe/article/view/v3n3\\_03](https://periodicos.ufes.br/index.php/bjpe/article/view/v3n3_03). – Acesso em: 19 abr. 2022

HOLT, C. E.: *Forecasting seasonals and trends by exponentially weighted averages*. Pittsburgh USA : Carnegie Institute of Technology, 1957

HYNDMAN, R. J. ; ATHANASOPOULOS, G.: *Forecasting: principles and practice*. 3. Melbourne, Australia : OTexts, 2021. – Disponível em: <https://otexts.com/fpp3/>. – Acesso em: 10 abr. 2022.

MORETTIN, Pedro A. ; TOLOI, Clélia M. C.: *Análise de Séries Temporais*. São Paulo, SP : Volume 1, 3ª edição. Blucher, 2018

NEVES, Gustavo Costa M. ; VIVAS, Hugo Catalão S.: *Previsão do Consumo Industrial de Energia Elétrica por Região do Brasil*. 2018

PHILLIPS, Peter C. B. ; PERRON, Pierre: Testing for a Unit Root in Time Series Regression. Em: *Biometrika* Vol. 75, No. 2, pp. 335-346 (1988)

SILVA, Felipe Leite C. da ; COSTA, Kleyton da ; RODRIGUES, Paulo C. ; SALAS, Rodrigo ; LÓPEZ-GONZALES, Javier L.: Statistical and Artificial Neural Networks Models for Electricity Consumption Forecasting in the Brazilian Industrial Sector. Em: *Energies* 15, no. 2: 588 (2022). – Disponível em: <https://doi.org/10.3390/en15020588>. – Acesso em: 21 abr. 2022

TASHMAN, Leonard J.: *Out-of-sample tests of forecasting accuracy: an analysis and review*. 2000

WINTERS, P. R.: *Forecasting sales by exponentially weighted moving averages*. Management Science, 6(3), 324–342, 1960