

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**Escola de Engenharia**  
**Programa de Pós-Graduação em Engenharia Mecânica**

Rodrigo de Almeida Nascimento

**ESTUDO DE MÉTODOS DE PREVISÃO DE SÉRIES TEMPORAIS APLICADOS  
AO PREÇO DA ENERGIA ELÉTRICA NO MERCADO DE CURTO PRAZO  
BRASILEIRO**

Belo Horizonte

2022

Rodrigo de Almeida Nascimento

**ESTUDO DE MÉTODOS DE PREVISÃO DE SÉRIES TEMPORAIS APLICADOS  
AO PREÇO DA ENERGIA ELÉTRICA NO MERCADO DE CURTO PRAZO  
BRASILEIRO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Minas Gerais como requisito para obtenção do título de mestre em Engenharia Mecânica.

Orientador: Ricardo Poley Martins Ferreira

Belo Horizonte

2022

N244e

Nascimento, Rodrigo de Almeida.

Estudo de métodos de previsão de séries temporais aplicados ao preço da energia elétrica no mercado de curto prazo brasileiro [recurso eletrônico] / Rodrigo de Almeida Nascimento. - 2022.

1 recurso online (107 f.: il., color.) : pdf.

Orientador: Ricardo Poley Martins Ferreira.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f.103-107.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia Mecânica - Teses. 2. Energia elétrica – Preços – Teses.  
3. Redes Neurais (Computação) Teses. I. Ferreira, Ricardo Poley Martins.  
II. Universidade Federal de Minas Gerais. Escola de Engenharia.  
III. Título.

CDU: 621 (043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE ENGENHARIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

## FOLHA DE APROVAÇÃO

### **ESTUDO DE MÉTODOS DE PREVISÃO DE SÉRIES TEMPORAIS APLICADOS AO PREÇO DA ENERGIA ELÉTRICA NO MERCADO DE CURTO PRAZO BRASILEIRO**

**RODRIGO DE ALMEIDA NASCIMENTO**

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Minas Gerais, constituída pelos Professores: Dr. Ricardo Poley Martins Ferreira (Orientador - Departamento de Engenharia Mecânica/UFMG), Dr. Ricardo Luiz Utsch de Freitas Pinto (Departamento de Engenharia Mecânica/UFMG) e Dr. Gustavo Rodrigues Lacerda Silva (Centro Universitário UNA), como parte dos requisitos necessários à obtenção do título de "Mestre em Engenharia Mecânica", na área de concentração de "Projeto e Sistemas".

Dissertação aprovada no dia 27 de janeiro de 2022.



Documento assinado eletronicamente por **Ricardo Poley Martins Ferreira, Professor do Magistério Superior**, em 28/01/2022, às 13:47, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ricardo Luiz Utsch de Freitas Pinto, Professor do Magistério Superior**, em 07/02/2022, às 11:44, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gustavo Rodrigues Lacerda Silva, Usuário Externo**, em 17/02/2022, às 00:23, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **1219601** e o código CRC **AF8CF74A**.

Dedico esse trabalho à minha família.

## RESUMO

Esse trabalho contribui com um estudo comparativo de desempenho de metodologias preditivas de séries temporais na previsão do preço da energia elétrica no mercado de curto prazo brasileiro. Foram aplicados métodos de suavização exponencial, ARIMA e SARIMA, modelos de redes neurais artificiais *Multilayer Perceptrons* (MLP) com entradas univariadas e multivariadas na previsão do Preço de Liquidações das Diferenças (PLD) do submercado de energia da região Sudeste/Centro-Oeste. Para o modelo *Multilayer Perceptrons* multivariado, foram selecionadas as variáveis preditoras mais relevantes para a previsão do PLD. Os desempenhos dos modelos foram quantificados por meio de métricas de avaliação de erro de previsão: *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE) e pela avaliação do desempenho em relação à previsão ingênua por meio do coeficiente U de Theil. Os resultados mostram a superioridade de previsão do modelo MLP multivariado em relação aos demais. Esse modelo apresenta ainda uma razoável capacidade de previsão de tendência dos valores do PLD para as semanas mais avançadas.

Palavras chaves: Previsão do preço de energia. PLD. Redes neurais artificiais. ARIMA. Suavização exponencial.

## ABSTRACT

This work contributes to a comparative study of the performance of time series predictive methodologies in forecasting the price of electricity in the Brazilian short-term market. Exponential smoothing methods, ARIMA and SARIMA, models of artificial neural networks Multilayer Perceptrons (MLP) with univariate and multivariate inputs were applied in predicting the PLD (settlement price for the differences) of the energy submarket in the Southeast/Center-West region of Brazil. For the multivariate Multilayer Perceptrons model, the most relevant predictor variables for PLD prediction were selected. The performance of the models was quantified through prediction error evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and by evaluating the performance in relation to the naive prediction by half of Theil's U coefficient. The results show the prediction superiority of the multivariate MLP model in relation to the others. This model also has a reasonable ability to forecast the trend of PLD values for more advanced periods.

Keywords: Electricity price forecasting. PLD. Artificial neural networks. ARIMA. Exponential smoothing.

## LISTA DE FIGURAS

Figura 1 – Decomposição de uma série temporal.....	21
Figura 2 – Modelo não linear de um neurônio $k$ .....	37
Figura 3 – Funções de ativação .....	40
Figura 4 – Arquitetura de uma rede neural MLP genérica.....	42
Figura 5 – Ambientes de contratação de energia.....	52
Figura 6 – Energia comercializada no Mercado de Curto Prazo.....	53
Figura 7 – Diagrama esquemático dos subsistemas (submercado) de energia elétrica.....	55
Figura 8 – Processo de decisão sistema hidrotérmico .....	57
Figura 9 – Funções de custo e o ponto ótima de operação.....	58
Figura 10 – Histórico PLD por submercado e patamar de carga.....	66
Figura 11 – <i>Boxplot</i> do PLD médio submercado Sudeste/Centro-Oeste.....	67
Figura 12 – Carga de Energia e Demanda Máxima.....	68
Figura 13 – Geração de energia hidráulica e térmica .....	69
Figura 14 – Intercâmbio de energia elétrica entre os subsistemas .....	70
Figura 15 – Energia Natural Afluyente e Armazenável .....	71
Figura 16 – Custo Marginal de Operação - patamar de carga médio.....	72
Figura 17 – Gráficos de dispersão entre PLD Médio SE/CO e variáveis predictoras.....	74
Figura 18 – ACF e PACF do PLD Médio SE/CO.....	77
Figura 19 – Distribuição de probabilidade dos erros residuais ARIMA(1,0,0) .....	78
Figura 20 – ACF e PACF dos erros residuais ARIMA(1,0,0) .....	79
Figura 21 – PLD Médio SE/CO - Separação dos subconjuntos.....	85
Figura 22 – Previsão do Modelo ETS para a primeira semana .....	86
Figura 23 – Previsão do modelo ARIMA para a primeira semana .....	88
Figura 24 – Previsão do modelo SARIMA para a primeira semana .....	89
Figura 25 – Previsão do modelo MLP para a primeira semana.....	91
Figura 26 – Previsão do modelo MLP multivariado para a primeira semana .....	93
Figura 27 – Comparação do desempenho dos modelos na métrica MAE.....	94
Figura 28 – Comparação do desempenho dos modelos na métrica RMSE.....	95
Figura 29 – Comparação do desempenho dos modelos na métrica MAPE.....	95
Figura 30 – Erros de previsão do modelo MLP multivariado .....	97
Figura 31 – Histogramas dos erros de previsão do modelo MLP multivariado .....	98
Figura 32 – Previsão do modelo MLP Multivariado para a 12ª semana .....	99



## LISTA DE TABELAS

Tabela 1 – Capacidade de geração do Brasil.....	54
Tabela 2 – Estatística descritiva PLD (R\$/MWh) no patamar médio.....	67
Tabela 3 – Correlações entre variáveis preditoras e o PLD.....	73
Tabela 4 – Hiperparâmetros do Modelo Holt-Winters.....	76
Tabela 5 – Variação da quantidade de neurônios por camada oculta.....	80
Tabela 6 – Configuração hiperparâmetros MLP.....	81
Tabela 7 – Medidas de dispersão para os subconjuntos de treino e teste (R\$/MWh).....	84
Tabela 8 – Métricas de desempenho do modelo ETS.....	86
Tabela 9 – Métricas de desempenho do modelo ARIMA(1,0,2).....	88
Tabela 10 – Métricas de desempenho do modelo SARIMA(1,0,2)(2,1,3) <sub>6</sub> .....	89
Tabela 11 – Métricas de desempenho do modelo MLP univariado.....	91
Tabela 12 – Configuração Entradas MLP multivariado.....	92
Tabela 13 – Métricas de desempenho do modelo MLP multivariado.....	93
Tabela 14 – U de Theil para cada um dos modelos de previsão.....	94

## LISTA DE ABREVIATURAS E SIGLAS

ACF	Função de Autocorrelação
ACL	Ambiente de Contratação Livre
ACR	Ambiente de Contratação Regulada
ADF	Dickey-Fuller Aumentado
AMFORP	American Foreign Power Company
ANEEL	Agência Nacional de Energia Elétrica
AR	Modelos Autorregressivos
ARIMA	Modelos Autorregressivos Integrados de Médias Móveis
CCEAR	Contratos de Comercialização de Energia Elétrica no Ambiente Regulado
CCEE	Câmara de Comercialização de Energia Elétrica
CEPEL	Centro de Pesquisas de Energia Elétrica
CMO	Custo Marginal de Operação
DP	Desvio Padrão
EAR	Energia Armazenada
Eletrobrás	Centrais Elétricas Brasileiras S.A.
ENA	Energia Natural Afluenta
EPE	Empresa de Pesquisa Energética
ETS	<i>Error, Trend, Seasonal</i>
FCF	Função de Custo Futuro
FCI	Função de Custo Imediato
FCT	Função de Custo Total
MA	Modelos Médias Móveis
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MCP	Mercado de Curto Prazo
MLP	<i>Multilayer Perceptron</i>
MLT	Média de Longo Termo
MME	Ministério de Minas e Energia
MP	Medida Provisória
MW	Megawatt
MWh	Megawatt-hora
N	Norte

NE	Nordeste
ONS	Operador Nacional do Sistema Elétrico
PAC	Programa de Aceleração do Crescimento
PACF	Funções de Autocorrelação parcial
PCA	<i>Principal Component Analysis</i>
PE	<i>Percentual Erro</i>
PLD	Preço de Liquidação das Diferenças
ReLU	<i>Rectified Linear Unit</i>
RMSE	<i>Root Mean Squared Error</i>
RNA	Redes Neurais Artificiais
RNN	Redes Neurais Naturais
RNR	Redes Neurais Recorrentes
S	Sul
SARIMA	Modelos Autorregressivos Integrados de Médias Móveis Sazonais
SE/CO	Sudeste/Centro-Oeste
SEB	Setor Elétrico Brasileiro
SIN	Sistema Interligado Nacional

## LISTA DE SÍMBOLOS

$Y_t$	Série temporal no período $t$
$S_t$	Sazonalidade da série temporal no período $t$
$T_t$	Tendência da série temporal no período $t$
$R_t$	Resíduos da série temporal no período $t$
$\hat{Y}_t$	Valor previsto para o instante $t$
$\alpha$	Parâmetro de suavização para o nível
$\beta$	Parâmetro de suavização para a tendência
$L_t$	Nível da série temporal no período $t$
$\varphi$	Parâmetro de amortecimento
$\gamma$	Parâmetro de suavização de sazonalidade
$M$	Frequência da sazonalidade
$Cov_{XZ}$	Covariância entre as variáveis $X$ e $Y$
$Corr_{XZ}$	Correlação entre as variáveis $X$ e $Y$
$\bar{X}$	Média da variável $X$
$S_X$	Desvio Padrão da variável $X$
$c_k$	Autocorrelação para a instância de atraso $k$
$\phi_{kk}$	Autocorrelação parcial para a instância de atraso $k$
$p$	Ordem modelo autorregressivo
$q$	Ordem modelo média móveis
$\phi_i$	Coefficientes do modelo autorregressivo
$\theta_j$	Coefficientes do modelo média móveis
$\varepsilon_t$	Processo aleatório
$\sigma^2$	Variância
$d$	Ordem de diferenciação modelo ARIMA
$B$	Operador de defasagem
$\mu$	Média do processo estocástico
$\Delta_s^D$	Operador de diferença sazonal $s$ de ordem $D$
$P$	Ordem modelo autorregressivo sazonal
$D$	Ordem modelo média móveis sazonal
$Q$	Ordem de diferenciação modelo SARIMA
$\Phi$	Coefficientes do modelo autorregressivo sazonal

$\Theta$	Coefficientes do modelo média móvel sazonal
$w_{kj}$	Peso sináptico j do neurônio k
$b_k$	Bias do neurônio k
$y_k$	Saída do neurônio k
$\varphi(\cdot)$	Função de ativação
$\bar{h}_k$	Vetor da saída da camada oculta k
$W$	Matriz de pesos sinápticos
$\bar{x}$	Vetor do sinal de entrada da rede neural
$\bar{y}$	Vetor do sinal de saída da rede neural
$E$	Função de erro
$\eta$	Taxa de aprendizagem da rede neural
$X_t$	Valor normalizado da série no momento t
$e$	Erro
$U$	Coefficiente U de Theil

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>14</b>
1.1	CONTEXTO E MOTIVAÇÃO .....	14
1.2	OBJETIVOS .....	16
1.3	ESTRUTURA DA DISSERTAÇÃO.....	17
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> .....	<b>18</b>
2.1	MÉTODOS PARA PREDIÇÃO DE SÉRIES TEMPORAIS .....	18
<b>2.1.1</b>	<b>Séries Temporais</b> .....	<b>18</b>
<b>2.1.2</b>	<b>Modelo de Médias Móveis Simples</b> .....	<b>22</b>
<b>2.1.3</b>	<b>Modelos de Suavização Exponencial</b> .....	<b>23</b>
2.1.3.1	Suavização Exponencial Simples .....	24
2.1.3.2	Suavização Exponencial Dupla.....	25
2.1.3.2.1	<i>Suavização Exponencial com Tendência Aditiva</i> .....	25
2.1.3.2.2	<i>Suavização Exponencial com Tendência Multiplicativa</i> .....	26
2.1.3.3	Suavização Exponencial Tripla.....	27
2.1.3.3.1	<i>Modelo Aditivo de Holt-Winters</i> .....	28
2.1.3.3.2	<i>Modelo Multiplicativo de Holt-Winters</i> .....	29
2.1.3.3.3	<i>Modelo Holt-Winters Amortecido</i> .....	29
<b>2.1.4</b>	<b>Modelos de Box-Jenkins</b> .....	<b>30</b>
2.1.4.1	Modelos Autorregressivos Integrados de Médias Móveis.....	33
2.1.4.2	Modelos Sazonais Autorregressivos Integrados de Médias Móveis.....	35
<b>2.1.5</b>	<b>Redes Neurais Artificiais</b> .....	<b>36</b>
2.1.5.1	Perceptron de Múltiplas Camadas.....	41
2.2	O SETOR ELÉTRICO BRASILEIRO .....	45
<b>2.2.1</b>	<b>Breve Histórico</b> .....	<b>45</b>
<b>2.2.2</b>	<b>Mercado de Energia Elétrica Brasileiro</b> .....	<b>51</b>
<b>2.2.3</b>	<b>O Sistema Interligado Nacional</b> .....	<b>54</b>
<b>2.2.4</b>	<b>Planejamento da Operação em Sistemas Hidrotérmicos</b> .....	<b>56</b>
<b>2.2.5</b>	<b>Cálculo do Preço de Liquidação das Diferenças</b> .....	<b>59</b>
2.3	PESQUISA BIBLIOGRÁFICA.....	60
<b>3</b>	<b>MATERIAIS E MÉTODOS DE PESQUISA</b> .....	<b>63</b>
3.1	DEFINIÇÃO DO PROBLEMA.....	64
3.2	COLETA E ANÁLISE PRELIMINAR DOS DADOS .....	65

3.3	SELEÇÃO E AJUSTE DO MODELO.....	75
<b>3.3.1</b>	<b>Modelos Holt-Winters .....</b>	<b>75</b>
<b>3.3.2</b>	<b>Modelos de Box-Jenkins .....</b>	<b>76</b>
3.3.2.1	Modelo ARIMA.....	76
3.3.2.2	Modelo SARIMA.....	79
<b>3.3.3</b>	<b>Modelos Multilayer Perceptron.....</b>	<b>80</b>
3.4	USO E VALIDAÇÃO DO MODELO .....	82
<b>4</b>	<b>RESULTADOS E ANÁLISES.....</b>	<b>84</b>
4.1	SEPARAÇÃO DOS SUBCONJUNTOS.....	84
4.2	MODELO HOLT-WINTERS.....	85
4.3	MODELOS BOX-JENKINS .....	87
<b>4.3.1</b>	<b>Modelo ARIMA.....</b>	<b>87</b>
<b>4.3.2</b>	<b>Modelo SARIMA.....</b>	<b>89</b>
4.4	MODELOS MULTILAYER PERCEPTRON.....	90
<b>4.4.1</b>	<b>Modelo MLP Univariado.....</b>	<b>91</b>
<b>4.4.2</b>	<b>Modelo MLP Multivariado .....</b>	<b>92</b>
4.5	COMPARAÇÃO DE DESEMPENHO DOS MODELOS.....	94
4.6	ANÁLISE DOS ERROS DE PREVISÃO.....	96
<b>5</b>	<b>CONCLUSÕES E SUGESTÕES.....</b>	<b>100</b>
5.1	CONCLUSÕES .....	100
5.2	SUGESTÕES DE TRABALHOS FUTUROS .....	102
	<b>REFERÊNCIAS .....</b>	<b>103</b>

# 1 INTRODUÇÃO

Esse capítulo contextualiza o trabalho, apresentando a proposta de estudo comparativo de métodos preditivos de séries temporais na previsão do preço da energia no mercado de curto prazo brasileiro, trazendo a motivação para essa aplicação, o objetivo geral e específicos, e também a estrutura do trabalho.

## 1.1 CONTEXTO E MOTIVAÇÃO

Métodos capazes de prever quantitativamente uma variável e mensurar o valor dessa previsão têm sido cada vez mais aplicados nos diversos setores da sociedade. Por meio desses instrumentos podem-se prever situações futuras para planejar, determinar e dimensionar recursos de maneira a tentar diminuir gastos com medidas equivocadas.

As previsões são problemas importantes que abrangem muitos campos, incluindo negócios e indústria, governo, economia, ciências ambientais e sociais, medicina, engenharia, política e finanças (MONTGOMERY; JENNINGS; KULAHCI, 2015).

Dentre as diversas aplicações possíveis, as previsões se diferenciam pelo horizonte de tempo de projeção. Problemas de previsão de curto prazo envolvem a previsão de eventos apenas em alguns períodos de tempo (dias, semanas e meses) no futuro. Já as previsões de médio prazo podem se estender até dois anos no futuro, e os problemas de previsão de longo prazo são considerados aqueles que se estendem por anos (MONTGOMERY; JENNINGS; KULAHCI, 2015).

As previsões de curto e médio prazo são necessárias para atividades que vão desde o gerenciamento de operações até o orçamento e a seleção de novos projetos de pesquisa e desenvolvimento. Já as previsões de longo prazo impactam questões como o planejamento estratégico (MONTGOMERY; JENNINGS; KULAHCI, 2015).

Não raro, resultados de previsões de longo prazo requerem investimentos em infraestrutura, principalmente àqueles referentes ao crescimento de algum tipo de demanda. Um exemplo é a previsão do mercado potencial de veículos elétricos (JENSEN et al., 2017). É preciso garantir infraestrutura adequada de atendimento à essa expectativa de demanda



crecente, como, por exemplo, o desenvolvimento de metodologias para locação de facilidades, com a definição geográfica dos locais favoráveis à implementação de postos de carregamento de veículos elétricos (NASCIMENTO; FERREIRA, 2020), prever o impacto desse crescimento no sistema elétrico (KAPUSTIN; GRUSHEVENKO, 2020) e, ainda, considerar o arcabouço regulatório necessário (WADY, 2021).

As previsões de curto e médio prazo normalmente baseiam-se na identificação, modelagem e extrapolação dos padrões encontrados em dados históricos, dos quais a utilização de métodos estatísticos são úteis para esse tipo de previsão (MONTGOMERY; JENNINGS; KULAHCI, 2015). A maior parte desses problemas, portanto, envolve o uso de dados de séries temporais, que corresponde à uma sequência cronológica ou orientada no tempo de observações de uma variável de interesse.

Há uma variedade de casos para os quais modelos de previsão por séries temporais se fazem úteis e ainda diversos métodos preditivos disponíveis, cada qual para um comportamento de dados específico, bem como para janelas de previsões diferentes, de tal forma que a escolha de um modelo em detrimento a outro pode gerar resultados distintos.

Nesse contexto, este trabalho se propõe ao estudo de métodos de previsão de séries temporais aplicados na previsão do preço de energia elétrica no mercado de curto prazo brasileiro.

O mercado de energia elétrica é essencial para o adequado funcionamento e desenvolvimento de qualquer economia, devido sua relevância em boa parte dos setores econômicos de um país (MEDEIROS, 2003). Modelos de previsão dos preços de energia elétrica devem levar em consideração a estrutura e o funcionamento do mercado de energia em que estão inseridos. Cada mercado possui uma particularidade e diferentes variáveis a serem consideradas na formação dos preços.

O mercado de curto prazo brasileiro é responsável por proceder à contabilização e liquidação financeira das diferenças apuradas entre os montantes de energia elétrica contratados e consumidos ou gerados pelos agentes e registrados na Câmara de Comercialização de Energia Elétrica. Essas liquidações são parametrizadas pelo Preço de Liquidação das Diferenças (PLD).

O valor do PLD é derivado do planejamento da operação de geração de energia do sistema elétrico brasileiro, majoritariamente hidrotérmico, em que os modelos matemáticos utilizados para calcular os despachos ótimos das usinas baseiam-se nas condições hidrológicas presente e futura, na demanda de energia, nos preços de combustível, no custo de déficit, na entrada de novos projetos e na disponibilidade dos sistemas de geração e transmissão (CCEE, 2021).

A formação do PLD, portanto, está atrelada à minimização dos custos marginais de operação, onde os preços representam os custos do último megawatt a ser suprido e não necessariamente uma formação de preço por oferta e demanda. Assim, toda a estrutura de determinação dos valores do PLD e as características da matriz energética brasileira trazem uma elevada volatilidade aos preços, levando os agentes do mercado a investirem em alternativas de gerenciamento de risco de seu portfólio de contratos.

Com isso, a previsão dos preços de energia auxilia, portanto, os tomadores de decisão a gerenciar seus contratos a curto, médio e longo prazos de compra e venda de energia elétrica.

## 1.2 OBJETIVOS

O objetivo geral desse trabalho é comparar o desempenho de metodologias preditivas de séries temporais na previsão do valor do Preço de Liquidações das Diferenças para um horizonte de até 12 semanas.

Foram considerados os modelos de suavização exponencial ou Holt-Winters, modelos Box-Jenkins ARIMA e SARIMA, modelos de redes neurais artificiais *Multilayer Perceptrons* univariados e multivariados.

Como objetivos específicos tem-se:

- a) identificar, caracterizar e ajustar os modelos de previsão abordados no trabalho;
- b) comparar os resultados dos modelos em relação à previsão ingênua ou trivial e avaliar seus desempenhos por meio de métricas de erros de previsão;
- c) identificar as variáveis de entrada ótimas para o modelo de rede neural artificial multivariado para cada uma das 12 semanas de previsão;
- d) realizar análise dos erros de previsão do modelo com melhor desempenho.

### 1.3 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está estruturada em 5 capítulos.

O Capítulo 2 apresenta o referencial teórico dos modelos de previsão de séries temporais utilizados no trabalho. Nesse capítulo, é feita uma contextualização do setor elétrico brasileiro, apresentando seu histórico e reformas estruturais com vistas a trazer as bases para o mercado de energia e a precificação do Preço de Liquidação das Diferenças. Por fim, é exibida a pesquisa bibliográfica a respeito da previsão do preço da energia por séries temporais, incluídos trabalhos específicos de predição do PLD.

No Capítulo 3 são promovidas a caracterização e contextualização do estudo no que se refere aos materiais e métodos de pesquisa utilizados na obtenção dos objetivos do trabalho. No Capítulo 3 também são apresentadas as etapas definidas para previsão de séries temporais, contemplando a definição do problema, tratamento dos dados, ajustes dos parâmetros dos modelos e apresentação das métricas de avaliação de desempenho de previsão.

No Capítulo 4 são apresentados e analisados os resultados obtidos por cada um dos modelos, a comparação do desempenho dos modelos para cada uma das métricas de erro e ainda é promovida a avaliação dos erros do modelo com melhor desempenho.

Por fim, no Capítulo 5 é feito o relato das conclusões mais relevantes e sugestões para a continuidade de estudo sobre o tema.

## 2 REFERENCIAL TEÓRICO

Este capítulo apresenta a fundamentação teórica dos métodos estatísticos de previsão de séries temporais, tanto de modelos univariados como a suavização exponencial (HYNDMAN; ATHANASOPOULOS, 2018) e modelos Box-Jenkins (BOX et al., 2015), bem como o multivariado com a utilização de redes neurais artificiais (HAYKIN, 2010).

Além disso, como forma de dar embasamento para o estudo de caso da aplicação dos modelos na previsão do Preço de Liquidações das Diferenças (preço *spot* da energia) no mercado de curto prazo de energia elétrica, é apresentada a contextualização do sistema elétrico brasileiro.

### 2.1 MÉTODOS PARA PREDIÇÃO DE SÉRIES TEMPORAIS

#### 2.1.1 Séries Temporais

Uma série temporal é uma coleção de observações feitas sequencialmente no tempo (BOX et al., 2015). Portanto, uma série temporal adiciona uma dependência de ordem explícita entre as observações: a dimensão de tempo. Essa dimensão adicional é uma restrição e uma estrutura que fornece uma fonte de informações. A característica mais importante deste tipo de dados é que as observações adjacentes são dependentes e há, portanto, o interesse em analisar e modelar esta dependência.

Segundo Morettin e Tolo (2006), ao se estudar uma série temporal, busca-se a análise e modelagem da série temporal (descrever a série, verificar suas características mais relevantes e suas possíveis relações com outras séries) por meio de seus valores históricos, procura-se estimar as previsões (*forecast*). O número de instantes à frente para o qual é feita a previsão é chamado de horizonte de previsão.

Segundo Shmueli e Lichtendahl Jr (2016) na modelagem descritiva ou análise de série temporal, uma série temporal é modelada para determinar seus componentes em termos de

padrões sazonais, tendências, relação com fatores externos e semelhantes. Em contraste, a previsão de série temporal usa as informações para prever os valores futuros dessa série.

Ao se usar estatística clássica, a principal preocupação é a análise de séries temporais. A análise de séries temporais envolve o desenvolvimento de modelos que melhor capturam ou descrevem o comportamento da série. Isso geralmente envolve fazer suposições sobre a forma dos dados e decompor as séries temporais em componentes de constituição. A qualidade de um modelo descritivo é determinada por quão bem ele descreve todos os dados disponíveis e pela interpretação que ele fornece sobre o domínio do problema.

O objetivo principal da análise de série temporal é desenvolver modelos matemáticos que forneçam descrições plausíveis dos dados da amostra (SHUMWAY; STOFFER; STOFFER, 2000).

Fazer previsões sobre o futuro é chamada de extrapolação no tratamento estatístico clássico de dados. A previsão envolve o ajuste de modelos em dados históricos e o uso deles para prever observações futuras. Uma distinção importante na previsão é que o futuro está completamente indisponível e só deve ser estimado a partir do que já aconteceu.

A habilidade de um modelo de previsão de série temporal é determinada por seu desempenho em prever o futuro. Muitas vezes, isso ocorre às custas de ser capaz de explicar por que uma previsão específica foi feita, os intervalos de confiança e ainda compreender as causas subjacentes ao problema.

Nos estudos de séries temporais é importante a identificação e análise de padrões comuns a esses tipos de dados. De acordo com Box et al. (2015), um método tradicional de análise de séries temporais é decompor a série arbitrariamente em três componentes: tendência, componente sazonal (normalmente analisado juntamente com a ciclicidade) e um componente residual ou aleatório, a saber:

- a) tendência: indica o seu comportamento de longo prazo, isto é, se o valor da variável cresce, decresce ou permanece estável, podendo ser ou não linear;
- b) sazonalidade: está ligada às oscilações de subida ou quedas e também à representatividade ocorrida em determinados períodos, como anual, mensal, diários, etc. (a diferença entre o efeito sazonal e cíclico é que o primeiro indica a repetição de um padrão de uma série dentro do período de um ano, já os ciclos indicam padrões que se repetem em períodos superiores a um ano);
- c) residual: baseadas em mudanças irregulares na série temporal, que não são causadas por outro componente em análise; considerados aleatórios e

imprevisíveis, esses efeitos são representados pela diferença entre o dado original e os dois outros componentes.

Portanto, a decomposição de uma série temporal ( $Y_t$ ) nos componentes tradicionais apresentados, pode ser representada matematicamente pela seguinte função:

$$Y_t = f(S_t, T_t, R_t), \quad (1)$$

em que,  $S_t$  representa a sazonalidade,  $T_t$  a tendência e  $R_t$  a componente residual, no instante de tempo  $t$ . Essa decomposição pode apresentar uma relação aditiva ou multiplicativa entre os termos padrões, conforme Equações (2) e (3), respectivamente.

$$Y_t = S_t + T_t + R_t, \quad (2)$$

$$Y_t = S_t \times T_t \times R_t. \quad (3)$$

Pode-se pensar que estes componentes constituintes se combinam para formar as séries temporais observadas. Suposições podem ser feitas sobre estes componentes, tanto no que se refere ao comportamento quanto em como eles são combinados, o que permite que eles sejam modelados usando métodos estatísticos tradicionais. Estes componentes também podem ser a maneira mais eficaz de fazer previsões sobre valores futuros. Nos casos em que esses métodos clássicos não resultam em um desempenho eficaz, estes componentes ainda podem apresentar conceitos úteis e até mesmo entradas para métodos alternativos.

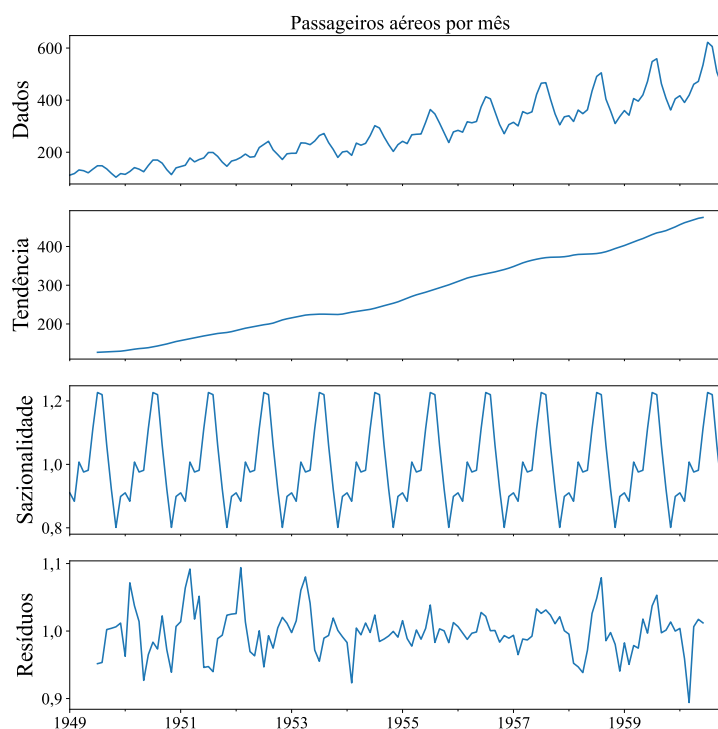
Um exemplo de uma decomposição aditiva de uma série temporal é apresentado na Figura 1. O conjunto de dados<sup>1</sup> apresenta a quantidade total mensal de passageiros em voos internacionais nos Estados Unidos entre 1949 a 1960 (EHLERS, 2007). Esses dados foram originalmente usados por Box e Jenkins em seu trabalho em 1976.

É possível observar que a decomposição permite realizar a separação de cada um dos componentes citados: tendência, sazonalidade e os componentes irregulares ou aleatórios. A série original (dados) é a somatória de cada um desses componentes, conforme Equação (2).

---

<sup>1</sup> Dados obtidos em: <https://c4science.ch/diffusion/2096/browse/master/input/international-airline-passengers.csv> acessado em 22 de junho de 2021.

Figura 1 – Decomposição de uma série temporal



Fonte: Elaboração Própria.

Os modelos utilizados para descrever séries temporais são processos estocásticos, portanto, o controle é realizado por meio de leis probabilísticas (MORETTIN; TOLOI, 2006).

Quanto à variabilidade das observações, podem-se classificar as séries temporais em estacionárias, quando as suas estatísticas não são afetadas por variações no tempo, e ergódicas, se apenas uma realização do processo estocástico é suficiente para se obter todas as estatísticas da mesma (MORETTIN; TOLOI, 2006).

De acordo com Hyndman e Athanasopoulos (2018), para o estudo de séries temporais, são necessárias as seguintes etapas<sup>2</sup>: definição do problema, coleta de informações, análise preliminar dos dados, escolha, ajuste, uso e avaliação do modelo de previsão.

A maioria dos métodos de previsão baseiam-se na ideia de que as observações passadas contêm informações sobre o padrão de comportamento da série temporal. O propósito dos métodos é distinguir o padrão de qualquer ruído que possa estar contido nas observações e então usar esse padrão para prever valores futuros da série. Técnicas específicas desse tipo assumem que os valores extremos da série representam aleatoriedade e assim, por meio da suavização desses extremos, pode-se identificar o padrão básico (MORETTIN; TOLOI, 2006).

<sup>2</sup> Essas etapas serão melhor detalhadas no Capítulo 3 MATERIAIS E MÉTODOS DE PESQUISA

A seguir são apresentados alguns modelos probabilísticos de previsão de séries temporais.

### 2.1.2 Modelo de Médias Móveis Simples

O modelo de médias móveis simples é um tipo comum de suavização usado em análises e previsões de séries temporais. Uma média móvel requer que você especifique um tamanho de janela que define o número de observações brutas usadas para calcular o valor da média. A parte móvel refere-se ao fato de que essa janela é deslizada ao longo da série temporal para calcular os valores médios.

A média móvel pode ser usada como uma técnica de preparação de dados para criar uma versão suavizada do conjunto de dados original. A suavização é útil na preparação dos dados, pois pode reduzir a variação aleatória nas observações e expor melhor a estrutura dos processos causais subjacentes.

O valor da média móvel também pode ser usado diretamente para fazer previsões, nesse caso é conhecido como método da persistência. Nesse modelo assume-se que os componentes de tendência e sazonalidade da série histórica já foram removidos ou ajustados. O modelo de média móvel para previsões pode ser facilmente usado de maneira progressiva: conforme novas observações são disponibilizadas o modelo pode ser atualizado para ser utilizado na próxima previsão (BRESEGHELLO, 2005).

Os modelos de médias móveis geram previsões médias com menor variabilidade que os dados originais, devido ao processo de combinação entre as observações com valores altos e baixos.

Segundo Da Silva et al. (2018), a aplicação de média móvel sobre uma série temporal  $Y$  para a previsão no período  $t$  ( $\hat{Y}_t$ ) pode ser obtida pela média de  $n$  elementos da série anteriores ao período  $t$ , conforme apresentado na Equação (4) a seguir:

$$\hat{Y}_t = \frac{1}{n} \sum_{j=1}^n Y_{t-j}. \quad (4)$$

De acordo com Breseghello (2005), o tamanho da janela  $n$  de períodos incluídos no cálculo da média móvel determina sua sensibilidade com relação aos dados mais recentes.



Períodos pequenos proporcionam uma reação maior a possíveis mudanças no padrão dos dados e grandes períodos, por sua vez, produzem uma média mais homogênea.

No caso particular em que  $n$  é igual a  $1$ , o método da persistência corresponde ao método de previsão trivial ou ingênuo (*naive*).

O modelo de previsão ingênua ou trivial supõe simplesmente que a previsão para o período  $t$  é exatamente o valor observado no período  $t - 1$ . Tal modelo é conhecido, também como modelo do caminho aleatório (*random walk*) e é formulado de acordo com a Equação (5).

$$\hat{Y}_t = Y_{t-1}, \quad (5)$$

onde  $\hat{Y}_t$  é a previsão para o período  $t$  e  $Y_{t-1}$  é o valor real observado no período  $t - 1$ .

Esse método pode servir de parâmetro para comparação em relação ao coeficiente  $U$  de Theil, utilizado como análise de decisão sobre a adequação de um modelo em relação aos seus erros, além de ser útil como uma medida que compara os modelos com as metodologias triviais de previsão.

### 2.1.3 Modelos de Suavização Exponencial

A suavização exponencial é um método de previsão de série temporal para dados univariados. Enquanto na média móvel simples as observações anteriores são ponderadas igualmente, as funções exponenciais são usadas para atribuir pesos decrescentes exponencialmente ao longo do tempo.

Portanto, as previsões produzidas usando métodos de suavização exponencial são médias ponderadas de observações anteriores, com os pesos decaindo exponencialmente à medida que as observações ficam mais antigas. Em outras palavras, quanto mais recente a observação, maior o peso associado (HYNDMAN; ATHANASOPOULOS, 2018).

Os métodos de suavização exponencial podem ser considerados uma alternativa à popular classe de métodos ARIMA<sup>3</sup> de Box-Jenkins para previsão de séries temporais. Coletivamente, os métodos às vezes são chamados de modelos ETS (*Error, Trend, Seasonal*), referindo-se à modelagem explícita de erro, tendência e sazonalidade.

---

<sup>3</sup> O modelo ARIMA será apresentado na seção 2.1.4.1

Existem três tipos principais de métodos de previsão de séries temporais de suavização exponencial. Um método simples que não assume nenhuma estrutura sistemática, uma extensão que lida explicitamente com as tendências e a abordagem mais avançada que adiciona suporte para sazonalidade.

### 2.1.3.1 Suavização Exponencial Simples

A suavização exponencial simples, é um método de previsão de série temporal para dados univariados sem tendência ou sazonalidade. Requer um único parâmetro  $\alpha$ , denominado fator (ou coeficiente) de suavização. Este parâmetro controla a taxa na qual a influência das observações em etapas de tempo anteriores decai exponencialmente.

Segundo Hyndman e Athanasopoulos (2018), a representação matemática para suavização exponencial simples pode ser dada pela Equação (6).

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha) \hat{Y}_t, \quad (6)$$

onde  $\hat{Y}_{t+1}$  é a previsão da série temporal para o período  $t+1$ ,  $\hat{Y}_t$  e  $Y_t$  são, respectivamente, o valor da previsão e o valor observado para a série temporal, ambas no período  $t$ . O peso utilizado na ponderação é  $\alpha$ , ( $0 < \alpha < 1$ ).

A fim de demonstrar a distribuição de peso entre os valores mais recentes e os valores mais antigos da série temporal, a Equação (6) pode ser expandida para:

$$\begin{aligned} \hat{Y}_{t+1} = & \alpha Y_t + \alpha(1 - \alpha)Y_{t-1} + \alpha(1 - \alpha)^2 Y_{t-2} + \alpha(1 - \alpha)^3 Y_{t-3} + \\ & \alpha(1 - \alpha)^4 Y_{t-4} + \dots + \alpha(1 - \alpha)^{t-1} Y_1 + \alpha(1 - \alpha)^t Y_0. \end{aligned} \quad (7)$$

Segundo Shmueli e Lichtendahl Jr (2016), valores altos de  $\alpha$  (próximo de 1) produzem previsões que dependem mais das observações mais recentes (indica aprendizado rápido), enquanto que valores baixos de  $\alpha$  (próximo de 0) produzem previsões que dependem mais das observações mais antigas (indica aprendizado lento). Se o valor de  $\alpha$  for igual a 1, temos o método de previsão ingênuo, onde a previsão é dada apenas pela última observação.

Tanto para esta suavização quanto para as demais ao longo desta seção, os valores das constantes ótimas são encontrados iterando-se valores até minimizar algum tipo de erro, os quais serão discutidos a posteriori.

### 2.1.3.2 Suavização Exponencial Dupla

A suavização exponencial dupla é uma extensão da suavização exponencial que adiciona explicitamente suporte para tendências na série temporal univariada. Além do parâmetro  $\alpha$  para controlar o fator de suavização para o nível, o fator de suavização  $\beta$  é adicionado para controlar a queda da influência da mudança na tendência. O nível é uma estimativa suavizada do valor dos dados no final de cada período. Já a tendência é uma estimativa suavizada do crescimento médio no final de cada período.

O método suporta tendências que mudam de maneiras diferentes: um aditivo e um multiplicativo, dependendo se a tendência é linear ou exponencial, respectivamente.

#### 2.1.3.2.1 Suavização Exponencial com Tendência Aditiva

A suavização exponencial dupla com uma tendência aditiva é classicamente chamada de modelo de tendência linear de Holt, em homenagem ao desenvolvedor do método Charles Holt. Segundo Hyndman e Athanasopoulos (2018), esse método envolve três equações, a Equação (8) para a previsão da série temporal para os próximos  $k$  períodos e duas equações de suavização, a Equação (9) para a estimativa do nível  $L_t$  (*level*) e a Equação (10) para a tendência  $T_t$  (*trend*), ambas aditivas suavizadas no final do período  $t$ .

$$\hat{Y}_{t+k} = L_t + kT_t, \quad (8)$$

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (9)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (10)$$

em que  $\beta$  é a constante de suavização para a tendência,  $0 \leq \beta \leq 1$ ,  $\alpha$  é a constante de suavização para o nível,  $0 \leq \alpha \leq 1$ ;  $T_{t-1}$  e  $L_{t-1}$  representam, respectivamente, a tendência e o nível suavizados no período anterior ( $t-1$ ).

Assim como na suavização exponencial simples, o método de Holt requer valores iniciais, nesse caso  $L_0$  e  $T_0$ .

Um problema evidente da Equação (8) é que os resultados das previsões aumentam indefinidamente conforme o horizonte de tempo aumenta. Contudo, na prática normalmente existe um limite para o crescimento de determinado dado. Portanto, é possível adaptar o método linear de Holt para que ele cause um amortecimento com o passar do horizonte de previsão, de modo a representar de maneira mais realista o comportamento da série.

De acordo com Hyndman e Athanasopoulos (2018), as equações de Holt com a inclusão do parâmetro de amortecimento ( $0 < \phi < 1$ ) são:

$$\hat{Y}_{t+k} = L_t + (\phi + \phi^2 + \dots + \phi^k)T_t, \quad (11)$$

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + \phi T_{t-1}), \quad (12)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\phi T_{t-1}. \quad (13)$$

Se  $\phi = 1$ , o método é idêntico ao modelo linear de Holt. Para valores entre 0 e 1,  $\phi$  amortece a tendência para que se aproxime de uma constante em algum momento futuro. Na verdade, as previsões convergem para  $L_t + \phi T_t / (1 - \phi)$  quando  $k \rightarrow \infty$  para qualquer valor de  $\phi$  entre 0 e 1. Isso significa que as previsões de curto prazo possuem tendência, enquanto as previsões de longo prazo são constantes.

#### 2.1.3.2.2 Suavização Exponencial com Tendência Multiplicativa

A suavização exponencial dupla com uma tendência multiplicativa assume que o nível muda de um período para outro por um determinado fator, em contraste ao modelo de tendência aditiva que assume que o nível muda de um período para outro por valor fixo.

De acordo com Shmueli e Lichtendahl Jr (2016), a suavização exponencial com uma tendência multiplicativa produz, portanto, previsões  $k$  passos à frente de acordo com a Equação (14).

$$\hat{Y}_{t+k} = L_t \times T_t^k, \quad (14)$$

em que, as Equações (15) e (16) representam, respectivamente, a atualização para o nível e a tendência ao final do período  $t$ .

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} \times T_{t-1}), \quad (15)$$

$$T_t = \beta(L_t/L_{t-1}) + (1 - \beta)T_{t-1}. \quad (16)$$

### 2.1.3.3 Suavização Exponencial Tripla

A suavização exponencial tripla é uma extensão da suavização exponencial que adiciona explicitamente suporte para sazonalidade à série temporal univariada. Esse método às vezes é chamado de Suavização exponencial de Holt-Winters, em homenagem a dois contribuintes do método: Charles Holt e Peter Winters. Além dos fatores de suavização  $\alpha$  e  $\beta$ , é adicionado o parâmetro  $\gamma$  que controla a influência no componente sazonal.

Tal como acontece com a tendência, a sazonalidade pode ser modelada como um processo aditivo ou multiplicativo para uma mudança linear ou exponencial na sazonalidade.

De acordo com Hyndman e Athanasopoulos (2018), o método aditivo é preferido quando as variações sazonais são aproximadamente constantes ao longo da série, enquanto o método multiplicativo é preferido quando as variações sazonais estão mudando proporcionalmente ao nível da série. Com o método aditivo, a componente sazonal é expressa em termos absolutos na escala das séries observadas e na equação de nível a série é ajustada sazonalmente subtraindo a componente sazonal. Com o método multiplicativo, a componente sazonal é expressa em termos relativos (percentagens) e a série é ajustada sazonalmente pela divisão da componente sazonal.

### 2.1.3.3.1 Modelo Aditivo de Holt-Winters

Assumindo uma série temporal com sazonalidade com frequência  $M$  para denotar o número de temporadas em um ano (por exemplo, para dados trimestrais  $M = 4$  e para dados mensais  $M = 12$ ), a previsão para o modelo aditivo de Holt-Winters é dado pela Equação (17) (SHMUELI; LICHTENDAHL JR, 2016).

$$\hat{Y}_{t+k} = L_t + kT_t + S_{t-M+k}. \quad (17)$$

Observa-se que, para produzir previsões usando esta fórmula, a série deve ter incluído pelo menos um ciclo completo de estações no tempo de previsão  $t$ , ou seja,  $t > M$ .

Por ser um método adaptativo, a suavização exponencial de Holt-Winters permite que os padrões de nível, tendência e sazonalidade mudem ao longo do tempo. Os três componentes são estimados e atualizados à medida que novas informações chegam. As equações de atualização para o nível, tendência e sazonalidade aditiva são fornecidas pela Equações (18), (19) e (20), respectivamente (SHMUELI; LICHTENDAHL JR, 2016).

$$L_t = \alpha(Y_t - S_{t-M}) + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (18)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (19)$$

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-M}. \quad (20)$$

A equação de nível mostra uma média ponderada entre a observação ajustada sazonalmente ( $Y_t - S_{t-M}$ ) e a previsão não sazonal ( $L_{t-1} + T_{t-1}$ ) para o tempo  $t$ .

A equação de tendência é idêntica ao método linear de Holt. A equação sazonal mostra uma média ponderada entre o índice sazonal atual,  $\gamma(Y_t - L_t)$ , e o índice sazonal da mesma temporada do período anterior.

### 2.1.3.3.2 Modelo Multiplicativo de Holt-Winters

Para a sazonalidade multiplicativa, tem-se:

$$Y_{t+k} = (L_t + kT_t)S_{t+k-M}, \quad (21)$$

$$L_t = \alpha Y_t / S_{t-M} + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (22)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}, \quad (23)$$

$$S_t = \gamma(Y_t / L_t) + (1 - \gamma)S_{t-M}. \quad (24)$$

A Equação (22) é semelhante à Equação (9) da suavização exponencial dupla, exceto que ela usa o valor ajustado sazonalmente no tempo  $t$  em vez do valor bruto. Isso é feito dividindo  $Y_t$  pelo seu índice sazonal, estimado no último ciclo. A Equação (23) é idêntica à Equação (10) de suavização exponencial dupla. A Equação (24) significa que o índice sazonal é atualizado tirando uma média ponderada do índice sazonal do ciclo anterior e o valor atual ajustado pela tendência.

### 2.1.3.3.3 Modelo Holt-Winters Amortecido

Assim como no método amortecido para o modelo linear de Holt, pelos mesmos motivos se justificam a aplicação para o modelo de Holt-Winters.

Utiliza-se a mesma constante de amortecimento  $\phi$  aplicada ao método Holt-Winters multiplicativo, de forma que a previsão é dada pela Equação (25), enquanto que as estimativas para o nível, tendência e sazonalidade são dadas pelas Equações (26), (27) e (28) respectivamente.

$$\hat{Y}_{t+k} = (L_t + (\phi + \phi^2 + \dots + \phi^k)kT_t)S_{t+k-M}, \quad (25)$$

$$L_t = \alpha Y_t / S_{t-M} + (1 - \alpha)(L_{t-1} + \phi T_{t-1}), \quad (26)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\phi T_{t-1}, \quad (27)$$

$$S_t = \gamma(Y_t / L_t) + (1 - \gamma)S_{t-M}. \quad (28)$$

#### 2.1.4 Modelos de Box-Jenkins

Os modelos de Box-Jenkins foram propostos por George Box e Gwilym Jenkins no início dos anos 70 (BOX et al., 2015).

Os modelos de Box-Jenkins partem da ideia de que os valores de uma série temporal são altamente dependentes, ou seja, cada valor pode ser explicado por valores prévios da série.

A estratégia para construção deste modelo é baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo baseia-se nos próprios dados (SILVA, 2010).

Antes de apresentar os modelos, alguns conceitos são necessários, como a definição de um processo estocástico, conforme apresentados na sequência.

Pode definir-se um processo estocástico como um conjunto de variáveis aleatórias indexadas a uma variável (geralmente a variável tempo), sendo representado por  $\{Z(t), t \in T\}$ . Para cada instante de tempo  $t$  tem-se que  $Z(t)$  é uma variável aleatória definida sobre o conjunto universo de probabilidades  $\Omega$ . Estabelecendo o paralelismo com o caso determinístico, onde uma função  $f(t)$  toma valores bem definidos ao longo do tempo, um processo estocástico toma valores aleatórios ao longo do tempo. Aos valores que  $Z(t)$  pode assumir chamam-se estados e ao seu conjunto  $Z$  espaço de estados.

Uma importante classe de modelos estocásticos utilizados na representação das séries temporais são os modelos estacionários. Tais modelos pressupõem um processo em equilíbrio, onde a família de variáveis se mantém em um nível médio constante (BOX et al., 2015).

Outra definição importante na adoção de um modelo de série temporal diz respeito à função de autocorrelação, por meio da qual é possível avaliar a relação entre as observações atuais e anteriores.



Para isso, utilizar-se-á o conceito e cálculo de covariância e correlação de dados. Ambos medem a relação linear entre variáveis. Assumindo-se duas variáveis como  $X$  e  $Z$ , tem-se, para a covariância, a seguinte expressão:

$$Cov_{XZ} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Z_i - \bar{Z}), \quad (29)$$

onde  $\bar{X}$  e  $\bar{Z}$  são as médias de  $X$  e  $Z$ , respectivamente, e  $n$  é o número de observações de cada variável.

Contudo, a covariância não é um valor parametrizado, de tal forma que seus valores podem ser de difícil interpretação e variam dependendo da base de dados analisada. Para contornar esta questão existe a correlação expressa a seguir:

$$Corr_{XZ} = \frac{Cov_{XZ}}{S_X S_Z} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Z_i - \bar{Z})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Z_i - \bar{Z})^2}}, \quad (30)$$

onde  $S_X$  e  $S_Z$  são os desvios padrão das observações. Com isso, obtêm-se uma escala de relação linear com valores entre  $-1$  e  $+1$ .

Assim como a covariância e a correlação são usadas para analisar a relação entre duas variáveis, a autocorrelação estuda a correlação entre os dados da mesma variável. Tal estudo é de extrema importância na análise de séries temporais nos cálculos de projeções, e é expressa por  $c_k$  para a instância de atraso  $k$ , de acordo com a Equação (31) (BOX et al., 2015).

$$c_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}. \quad (31)$$

A autocorrelação permite compreender melhor o comportamento dos dados temporais, podendo, inclusive, ser utilizado para encontrar sazonalidades na série.

Séries temporais que não apresentam autocorrelação são denominadas como ruído branco. Tal categorização se dá quando o valor da autocorrelação está dentro dos limites definidos como  $\pm 1,96/\sqrt{n}$ , de acordo com Hyndman e Athanasopoulos (2018).

Semelhante à autocorrelação, a autocorrelação parcial também permite a análise da relação entre os valores de uma série temporal. No entanto, a autocorrelação parcial,

representado por  $\phi_{kk}$ , mede o grau de associação entre  $Y_t$  e  $Y_{t-k}$  quando os efeitos dos outros atrasos são removidos (BOX et al., 2015).

Com os conceitos básicos, é possível apresentar as etapas do processo iterativo de construção do modelo estocástico de Box-Jenkins, conforme Box et al. (2015):

- a) identificação: certificando-se de que as variáveis são estacionárias, identificando sazonalidade na série dependente (diferenciando-a sazonalmente, se necessário), usa-se os dados e todas as informações relacionadas para ajudar a selecionar uma subclasse de modelo que possa resumir melhor os dados; as principais ferramentas utilizadas no processo de identificação do modelo são a de funções de autocorrelação (ACF) e autocorrelação parcial (PACF);
- b) estimativa: use os dados para treinar os parâmetros do modelo, utilizando métodos numéricos para chegar aos coeficientes que melhor se ajustem ao modelo selecionado, os métodos mais comuns utilizam estimativa de máxima verossimilhança<sup>4</sup> ou estimativa de mínimos quadrados não linear;
- c) verificação de diagnóstico: avalie o modelo ajustado no contexto dos dados disponíveis e verifique as áreas onde o modelo pode ser melhorado. Um método para a verificação do modelo utiliza o cálculo dos resíduos ( $e_t = Y_t - \hat{Y}_t$ ). Em particular, os resíduos devem ser independentes uns dos outros e constantes na média e na variância ao longo do tempo (traçar a autocorrelação e a autocorrelação parcial dos resíduos são úteis para identificar erros de especificação). Se a estimativa for inadequada, volta-se à etapa de identificação para tentar construir um modelo melhor.

Uma vez determinado o melhor modelo para a série temporal, pode-se utilizá-lo para gerar previsões. Partindo-se de um período  $t$ , e supondo que se deseja prever a série em um período futuro,  $\hat{Y}_{t+k}$  representa a previsão para o período  $k$  feita a partir de  $t$ .

Dentre os processos estocásticos importantes da metodologia de Box-Jenkins, destaca-se os modelos autorregressivos (AR), médias móveis (MA) e combinações mistas desses, tais como os modelos ARMA, ARIMA e SARIMA. Todos eles são paramétricos, ou seja, possuem um número finito de parâmetros cujos valores são estimados a partir do sinal ou série temporal sob análise (FILHO, 2014).

---

<sup>4</sup> ELIASON, S. R. Maximum Likelihood Estimation: Logic and Practice. Newbury Park: Sage, 1993.

### 2.1.4.1 Modelos Autorregressivos Integrados de Médias Móveis

Um modelo autorregressivo integrado de médias móveis (ARIMA) é uma classe de modelo estatístico para análise e previsão de dados de séries temporais. ARIMA é uma sigla que significa *Autoregressive Integrated Moving Average*. Este acrônimo é descritivo, capturando os principais aspectos do próprio modelo. Resumidamente:

- a) AR – autorregressão: um modelo que usa a relação dependente entre uma observação e um certo número de observações defasadas;
- b) I – integrado: o uso de diferenciação de observações brutas (ou seja, subtraindo uma observação de uma observação na etapa de tempo anterior) a fim de tornar a série temporal estacionária;
- c) MA – média móvel: um modelo que usa a dependência entre uma observação e erros residuais de um modelo de média móvel aplicado a observações defasadas.

Portanto, um modelo ARIMA é uma conjunção de modelos autorregressivos  $AR(p)$  e modelos de média móvel  $MA(q)$ , onde  $p$  e  $q$  representam as respectivas defasagens da série temporal a partir da qual se anulam (ou aproximadamente se anulam) as funções de autocorrelação (ACF) e autocorrelação parcial (PACF), respectivamente. Além da série ser integrada com  $d$  diferenças, para que a mesma se torne estacionária.

Em conformidade com (BOX et al., 2015), um modelo autorregressivo de médias móveis ARMA ( $p, q$ ) é definido como:

$$Y_t(t) = \sum_{i=1}^p \phi_i Y_{t-1} + \sum_{j=0}^q \theta_j \varepsilon_{t-j}, \quad (32)$$

onde  $Y_t$  é o valor da série temporal defasada por momentos  $t = 1, 2, \dots, l$ ,  $\phi_i$  e  $\theta_j$  são os coeficientes do modelo autorregressivo e de média móvel respectivamente, que, juntamente com as respectivas ordens  $p$  e  $q$ , constituem os parâmetros do modelo e  $\varepsilon_t$  é um processo puramente aleatório com média zero e variância  $\sigma^2$ .

Uma das condições necessárias para a aplicação do modelo ARMA é a estacionariedade das séries temporais, que na prática raramente é atendida. Por este motivo, existe a extensão do modelo ARMA, que permite aplicar o modelo mesmo em séries temporais não estacionárias,

denominado processo de média móvel integrado autorregressivo (ARIMA). Esta extensão transforma as séries temporais diferenciando-as pela ordem de  $d$ , garantindo assim uma série temporal estacionária. Em outras palavras, se a série  $Y_t$  não é estacionária, mas a  $d$ -ésima diferença, Equação (33), é estacionária (BOX et al., 2015).

$$\Delta^d Y_t = (1 - B)^d Y_t, \quad (33)$$

onde  $B$  é o operador de defasagem dado por:

$$B Y_t = Y_{t-1} \quad \therefore \quad B^j Y_t = Y_{t-j}. \quad (34)$$

O modelo ARIMA  $(p, d, q)$  é dado por:

$$\phi(B)[(1 - B)^d Y_t - \mu] = \theta(B)\varepsilon_t, \quad (35)$$

onde o parâmetro  $\mu$  é a média sobre a qual o processo estocástico estacionário varia, e os operadores polinomiais em  $B$ ,  $\phi(B)$  e  $\theta(B)$ , de graus  $p$  e  $q$ , respectivamente, são dados por:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p, \quad (36)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \quad (37)$$

Assim, o modelo ARIMA de ordem  $(p, d, q)$ , muitas vezes simplesmente denotado como ARIMA  $(p, d, q)$ , pode ser usado para prever os valores de séries temporais não estacionárias. No entanto, o modelo ARIMA mencionado acima não pode representar séries temporais com padrões sazonais com sucesso, especialmente quando a sazonalidade apresenta um padrão dinâmico. Portanto, o modelo ARIMA  $(p, d, q)$  pode ser estendido, formando o modelo ARIMA sazonal  $(p, d, q) (P, D, Q)$  de séries temporais.

#### 2.1.4.2 Modelos Sazonais Autorregressivos Integrados de Médias Móveis

O modelo de Média Móvel Integrada Autorregressiva Sazonal, ou SARIMA, é uma abordagem para modelar dados de séries temporais univariados que podem conter tendência e componentes sazonais.

O comportamento autorregressivo sazonal apresenta os atrasos no gráfico PACF abaixo do intervalo de confiança, exceto nos atrasos de tempo (*lags*)  $m = s, 2s, \dots, Ps$ , nesse caso sugere-se um modelo  $AR_s(P)$ , onde  $s$  representa o período sazonal. O comportamento autorregressivo se refere à correlação existente entre os valores atuais e anteriores.

O comportamento de média móvel sazonais apresenta os atrasos no gráfico ACF abaixo do intervalo de confiança, exceto nos *lags*  $m = s, 2s, \dots, Qs$ , nesse caso sugere-se um modelo  $MA_s(Q)$ . O comportamento de médias móveis se refere a correlação existente entre os valores atuais e os resíduos obtidos de valores anteriores.

O operador de diferença sazonal é definido de acordo com a Equação (38).

$$\Delta_s = 1 - B^s. \quad (38)$$

Assim  $\Delta_s Y_t = (1 - B^s)Y_t = Y_t - Y_{t-s}$  é a primeira diferenciação sazonal. Em geral,  $D$  diferenciações sazonais podem ser requeridas para produzir uma série estacionária. Nesse caso, o operador de diferenciação sazonal de ordem  $D$  é  $\Delta_s^D = (1 - B^s)^D$ .

A forma geral do modelo sazonal autorregressivo integrado de média móveis de ordem  $(P, D, Q)$  é (BOX et al., 2015):

$$\Phi(B^s)\Delta_s^D Y_t = \Theta(B^s)\varepsilon_t, \quad (39)$$

onde  $\Phi(B^s)$  e  $\Theta(B^s)$  são polinômios em  $B^s$  de graus  $P$  e  $Q$  e coeficientes  $\Phi$  e  $\Theta$ , respectivamente, que satisfazem as condições de série não-estacionária e  $\varepsilon_t$  é um processo de ruído aleatório que, geralmente, está correlacionado com  $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ . Para tratar tal correlação, introduz-se um segundo modelo:

$$\phi(B)\Delta^d \varepsilon_t = \theta(B)\varepsilon_t. \quad (40)$$

Conforme Equações (36) e (37),  $\phi(B)$  e  $\theta(B)$ , são polinômios em  $B$  de graus  $p$  e  $q$ , respectivamente, que satisfazem as condições de série não-estacionária.

Substituindo a Equação (40) na Equação (39), obtém-se o modelo SARIMA:

$$\phi_p(B)\Phi_P(B^s)\Delta^d\Delta_s^D Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t, \quad (41)$$

chamado de processo multiplicativo de ordem  $(p, d, q) \times (P, D, Q)_s$ .

O modelo SARIMA é uma abordagem eficaz para a previsão de séries temporais, embora exija uma análise cuidadosa e experiência de domínio para configurar, previamente, os sete parâmetros do modelo. Uma abordagem alternativa para configurar o modelo é pesquisar em grade um conjunto de configurações desses parâmetros para descobrir o que funciona melhor. Frequentemente, esse processo pode revelar configurações de modelo não intuitivas que resultam em erros de previsão mais baixos do que as configurações especificadas por meio de uma análise cuidadosa.

### 2.1.5 Redes Neurais Artificiais

Uma Rede Neural Artificial (RNA) é um modelo matemático que, basicamente, tenta reproduzir o modelo de funcionamento do cérebro, ou seja, de redes neurais naturais (RNN). Todos os cérebros, independentes de sua natureza, são formados por neurônios, os quais comunicam-se entre si por junções conhecidas como sinapses. A definição de uma RNA está ligada a algoritmos para desempenho de tarefas cognitivas, como aprendizagem, otimização e modelagem, assim como as RNN (FLORES, 2009).

Apesar de alguns modelos terem suposições quanto a distribuição de probabilidade dos dados, os modelos de RNA são considerados não-paramétricos ou de estatística robusta, por serem capazes de comportar a não linearidade. Outra característica importante é que modelos de RNA tem capacidade de lidar com sistemas dinâmicos, como séries temporais. Por sistema dinâmico entende-se por um sistema caracterizado por estados que mudam com o tempo. Uma série temporal não estacionária é um exemplo de sistema dinâmico.

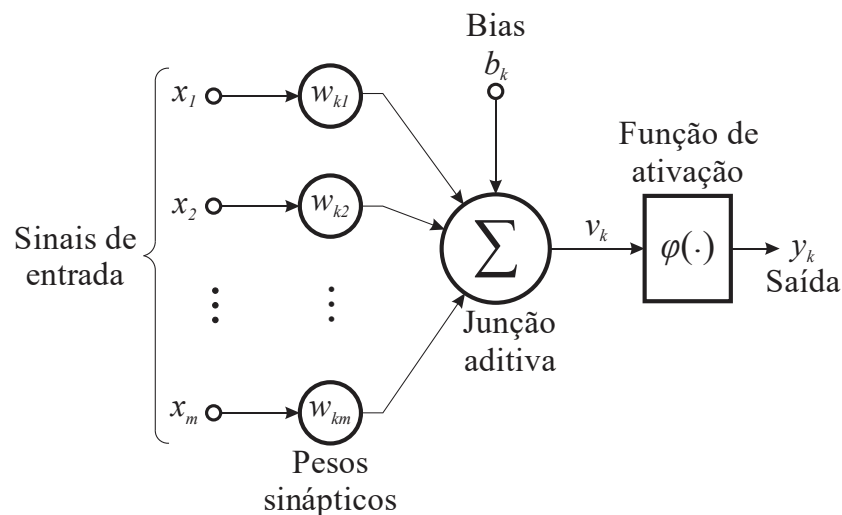
Outra característica das redes neurais artificiais é o fato delas poderem ser utilizadas para modelar séries temporais multivariadas, ou seja, séries que possuem mais de uma observação para cada intervalo de tempo.

Portanto, os modelos de redes neurais artificiais podem ser tratados como um método estatístico multivariado, não-paramétrico e não-linear.

A principal limitação de RNA reside no fato de não haver uma relação explícita entre as variáveis de entrada e saída, assim como são obscuros os detalhes de como são processados os dados para se chegar aos mapeamentos de resposta (FILHO, 2014).

De acordo com Haykin (2010), um neurônio é uma unidade de processamento de informações que é fundamental para a operação de uma rede neural. O diagrama em blocos da Figura 2 mostra o modelo de um neurônio, que constitui a base para o projeto de redes neurais artificiais. Identificam-se três elementos básicos do modelo neural:

Figura 2 – Modelo não linear de um neurônio  $k$



Fonte: Adaptado de (HAYKIN, 2010).

- conjunto de sinapses ou elos de ligação: cada uma caracterizada por um peso, especificamente, um sinal  $x_j$  na entrada da sinapse  $j$  conectada ao neurônio  $k$  é multiplicado pelo peso sináptico  $w_{kj}$ ;
- junta aditiva: para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio;
- função de ativação: para restringir a amplitude da saída de um neurônio. Tipicamente, o intervalo normalizado da amplitude de saída de um neurônio é  $[0, 1]$  ou alternativamente  $[-1, 1]$ .

O modelo neural da Figura 2 inclui também um *bias* ou viés aplicado externamente, representado por  $b_k$ . O *bias*  $b_k$  tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente (HAYKIN, 2010).

Em termo matemáticos, pode-se escrever um neurônio  $k$  apresentado na Figura 2 com o seguinte par de equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j, \quad (42)$$

$$y_k = \varphi(u_k + b_k), \quad (43)$$

onde  $x_1, x_2, \dots, x_m$  são os sinais de entrada;  $w_{k1}, w_{k2}, \dots, w_{km}$  são os pesos sinápticos do neurônio  $k$ ;  $u_k$  é a saída do combinador linear devido aos sinais de entrada;  $b_k$  é o *bias*;  $\varphi(\cdot)$  é a função de ativação; e  $y_k$  é o sinal de saída do neurônio. O *bias*  $b_k$  tem o efeito de aplicar uma transformação à saída  $u_k$  do combinador linear, conforme Equação (44).

$$v_k = u_k + b_k. \quad (44)$$

De acordo com Aggarwal (2018), o valor calculado antes de aplicar a função de ativação  $\varphi(\cdot)$  é referido como o valor de pré-ativação, enquanto o valor calculado após a aplicação da função de ativação é referido como o valor de pós-ativação. A saída de um neurônio é sempre o valor de pós-ativação, embora as variáveis de pré-ativação sejam frequentemente usadas em diferentes tipos de análises, como os cálculos do algoritmo de retropropagação discutido posteriormente nesta seção.

A função de ativação  $\varphi(\cdot)$  mais básica é a identidade ou ativação linear, que não fornece não linearidade, conforme apresentado na Equação (45).

$$\varphi(v) = v. \quad (45)$$

Conforme Aggarwal (2018), a função de ativação linear é frequentemente usada no nó de saída, quando o alvo é um valor real.

As funções de ativação clássicas que foram usadas no início do desenvolvimento de redes neurais foram as funções de sinal, sigmoide e tangente hiperbólica ( $\tanh$ ), conforme apresentada na Equação (46), (47) e (48), respectivamente.

$$\varphi(v) = \text{signal}(v), \quad (46)$$



$$\varphi(v) = \frac{1}{1 + e^{-v}}, \quad (47)$$

$$\varphi(v) = \frac{e^{2v} - 1}{e^{2v} + 1}. \quad (48)$$

Embora a função ativação de sinal, Equação (46), possa ser usada para mapear saídas binárias no momento da previsão, sua não diferenciabilidade impede seu uso para criar a função de perda no momento do treinamento. A função tanh tem uma forma semelhante à da função sigmoide, exceto que é redimensionada para  $[-1, 1]$ . Portanto a função tanh é preferível ao sigmoide quando os resultados dos cálculos são desejados para serem positivos e negativos.

Segundo Aggarwal (2018) as funções sigmoide e tanh têm sido as ferramentas históricas de escolha para incorporar a não linearidade na rede neural. Nos últimos anos, no entanto, uma série de funções de ativação linear por partes se tornaram mais populares, como a função *Rectified Linear Unit* (ReLU), Equação (49), e a *hard tanh*, Equação (50).

$$\varphi(v) = \max\{v, 0\}, \quad (49)$$

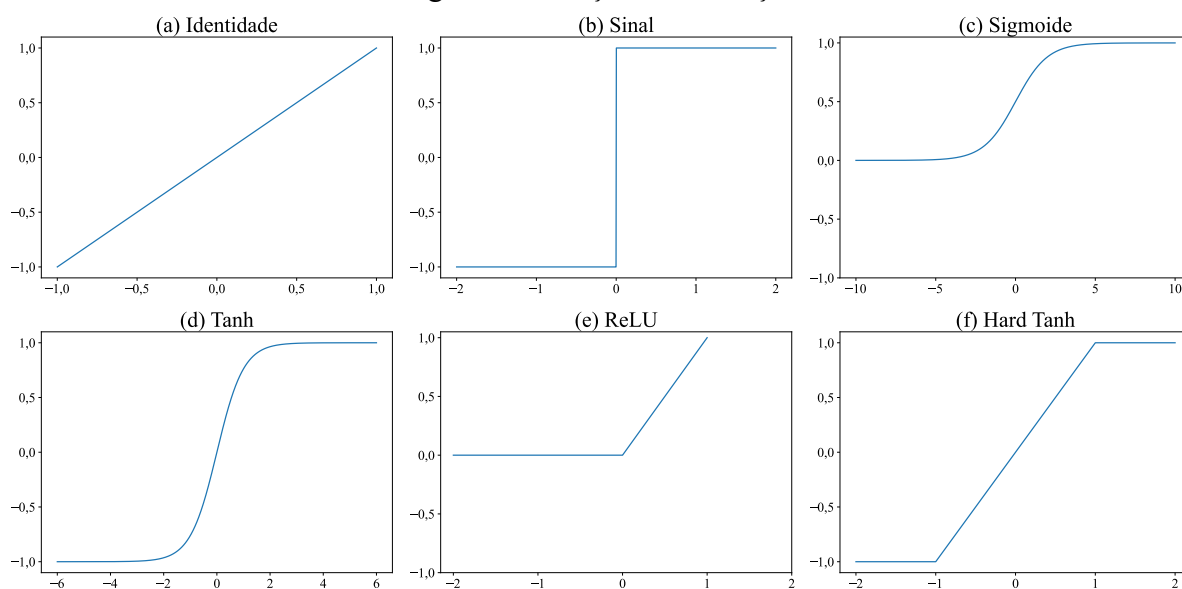
$$\varphi(v) = \max\{\min[v, 1], -1\}. \quad (50)$$

As funções de ativação não-lineares ReLU e *hard tanh* substituíram amplamente as funções de ativação sigmoide e tanh nas redes neurais modernas devido à facilidade de treinamento, dessas funções, nas redes multicamadas.

O uso de uma ativação não linear desempenha um papel fundamental no aumento do poder de modelagem de uma rede neural artificial. Se uma rede utilizar apenas ativações lineares, ela não irá fornecer um melhor poder de modelagem do que uma rede linear de camada única.

As representações de todas as funções de ativação mencionadas (Equação (45) a Equação (50)) são ilustradas na Figura 3.

Figura 3 – Funções de ativação



Fonte: Elaboração própria.

As funções de ativação mostradas na Figura 3 são monotônicas. Além da função de ativação de identidade, a maioria das outras funções de ativação saturam para grandes valores absolutos do argumento de tal forma que acréscimos adicionais não alteram significativamente a ativação (AGGARWAL, 2018).

Os neurônios costumam ser organizados em camadas. Com base no padrão de conexão entre os neurônios (arquitetura), as RNAs podem ser agrupadas em duas categorias principais: redes *feedforward* e redes recorrentes.

Em uma rede *feedforward*, os neurônios de uma camada conectam-se apenas aos neurônios das camadas adjacentes e o fluxo de informações ocorre apenas na direção da entrada para a saída da rede. Não há, portanto, nesse tipo de arquitetura, ciclos ou *loops* na rede. A camada de neurônios que recebe dados externos é chamada de camada de entrada, enquanto a camada que produz a saída é chamada de camada de saída. Entre essas duas camadas, pode haver diferentes quantidades de camadas intermediárias, conhecidas como camadas ocultas. As camadas intermediárias são referidas como camadas ocultas porque os cálculos realizados não são visíveis ao usuário (AGGARWAL, 2018).

Já as redes neurais recorrentes apresentam a característica de retroalimentação. As realimentações consistem em saídas de neurônios de determinada camada serem reintroduzidas como entradas de neurônios de camadas anteriores ou da própria camada. Um neurônio pode, inclusive, ser realimentado por sua própria saída. Essas possibilidades fazem com que a

arquitetura de uma rede recorrente possa tomar diversas formas, não havendo, portanto, uma arquitetura padrão para as redes recorrentes. (GOMES, 2005).

As redes neurais artificiais se adaptam à diferentes aplicações por meio do aprendizado que é alcançado ajustando-se os pesos das sinapses até que um conjunto de dados de entrada resulte nos valores de saída desejados. Ainda, as RNAs podem ser personalizadas para vários objetivos, usando diferentes topologias, funções de neurônios e estratégias de aprendizagem.

A capacidade de generalização de uma rede neural, ou seja, sua habilidade de fornecer saídas corretas diante de novos padrões de entrada, está relacionada com fatores como a boa escolha de uma arquitetura de rede e o tipo de algoritmo de treinamento utilizado.

#### 2.1.5.1 Perceptron de Múltiplas Camadas

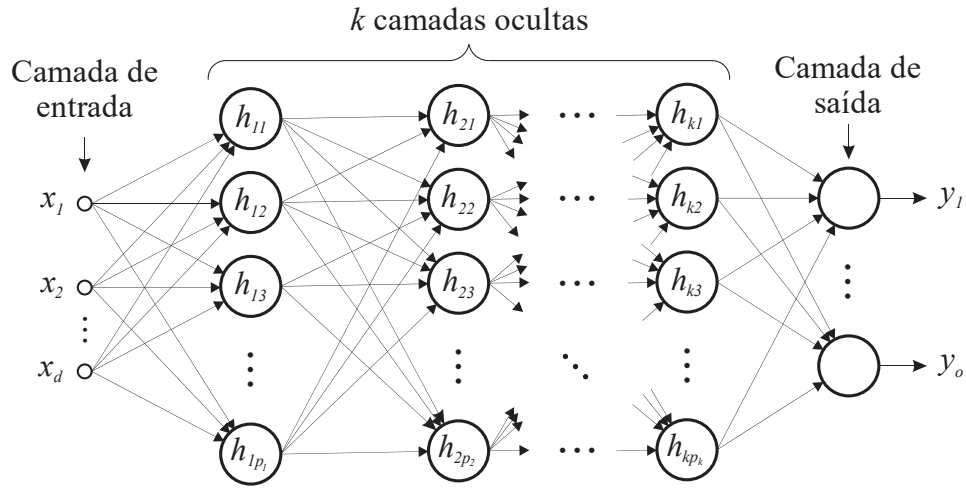
A rede neural mais simples é o *perceptron* e foi proposta por Frank Rosenblatt em 1958 (AGGARWAL, 2018; HAYKIN, 2010). A arquitetura básica do *perceptron* é a do modelo do neurônio mostrada na Figura 2. Essa rede neural contém uma única camada de entrada que transmite os dados para o nó de saída, da qual a camada de saída é a única camada de execução de computação e todos os cálculos são completamente visíveis ao usuário (AGGARWAL, 2018).

Já as redes neurais *perceptron* de múltiplas camadas, ou *Multilayer Perceptron* (MLP), contém mais de uma camada computacional, ou seja, a arquitetura mínima de uma MLP contém, pelo menos, uma camada intermediária (oculta) entre a camada de entrada e camada de saída (AGGARWAL, 2018).

A rede de conexão dos neurônios entre as camadas de MLP é do tipo *feed-forward* e o modelo de cada um dos neurônios inclui uma função de ativação não linear que é diferenciável (HAYKIN, 2010).

A arquitetura de uma rede MLP genérica com  $d$  entradas,  $k$  camadas ocultas com  $p_i$  neurônios em cada uma, com  $i = 1, 2, \dots, k$  e com  $o$  neurônios na camada de saída, é mostrada na Figura 4.

Figura 4 – Arquitetura de uma rede neural MLP genérica



Fonte: Elaboração própria.

A camada de entrada geralmente não é contada, porque ela simplesmente transmite os dados e nenhum cálculo é realizado nessa camada. O número de unidades de neurônios em cada camada é referido como a dimensionalidade dessa camada (AGGARWAL, 2018).

A rede neural apresentada na Figura 4 contém  $p_1, p_2 \dots p_k$  neurônios em cada uma de suas  $k$  camadas ocultas, então as representações vetoriais (coluna) dessas saídas, denotadas por  $\bar{h}_1, \bar{h}_2, \dots, \bar{h}_k$ , tem dimensionalidades  $p_1, p_2, \dots, p_k$ , respectivamente.

Os pesos das conexões entre a camada de entrada e a primeira camada oculta estão contidos em uma matriz  $\mathbf{W}_1$  com tamanho  $d \times p_1$ , onde  $d$  é a quantidade de sinal de entradas e  $p_1$  a quantidade de neurônios da primeira camada oculta. Já os pesos sinápticos entre a camada oculta  $(r - 1)$  e a camada oculta  $r$  são denotados pela matriz  $\mathbf{W}_r$  de dimensão  $p_{r-1} \times p_r$ . Se a camada de saída contém  $o$  nós, então a matriz final  $\mathbf{W}_{k+1}$  tem o tamanho  $p_k \times o$ .

Portanto, o vetor de entrada  $d$ -dimensional  $\bar{x}$  é transformado nas saídas pelas seguintes equações recursivas (AGGARWAL, 2018):

$$\bar{h}_1 = \varphi(\mathbf{W}_1^T \bar{x}), \quad (51)$$

$$\bar{h}_{p+1} = \varphi(\mathbf{W}_{p+1}^T \bar{h}_p) \quad \forall p \in \{1, 2, \dots, k - 1\}, \quad (52)$$

$$\bar{y} = \varphi(\mathbf{W}_{k+1}^T \bar{h}_k), \quad (53)$$

onde  $\bar{y}$  é a representação vetorial da saída de dimensão  $o \times 1$  e os ajustes de todos os pesos sinápticos da rede é realizado por meio de processo de treinamento ou aprendizagem.

Em um sentido implícito, o processo de aprendizagem deve decidir quais características do padrão de entrada devem ser representadas pelos neurônios ocultos. O processo de aprendizagem é, portanto, dificultado nos casos de redes multicamadas porque o erro (ou função de perda) não pode ser calculado como uma função direta dos pesos, e sim por meio de composição complexa dos pesos nas camadas anteriores. Um método popular para o treinamento de *perceptrons* multicamadas é o algoritmo de retropropagação do erro ou *backpropagation* (AGGARWAL, 2018).

O algoritmo de retropropagação é uma aplicação da programação dinâmica. Ele contém duas fases principais, conhecidas como fases de avanço e retrocesso, respectivamente.

Primeiro, na fase de avanço (direta), um padrão é apresentado à camada de entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo, na fase reversa, a saída obtida é comparada à saída desejada para esse padrão particular. Se esta não estiver correta, o erro é calculado. O erro é propagado a partir da camada de saída até a camada de entrada, e ajustes sucessivos são feitos nos pesos sinápticos da rede conforme o erro é retropropagado.

O algoritmo pode ser iniciado atribuindo valores aleatórios para os pesos sinápticos. Quando o vetor de dados de entrada  $\bar{x}(t)$  é apresentado para a rede, a rede neural produz uma saída  $\hat{y}(t)$ , que em geral é diferente do vetor alvo  $\bar{y}(t)$ . O que se quer é fazer com que  $\hat{y}(t)$  e  $\bar{y}(t)$  tenham a menor diferença possível, para  $i = 1, 2, \dots, P$  através do uso do algoritmo de aprendizagem. Portanto, a intenção é diminuir a função de erro da rede (FILHO, 2014), que pode ser dado pela Equação (54).

$$E = \frac{1}{2} \sum_{i=1}^P |\hat{y}(t) - \bar{y}(t)|^2. \quad (54)$$

Na fase de avanço, calcula-se o erro de saída de cada neurônio e o valor do erro é propagado na rede, de modo que os neurônios das camadas escondidas e das camadas de saídas recebem os erros das camadas anteriores. Cada neurônio soma os valores do erro dos neurônios de camadas anteriores ligados a ele através de sinapse e, ao final, nos neurônios de saída da rede, tem-se o valor total do erro.

Os pesos da rede são os parâmetros que podem ser modificados para tornar o valor da função de erro o menor possível. A função de erro  $E$  é uma função contínua e diferenciável dos pesos  $w_1, w_2, \dots, w_l$  da rede. É possível, portanto, minimizar  $E$  usando um processo iterativo do gradiente descendente (descida do gradiente), conforme Equação (55).

$$\nabla E = \left( \frac{\partial E}{\partial w_1} + \frac{\partial E}{\partial w_2} + \dots + \frac{\partial E}{\partial w_l} \right). \quad (55)$$

Cada peso é ajustado utilizando-se o seguinte incremento:

$$\nabla w_i = -\eta \frac{\partial E}{\partial w_i} \text{ para } i = 1, 2, \dots, l, \quad (56)$$

onde  $\eta$  representa a taxa de aprendizagem, ou seja, um parâmetro de proporcionalidade que define o comprimento do passo de cada iteração na direção negativa do gradiente. Uma vez que se tem um método para calcular o gradiente, pode-se ajustar os pesos da rede de forma iterativa de maneira a encontrar o valor mínimo da função erro e, conseqüentemente, obter saídas do modelo o mais próximo possível da resposta esperada.

Os dados que são utilizados para o aprendizado da rede são chamados de dados de treinamento e os modelos que respondem a esse conjunto é dito que apresentam capacidade de convergência.

Após o modelo treinado, com os pesos sinápticos todos definidos, ele deve ser submetido aos dados do conjunto de teste que deve ser distinto dos dados do conjunto de treinamento. Espera-se que o modelo apresente habilidade de generalização que é a capacidade de um modelo de aprendizado responder corretamente a padrões novos que lhe são apresentados. Um modelo com boa generalização é aquele que responde corretamente a padrões contidos na base de dados, e também aos padrões novos contidos da base de teste. A capacidade de generalização é o principal objetivo no processo de aprendizado.

Afim de dar embasamento para o estudo de caso da aplicação dos modelos na previsão do preço das liquidações da diferença (preço *spot* da energia) no mercado de curto prazo de energia elétrica, é apresentada, na próxima seção, uma contextualização do setor elétrico brasileiro.

## 2.2 O SETOR ELÉTRICO BRASILEIRO

### 2.2.1 Breve Histórico

A eletricidade foi trazida para o Brasil na segunda metade do século XIX, por iniciativa do então imperador D. Pedro II, inicialmente na implementação de iluminação pública alimentada por dínamos na cidade do Rio de Janeiro e posteriormente com inaugurações de algumas usinas hidrelétricas com vistas ao fornecimento de energia aos serviços públicos (iluminação e transportes coletivos) e força motriz nas indústrias (ELETROBRÁS, 2015).

A crescente demanda por energia elétrica, aliada às dificuldades da administração pública em atendê-la, determinaram a instituição do regime de concessões ao capital privado para a produção e distribuição de eletricidade. Esses primeiros concessionários, em geral fazendeiros, comerciantes e empresários locais, se organizavam em empresas de âmbito municipal (JANNUZZI, 2007).

Nos primórdios da República, o Brasil passou a conviver com um significativo aumento do capital estrangeiro, fato que norteou a monopolização e desnacionalização do setor elétrico. Nesse processo o Grupo Light, de origem canadense, e a American Foreign Power Company (AMFORP) incorporaram outras concessionárias e dominaram o mercado local de eletricidade até meados dos anos 1960 (JANNUZZI, 2007; WALVIS; GONÇALVES, 2014).

Devido ao elevado preço da energia, esse modelo privado do setor elétrico começou a ser questionado a partir da década de 1920, sustentado pela crise definida pela contestação das articulações políticas entre as oligarquias regionais do Brasil que sobrepujam os seus interesses particulares aos interesses do Estado e da Nação como um todo.

Com o êxito da Revolução de 1930 o setor elétrico começou a seguir novos rumos liderado por um governo federal forte, nacionalista, que buscava acelerar o desenvolvimento industrial. O Código das Águas (Decreto n.º 26.234), promulgado em 1934 e regulamentado em 1957, iniciou o processo de intervenção do Estado nos serviços de eletricidade e se tornou uma peça central na regulação da energia elétrica brasileira. Com o decreto, todos os recursos hídricos foram incorporados ao patrimônio da União, com a concessão de uso tornando-se uma prerrogativa do governo federal, ampliando o domínio público sobre as águas e regulamentando a indústria hidrelétrica, restringindo a concessão a brasileiros ou a empresas organizadas no Brasil (ALBUQUERQUE, 2008; BRASIL, 2016a; WALVIS; GONÇALVES, 2014).

Na década de 1950 houve um elevado crescimento na demanda por energia elétrica em virtude da política econômica implementada no Brasil, pautada pelo modelo de substituição de importações que favorecia o crescimento liderado pelo Estado, responsável pelos investimentos em infraestrutura com características de monopólios naturais. Nesse modelo o governo federal era o principal agente, assumindo o controle da geração e transmissão de energia, que cada vez mais exigiam vultuosos investimentos e com longos períodos de maturação. A distribuição, inicialmente controlada por empresas estrangeiras, foi progressivamente transferida para os governos estaduais (DE OLIVEIRA, 2000; WALVIS; GONÇALVES, 2014).

Ao longo das décadas de 1940 e 1950, foram constituídas várias empresas estatais que investiram na geração de energia elétrica. No âmbito federal, foram constituídas a Centrais Elétricas S.A. – FURNAS (1957), a Companhia Hidrelétrica do São Francisco – CHESF (1945), no âmbito estadual a Centrais Elétricas de Minas Gerais – CEMIG (1952), a Companhia Estadual de Energia Elétrica – CEEE (1943), a Companhia Paranaense de Energia – COPEL (1954), dentre outras (WALVIS; GONÇALVES, 2014).

A partir do fim da Segunda Guerra Mundial em 1945, um desequilíbrio entre o crescimento da demanda por energia elétrica e a capacidade instalada de geração determinou um período de crise no setor, onde os principais centros econômicos conviveram, desde então, com uma instabilidade na oferta de eletricidade causando prejuízos aos grandes consumidores.

Em meados de 1960, foi criado o Ministério das Minas e Energia (MME)<sup>5</sup>, cada vez mais necessário diante da complexidade desenvolvida no setor energético, encarregado do estudo e despacho de todos os assuntos relativos à produção mineral e à energia.

Em 1962 foi criada a Centrais Elétricas Brasileiras S.A. (Eletrobrás), que passou a atuar como holding das concessionárias federais de energia elétrica e no planejamento da expansão do sistema elétrico brasileiro, promovendo estudos, projetos de construção e operação de usinas geradoras, linhas de transmissão e subestações destinadas ao suprimento de energia elétrica do país. A frequência elétrica em todo país foi padronizada permitindo a interligação dos grandes troncos de transmissão de energia englobando regiões cada vez mais distantes e integrando as atividades (ELETROBRÁS, 2015).

A Eletrobrás exerceu um papel importante na obtenção de investimento de capital estrangeiro com taxas de juros baixas e amortização a longo prazo, causando euforia no setor elétrico na década de 1970.

---

<sup>5</sup> Em 12 de abril de 1990 foi extinto o Ministério das Minas e Energia e criado o Ministério da Infraestrutura pela Lei n.º 8.028. Em 13 de maio de 1992 foi extinto o Ministério da Infraestrutura e criado o Ministério de Minas e Energia pela Lei n.º 8.422.



No início dos anos 80, com o grave aumento das taxas de juros no mercado internacional e escassez de crédito nacional, os investimentos no setor elétrico foram praticamente interrompidos. Aliado ao fato de as concessionárias apresentarem uma crítica situação econômico-financeira, de um sistema de tarifas equalizadas que estimulava a ineficiência administrativa por transferir recursos para empresas deficitárias e uma política de compensação tarifária de combate à inflação adotada pelo governo federal, contribuíram para que o modelo estatizante começasse a entrar em crise (DE OLIVEIRA, 2000; JANNUZZI, 2007).

O setor elétrico entra na década de 1990 em uma situação bastante delicada. Com as concessionárias endividadadas e o Estado sem condições de investir na continuidade dos planos de expansão, a privatização se apresenta como uma alternativa viável na resolução dos problemas financeiros e ajuste patrimonial das empresas de eletricidade. Nesse cenário, em 1990, o Governo lança o Programa Nacional de Desestatização (PND) com o objetivo de privatizar as empresas estatais de forma a obter recursos para o tesouro nacional (LORENZO, 2001).

Na década de 1990, o setor elétrico brasileiro passou por profundas mudanças estruturais de cunho institucional e operacional. A Lei da Reforma Tarifária de 1993 (Lei n.º 8.631) que dispõe sobre a fixação dos níveis das tarifas para o serviço público de energia elétrica e extingue o regime de remuneração garantida além de outras providências, é reconhecida como um dos primeiros passos em direção à reestruturação do setor elétrico. Essa medida foi uma tentativa de dar às concessionárias condições de recuperação financeira e torná-las mais atrativas para os investimentos privados (TEODORO, 2006; WALVIS; GONÇALVES, 2014).

As mudanças no contexto institucional tiveram início efetivo em 1995, após a promulgação da Lei das Concessões (Lei n.º 8.987) que dispôs sobre o regime de concessões e de permissão de prestação de serviços públicos, que transferiu para o Poder Público a responsabilidade da prestação de serviços de interesse geral e da Lei n.º 9.074 que estabeleceu as normas para outorga e prorrogação de concessões e permissões de serviços públicos. Em conjunto, essas leis introduziram importantes alterações nas normas do setor elétrico, tais como: desmembramento contábil das atividades de geração, transmissão e distribuição; licitação de novos potenciais hídricos visando estimular a competição na geração; livre acesso à rede de transmissão e definição de uma malha básica; previsão da revisão tarifária nos contratos de concessão; criação da figura jurídica do produtor independente de energia e abertura da possibilidade para o surgimento dos consumidores livres (com carga igual ou superior a 10 MW, atendidos em tensão igual ou superior a 69 KV) terem direito à contratação de energia de produtores independentes (TEODORO, 2006; WALVIS; GONÇALVES, 2014).

Em 1996 o governo federal contratou a consultora internacional Coopers & Lybrand para liderar o consórcio responsável por implementar a reforma setorial no âmbito do Projeto de Reestruturação do Setor Elétrico Brasileiro (RESEB). O conjunto de recomendações desenvolvidas pelo consórcio, que utilizou como paradigma a reforma inglesa, visava à implementação de um modelo para a indústria de energia elétrica nacional desverticalizado, com competição nos seguimentos de geração e comercialização e forte regulamentação nos segmentos de transmissão e distribuição, além de propor mudanças que visavam dar condições para desenvolver o programa de privatizações das empresas do setor (TEODORO, 2006).

Para dar prosseguimento à execução das mudanças estruturais e específicas, foram constituídos diversos agentes setoriais. Em 1996 foi criada a Agência Nacional de Energia Elétrica (ANEEL), autarquia vinculada ao Ministério de Minas e Energia (MME)<sup>6</sup> responsável pela regulação do Setor Elétrico Brasileiro (SEB) (ANEEL, 2019). Em 1998 foi criado o Mercado Atacadista de Energia, destinado à livre negociação de energia elétrica, e institucionalizou a figura do comercializador de energia, agente autorizado a vender energia elétrica a consumidores finais e comercializá-la no âmbito do Mercado Atacadista de Energia (CCEE, 2015). Nesse mesmo ano foi regulamentado o Operador Nacional do Sistema Elétrico (ONS), órgão responsável pela coordenação e controle da operação das instalações de geração e transmissão de energia elétrica no Sistema Interligado Nacional (SIN)<sup>7</sup> e pelo planejamento da operação dos sistemas isolados do país, sob a fiscalização e regulação da ANEEL (ONS, 2018). Em 1999 constituiu-se o Comitê Coordenador de Expansão de Sistemas Elétricos (CCPE), responsável por coordenar a elaboração do planejamento da expansão dos sistemas elétricos brasileiros, de caráter indicativo para a geração, consubstanciado nos Planos Decenais de Expansão e nos Planos Nacionais de Energia Elétrica de longo prazo (ANEEL, 1999).

Diante de tantas mudanças, os agentes do setor decidiram adiar os investimentos no aguardo de uma regulamentação estável que sinalizasse com clareza as oportunidades para investir. Com isso, o que se viu foi um crescimento da demanda por energia sem a devida correspondência da oferta, provocando um desequilíbrio nos reservatórios (TEODORO, 2006).

Esse cenário marca o início dos anos 2000 como o ano da crise de abastecimento de energia elétrica com profundos impactos na economia e na vida da sociedade brasileira. Apesar da escassez de energia ter sido atribuída ao quadro hidrológico extremamente desfavorável ao funcionamento pleno das usinas e reservatórios, a falta de investimentos na expansão da

---

<sup>6</sup> A partir de maio de 1992, MME se refere ao Ministério de Minas e Energia e não mais ao Ministério das Minas e Energia criado em 1960 e que vigorou até 1990.

<sup>7</sup> Na seção 2.2.3 serão fornecidas mais informações sobre o SIN.

capacidade de geração e na transmissão levou à degradação da operação do sistema comprometendo a oferta de energia (TEODORO, 2006).

Em abril de 2001, com o fim do período das chuvas e os níveis insatisfatórios dos reservatórios, a ANEEL reconheceu a crise do fornecimento de energia elétrica. Como medida emergencial foi lançado pelo governo o Plano de Racionalização de Energia Elétrica, obrigando consumidores à redução de consumo.

Evidenciado pela crise, a reestruturação da década de 90 foi incapaz de criar condições favoráveis à expansão da geração, que ficou sob responsabilidade dos agentes privados, resultando em dois apagões, seguidos de um racionamento de nove meses, entre junho de 2001 e fevereiro de 2002, acompanhado de uma queda de 20% no consumo de energia elétrica.

Assim, partindo dessa premissa da insuficiência do modelo anterior, o ano de 2003 marca um período de transição do SEB para um Novo Modelo.

Em 2004, foram instituídas as Leis n.º 10.847 e n.º 10.848, com papel de estabelecer novas bases regulatórias e institucionais para o setor elétrico, sendo essa função aplicada por meio da reorganização das atividades de planejamento, alicerçado na retomada dos investimentos para expansão do sistema, garantindo segurança de suprimento e assim, afastando o risco de uma nova crise de abastecimento e assegurando um marco regulatório estável, além de regular o processo de comercialização da energia dentro do mercado cativo, tentando reorganizar a política setorial mediante os objetivos de modicidade tarifária e promover a inserção social através de um Programa de Universalização do Acesso e do Uso dos Serviços de Eletricidade. Dessa forma, a regulação do Novo Modelo, privilegia a centralização do planejamento no Estado, abdicando de algumas relações de mercado (BEZERRA, 2020; TEODORO, 2006).

Os princípios básicos no Novo Modelo do SEB podem ser definidos como: estabelecimento de novas regras e ambientes (um livre e outro regulado) para contratação de energia elétrica; desverticalização das distribuidoras; definição de regras para novos empreendimentos na geração; determinação de prazos para que os consumidores potencialmente livres possam adquirir energia de outro fornecedor que não a distribuidora que o atende; redefinição das funções, atividades e responsabilidades do MME, ANEEL, ONS e Eletrobrás; criação de novas instituições (TEODORO, 2006).

A proposta do Novo Modelo inclui a criação da Empresa de Pesquisa Energética (EPE), vinculada ao MME, cujo papel é gerar subsídios às entidades governamentais e agentes setoriais para definição de políticas e ações necessárias à garantia do suprimento de energia; a Câmara de Comercialização de Energia Elétrica (CCEE) em substituição ao Mercado Atacadista de

Energia, incorporando suas estruturas organizacionais e operacionais; e o Comitê de Monitoramento do Setor Elétrico (CMSE), com a função de acompanhar e avaliar permanentemente a continuidade e a segurança do suprimento eletroenergético em todo o território nacional (BRASIL, 2004; CCEE, 2015; EPE, 2017).

Em janeiro de 2007, o governo federal lançou o Programa de Aceleração do Crescimento (PAC) para execução de grandes obras de infraestrutura logística, energética e social e urbana do país. Em nove anos, o PAC promoveu, no setor elétrico, o aumento da capacidade do parque gerador em 31.976 megawatts (MW), com a entrada em operação de grandes empreendimentos como as usinas hidrelétricas de Belo Monte – PA, Santo Antônio – RO, Jirau – RO e Teles Pires – MT; além da expansão e reforço do SIN, construindo subestações e linhas de transmissão que totalizam 26.029 quilômetros (km) de extensão, incluindo grandes obras como a interligação Tucuruí – Macapá – Manaus, que conectou os estados do Amazonas e Amapá ao SIN (BRASIL, 2016b).

Em 2012, um novo capítulo na história do setor elétrico brasileiro começou com a Medida Provisória (MP) 579 editada para reduzir o custo final da energia elétrica. Convertida na Lei n.º 12.783/2013, a MP renovou antecipadamente as concessões de empresas geradoras e transmissoras que venceriam em 2015, desde que elas aceitassem ter os preços da energia definidos pela ANEEL, estabelecendo preços e tarifas de energia ao consumidor final em valores muito abaixo aos efetivamente praticados no mercado energético. Com a redução tarifária, o consumo foi estimulado, impulsionando um maior desencontro entre a demanda do mercado e a quantidade de energia produzida pelas geradoras hídricas. Houve a necessidade de deslocamento do bloco de geração hidráulica pelo acionamento fora de ordem das usinas de geração termelétrica que, por sua vez, são matrizes energéticas mais caras, seja sob a ótica financeira ou ambiental (FERREIRA, 2020).

Os resultados da MP 579/2012 não foram os esperados, enfraquecendo as geradoras públicas e tornando o setor elétrico submisso a recursos do governo federal. Diante da situação, diversos representantes e associações das geradoras insurgiram junto ao MME e à ANEEL, alegando que o cenário crítico se devia, além das condições climáticas, ao aumento do despacho por usinas térmicas devido ao crescimento da demanda, ampliação da matriz renovável e políticas públicas implementadas de formas desacertadas, pleiteando, portanto, o repasse do prejuízo aos consumidores. A ausência de uma resposta setorial da ANEEL ocasionou uma onda de judicialização do setor elétrico por parte das geradoras hídricas. (FERREIRA, 2020).

Em janeiro de 2015, a ANEEL passou a aplicar as bandeiras tarifárias mediante condições climáticas desfavoráveis, repassando aos consumidores os custos de aquisição pelas

distribuidoras de energia elétrica de usinas térmicas. Um dos efeitos da bandeira tarifária é a inflação de custos, gerado pelo aumento do custo de produção de bens de consumo.

Em 2017, o MME instaurou a Consulta Pública n.º 33 referente a proposta de medidas legais de aprimoramento do marco legal do setor elétrico, viabilizando seu futuro com sustentabilidade a longo prazo. Os resultados apresentaram as seguintes sugestões para o aperfeiçoamento do setor: ampliação do acesso ao mercado livre; ajuste na autoprodução; adequações na formação de preço; alinhamento de custos na transação de transmissão e geração; entre outras. A consulta suportou a publicação da Portaria n.º 465/2019/MME que estabeleceu, com base na Lei n.º 9.074/1995, um cronograma de ampliação do mercado livre para consumidores com carga acima de 500 quilowatt (kW). Complementarmente, o projeto de Lei do Senado n.º 232 de 2016 dispôs sobre o modelo comercial do setor elétrico e as concessões de geração de energia elétrica, com sinalização de novas licitações e regras para produção independente (MME, 2017; SENADO, 2020).

### **2.2.2 Mercado de Energia Elétrica Brasileiro**

Com a reestruturação do Setor Elétrico Brasileiro (SEB) foram instituídos dois ambientes de mercado para a celebração de contratos de compra e venda de energia elétrica: o Ambiente de Contratação Regulada (ACR) e o Ambiente de Contratação Livre (ACL). Cada um desses ambientes possui suas particularidades quanto aos tipos e flexibilidades dos contratos e a forma de definição dos seus preços.

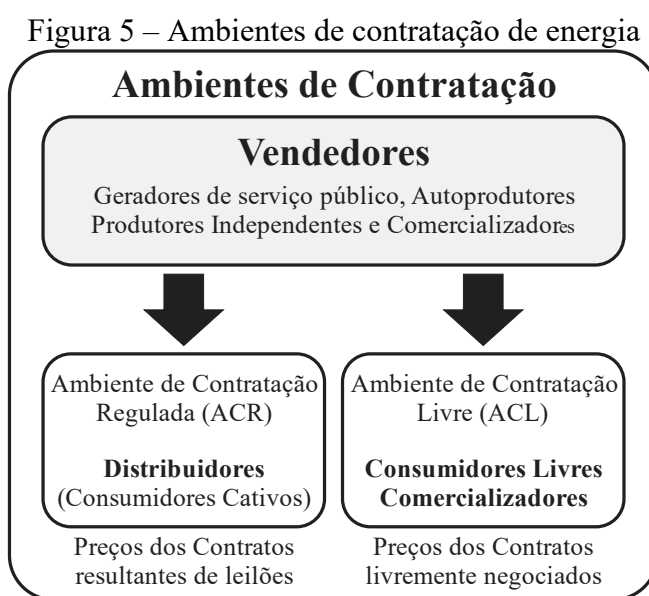
No ACR a contratação de energia elétrica é realizada por meio de leilões de energia promovidos pela CCEE, sob delegação da ANEEL, de forma a aumentar a competitividade na busca de promover uma menor tarifa de energia elétrica para o consumidor final. A contratação é formalizada através de contratos bilaterais regulados pela ANEEL, denominados Contratos de Comercialização de Energia Elétrica no Ambiente Regulado (CCEAR) celebrados entre os agentes vendedores (geradores, produtores independentes, autoprodutores e comercializadores) e as distribuidoras para o suprimento de seu mercado cativo.

O ACL é o ambiente de aquisição de energia elétrica dos consumidores livres por meio de contratos celebrados com os agentes de geração, comercializadores ou importadores e exportadores de energia elétrica. Nesse ambiente há liberdade para se estabelecer volumes de compra e venda de energia e seus respectivos preços, sendo as transações pactuadas através de

contratos bilaterais. Isto envolve o montante de energia contratada, o período de suprimento, o preço (dado em reais por megawatt-hora), a localização da entrega, e algumas flexibilidades que podem ser incorporadas, como por exemplo, a sazonalização e modulação (FILHO, 2014).

Por sazonalização, entende-se a distribuição da energia ao longo do ano que, na prática, é a distribuição mensal do montante anual contratado. Já modulação é a distribuição horária dos montantes mensais definidos. Tais flexibilidades permitem atender a um maior consumo em um determinado período do ano ou em determinadas horas do dia.

A Figura 5 resume a relação entre os agentes e os ambientes de contratação.



Fonte: Adaptado de CCEE [www.ccee.org.br](http://www.ccee.org.br).

Todos os contratos de compra e venda de energia celebrados no mercado – tanto no ACR como no ACL – devem ser registrados na CCEE e servem de base para a contabilização e liquidação das diferenças no Mercado de Curto Prazo (MCP). Esse registro inclui, apenas, as partes envolvidas, os montantes de energia e o período de vigência. Os preços de energia firmados nos contratos não são registrados na CCEE, mas, utilizados especificamente pelas contrapartes em suas liquidações bilaterais, fora do ambiente de operações da CCEE e de acordo com as condições contratuais específicas (CCEE, 2013).

Da mesma forma, os valores de medições verificados de geração e consumo de energia são registrados na CCEE pelos agentes ou coletados diretamente através do Sistema de Coleta de Dados de Energia (SCDE), administrado pela CCEE e responsável pela coleta e tratamento dos dados de geração e consumo de energia do SIN. Com base nos contratos e dados de medição registrados, contabilizam-se as diferenças entre o que foi produzido ou consumido e o que foi

contratado, sendo as diferenças positivas ou negativas liquidadas ao Preço de Liquidação das Diferenças (PLD). A contabilização é realizada mensalmente, sendo calculados os montantes negociados no Mercado de Curto Prazo (à vista), que pode ser visto como o mercado das diferenças, conforme Figura 6. No MCP não existem contratos, ocorrendo a contratação multilateral, conforme as Regras de Comercialização (CCEE, 2013).

Figura 6 – Energia comercializada no Mercado de Curto Prazo



Fonte: Adaptado de CCEE [www.ccee.org.br](http://www.ccee.org.br).

Cabe ainda ressaltar que, em acordo com o Decreto nº 5.163/2004, todos os agentes de mercado devem possuir 100% dos seus contratos de venda lastreados em contratos de compra ou garantia física<sup>8</sup> de usinas. Além disso, todo o consumo deve estar lastreado em contratos de compra ou autoprodução de energia. Em ambos cenários os agentes estão sujeitos a penalidades caso não comprovem a existência de lastro ou cobertura junto à CCEE.

Conforme apresentado, o PLD é o valor utilizado para liquidar as diferenças apuradas pela CCEE entre compra (consumo e contratos de compra) e venda (geração e contratos de venda) de energia no mercado de curto prazo (mercado *spot*) por submercado<sup>9</sup>. Ressalta-se, contudo, que o PLD não é um valor de mercado obtido pela livre comercialização de energia entre os agentes e, tampouco, o preço à vista negociado no mercado.

O cálculo e divulgação do PLD é realizado pela CCEE, a partir de dados considerados pela ONS nas decisões de planejamento da operação centralizada do SIN, baseado em modelos de otimização (um modelo de curto prazo acoplado a um de longo prazo), cujo objetivo é minimizar o custo total de operação do sistema hidrotérmico ao longo de um horizonte de

<sup>8</sup> A garantia física de cada usina é calculada conforme metodologia descrita em portarias do Ministério de Minas e Energia (MME) e é um certificado que garante um montante de energia que uma determinada usina pode comercializar por toda sua operação comercial (DOS SANTOS, 2013).

<sup>9</sup> Submercados ou subsistemas de energia são as regiões eletro-energéticas do Brasil (Norte, Sul, Nordeste e Sudeste/Centro-Oeste) separadas devido restrições na transmissão e intercâmbio de energia. A divisão dos submercados é apresentada na Figura 7 da próxima seção.

planejamento. Esse assunto será abordado com um pouco mais de detalhes nas próximas seções, mas antes é importante entender um pouco mais sobre o Sistema Interligado Nacional.

### 2.2.3 O Sistema Interligado Nacional

O Brasil possui um sistema elétrico peculiar, tanto no que se refere ao seu tamanho, quanto às suas características. A produção e transmissão de energia elétrica no Brasil pode ser considerada um sistema hidro-termo-eólico de grande porte com 176.061 MW de potência instalada, constituído, predominantemente, por usinas hidrelétricas que representam 58,5% da capacidade de geração do país. A Tabela 1 apresenta a matriz elétrica brasileira, com informações sobre a quantidade e potência instalada de cada um dos tipos de unidades geradoras.

Tabela 1 – Capacidade de geração do Brasil

Tipo	Quantidade	Potência Instalada (MW)	% (Potência Instalada)
Usina Hidrelétrica	223	103.026,52	58,52%
Usina Termoelétrica	3.179	42.916,61	24,38%
Central Geradora Eólica	1.067	18.506,29	10,51%
Pequena Central Hidrelétrica	542	5.473,57	3,11%
Central Geradora Solar Fotovoltaica	4.872	3.297,14	1,87%
Usina Termonuclear	3	1.990,00	1,13%
Central Geradora Hidrelétrica	731	851,39	0,48%
Central Geradora Undi-elétrica	1	0,05	0,00%
<b>Total</b>	<b>10.618</b>	<b>176.061,56</b>	<b>100%</b>

Fonte: SIGA (2021).

O sistema elétrico brasileiro é constituído pelo Sistema Interligado Nacional (SIN) que engloba o conjunto de instalações e de equipamentos que possibilitam o suprimento de energia elétrica nas regiões do país interligadas eletricamente, e por sistemas isolados independentes, localizados principalmente na Região Amazônica, representando menos de 1% do consumo da carga total do país, sendo suprida, principalmente, por térmicas a óleo diesel.

Em função da magnitude deste sistema interligado e à existência de restrições físicas associadas à transmissão de energia elétrica por longas distâncias, o SIN é dividido em quatro



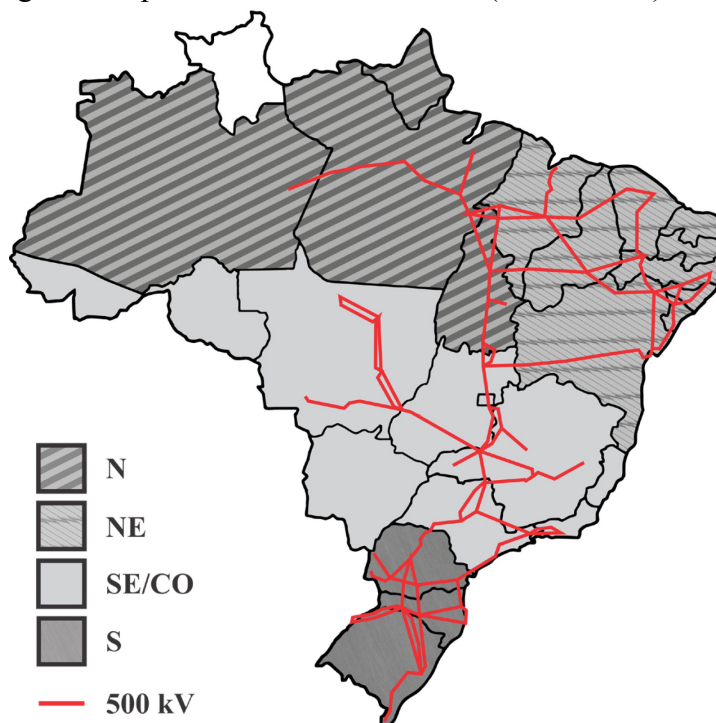
subsistemas<sup>10</sup>, cada qual com seu respectivo mercado *spot* de energia: Norte (N), Nordeste (NE), Sudeste/Centro-Oeste (SE/CO) e Sul (S). Vale destacar que Roraima é o único estado brasileiro que não está integrado ao SIN e seu abastecimento elétrico é feito parcialmente pela Venezuela.

A interconexão dos sistemas elétricos, por meio da malha de transmissão, propicia a transferência de energia entre subsistemas, permite a obtenção de ganhos sinérgicos e explora a diversidade entre os regimes hidrológicos das bacias. A integração dos recursos de geração e transmissão permite o atendimento ao mercado com segurança e economicidade (ONS, 2017).

Segundo dados do ONS (2021a), a extensão da rede básica de transmissão (tensões acima de 230 kV) do SIN em 2020 era de 145.600 km, dos quais cerca de 38,1% correspondem à classe de tensão de 230 kV e 36,5% de 500 kV sendo, conjuntamente, as principais responsáveis pela interligações entre os subsistemas.

A Figura 7 apresenta, de forma esquemática, a divisão dos subsistemas no SIN e o diagrama unifilar da rede básica de transmissão de 500 kV de tensão (principal rede de interligação entre os subsistemas).

Figura 7 – Diagrama esquemático dos subsistemas (submercado) de energia elétrica



Fonte: Elaboração própria com informações do ONS.

<sup>10</sup> Apesar de subsistema e submercado representarem a mesma região geográfica, cada um dos termos é utilizado, respectivamente, no âmbito da ONS e CCEE.

A capacidade instalada de geração do SIN é composta, principalmente, por usinas hidrelétricas que apresentam os menores custos operacionais e com possibilidade de armazenar energia na forma de água acumulada nos reservatórios. No entanto, são as termelétricas que garantem a segurança de fornecimento, porém, apresentam custos operacionais mais elevados, diretamente relacionados ao consumo de seus combustíveis. A instalação de usinas eólicas, principalmente nas regiões Nordeste e Sul, apresentou um forte crescimento, praticamente dobrando a capacidade instalada nos últimos cinco anos e aumentando a importância dessa geração para o atendimento do mercado (ONS, 2021b).

A decisão sobre o despacho e a transmissão de energia elétrica é feita de forma centralizada pelo Operador Nacional do Sistema (ONS), baseada em um critério de custo mínimo de operação. Ou seja, a decisão sobre a quantidade a ser ofertada de energia por cada usina, em cada instante do tempo, não cabe ao proprietário desta, e sim ao ONS que libera o despacho de cada uma seguindo uma ordem de mínimo custo marginal de operação até que a demanda seja plenamente atendida.

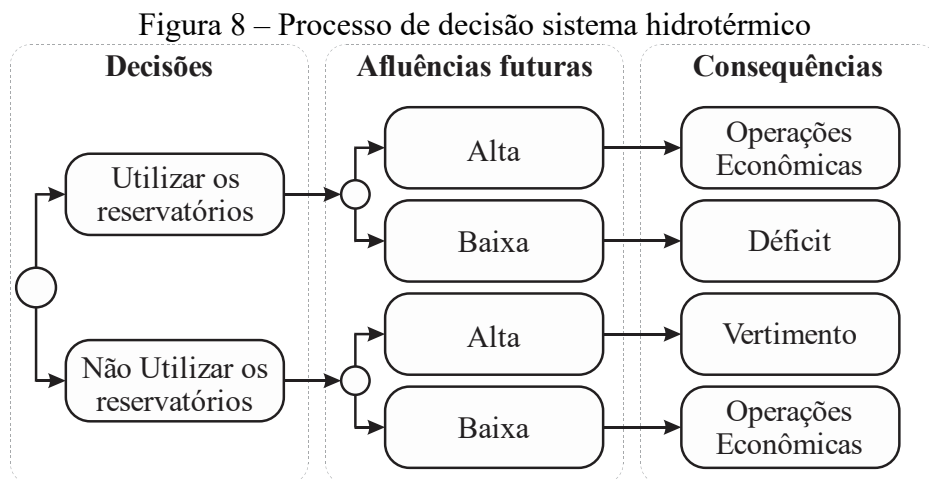
#### **2.2.4 Planejamento da Operação em Sistemas Hidrotérmicos**

O objetivo básico do planejamento da operação de um sistema hidrotérmico é calcular a política de operação que permite determinar metas de geração para cada usina do sistema que atendam à demanda e minimizem o valor esperado do custo de operação ao longo do período de planejamento e atendendo um critério de aversão ao risco. Esse custo é composto pelo custo variável de combustível das usinas termelétricas e pelo custo associado a eventuais déficits no suprimento de energia, representado por uma função de penalização (CEPEL, 2018).

Do ponto de vista econômico, o aspecto mais relevante desses sistemas é a possibilidade do gerenciamento da água armazenada nos reservatórios para geração de energia, minimizando os custos com combustíveis. Contudo, a geração hidroelétrica é limitada pela capacidade e o nível de armazenamento dos reservatórios, o que gera uma dependência entre a decisão operativa do presente e os custos operativos no futuro (LUZ, 2011).

Dessa forma, a decisão sobre quando utilizar a água armazenada nos reservatórios, está intrinsecamente ligada à incerteza quanto às aflúncias futuras, devendo resultar de uma análise probabilística de seu comportamento.

A utilização prioritária de geração hidroelétrica máxima disponível em cada período é a premissa mais econômica do ponto de vista imediato, por minimizar os custos de combustíveis das termoelétricas. No entanto, essa medida apresenta maiores riscos de déficits futuros, com possíveis cortes de carga ou medidas de racionamento de energia caso se configure, por exemplo, condições hidrológicas desfavoráveis. Por sua vez, manter o nível dos reservatórios o mais elevado possível, no intuito de garantir a máxima confiabilidade de fornecimento, significa utilizar mais termoelétricas na geração de energia, aumentando assim os custos da operação. Contudo, caso as aflúências futuras sejam elevadas, os reservatórios das hidrelétricas podem verter (escoamento de água acumulada em excesso), representando um desperdício do potencial de geração de energia (LUZ, 2011). A Figura 8 ilustra esse processo de decisão de utilização de reservatórios com os possíveis cenários de aflúência e as respectivas consequências.



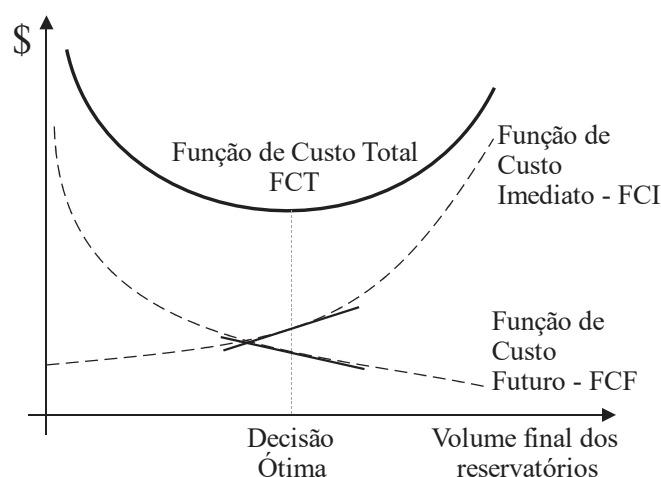
Fonte: Adaptado de (FILHO, 2014).

Portanto, na operação de um sistema hidrotérmico, deve-se comparar o benefício presente (imediato) do uso da água e o benefício futuro de seu armazenamento. Matematicamente, esta lógica pode ser expressa pela Função de Custo Imediato (FCI), representando os custos de geração térmica no estágio  $t$ . O custo imediato aumenta à medida que diminui a utilização de recursos hídricos disponíveis. Já a Função de Custo Futuro (FCF) está associada ao custo esperado de geração térmica e ao racionamento desde o final do estágio  $t$  (início do próximo estágio,  $t+1$ ) até o final do período de estudo. A FCF diminui à medida que aumenta o volume final armazenado, em função da maior disponibilidade de energia hidroelétrica no futuro.

Nesse contexto, o uso ótimo da água armazenada é aquele que minimiza a soma dos custos: de geração térmica no presente e o valor esperado do custo de geração até o fim do período de estudo.

Observa-se na Figura 9 que o problema de encontrar o menor custo global se resume a encontrar o ponto de mínimo da Função de Custo Total (FCT) – dada pela soma das FCF e FCI. Este ponto é onde a derivada de FCT em relação ao volume final é igual a zero ou, ainda, onde as derivadas de FCF e FCI em relação ao volume final são iguais em módulo. Assim, o chamado “valor da água” pode ser calculado pela inclinação da FCI (ou da FCF) no ponto de decisão ótimo correspondente ao nível de armazenamento ao final do estágio  $t$  (MEDEIROS, 2003).

Figura 9 – Funções de custo e o ponto ótima de operação



Fonte: Adaptado de (MEDEIROS, 2003).

As funções de custo imediato e futuro são compostas por inúmeras variáveis, o que torna necessária a adoção de um modelo matemático para obtenção da decisão ótima da operação.

Este valor não se mede de maneira isolada em cada usina, pois depende da operação conjunta do sistema. Para se obter ganhos sinérgicos em um sistema hidrotérmico interligado, é necessário operar o sistema de maneira integrada, otimizando conjuntamente a operação de todas as usinas - térmicas, hidroelétricas, biomassa, eólica e solar, e as decisões de intercâmbios de energia, com o objetivo de minimizar o custo total de operação (CEPEL, 2018).

Faz-se, então, um desdobramento do problema em vários subproblemas, utilizados modelos com diferentes graus de detalhamento para a representação do sistema, com horizontes de planejamento (ou de otimização) distintos. O Operador Nacional do Sistema (ONS) utiliza desses modelos para o planejamento e a programação da operação do Sistema Interligado Nacional.

A cadeia completa desses modelos com diferentes horizontes, desenvolvidos pelo Centro de Pesquisas de Energia Elétrica (CEPEL), é composta pelos programas NEWAVE (planejamento da operação de longo e médio prazo), DECOMP (planejamento da operação de curto prazo) e DESSEM (despacho de curto prazo).

O programa NEWAVE realiza a otimização energética do sistema para um horizonte de cinco anos, mês a mês, trabalhando com o conjunto de reservatórios agregado por subsistema, considerando somente as restrições de intercâmbio de energia entre os subsistemas. O programa DECOMP realiza a otimização da operação no curto prazo, períodos semanais, detalhando o sistema que foi representado de forma agregada no modelo NEWAVE, considerando as usinas de forma individualizada e todas restrições operativas do sistema. O DESSEM é um modelo com representação das usinas termelétricas por unidade geradora, considerando as restrições de *unit commitment*<sup>11</sup> e a operação em ciclo combinado, enquanto as usinas hidrelétricas são representadas de forma individualizada. O DESSEM atua de forma integrada aos modelos DECOMP e NEWAVE, e vem sendo utilizado oficialmente pelo ONS desde janeiro de 2020 para a programação diária da operação do SIN (CCEE, 2021).

Esses programas utilizam como dados de entrada as características das usinas hidráulicas e térmicas do SEB. Possuem um modelo para geração de cenários de afluência, um cenário de projeção de demanda, limites de intercâmbio entre os subsistemas, custo de déficit de energia, algumas restrições operacionais, dentre outras. Eles também compartilham da mesma modelagem matemática para a solução do problema de otimização: a programação dinâmica dual estocástica (PDDE).

### 2.2.5 Cálculo do Preço de Liquidação das Diferenças

Conforme já mencionado, o planejamento da operação de um sistema hidrotérmico visa determinar, para cada etapa do período de estudo, as metas de geração para cada usina, de modo a atender a demanda e minimizar o custo esperado de operação ao longo do período.

Os modelos matemáticos utilizados para calcular os despachos ótimos baseiam-se nas condições hidrológicas presente e futura, na demanda de energia, nos preços de combustível,

---

<sup>11</sup> *Unit Commitment* é o conjunto de restrições operativas que engloba a representação da rampa de acionamento, tempo mínimo de acionamento, rampa de desligamento, tempo mínimo desligamento até um novo acionamento, além de taxa e frequência de tomada de carga, dentre outras.

no custo de déficit, na entrada de novos projetos e na disponibilidade dos sistemas de geração e transmissão (CCEE, 2021). Tem-se, portanto, o despacho ótimo para ambos os tipos de geração, hidráulica e térmica, e para cada subsistema (norte, nordeste, centro-oeste/sudeste e sul). Como resultados são obtidos os Custos Marginais de Operação (CMO) do sistema, associado à restrição de atendimento à demanda, significando o custo de produção de 1 (um) megawatt-hora (MWh) adicional no ponto ótimo de minimização de custos (em R\$/MWh).

O cálculo do CMO é realizado por subsistema e baseia-se no despacho *ex-ante*, ou seja, é apurado com base em informações previstas, anteriores à operação real do sistema, considerando-se os valores declarados e previstos de disponibilidade de geração e consumo de cada subsistema. O processo completo de cálculo consiste na utilização dos modelos computacionais NEWAVE, DECOMP e DESSEM, que produzem como resultado o CMO de cada subsistema, em base mensal, semanal e diária, respectivamente (CCEE, 2021).

Na determinação do PLD, a CCEE utiliza os mesmos modelos adotados pelo ONS para determinação da programação e despacho de geração do sistema, porém com as restrições internas aos subsistemas suprimidas, de forma que a energia comercializada seja considerada igualmente disponível em todos os pontos, e conseqüentemente seu preço uniforme em cada submercado (as restrições entre os subsistemas são mantidas).

Anteriormente à entrada do cálculo do PLD com granularidade horário, até dezembro de 2020, o PLD era calculado pela CCEE semanalmente para cada patamar de carga com base no Custo Marginal de Operação, limitado por um preço máximo e mínimo definidos pela ANEEL vigentes para cada período de apuração e para cada submercado (CCEE, 2021).

O cálculo do PLD por modelos computacionais que dependem de cenários futuros de operação como o NEWAVE, DECOMP e DESSEM requer um processamento computacional robusto. Alternativamente, pesquisas utilizando métodos de séries temporais têm apresentado resultados satisfatórios para a previsão de curto prazo, além de facilidade de implementação.

A próxima seção apresenta a pesquisa bibliográfica referente à trabalhos de previsão do preço de energia elétrica por modelos de séries temporais.

### 2.3 PESQUISA BIBLIOGRÁFICA

Cada mercado de energia possui uma particularidade e diferentes variáveis a serem consideradas na formação de preços. No Brasil, a matriz energética é predominantemente

hidráulica e, como esta depende de fatores naturais, com baixa acuidade em termos de previsibilidade para a disponibilidade de recursos, os preços de energia elétrica no mercado de curto prazo variam significativamente.

Nas aplicações de modelos de previsão dos preços de energia elétrica devem considerar o contexto do qual estão inseridos. Portanto, nem sempre a importação de um determinado modelo de previsão aplicado em contexto diferente do brasileiro irá gerar bons resultados.

De qualquer forma, os estudos sobre a previsibilidade de preços no setor energético no Brasil e no mundo apresentam distintos resultados de acordo com o mercado avaliado e os métodos de previsão considerados.

Dentre os diferentes modelos e métodos, a pesquisa se restringiu àqueles fundamentados em séries temporais, que inclui os modelos baseados em regressão, tais como suavização exponencial, *Autoregressive Integrated Moving Average* (ARIMA), redes neurais artificiais (RNA), *Generalized Autorregressive Conditional Heteroskedasticity Process* (GARCH), lógica fuzzy, dentre outros.

Os trabalhos de Bowden e Payne (2008); Contreras et al. (2003) e Ziel e Steinert (2018) abordam o problema da previsão do preço da energia, a partir da replicação das propriedades estatísticas de seus dados histórico. Nesses trabalhos foram desenvolvidos modelos estocásticos ou econométricos, utilizando metodologias de previsão de séries temporais. Em Soares (2013) foi realizado um estudo, análise e comparação de modelos de suavização exponencial, também conhecidos como métodos de Holt-Winters, na previsão do preço de energia para o Mercado Ibérico de Energia Elétrica e aferidas por meio de métricas de erros.

Já as técnicas baseadas em inteligências artificial, modeladas a partir de uso de redes neurais, lógica difusa, combinação de técnicas, dentre outros; podem ser citados nos trabalhos de Panapakidis e Dagoumas (2016), Singhal e Swarup (2011) e Wang et al. (2017). Já em Lago, De Ridder e De Schutter (2018) são propostos quatro modelos de previsão do preço *spot* da energia por redes neurais de aprendizado profundo e as comparam com diversos métodos estatísticos tradicionais.

Em relação ao mercado de energia nacional podem ser citados alguns trabalhos que abordam propostas de previsão do PLD.

Em Medeiros (2003), por exemplo, uma metodologia foi proposta para a previsão do preço *spot* da energia elétrica em sistemas hidrotérmicos com base hidráulica, como o brasileiro, baseada em sistemas neuro-fuzzy e programas do planejamento da operação. O modelo utilizou da simulação feita pelo NEWAVE (Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Longo e Médio Prazo) do Centro de Pesquisas de

Energia Elétrica (CEPEL) considerando variáveis como energia natural afluyente, energia armazenada e custo marginal de operação.

Em Rodrigues (2009) é proposta a projeção de preços para o mercado de curto prazo brasileiro por meio de Redes Neurais Artificiais (RNA). O modelo utiliza da arquitetura RNA do tipo *Multilayer Perceptron* (MLP) e dados de entrada baseados no histórico da operação do Sistema Elétrico Brasileiro.

Em Almeida (2012) foi implementada uma rede neural artificial do tipo *feed-forward* para a previsão do preço da energia elétrica do mercado da Califórnia e Brasil. Foram testados diferentes parâmetros de treinamento e com as diferentes combinações das variáveis, o que se mostrou, de um modo geral, eficiente, tendo em vista que as redes convergiram para a previsão das tendências do preço da energia.

O trabalho de Filho (2014) propõe uma combinação híbrida de filtro ARIMA e RNA para a previsão do PLD usando como variáveis explanatórias a geração hidráulica, geração térmica, carga de energia, energia natural afluyente e energia armazenada. Foi proposto também um modelo classificador dos preços futuros da energia elétrica, através do uso combinado de algoritmo de árvore de decisão e redes neurais. Os modelos foram implementados no software IBM SPSS *Modeler*. Modelos híbridos para previsão do PLD utilizando redes neurais também são propostos em (DE CAMPOS et al., 2020).

Em Junior (2016) são utilizadas redes neurais diretas e recorrentes (além da comparação com o modelo ARIMA) para a previsão do preço de energia elétrica de curto prazo brasileiro com uso da técnica de correlação para seleção das variáveis de entrada da rede e também para escolha dos atrasos nestas variáveis selecionadas. Mostra-se que, na previsão de um passo à frente, as redes neurais implementadas superam o desempenho do modelo ARIMA para esta série e, em geral, a rede direta apresenta melhor resultado que a recorrente.

Já Toshioka (2017) implementa redes neurais artificiais *Multilayer Perceptron* e Santos (2019) implementa as do tipo *Long Short Term Memory* (LSTM), ambas utilizando a biblioteca *Deep Learning Toolbox* do MATLAB para previsão do PLD em uma semana para todos os submercados e patamar de cargas. Redes neurais LSTM também foram utilizadas por Porto e Olivi (2020) para a previsão do PLD.



### 3 MATERIAIS E MÉTODOS DE PESQUISA

A pesquisa possui uma abordagem quantitativa de natureza aplicada, na medida em que foram desenvolvidas análises comparativas de modelos preditivos de séries temporais empregados na previsão do PLD, utilizando de recursos e técnicas estatísticas com ênfase na avaliação dos resultados obtidos.

Inicialmente foi realizada uma etapa exploratória por meio de revisão da literatura referentes à aplicação de modelos de previsão de séries temporais empregadas na predição do PLD. Nessa fase, objetivou-se o aprimoramento de conceitos e ideias de forma a tornar o problema mais explícito.

O conjunto de dados para o desenvolvimento da pesquisa, delineados pelos procedimentos técnicos de um estudo de caso, foram obtidos do repositório de séries históricas do Preço de Liquidação das Diferenças (PLD) fornecido pela Câmara de Comercialização de Energia Elétrica (CCEE, 2020) e do Planejamento da Operação do Sistema Elétrico Brasileiro disponibilizados pelo Operador Nacional do Sistema (ONS, 2021b).

É importante destacar a estruturação e tratamento dos dados, além de sua análise por estatística descritiva, de forma que possam ser utilizados pelos modelos de previsão.

Após tratados, os dados foram separados em dois diferentes subconjuntos: treinamento (80% da amostra) e validação (20% da amostra). O subconjunto de treinamento foi utilizado para estimação, configuração e ajuste dos parâmetros de acordo com metodologia específica para cada um dos modelos considerados. Após essas definições, o desempenho de previsão dos modelos foi comparado por meio de métricas de erro utilizando como teste o subconjunto destinado para essa finalidade.

Na análise dos resultados buscou-se identificar os fatores que contribuem para a melhoria da previsão para o caso em estudo. Os resultados comparativos apontam o modelo com melhor desempenho para a previsão do PLD, e ainda as variáveis de entrada, no modelo multivariado, que mais influenciam nessa predição.

As modelagens, análises estatísticas, carregamento dos dados em formato *comma-separated values* (CSV), gráficos e testes foram implementados no ambiente computacional *web Jupyter Notebook* (JUPYTER, 2020) utilizando da linguagem de programação Python na versão 3.7.9 (FOUNDATION, 2020).

Para o estudo de previsão das séries temporais, foram seguidas as etapas estabelecidas por Hyndman e Athanasopoulos (2018), a saber: definição do problema, coleta de informações

e análise preliminar dos dados, escolha e ajuste do modelo, uso e avaliação do modelo de previsão.

Alguns resultados preliminares foram antecipados do Capítulo 4 e apresentados nas etapas detalhadas na sequência, no intuito de manter o contexto do qual eles estão inseridos.

### 3.1 DEFINIÇÃO DO PROBLEMA

As questões que devem ser abordadas durante esta fase incluem o período e horizonte de previsão, com que frequência as previsões precisam ser revisadas e qual nível de precisão é requerido.

O nível de detalhe pretendido é uma consideração extremamente importante e é afetado por fatores como disponibilidade de dados, acurácia, custos de análise e preferências gerenciais.

Neste trabalho, a comparação de métodos de previsão de séries temporais foi promovida por meio da aplicação de modelos na predição do valor do preço *spot* de energia elétrica no mercado de curto prazo, ou Preço da Liquidação das Diferenças (PLD). Conforme já discutido, o PLD é divulgado semanalmente<sup>12</sup> pela Câmara de Comercialização de Energia Elétrica (CCEE), calculado conforme metodologia de otimização do planejamento da operação energética do Sistema Interligado Nacional (SIN).

Portanto, para o caso em estudo, o período de previsão (a unidade básica de tempo) é expresso em semanas. Especificamente, trata-se da semana operativa, que é a semana utilizada pelo Operador Nacional do Sistema (ONS) que se inicia à zero hora de sábado e termina às 23h59 da sexta-feira da semana subsequente.

O PLD é definido por submercado (Norte, Sul, Nordeste, Sudeste/Centro-Oeste) e por patamar de carga (leve, média e pesada) que representa o nível de consumo de energia elétrica.

Especificamente neste estudo, utilizou-se a base histórica de PLD semanal do submercado Sudeste/Centro-Oeste – por representar cerca de 41,7% da potência instalada no SIN de acordo com SIGA (2021) – no patamar de carga médio – horários em que o consumo de energia elétrico é mediano.

---

<sup>12</sup> A partir de 1º de janeiro de 2021, o PLD passou a ser calculado oficialmente para cada submercado em base horária, conforme proposto pela Comissão Permanente para Análise de Metodologias e Programas Computacionais do Setor Elétrico (CPAMP), com cronograma de implantação definido pela Portaria MME 301/2019.

Estipulou-se o horizonte de previsão de até 12 semanas (períodos) com o intuito de avaliar o desempenho de modelos tanto nas previsões de períodos mais imediatos (primeiras semanas de previsão) quanto em períodos de semanas um pouco mais adiantadas.

A frequência de revisão das previsões também foi realizada de forma semanal, ou seja, a cada período novas previsões foram preparadas para um horizonte de 12 semanas à frente. Os parâmetros dos modelos preditivos também foram revistos para cada previsão realizada, exceto os modelos MLP que, devido aos custos computacionais, tiveram seus parâmetros estimados e ajustados uma única vez para a previsão de todos os períodos do subconjunto de teste.

Nos modelos univariados, a base histórica do PLD é a única série temporal utilizada para estimação e ajuste dos parâmetros de previsão. Já no modelo multivariado outras séries temporais também foram utilizadas na entrada do modelo como variáveis explicativas para a previsão do PLD.

A base de dados utilizada é detalhada na seção seguinte.

### 3.2 COLETA E ANÁLISE PRELIMINAR DOS DADOS

A coleta de dados consiste em obter o histórico relevante para as variáveis que devem ser previstas, incluindo informações históricas sobre variáveis preditoras potenciais. Nesses conjuntos, geralmente, é necessário lidar com valores ausentes, discrepantes e outros problemas relacionados aos dados.

A análise de dados é uma etapa preliminar e importante para a seleção do modelo de previsão. Os gráficos de séries temporais dos dados devem ser construídos e inspecionados visualmente para padrões reconhecíveis, como tendências e componentes sazonais ou outros componentes cíclicos. Os resumos numéricos dos dados, como média da amostra, desvio padrão, percentis e autocorrelações, também devem ser avaliados. Se variáveis preditoras potenciais estiverem disponíveis, os gráficos de dispersão de cada par de variáveis devem ser examinados.

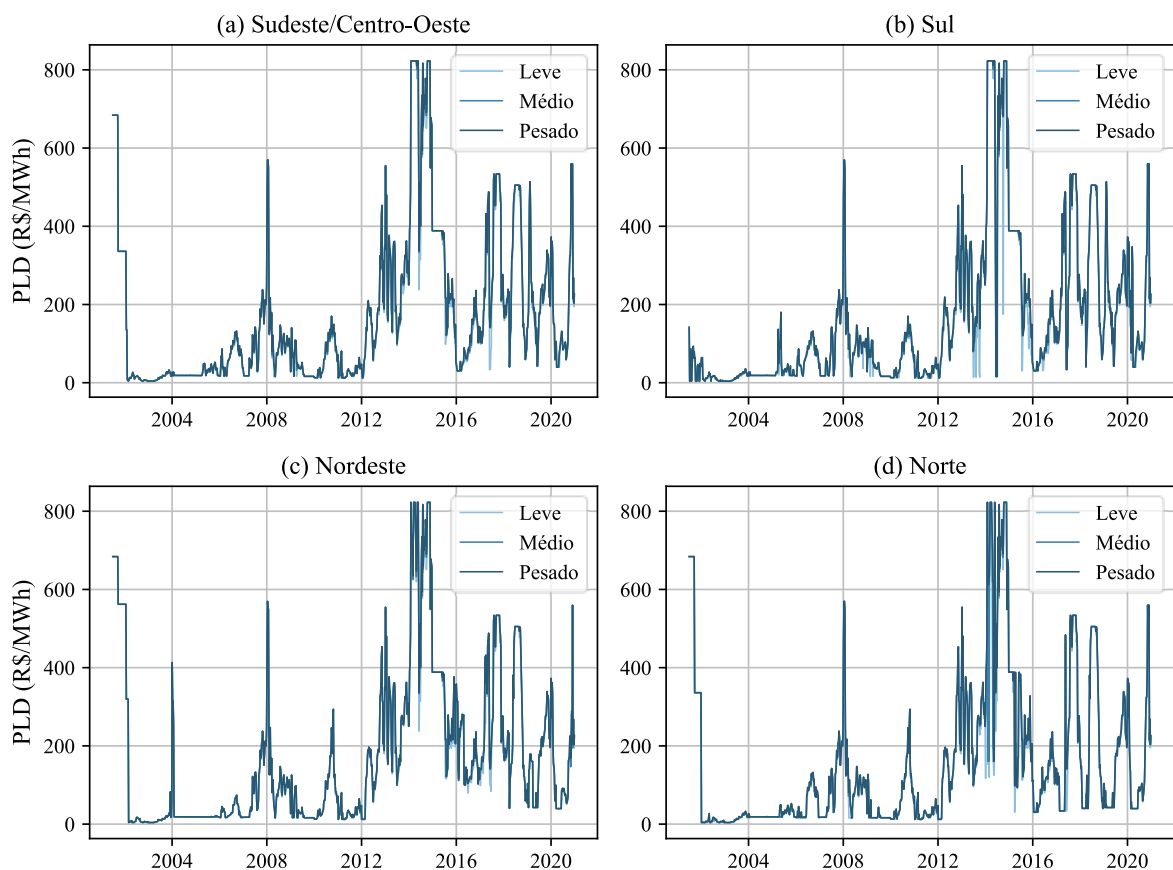
Essas informações geralmente sugerem os tipos iniciais de métodos e modelos de previsão quantitativa a serem explorados.

O conjunto de dados referente à série histórica do PLD (variável a ser predita) foi obtida em CCEE (2020). O conjunto apresenta os valores do PLD, para cada submercado e patamar de carga, correspondente às semanas operativas (de sábado à sexta-feira subsequente) de 30 de

junho de 2001 à 26 de dezembro de 2020. Foram expurgados os valores de PLD que foram divulgados fora da semana operativa. Essa divulgação fora da ordem regular ocorreu em 7 oportunidades e sempre para o primeiro dia de alguns meses.

Cada uma das séries temporais referente à divulgação do PLD, por submercado e patamar de carga (leve, médio e pesado), é apresentada graficamente na Figura 10.

Figura 10 – Histórico PLD por submercado e patamar de carga



Fonte: Elaboração própria a partir de dados da (CCEE, 2020).

O PLD é mensurado em reais por megawatt-hora (R\$/MWh). Os limites máximos e mínimos de PLD são definidos para cada período de apuração e submercado. Os patamares de carga são definidos pelo ONS, também para cada período de apuração.

A Tabela 2 apresenta, por meio de estatística descritiva, as medidas de dispersão dos dados históricos do PLD para cada submercado no patamar de carga médio.

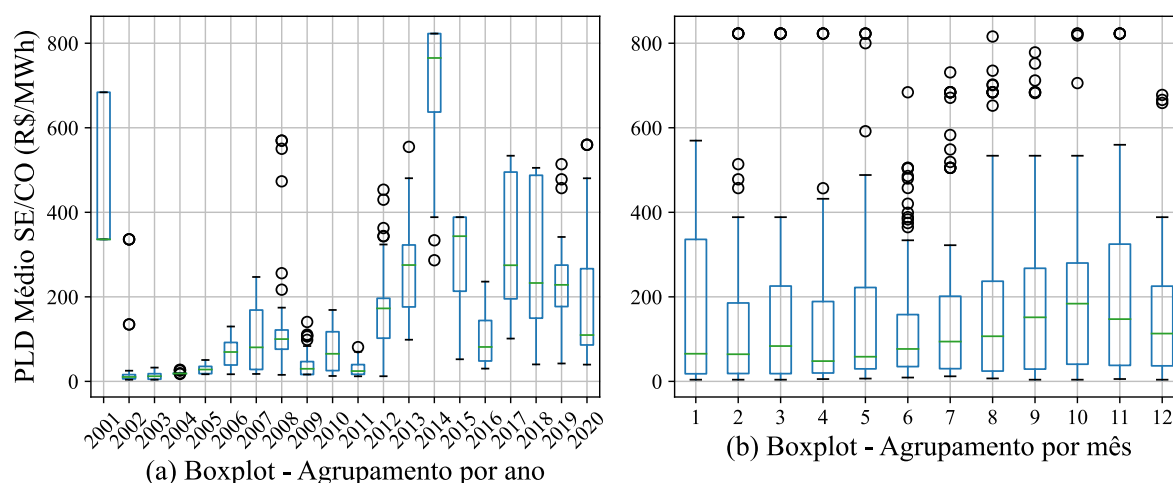
Tabela 2 – Estatística descritiva PLD (R\$/MWh) no patamar médio

Medida Estatística	Sudeste/Centro-Oeste	Sul	Nordeste	Norte
Média	171,08	157,16	173,67	151,40
Desvio Padrão	194,55	185,58	197,53	184,95
Mínimo	4,00	4,00	4,00	4,00
25%	25,03	22,90	18,59	18,59
50%	99,51	86,97	102,67	79,15
75%	237,49	215,81	254,35	206,86
Máximo	822,83	822,83	822,83	822,83

Fonte: Elaboração própria a partir de dados da (CCEE, 2020).

Cada conjunto considerado contém 1.018 períodos (semanas operativas). Em uma análise preliminar, observa-se que os dados históricos do PLD apresentam elevada volatilidade e dispersão, com o desvio padrão acima da média e, ainda, um distanciamento, relativamente considerável entre os valores extremos. Verifica-se também que mais de 75% dos dados estão abaixo dos valores de 255,0 R\$/MWh.

Utilizando os dados históricos do PLD do submercado Sudeste/Centro-Oeste (SE/CO) no patamar de carga médio, separando o conjunto por mês e ano é possível apresentar a variação dos dados por meio de um diagrama de caixa (*boxplot*), conforme mostrado na Figura 11.

Figura 11 – *Boxplot* do PLD médio submercado Sudeste/Centro-Oeste

Fonte: Elaboração própria a partir de dados da (CCEE, 2020).

Em uma análise preliminar, verifica-se que a definição de qualquer padrão sazonal e cíclico pode ser contestada pela elevada dispersão dos dados. De qualquer forma, é possível observar pela Figura 11 (b) que, na média, os valores mais elevados do PLD tendem a ocorrer no último quadrimestre do ano. É esperado que nesse período os reservatórios das hidrelétricas

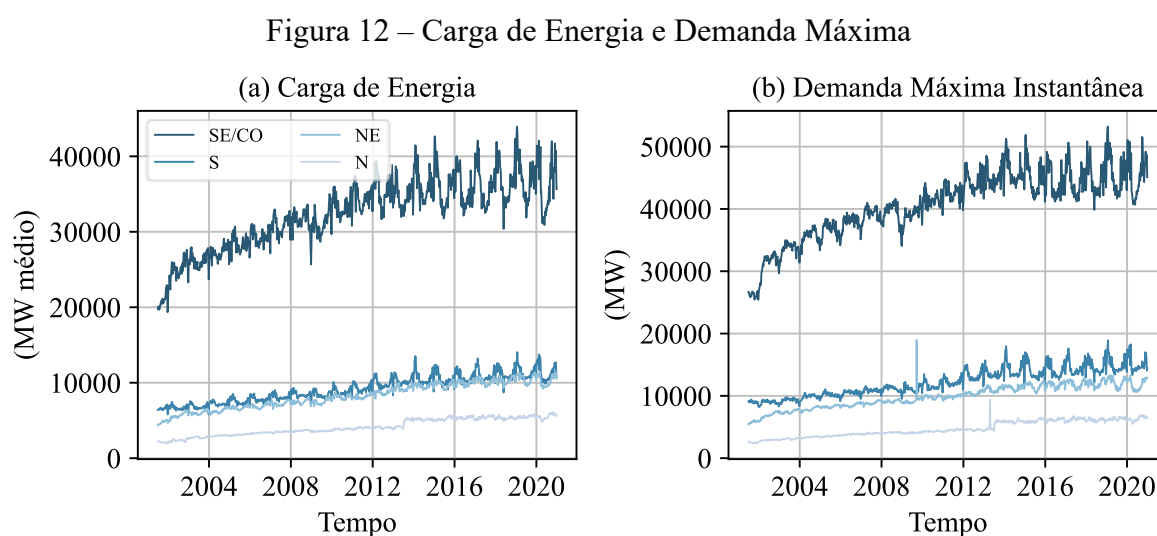
estejam em seus níveis mais baixos ao se considerar o período de estiagem que, para a região sudeste, normalmente ocorre entre os meses de abril a setembro<sup>13</sup>.

Como variáveis predictoras potenciais na previsão dos valores do PLD utilizou-se os dados históricos da operação do sistema elétrico disponibilizados pelo Operador Nacional do Sistema em ONS (2021b). As variáveis utilizadas nos modelos de previsão são apresentadas na sequência.

Carga de Energia, dado em MW médio<sup>14</sup>, é a demanda média requerida por determinado subsistema durante um período de referência. Representa a média de energia elétrica consumida em MWh por um determinado intervalo de tempo (expresso em horas).

Já a Demanda Máxima Instantânea, dada em megawatt (MW) corresponde ao maior valor de demanda requerida num determinado intervalo de tempo. Condizente com a divulgação do PLD, utilizou-se a semana operativa como intervalo de tempo.

A Figura 12 apresenta o histórico da Carga de Energia e Demanda Máxima para as semanas operativas de 30 de junho de 2001 a 26 de dezembro de 2020 para os quatro subsistema de energia: Sudeste/Centro-Oeste (SE/CO), Nordeste (NE), Sul (S) e Norte (N).



Fonte: Elaboração própria a partir de dados da (ONS, 2021b).

Os gráficos da Figura 12 compartilham uma só legenda de identificação dos subsistemas.

<sup>13</sup> Segundo Minuzzi, R. B., Sediya, G. C., Barbosa, E. M., Júnior, J. C. F. M. Climatologia do comportamento do período chuvoso da região sudeste do Brasil, Revista Brasileira de Meteorologia, 2007

<sup>14</sup> O MW médio é calculado por meio da razão MWh/h, onde MWh representa a energia média produzida/consumida e h representa a quantidade de horas do período de tempo no qual a referida quantidade de energia foi produzida/consumida.

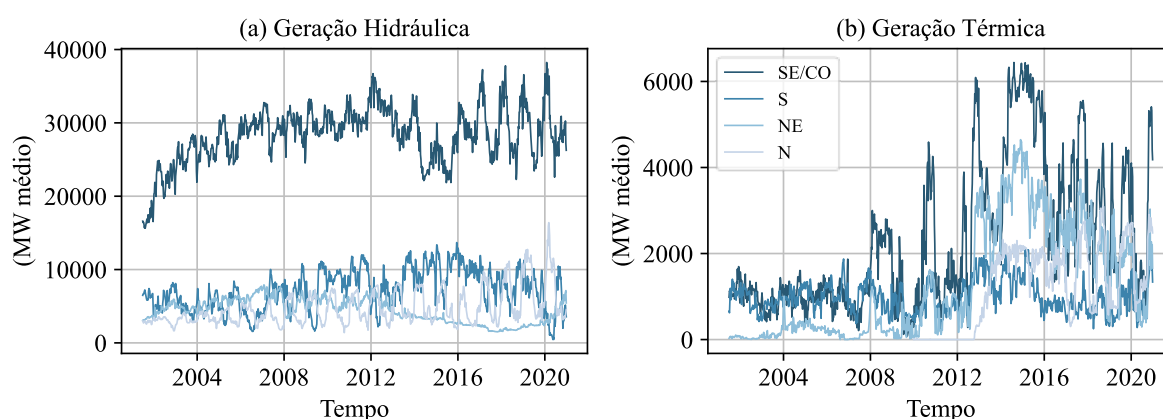
Em suma, a Carga de Energia representa o consumo médio de energia elétrica do subsistema durante uma semana operativa e a Demanda Máxima é a maior demanda requerida durante esse mesmo período.

Observa-se, pela Figura 12, o maior consumo de energia elétrica do subsistema SE/CO em comparação aos demais subsistemas. Além disso, os dados apresentam sazonalidade e uma nítida tendência de crescimento.

Assim como os dados de consumo apresentados, o ONS também disponibiliza dados históricos de geração de energia elétrica que podem ser discriminados por tipo de usina (hidrelétrica, térmica, eólica, solar e nuclear), tipos de combustíveis (no caso de usina térmica), subsistema e modalidade de operação.

A Figura 13 apresenta a geração de energia elétrica, em MW médio por semana operativa, para o conjunto de usinas (a) hidráulicas e (b) térmicas para cada subsistema, do período de 30 de junho de 2001 a 26 de dezembro de 2020.

Figura 13 – Geração de energia hidráulica e térmica



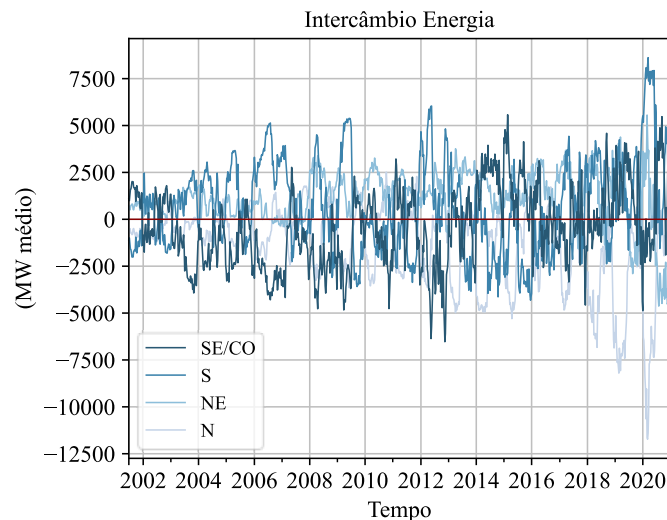
Fonte: Elaboração própria a partir de dados da (ONS, 2021b).

Algumas observações preliminares são apresentadas em relação à geração de energia mostradas na Figura 13. Apesar da aparente natureza sazonal da geração de energia por usinas hidrelétricas, há uma diminuição da produção, no subsistema SE/CO, entre os anos de 2014 a 2016, conforme Figura 13 (a). Em comparação aos valores do PLD apresentados na Figura 10, percebe-se que, nesse período, o PLD apresentou elevados valores. Outra observação é que há, para todos os subsistemas, um crescimento do volume de geração de energia por usinas térmicas a partir de 2008, ficando mais pronunciado esse aumento a partir de 2012, conforme é mostrado na Figura 13 (b). A partir desse período observa-se pela Figura 10 que o PLD exibe uma maior volatilidade.

Além da produção de energia o ONS fornece dados referente ao intercâmbio de energia entre os subsistemas, dado em MW médio. As grandezas representam a soma das medidas de fluxo médio de potência ativa nas linhas de transmissão de fronteira entre os subsistemas. Devido às condições físicas do SIN, esse intercâmbio pode ocorrer apenas entre o subsistema SE/CO com todos os demais subsistemas e entre o subsistema Norte com o Nordeste.

A Figura 14 apresenta o intercâmbio líquido de energia média para cada subsistema, por semana operativa de 30 de junho de 2001 a 26 de dezembro de 2020. Um valor positivo para essa grandeza significa que o referido subsistema recebeu aquela quantidade de energia líquida média transferida de outro(s) subsistema(s). Já um valor negativo significa, portanto, que esse subsistema forneceu energia para outro(s) subsistema(s).

Figura 14 – Intercâmbio de energia elétrica entre os subsistemas



Fonte: Elaboração própria a partir de dados da (ONS, 2021b).

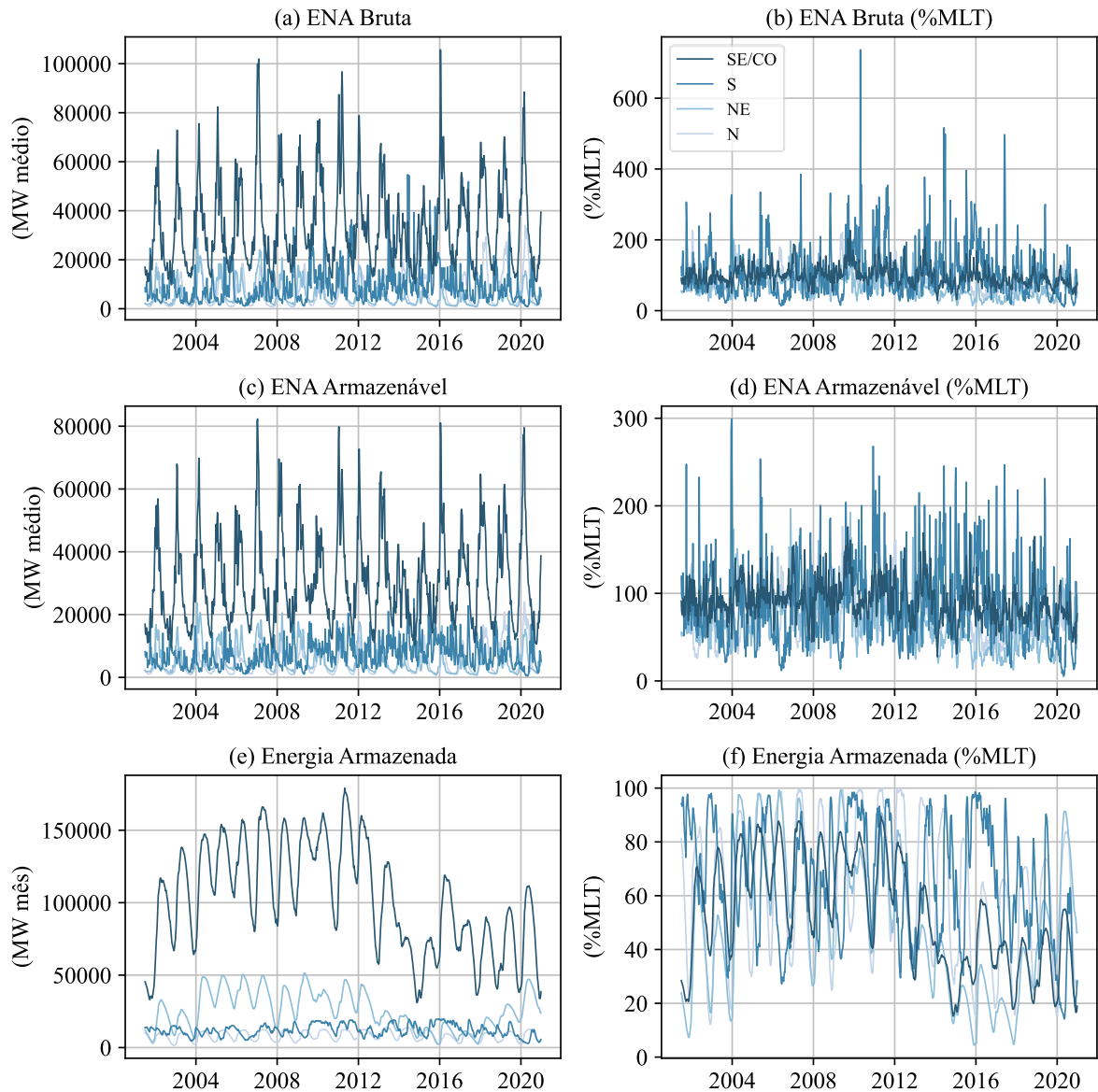
Além dos intercâmbios de energia entre subsistemas, há o intercâmbio internacional de energia entre o SIN com a Argentina e o Uruguai, que não foram aqui apresentados.

Ao somar, para determinado subsistema, a geração de energia elétrica de todas suas fontes juntamente com o intercâmbio líquido de energia realizado nesse mesmo subsistema, obtém-se um montante de energia tal que corresponde, exatamente, à Carga de Energia apresentado na Figura 12 (a) para o subsistema. Isso era esperado já que a energia produzida e importada de outros subsistemas deve coincidir com a energia consumida pelas cargas desse subsistema.



O ONS também fornecem dados históricos relacionados às energias produzíveis pelas afluências<sup>15</sup> das vazões naturais, bem como àquelas associadas ao volume de água nos reservatórios que pode ser convertida em geração de energia. A Figura 15 apresenta esses dados.

Figura 15 – Energia Natural Afluente e Armazenável



Fonte: Elaboração própria a partir de dados da (ONS, 2021b).

A Energia Natural Afluente (ENA) Bruta, Figura 15 (a) e (b), representa a energia produzível pela usina e é calculada pelo produto das vazões naturais aos reservatórios com as produtividades a 65% dos volumes úteis. A ENA Armazenável, Figura 15 (c) e (d), considera as vazões naturais descontadas das vazões vertidas nos reservatórios.

<sup>15</sup> Afluência é o volume de água que passam numa dada seção durante um período de tempo determinado.

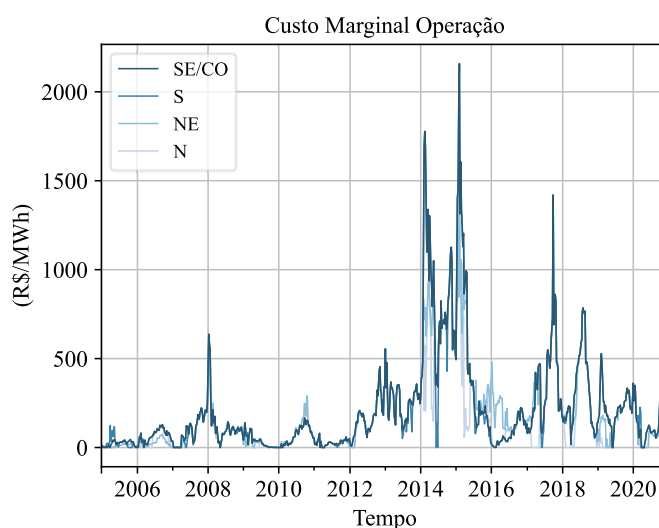
Cada uma dessas variáveis (ENA Bruta e Armazenável) é apresentada tanto em MW médio – representando a energia média produzível pelo período de tempo considerado (semana operativa) – e também em porcentagem da Média de Longo Termo (MLT), que corresponde à média de um mesmo período, verificadas durante a série histórica de observações.

A Energia Armazenada (EAR), Figura 15 (e) e (f), representa a energia associada ao volume de água disponível nos reservatórios que pode ser convertido em geração na própria usina e em todas as usinas à jusante na cascata. A grandeza de EAR leva em conta o nível verificado nos reservatórios na data de referência, dada em MW mês, que representa a média mensal da energia armazenada.

O ONS divulga também o Custo Marginal de Operação (CMO) que representa, conforme já apresentado, o valor do custo, por unidade de energia produzida, para atender ao incremento de uma unidade de carga no SIN. São divulgados os valores para cada semana operativa por subsistema, e por patamar de carga, estimados pelo modelo DECOMP, conforme apresentado na seção 2.2.5.

A Figura 16 apresenta o Custo Marginal de Operação no patamar de carga médio para cada um dos subsistemas das semanas operativas de 01 de janeiro de 2005 a 26 de dezembro de 2020.

Figura 16 – Custo Marginal de Operação - patamar de carga médio



Fonte: Elaboração própria a partir de dados da (ONS, 2021b).

Como o CMO é a base para a determinação do PLD, essas variáveis apresentam uma relação intrínseca, com a diferença do PLD apresentar limites pré-determinados.

As correlações entre as variáveis preditoras e o PLD para o submercado SE/CO, calculadas pela Equação (30), estão apresentadas na Tabela 3.

Tabela 3 – Correlações entre variáveis preditoras e o PLD

Variável preditora	Correlação com o PLD
Carga de Energia	0,31
Demanda Máxima	0,32
ENA Bruta	-0,25
ENA Bruta (%MLT)	-0,41
ENA Armazenável	-0,22
ENA Armazenável (%MLT)	-0,36
Energia Armazenada	<b>-0,64</b>
Energia Armazenada (%MLT)	<b>-0,66</b>
Geração Hidráulica	-0,24
Geração Térmica	<b>0,71</b>
Intercâmbio Energia	0,36
Custo Marginal de Operação (CMO)	<b>0,87</b>

Fonte: Elaboração própria a partir de dados da (ONS, 2021b) e (CCEE, 2020).

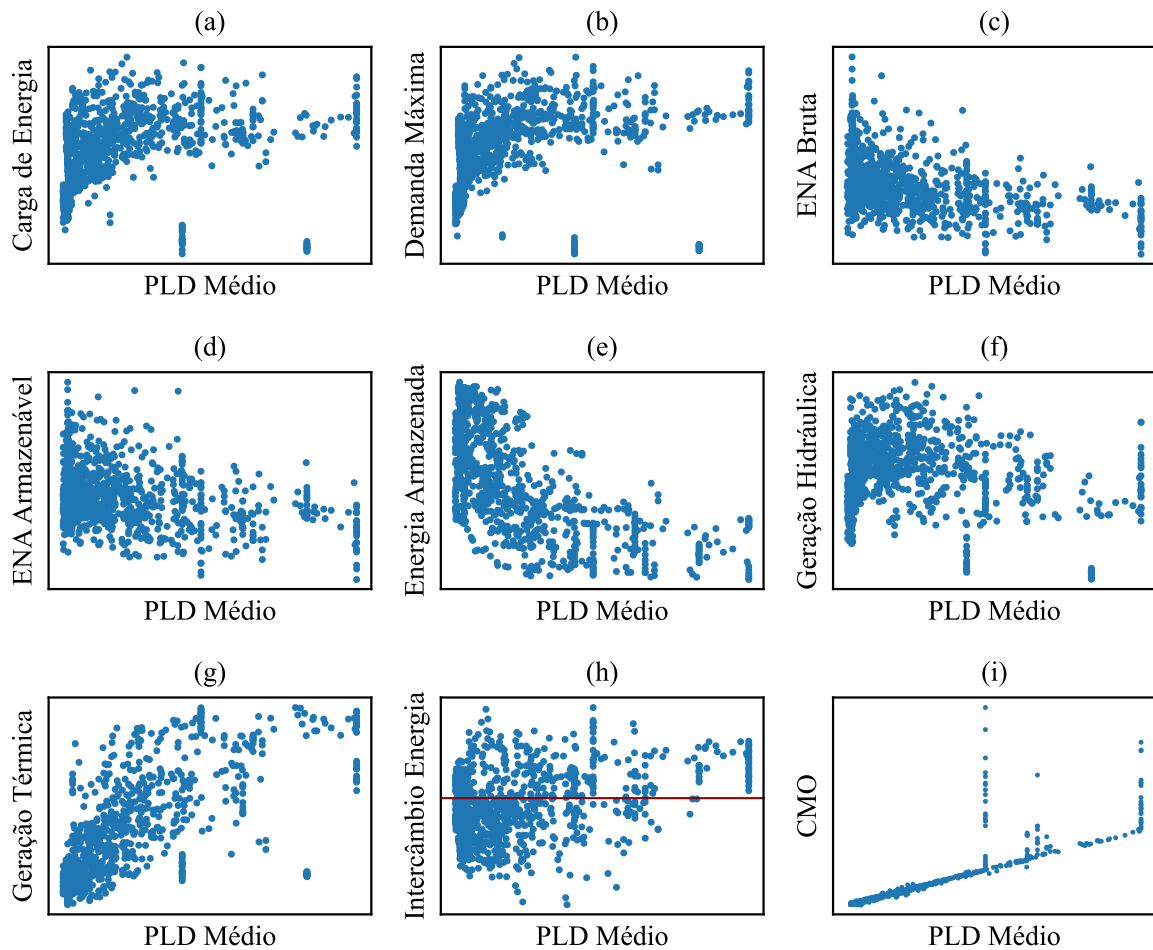
Como já era esperado pelas explicações já apresentadas, há uma forte correlação (maior que 0,7) entre os dados de PLD com os do CMO. Os dados de Geração Térmica também apresentam uma forte correlação com dos dados de PLD. Essa correlação também está atrelada aos custos operativos que, para esse tipo de geração, apresentam valores mais elevados devido aos gastos com combustíveis. Portanto, quanto maior for a participação de térmicas no suprimento da demanda de energia elétrica, mais elevado é esperado que seja o valor do PLD.

Por outro lado, maior participação de usinas hidrelétricas ao atendimento à demanda representa menores custos operacionais e, conseqüentemente, valores menores para o PLD. No entanto, conforme é possível verificar na Tabela 3, essa correlação não é tão pronunciada. O sinal negativo indica a relação inversamente proporcional entre as variáveis.

Por fim, é interessante notar que as variáveis que também apresentam seus registros referente à Média de Longo Termo (ENA e Energia Armazenável) apresentam correlações ligeiramente mais fortes nesse tipo de medida do que àquelas referente à unidade de energia produzível.

As correlações entre o PLD e as demais variáveis também foram analisadas graficamente por diagramas de dispersão apresentados na Figura 17.

Figura 17 – Gráficos de dispersão entre PLD Médio SE/CO e variáveis predictoras



Fonte: Elaboração própria a partir de dados da (ONS, 2021b) e (CCEE, 2020).

Os dados PLD foram representados no eixo das abscissas e as variáveis predictoras nas ordenadas. Todas as variáveis são referentes ao submercado SE/CO. A linha horizontal demarcada da Figura 17 (h) corresponde ao eixo  $y = 0$  para o intercâmbio de energia.

Por meio desses gráficos é possível observar as correlações de linearidade entre as variáveis. Grande parte das correlações se apresentam de forma não linear (aparentemente exibindo uma tendência polinomial). Exceção se faz ao Custo Marginal de Operação que apresenta uma correlação linear.

Pela Figura 17 (i) é possível observar que, a grande diferença entre o PLD e o CMO reside no fato do primeiro apresentar limites mínimos e máximos definidos pelo CCEE, enquanto o segundo não apresenta essa mesma restrição imposta pelo ONS.

### 3.3 SELEÇÃO E AJUSTE DO MODELO

Essa etapa consiste em fazer a seleção dos modelos que serão utilizados nas previsões, bem como promover o ajuste de seus parâmetros. Para isso os dados tratados na etapa anterior foram separados em dois subconjuntos. O subconjunto de treinamento, correspondendo a 80% dos dados da amostra, é responsável pela definição, estimação e ajuste dos parâmetros dos modelos, enquanto que os dados restantes são destinados à sua validação.

Em cada tipo de modelo considerado há parâmetros que são estimados, pelo próprio modelo, na fase de treinamento e outros que são configurados antecipadamente e fornecidos antes que ele seja exposto ao conjunto de treino. Esses parâmetros definidos previamente são chamados de hiperparâmetros.

Para a definição dos hiperparâmetros utilizou-se da técnica conhecida como pesquisa em grade. Nessa técnica, a definição dos hiperparâmetros é realizada por meio da avaliação do desempenho do modelo configurado para cada combinação possível obtido da grade de hiperparâmetros considerados no espaço de pesquisa.

A seguir são apresentados os modelos e a configuração dos hiperparâmetros utilizados na previsão do PLD.

#### 3.3.1 Modelos Holt-Winters

Conforme apresentado na seção 2.1.3, esse modelo possui hiperparâmetros que controlam o nível de suavização ( $\alpha$ ), de tendência ( $\beta$ ) com seu amortecimento ( $\phi$ ) e a sazonalidade ( $\gamma$ ). Todos esses hiperparâmetros podem ser especificados ao definir o modelo. Do contrário, a biblioteca utilizada ajustará automaticamente o modelo e encontrará seus valores ideais.

No entanto, há outros hiperparâmetros que o modelo não ajustará automaticamente e que podem ser especificados. São eles: o tipo de tendência (aditivo, multiplicativo ou nenhuma), se o componente de tendência deve ou não ser amortecido; o período e o tipo de sazonalidade (aditivo, multiplicativo ou nenhuma); se deve ou não usar uma transformação de potência da série e se deve ou não remover o viés ao ajustar o modelo.

Foi realizada uma pesquisa em grade desses hiperparâmetros considerando os seguintes períodos (semanas) para a sazonalidade: 1, 2, 4, 5, 6, 7, 11, 19, 21, 35 e 52. Com essas combinações, foram testados 792 modelos Holt-Winters.

A configuração dos hiperparâmetros do modelo que apresentou o melhor desempenho para a previsão do PLD Médio SE/CO está mostrada na Tabela 4.

Tabela 4 – Hiperparâmetros do Modelo Holt-Winters

Hiperparâmetro	Configuração
Tipo de tendência	Nenhuma
Amortecimento da tendência	Não
Tipo de sazonalidade	Aditivo
Período de sazonalidade	6 semanas
Transformação de potência	Não
Remoção de viés	Não

Fonte: Elaboração própria.

### 3.3.2 Modelos de Box-Jenkins

Foram configurados os parâmetros dos modelos ARIMA e SARIMA, conforme apresentado na sequência.

#### 3.3.2.1 Modelo ARIMA

O modelo ARIMA resulta da combinação de três filtros: o componente Auto-regressivo (AR), o filtro de Integração (I) e o componente de Média Móveis (MA). Este modelo é comumente representado pela notação  $ARIMA(p,d,q)$ , sendo  $(p,d,q)$  a ordem do modelo para os componentes AR, I e MA respectivamente, sendo:

- a)  $p$ : o número de observações de defasagem incluídas no modelo, também chamado de ordem de defasagem;
- b)  $d$ : o número de vezes que as observações brutas são diferenciadas, também chamado de grau de diferenciação;

- c)  $q$ : o tamanho da janela da média móvel, também chamada de ordem da média móvel.

Conforme apresentado na seção 2.1.4 o processo de construção do modelo é uma abordagem iterativa que consiste nas etapas de identificação, estimativa e verificação.

Na etapa de identificação é verificada se a série temporal é estacionária e, se não, quantas diferenças são necessárias para torná-la estacionária. Isso pode ser realizado por meio de testes estatísticos de raiz unitária na série temporal. Especificamente, foi realizado o teste Dickey-Fuller Aumentado (ADF).

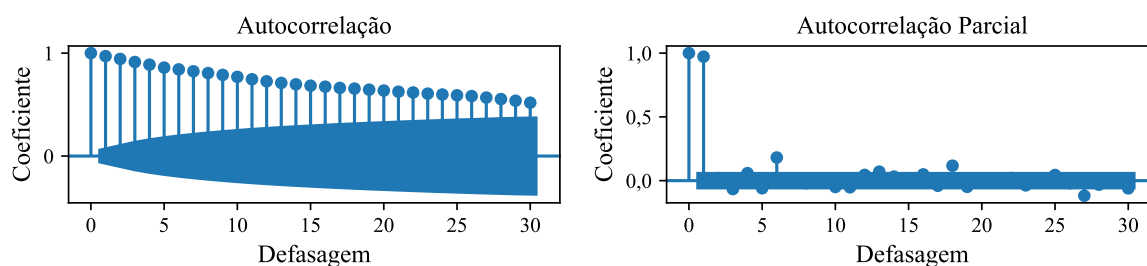
De forma preliminar, os resultados do ADF para o conjunto de dados do PLD Médio SE/CO mostram que o valor da estatística de teste (-3,37) é menor do que o valor crítico (-2,87) a 5%. Isso sugere que pode-se rejeitar a hipótese nula com um nível de significância inferior a 5% (ou seja, uma baixa probabilidade de que o resultado seja um acaso estatístico). Rejeitar a hipótese nula significa que o processo não possui raiz unitária e, por sua vez, que a série temporal é estacionária.

Nessa etapa também são identificados os parâmetros  $p$  e  $q$  do modelo ARIMA. Dois gráficos de diagnóstico podem ser usados para essa finalidade. São eles:

- função de autocorrelação ou *Autocorrelation Function* (ACF): o gráfico resume a correlação de uma observação com os valores de atraso;
- função de autocorrelação parcial ou *Partial Autocorrelation Function* (PACF): o gráfico resume as correlações de uma observação com valores defasados que não são contabilizados por observações defasadas anteriores.

A Figura 18 apresenta os gráficos ACF e PACF para os dados do PLD Médio SE/CO.

Figura 18 – ACF e PACF do PLD Médio SE/CO



Fonte: Elaboração própria a partir de dados da (CCEE, 2020).

Conforme (BOX et al., 2015):

- a) o modelo é AR se o ACF é declinante e tiver um encerramento forçado no PACF após um intervalo, nesse atraso é considerado o valor de  $p$ ;
- b) o modelo é MA se o PACF é declinante e tem um corte abrupto no ACF após o intervalo, neste valor de atraso é considerado o valor de  $q$ ;
- c) o modelo é uma mistura de AR e MA se o ACF e o PACF forem interrompidos.

Portanto, pela análise da Figura 18 é possível identificar um modelo AR com  $p = 1$ . Assim, o modelo ARIMA(1,0,0) é o modelo de partida.

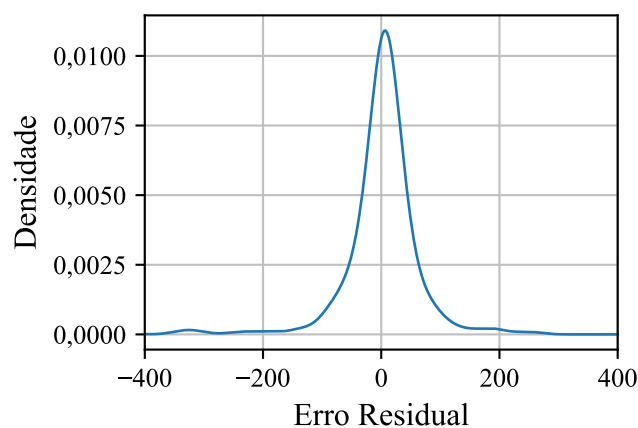
A etapa de estimativa envolve o uso de métodos numéricos para minimizar o termo de erro. Essa otimização é tratada pela biblioteca utilizada para o modelo.

Na etapa de diagnóstico é verificado se o modelo sobreajusta aos dados (*overfitting*), ou seja, apesar do modelo se ajustar bem ao conjunto de dados anteriormente observado, esse se mostra ineficaz para prever novos resultados. Isso é verificado por meio da análise do desempenho do modelo no conjunto de testes. O modelo ARIMA(1,0,0) apresentou bom resultado de desempenho.

Além disso, resíduos de previsão fornecem uma grande oportunidade para diagnósticos. Uma revisão da distribuição de erros pode ajudar a descobrir o viés do modelo. Uma distorção na distribuição ou uma média diferente de zero pode sugerir um viés nas previsões.

A função de probabilidade dos erros residuais da previsão pelo modelo ARIMA(1,0,0) do PLD Médio SE/CO é apresentada na Figura 19.

Figura 19 – Distribuição de probabilidade dos erros residuais ARIMA(1,0,0)



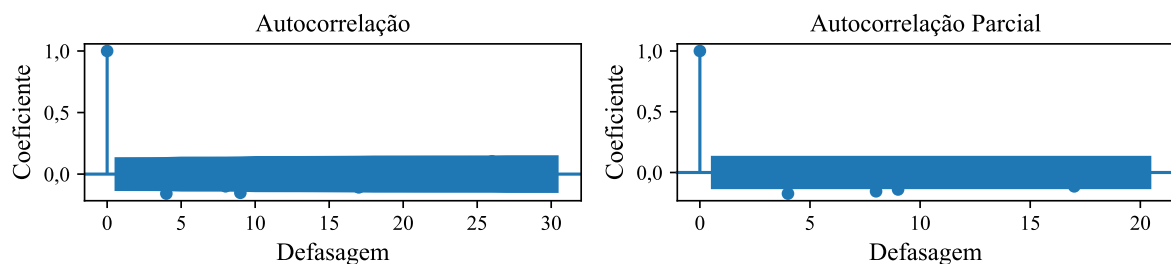
Fonte: Elaboração própria.



A Figura 19 sugere uma distribuição gaussiana com média próxima de zero com uma distribuição simétrica, indicando um modelo que não apresenta viés.

Ainda, um modelo ideal não deixaria estrutura temporal nos resíduos de previsão. A presença de correlação serial nos erros residuais sugere maior oportunidade de uso dessa informação no modelo. Isso pode ser verificado criando gráficos ACF e PACF das séries de erros residuais, conforme apresentado na Figura 20.

Figura 20 – ACF e PACF dos erros residuais ARIMA(1,0,0)



Fonte: Elaboração própria.

Os resultados confirmam a ausência de correlação temporal nos resíduos de previsão, sustentando que o modelo ARIMA(1,0,0) é um bom modelo de previsão.

Além dos gráficos de diagnósticos da série temporal, os hiperparâmetros também foram determinados por meio de pesquisa em grade. Por um processo automatizado, diferentes combinações de hiperparâmetros de modelo ARIMA foram treinados e avaliados.

Para a pesquisa em grade, os hiperparâmetros  $p$  e  $q$  variaram entre 0 e 10 e  $d$  variou entre 0 e 3 totalizando 484 modelos ARIMA distintos.

De forma preliminar, o modelo ARIMA(1,0,2) foi o que apresentou melhor desempenho das configurações consideradas, com previsões 0,21% melhores que o modelo ARIMA(1,0,0) obtido por meio dos gráficos de diagnósticos.

### 3.3.2.2 Modelo SARIMA

A média móvel integrada autoregressiva sazonal, SARIMA ou ARIMA sazonal, é uma extensão do ARIMA que oferece suporte explícito a dados de séries temporais univariadas com um componente sazonal. Adiciona-se três novos hiperparâmetros para especificar a

autorregressão (AR), diferenciação (I) e média móvel (MA) para a componente sazonal da série, bem como um parâmetro adicional para o período da sazonalidade.

A configuração de um SARIMA requer a seleção de hiperparâmetros tanto para a tendência ( $p, d, q$  do modelo ARIMA) quanto para os elementos sazonais  $P, D, Q, m$  que representam, respectivamente, as ordens autorregressivas, das diferenças, de média móvel e o número  $m$  de intervalos de tempo do período sazonal. A notação para um modelo SARIMA é especificada como: SARIMA ( $p, d, q$ ) ( $P, D, Q$ )  $m$ .

Os hiperparâmetros do modelo SARIMA foram especificados por meio de pesquisa em grade, considerando  $p, q, P$  e  $Q$  variando entre 0 e 3,  $d$  e  $D$  sendo 0 ou 1 e a sazonalidade  $m$  variando entre 1 e 6. Em tese, existem 6.144 configurações possíveis, no entanto, na prática algumas delas geram erro e não são avaliadas.

A configuração SARIMA(1,0,2)(2,1,3)<sub>6</sub> apresentou o melhor desempenho de previsão para o PLD Médio SE/CO.

### 3.3.3 Modelos Multilayer Perceptron

Os modelos de Multilayer Perceptron (MLP) possuem diversos hiperparâmetros para serem definidos, como o número de camadas ocultas e a quantidade de neurônios em cada uma, o número de observações de defasagem a serem usadas como entrada para o modelo, o número de épocas que é a quantidade de vezes que o modelo é exposto ao conjunto de dados de treinamento, o número de amostras (lotes) em uma época após a qual os pesos são atualizados.

Os hiperparâmetros foram definidos por meio da realização de pesquisa em grade mantendo o valor das épocas em 3.000 e o lote em 150. Considerou-se que as entradas defasadas do PLD Médio SE/CO poderiam variar entre 1 e 4 e a quantidade de neurônios de acordo com a Tabela 5. Foi considerado, a previsão de até 12 semanas (*step*) à frente do PLD Médio SE/CO.

Tabela 5 – Variação da quantidade de neurônios por camada oculta

Camada oculta	Quantidades de neurônios
1 <sup>a</sup>	6, 12, 25, 50, 100
2 <sup>a</sup>	0, 3, 12, 25, 50
3 <sup>a</sup>	0, 3, 12, 25
4 <sup>a</sup>	0, 3, 12

Fonte: Elaboração própria.

Com as combinações de entradas defasadas e quantidades de neurônio nas camadas ocultas, 1.200 modelos foram gerados (alguns redundantes) e testados.

A inicialização dos pesos dos modelos de rede neural é estocástica, ou seja, cada vez que o modelo é treinado, para a mesma configuração de hiperparâmetros, um conjunto de pesos diferentes é obtido, gerando um desempenho diferente. Portanto, para abordar a avaliação do modelo, cada configuração foi avaliada 10 vezes e o erro foi relatado como o erro médio.

Assim a configuração que apresentou o melhor desempenho é apresentada na Tabela 6.

Tabela 6 – Configuração hiperparâmetros MLP

Camada	Quantidades de neurônios
Entrada	1
1ª Oculta	6
2ª Oculta	3
3ª Oculta	12
Saída	12

Fonte: Elaboração própria.

Como os modelos MLP permitem entrada multivariável, foi realizado uma pesquisa em grade combinado as diversas variáveis preditoras apresentadas na seção 3.2 como entrada do modelo. No caso, foram consideradas todas as combinações possíveis entre as seguintes variáveis: carga de energia, demanda máxima, energia natural afluyente bruta e armazenável, energia armazenada, geração hidráulica e térmica, intercâmbio de energia, o mês do período considerado, além do próprio PLD. O custo marginal de operação não foi incluído nesse conjunto porque sua base de dados histórica é 3,5 anos menor que das outras variáveis.

Para entradas multivariáveis é preciso normalizar os dados de entrada. Assim, todas as variáveis supracitadas passaram por um processo de padronização descrito pela Equação (57)

$$X_t = \frac{x_t - \min(x)}{\max(x) - \min(x)}, \quad (57)$$

onde  $X_t$  é o valor normalizado da série no momento  $t$ ,  $x_t$  representa a variável original que passará pela transformação linear, os valores  $\max(x)$  e  $\min(x)$  representam os valores máximos e mínimos, respectivamente, da série original a ser normalizada.

O objetivo desta transformação linear é adequar os dados de entrada da rede à escala das funções de ativação, no caso a função ReLu. A transformação linear proposta pela Equação (57) não altera as propriedades dos dados de entrada, apenas os restringe ao conjunto  $[0,1]$ .

### 3.4 USO E VALIDAÇÃO DO MODELO

A validação do modelo consiste em avaliar suas previsões além daquelas realizadas durante a fase de ajuste dos parâmetros. Uma das principais maneiras de se avaliar a assertividade de um modelo de previsão é analisar o erro entre o que foi previsto e o valor real utilizando o conjunto de teste, conforme apresentado na Equação (58).

$$e = Y_i - \hat{Y}_i, \quad (58)$$

onde  $Y_i$  é o valor atual e a  $\hat{Y}_i$  é a previsão para o período  $i$  da série temporal.

Para quantificar o desempenho dos modelos em problemas de séries temporais é comum a utilização de métricas que comparam as séries de valores projetados e as séries de valores reais. Uma medida estatística que pode ser utilizada para essa finalidade é o Erro Médio Absoluto ou *Mean Absolute Error* (MAE) apresentado na Equação (59).

$$MAE = \frac{1}{n} \sum_{i=1}^n |e|. \quad (59)$$

Outra medida é a Raiz do Erro Médio Quadrático ou *Root Mean Squared Error* (RMSE) que é calculado conforme Equação (60).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e)^2}. \quad (60)$$

Para comparação de séries temporais com diferentes intervalos de tempo deve-se utilizar uma medida de erro relativa ou percentual.

O Erro Percentual ou *Percentual Error* (PE) pode ser calculado pela Equação (61), conforme apresentado a seguir:

$$PE_i = \frac{Y_i - \hat{Y}_i}{Y_i} \times 100\%, \quad (61)$$

por meio do qual é possível definir o Erro Médio Percentual Absoluto ou *Mean Absolute Percentage Error* (MAPE), Equação (62).

$$MAPE = \frac{1}{n} \sum_{i=1}^n |PE_i|. \quad (62)$$

Além do comparativo de erros entre metodologias distintas de previsão, faz-se útil ter uma medida que compara o quanto os resultados obtidos são melhores do que uma previsão ingênua ou trivial, ou seja, quando a melhor estimativa do próximo valor é o próprio valor atual. Matematicamente, o U de Theil é definido como:

$$U = \sqrt{\frac{\sum_{i=1}^N \left( \frac{Y_i - \hat{Y}_i}{Y_i} \right)^2}{\sum_{i=1}^N \left( \frac{Y_i - Y_{i-1}}{Y_i} \right)^2}} \quad (63)$$

através desse coeficiente pode-se analisar a qualidade de uma previsão através dos seguintes valores do coeficiente U de Theil:

- a)  $U > 1$ : o erro do modelo é maior que o erro na previsão ingênua;
- b)  $U < 1$ : o erro do modelo é menor que o erro na previsão ingênua.

Quanto mais próximo for de zero, melhor será o resultado da previsão.

Cabe destacar que, para toda a pesquisa em grade realizada, foi considerada a medida de erro RMSE para avaliação da performance dos modelos.

Os resultados da avaliação da performance dos modelos são apresentados no capítulo seguinte.

## 4 RESULTADOS E ANÁLISES

Este capítulo apresenta os resultados obtidos pelos modelos de previsão considerados neste trabalho aplicados na predição do PLD médio do submercado Sudeste/Centro-Oeste.

A capacidade de cada um dos modelos foi avaliada na predição de até doze semanas à frente (*step*). São expostos os gráficos comparativos entre os resultados projetados por cada um dos modelos e os valores reais para a primeira semana. Os resultados são avaliados de acordo com as métricas descritas na seção 3.4.

Ao final do capítulo é apresentada uma comparação do desempenho entre os modelos e uma análise de erros do modelo com melhor desempenho.

### 4.1 SEPARAÇÃO DOS SUBCONJUNTOS

O conjunto de dados foram separados em dois subconjuntos: 80% correspondente aos dados de treino e estimação dos parâmetros e 20% são destinados à validação do modelo.

Na separação dos subconjuntos, foi respeitada a sequência temporal dos dados, ou seja, dados de treinamento correspondem ao período inicial de todo histórico, equivalendo ao período de 2001 a 2016 e os de testes o restante do período sequencial, de 2017 a 2020.

A Tabela 7 apresenta estatística de dispersão de cada um dos subconjuntos, o que evidencia um comportamento distinto dos dados entre os subconjuntos.

Tabela 7 – Medidas de dispersão para os subconjuntos de treino e teste (R\$/MWh)

Medida	Conjunto Treino	Conjunto Teste
Média	148,87	247,97
Desvio Padrão	198,52	151,13
Mínimo	4,00	39,68
25%	18,59	127,57
50%	62,69	214,27
75%	186,66	339,47
Máximo	822,83	533,82

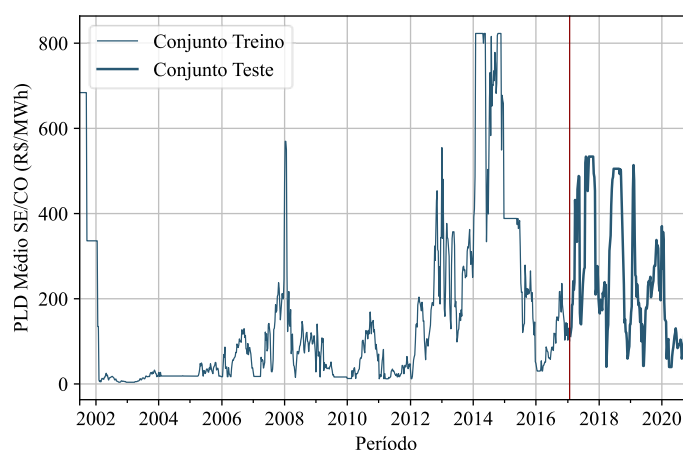
Fonte: Elaboração própria.

Com um desvio padrão maior e valores extremos mais distantes, o subconjunto de treino apresenta dados mais dispersos do que os do subconjunto de teste.

Apesar dessa distinção da estatística dos dados entre os subconjuntos, o balanceamento de amostras (procedimento comumente utilizado para treinamento de redes neurais) deve ser evitado quando se trata de séries temporais ao passo que, nesse caso, é preciso manter a sequência dos dados. Nesse estudo não foi utilizado nenhum pré-processamento dos dados que vise minimizar as discrepâncias entre os subconjuntos.

A Figura 21 apresenta o histórico semanal do PLD Médio SE/CO entre 30 de junho de 2001 e 26 de dezembro de 2020 com a distinção entre os subconjuntos de treino e teste.

Figura 21 – PLD Médio SE/CO - Separação dos subconjuntos



Fonte: Elaboração própria a partir de dados da (CCEE, 2020).

É possível verificar, pela figura, a distinção do comportamento da série entre os subconjuntos, conforme já evidenciado na Tabela 7.

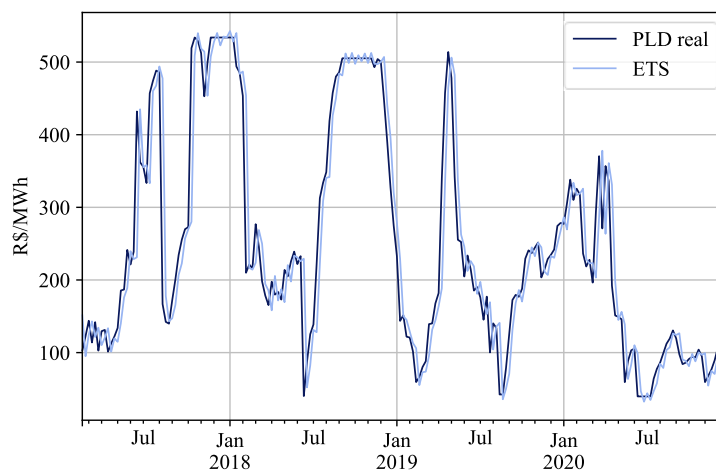
Cada um dos modelos configurados e apresentados no Capítulo 3 foi avaliado pelo subconjunto de teste com aplicação das métricas que comparam as séries de valores previstos e as séries de valores reais e os resultados exibidos na sequência.

## 4.2 MODELO HOLT-WINTERS

Os hiperparâmetros utilizados para o modelo de Holt-Winters, conhecidos também como modelo ETS (*Error, Trend, Seasonal*), foram apresentados na seção 3.3.1. Já os parâmetros do modelo foram revistos para cada previsão realizada.

O resultado da previsão do modelo ETS de uma semana à frente em comparação com os valores reais do conjunto de teste é exibido na Figura 22.

Figura 22 – Previsão do Modelo ETS para a primeira semana



Fonte: elaboração própria.

Apesar da Figura 22 apresentar apenas a curva de previsão da primeira semana, foram realizadas a previsão de até 12 semanas à frente por meio de um processo iterativo recursivo. Ou seja, a previsão de uma semana foi utilizada como entrada do modelo para a previsão da semana seguinte e assim sucessivamente até a 12ª semana. A Tabela 8 apresenta as métricas de erros do modelo ETS para as 12 semanas à frente (*step*), obtidas entre os valores previstos e os valores reais do conjunto de teste.

Tabela 8 – Métricas de desempenho do modelo ETS

Semana (Step)	MAE (R\$/MWh)	RMSE (R\$/MWh)	MAPE (%)
1	32,79	54,99	20,04
2	53,32	82,45	31,39
3	72,77	103,82	40,58
4	89,09	123,11	48,38
5	102,84	138,85	55,27
6	114,80	153,42	61,98
7	126,87	167,70	66,77
8	137,59	181,86	72,74
9	145,66	191,92	79,47
10	151,55	198,64	85,64
11	158,01	204,25	90,77
12	163,05	209,16	96,62

Fonte: elaboração própria.



Pela análise da Tabela 8 é possível verificar que os valores das métricas de erros são menores para as previsões mais recentes e vai aumentando para as previsões mais distantes. Isso era esperado na medida em que a previsão de uma determinada semana carrega os erros de previsão das semanas anteriores, devido ao processo iterativo recursivo de previsão a partir da segunda semana.

O valor do  $U$  de Theil, Equação (63), para a previsão da primeira semana para o modelo ETS é 0,996. Com  $U \approx 1$ , significa que o modelo ETS é, praticamente, tão eficiente quanto um modelo de previsão ingênua ou trivial, Equação (5), o qual considera que a previsão do próximo período é igual ao valor observado do período atual.

### 4.3 MODELOS BOX-JENKINS

Dois modelos Box-Jenkins foram avaliados, o modelo ARIMA(1,0,2) e o modelo SARIMA(1,0,2)(2,1,3)6.

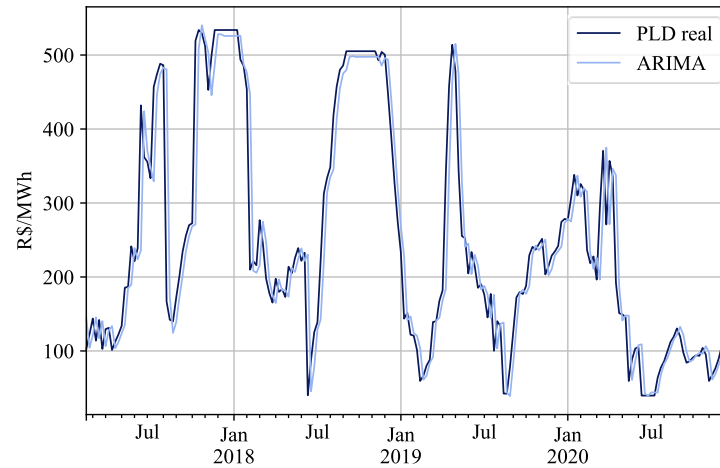
Assim como no modelo ETS, os modelos Box-Jenkins tiveram seus parâmetros ajustados para cada previsão realizada. Assim, após determinada previsão, seu respectivo valor real foi inserido no conjunto de estimação dos parâmetros do modelo para sua redefinição e então uma nova previsão é realizada para as 12 próximas semanas.

Os resultados das métricas de desempenho de previsão de cada um dos modelos Box-Jenkins são apresentados na sequência.

#### 4.3.1 Modelo ARIMA

O resultado da previsão do modelo ARIMA(1,0,2) de uma semana à frente em comparação com os valores reais do conjunto de teste é exibido na Figura 23.

Figura 23 – Previsão do modelo ARIMA para a primeira semana



Fonte: elaboração própria.

A Tabela 9 apresenta as métricas de erros do modelo ARIMA(1,0,2) para as 12 semanas à frente (*step*).

Tabela 9 – Métricas de desempenho do modelo ARIMA(1,0,2)

Semana (Step)	MAE (R\$/MWh)	RMSE (R\$/MWh)	MAPE (%)
1	32,52	54,61	19,76
2	52,92	80,57	31,03
3	71,82	100,98	39,81
4	87,81	118,38	47,32
5	100,13	132,24	53,61
6	111,02	144,64	59,35
7	121,39	156,85	63,34
8	130,45	168,93	67,96
9	136,59	176,80	72,97
10	140,72	181,08	76,80
11	144,70	184,49	79,75
12	147,55	187,07	83,10

Fonte: elaboração própria.

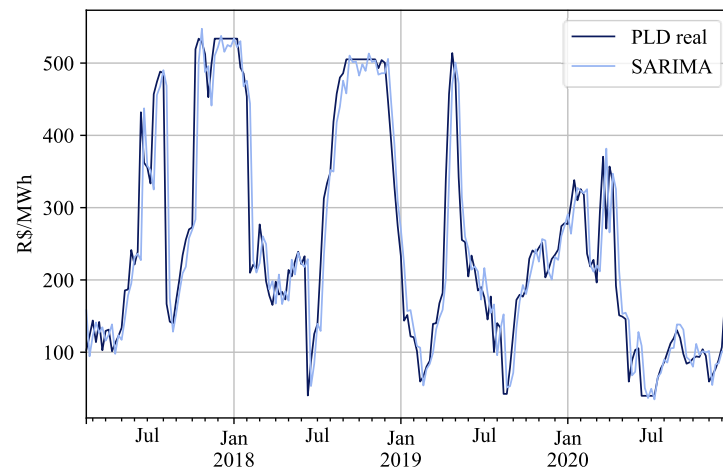
Em relação aos valores das métricas de desempenho do modelo ETS, o modelo ARIMA(1,0,2) apresentou uma pequena melhora para todas as previsões.

Para a primeira semana prevista o modelo ARIMA exibiu um U de Theil igual a 0,983, correspondendo a pequena melhora detectada.

### 4.3.2 Modelo SARIMA

O resultado da previsão do modelo SARIMA(1,0,2)(2,1,3)<sub>6</sub> de uma semana à frente em comparação com os valores reais do conjunto de teste é exibido na Figura 24.

Figura 24 – Previsão do modelo SARIMA para a primeira semana



Fonte: elaboração própria.

A Tabela 10 apresenta as métricas de erros do modelo SARIMA(1,0,2)(2,1,3)<sub>6</sub>.

Tabela 10 – Métricas de desempenho do modelo SARIMA(1,0,2)(2,1,3)<sub>6</sub>

Semana (Step)	MAE (R\$/MWh)	RMSE (R\$/MWh)	MAPE (%)
1	32,68	53,98	20,16
2	51,87	79,14	30,87
3	70,92	98,98	40,10
4	87,27	116,81	48,78
5	100,11	130,51	55,75
6	110,64	142,58	61,94
7	121,62	154,53	66,87
8	130,76	165,93	72,32
9	136,63	173,15	77,31
10	140,55	177,32	81,32
11	143,41	179,55	83,80
12	145,31	181,52	87,23

Fonte: elaboração própria.

Em relação ao modelo ARIMA, o modelo SARIMA apresentou melhoras para a previsão de algumas semanas e piora em outras nas métricas MAE e MAPE. Já o erro RMSE apresentou melhora para todas as previsões.

Ao comparar a Figura 23 com a Figura 24 é possível perceber o efeito da modelagem da sazonalidade na curva de predição, inserindo uma ínfima oscilação na curva. Essa pequena oscilação também pode ser percebida na curva do modelo ETS, Figura 22, que também tem um componente sazonal.

O valor de  $U$  de Theil para a previsão da primeira semana pelo modelo SARIMA apresentou uma pequena melhora ( $U = 0,952$ ) em relação aos demais modelos.

#### 4.4 MODELOS MULTILAYER PERCEPTRON

Os modelos de redes neurais, mais precisamente as redes *Multilayer Perceptron*, foram aplicados para a previsão do PLD em duas configurações: univariado, em que se utilizou como variável de entrada apenas o próprio PLD, e multivariado, em que se utilizou outras variáveis preditoras como entrada do modelo. Ambas configurações tiveram como saída a previsão do PLD.

Conforme já apresentado na seção 3.3.3, a pesquisa de grade realizada para a configuração dos hiperparâmetros do modelo MLP apontou que a entrada de um único período (sem defasagens anteriores) é a configuração de entrada com o melhor desempenho do conjunto de hiperparâmetros considerados.

Diferentemente dos modelos ETS e Box-Jenkins, os modelos MLP não tiveram seus parâmetros estimados (pesos sinápticos) para cada uma das previsões realizadas. Isso porque, do ponto de vista computacional, seria muito custoso promover o treinamento da rede para cada uma das previsões do conjunto de teste.

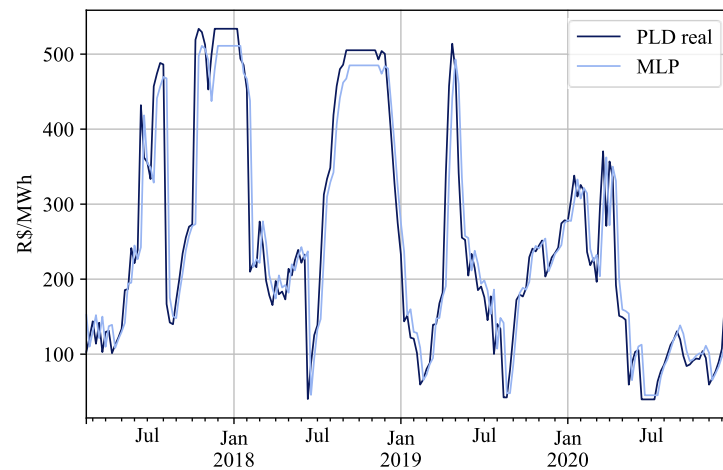
Alternativamente, os pesos sinápticos das redes MLP foram definidos por meio de um único processo de aprendizagem da rede pelo subconjunto de treino e, com a rede treinada, as previsões foram promovidas para cada um dos períodos do subconjunto de teste.

Na sequência são apresentados os resultados obtidos para cada um dos modelos MLP considerados: univariado e multivariado.

#### 4.4.1 Modelo MLP Univariado

A Figura 25 apresenta a curva de previsão da primeira semana do modelo MLP univariado em comparação com os dados reais do PLD.

Figura 25 – Previsão do modelo MLP para a primeira semana



Fonte: elaboração própria.

As medidas de erros para cada uma das 12 semanas são apresentadas na Tabela 11.

Tabela 11 – Métricas de desempenho do modelo MLP univariado

Semana (Step)	MAE (R\$/MWh)	RMSE (R\$/MWh)	MAPE (%)
1	33,44	54,18	20,16
2	54,17	79,74	31,65
3	72,66	99,70	40,22
4	88,88	116,95	47,24
5	101,18	130,64	53,12
6	112,07	143,99	58,20
7	122,16	156,42	62,10
8	131,03	169,22	65,88
9	137,17	177,09	70,70
10	142,36	182,45	74,69
11	147,30	186,51	77,58
12	151,10	190,05	80,95

Fonte: elaboração própria.

O modelo MLP univariado não apresentou melhoras nas previsões em relação ao modelo SARIMA. Isso é constatado, inclusive, pelo valor do coeficiente de U de Theil para a primeira semana prevista com valor de 0,955.

#### 4.4.2 Modelo MLP Multivariado

As variáveis preditoras apresentadas na seção 3.2 juntamente com a referência mensal foram utilizadas como entrada no modelo MLP multivariado. Utilizou-se de pesquisa em grade para promover as combinações das diferentes configurações das variáveis utilizadas na entrada do modelo no intuito de encontrar o conjunto que apresentasse o melhor desempenho de previsão para cada uma das semanas.

O conjunto de variáveis selecionadas pela pesquisa em grade depende da métrica de erro utilizada para avaliar a performance do modelo. Para a métrica RMSE a configuração de entradas que apresentou o melhor desempenho, para cada uma das semanas, é apresentada na Tabela 12.

Tabela 12 – Configuração Entradas MLP multivariado

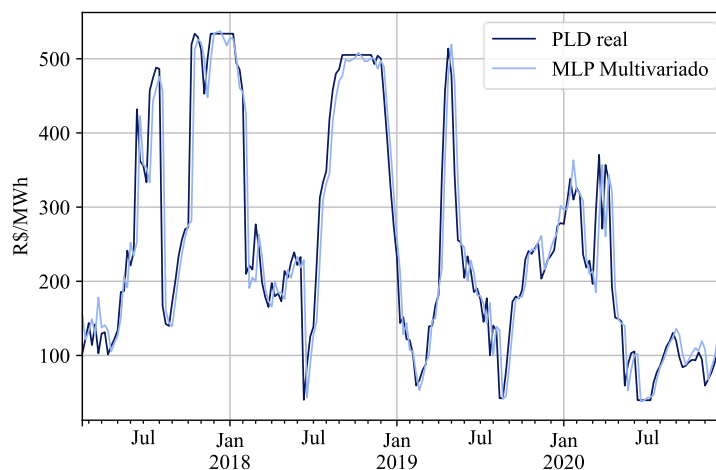
Semana (Step)	Conjunto de Entradas	RMSE (R\$/MWh)
1	PLD, ENA Armazenável, ENA Bruta, EAR, Geração Térmica, Mês	50,44
2	PLD, ENA Armazenável, EAR, Geração Hidráulica e Térmica	72,34
3	PLD, ENA Armazenável, ENA Bruta, EAR, Geração Hidráulica	88,80
4	PLD, ENA Bruta, EAR, Geração Térmica Mês	102,52
5	PLD, ENA Bruta, EAR, Geração Térmica	113,95
6	PLD, ENA Bruta, EAR, Geração Térmica	124,85
7	Carga, ENA Bruta, EAR, Geração Hidráulica e Térmica	129,02
8	Carga, ENA Bruta, EAR, Geração Hidráulica e Térmica	129,13
9	ENA Armazenável, EAR, Geração Térmica	130,55
1	ENA Armazenável, EAR, Geração Térmica	126,40
11	ENA Armazenável, EAR, Geração Térmica	121,92
12	ENA Armazenável, EAR, Geração Térmica	120,20

Fonte: elaboração própria.

A partir da 7ª semana prevista o PLD não faz parte do conjunto de variáveis preditoras de entrada. No entanto, em comparação aos modelos anteriores, é possível confirmar melhora do desempenho nas previsões para essas semanas.

A Figura 26 apresenta a curva de previsão da primeira semana do modelo MLP Multivariado.

Figura 26 – Previsão do modelo MLP multivariado para a primeira semana



Fonte: elaboração própria.

A Tabela 13 exibe os melhores resultados de cada uma das métricas de desempenho do modelo multivariado.

Tabela 13 – Métricas de desempenho do modelo MLP multivariado

Semana (Step)	MAE (R\$/MWh)	RMSE (R\$/MWh)	MAPE (%)
1	29,60	50,44	18,60
2	46,78	72,34	28,54
3	63,98	88,80	35,77
4	77,65	102,52	41,02
5	90,08	113,95	46,64
6	98,13	124,85	48,49
7	104,90	129,02	47,87
8	105,31	129,13	47,59
9	102,94	130,55	48,31
10	98,71	126,40	48,69
11	96,14	121,92	47,21
12	94,08	120,20	46,13

Fonte: elaboração própria.

O Valor do coeficiente U de Theil para a primeira semana prevista pelo modelo MLP é de 0,897, evidenciando a melhora de desempenho desse modelo em relação aos demais.

#### 4.5 COMPARAÇÃO DE DESEMPENHO DOS MODELOS

Para comparação do desempenho dos modelos de previsão, os resultados obtidos são apresentados em conjunto. A Tabela 14 exibe os valores do coeficiente U de Theil.

Tabela 14 – U de Theil para cada um dos modelos de previsão

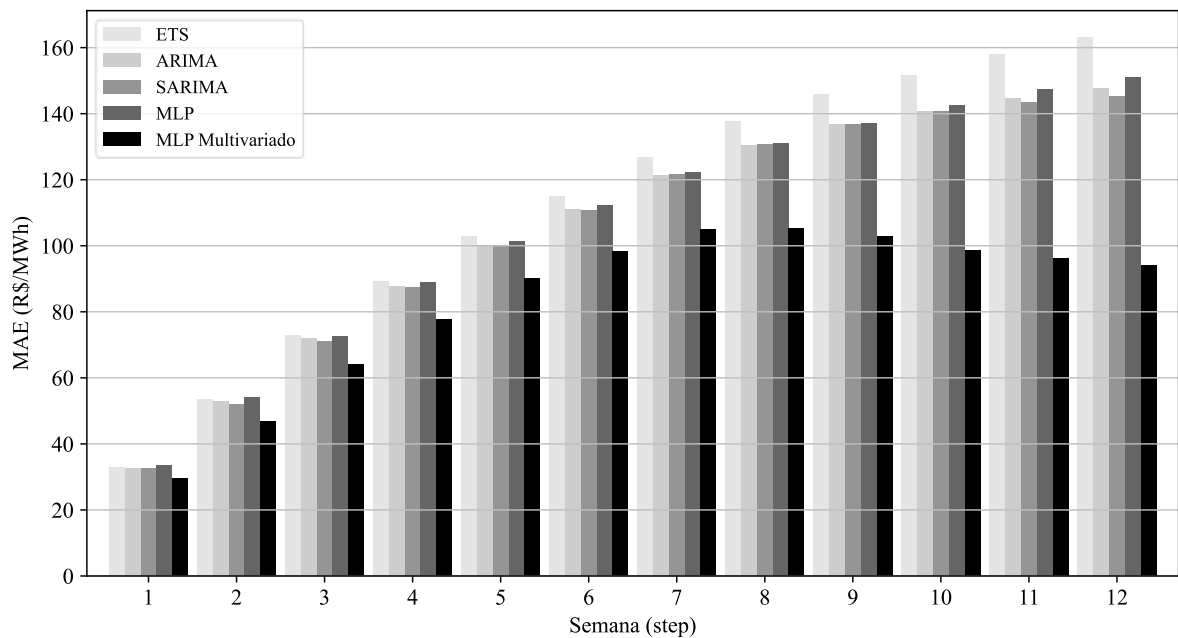
Métrica	ETS	ARIMA	SARIMA	MLP	MLP multivariado
U de Theil	0,996	0,983	0,952	0,955	0,897

Fonte: elaboração própria.

Com o menor valor de U de Theil, o modelo MLP multivariado é o que apresenta, dentre os modelos avaliados, a melhor acurácia de previsão.

A comparação dos valores das métricas de desempenho (MAE, RMSE e MAPE) por modelo e para cada uma das semanas previstas é mostrada na Figura 27, Figura 28 e Figura 29.

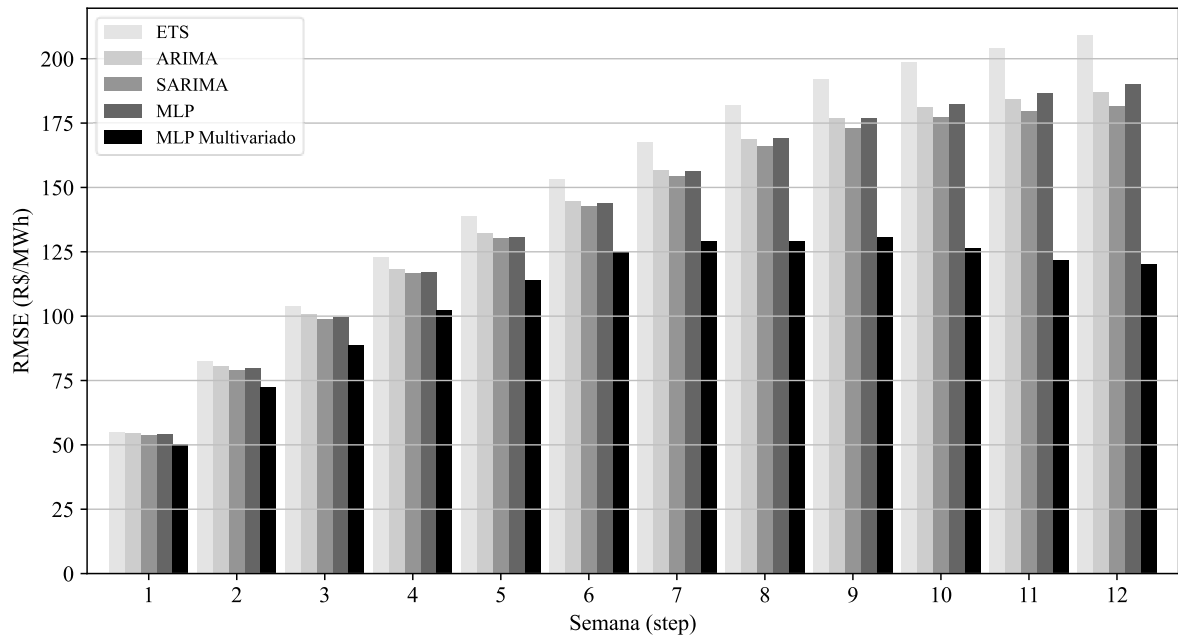
Figura 27 – Comparação do desempenho dos modelos na métrica MAE



Fonte: elaboração própria.

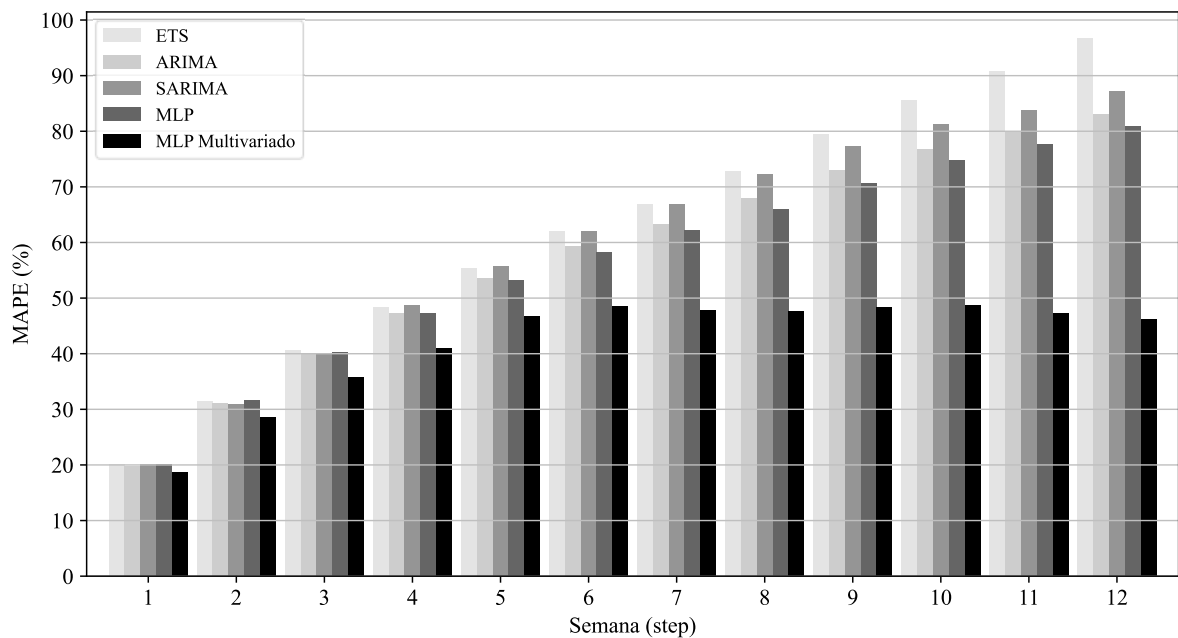


Figura 28 – Comparação do desempenho dos modelos na métrica RMSE



Fonte: elaboração própria.

Figura 29 – Comparação do desempenho dos modelos na métrica MAPE



Fonte: elaboração própria.

O modelo MLP multivariado apresentou os melhores resultados de previsão, dentre os modelos considerados, para todas as semanas e métricas de avaliação de desempenho. Essa

melhor capacidade de previsão do modelo MLP multivariado, em comparação com os demais modelos, fica cada vez mais pronunciada nas previsões de semanas mais avançadas.

Nas previsões dos períodos mais recentes, os modelos apresentaram as menores diferenças de desempenho. Na diferença mais expressiva para a previsão da primeira semana, o modelo MLP multivariado apresentou, na métrica MAE, um desempenho 11,49% melhor que o modelo MLP univariado. A melhoria mais modesta para a previsão da primeira semana ocorreu na métrica MAPE, em que o MLP multivariado apresentou um desempenho 5,87% melhor que o modelo ARIMA.

As medidas dos erros de previsão, indiferentemente da métrica utilizada, tendem a aumentar para previsões de semanas mais adiantadas. No entanto, para o modelo MLP multivariado, observa-se uma tendência de estabilização dos valores dos erros a partir das previsões da 7ª semana. Com isso, a diferença de erro de previsão entre o modelo MLP multivariado e os demais modelos se acentua ainda mais para as previsões a partir da 7ª semana.

As previsões para a 12ª semanas apresentaram as maiores diferenças de desempenho entre os modelos. Na medida de erro MAPE, o modelo MLP multivariado apresentou resultado abaixo de 50% enquanto que os demais modelos apresentaram erros acima de 80%. Avaliando os erros de previsão MAPE, ainda para a 12ª semana, verifica-se que houve melhorias entre 43,01% e 52,26% a favor do modelo MLP multivariado em comparação com os demais.

Já na métrica RMSE, o modelo MLP multivariado trouxe melhorias para a previsão da 12ª semana que variaram de 33,78% (em comparação com o modelo SARIMA) e 42,53% (em comparação com o modelo ETS).

Portanto, as variáveis preditoras inseridas como entrada no modelo MLP multivariado e apresentadas na Tabela 12 trouxeram melhorias significativas no desempenho das previsões em comparação com os modelos univariados utilizados. O histórico dessas variáveis preditoras carrega informações que a rede neural artificial é capaz de extrair para melhorar a previsão dos valores do PLD.

#### 4.6 ANÁLISE DOS ERROS DE PREVISÃO

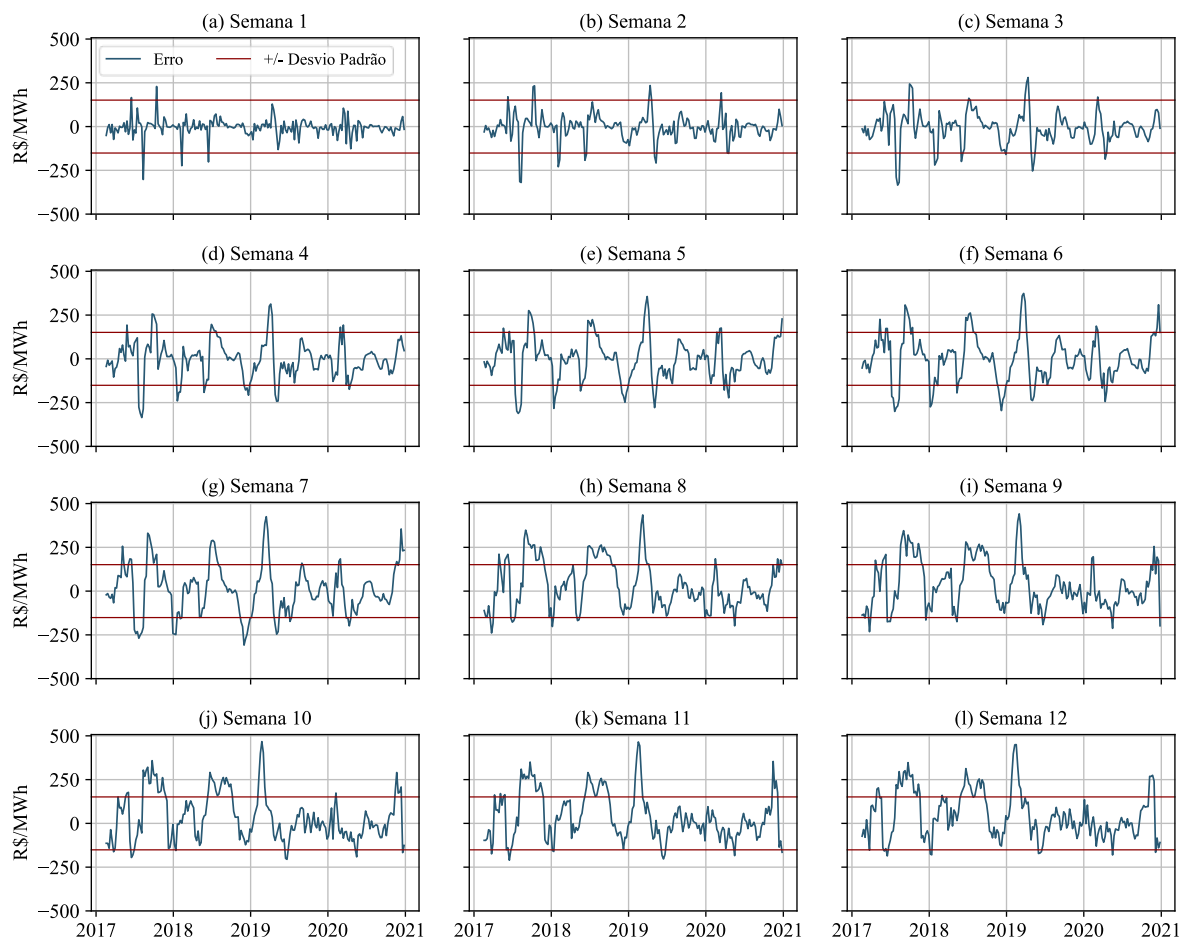
A avaliação do desempenho dos modelos é feita a partir de métricas calculadas por valores médios de um conjunto de determinada medida obtida a partir do erro de previsão.

Porém, além do valor médio da medida em si, é apropriado avaliar a distribuição desses erros. Isso pode ajudar a entender melhor o modelo de previsão, identificando os pontos de melhoria.

Conforme Equação (58), o erro de previsão no período  $i$  ( $e_i$ ) é definido como a diferença entre o valor real ( $Y_i$ ) e a previsão realizada pelo modelo ( $\hat{Y}_i$ ).

A curva de erro de previsão do PLD realizado pelo modelo MLP multivariado para cada uma das 12 semanas é apresentada na Figura 30.

Figura 30 – Erros de previsão do modelo MLP multivariado



Fonte: elaboração própria.

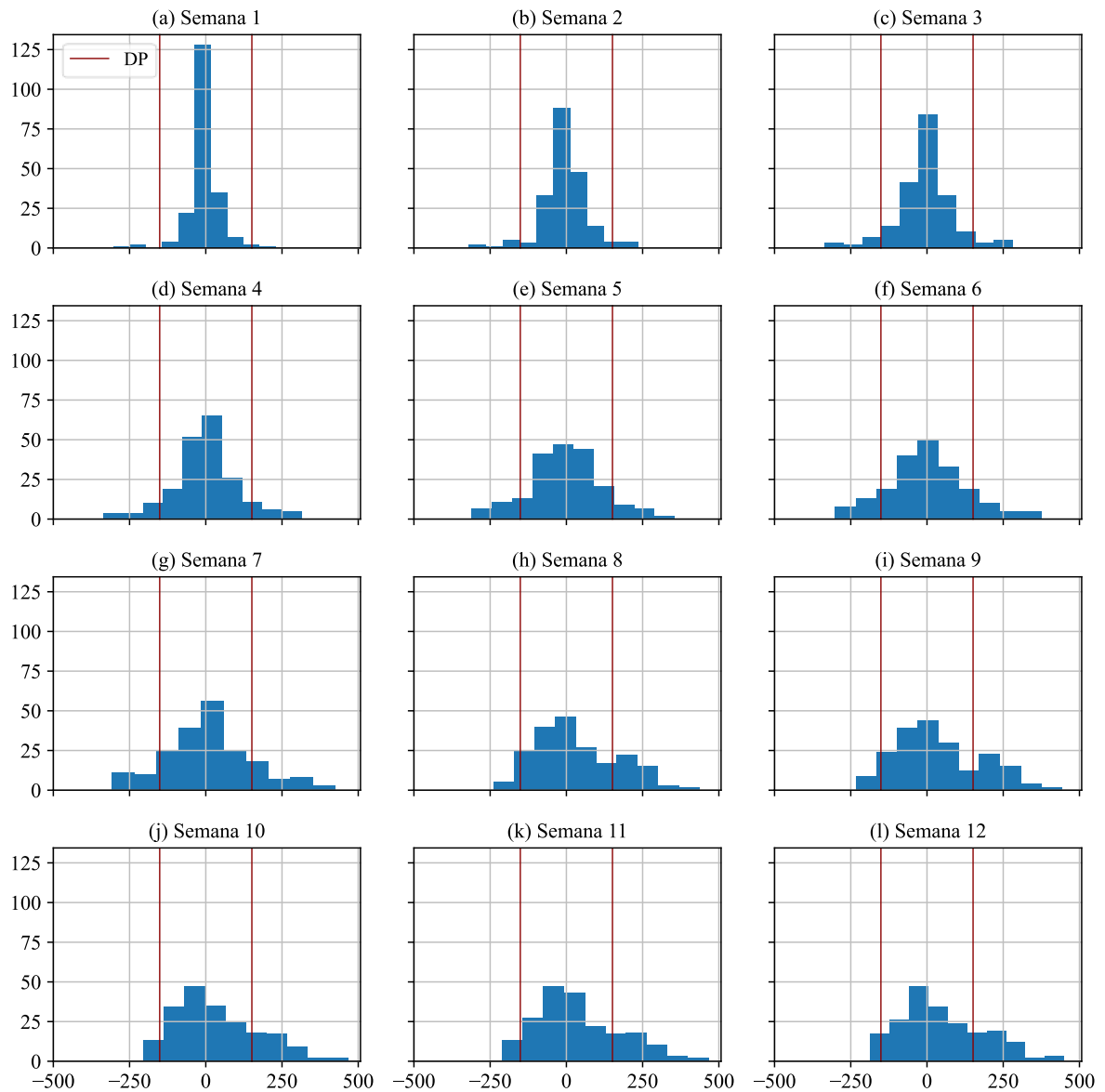
Pela Figura 30 é possível perceber o aumento da amplitude da curva de erros de previsão à medida que o modelo realiza previsões de semanas mais avançadas. Isso ocorre porque o modelo utiliza a previsão de uma semana para fazer a previsão da semana seguinte, carregando o erro de uma semana para outra.

As linhas horizontais em destaque em cada um dos gráficos da Figura 30 marca, como referência, os valores (positivo e negativo) do desvio padrão (DP) do subconjunto de teste (151,13 R\$/MWh conforme Tabela 7).

A medida em que as previsões se prolongam para semanas mais avançadas, mais a curva de erro ultrapassa os limites referenciados.

Os histogramas apresentados na Figura 31 evidenciam o aumento da frequência de erros com valores acima do desvio padrão na medida em que as previsões vão se avançando.

Figura 31 – Histogramas dos erros de previsão do modelo MLP multivariado



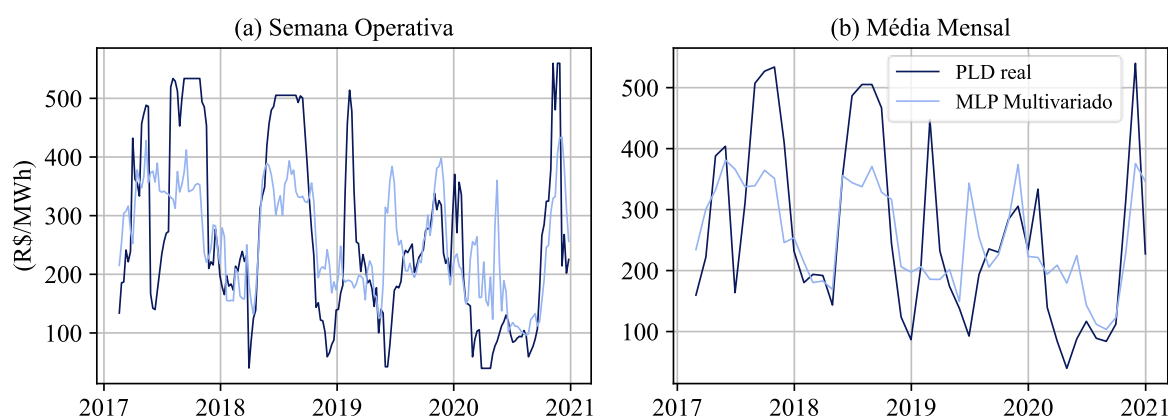
Fonte: elaboração própria.

As linhas verticais nos histogramas da Figura 31 indicam os valores (positivo e negativo) do desvio padrão do subconjunto de testes.

Ainda pela análise dos histogramas é possível perceber que há uma distribuição simétrica das frequências dos erros de previsão até a 7ª semana a partir da qual essa distribuição começa a apresentar uma assimetria com tendência positiva.

Essa assimetria fica cada vez mais evidente para as semanas mais avançadas. Isso é um indicativo de que o modelo apresenta previsões com amplitudes atenuadas em relação aos dados reais. Isso pode ser verificado na Figura 32, que exibe a curva de previsão para a 12ª semana do modelo MLP Multivariado em comparação com a curva real do PLD.

Figura 32 – Previsão do modelo MLP Multivariado para a 12ª semana



Fonte: elaboração própria.

A Figura 32 mostra que os valores previstos apresentam amplitudes menores que a curva real. Pelos valores médios mensais mostrados na Figura 32 (b) é possível perceber que, apesar do erro relativo à amplitude de previsão, o modelo responde razoavelmente bem com relação à previsão de variação da tendência. Principalmente ao se considerar que são previsões de 12 semanas à frente e ainda que as mesmas são realizadas de forma indireta, ou seja, sem que o modelo receba como variável de entrada o próprio valor do PLD.

Objetivamente, a previsão da tendência pode ser obtida pela comparação do sinal da derivada para cada um dos pontos da curva real e prevista. Considerando-se os valores médios mensais, Figura 32 (b), esse resultado é de 65,2%. Nos modelos univariados, a precisão da previsão da tendência para a 12ª semana é de, no máximo, 43,5%.

Portanto, o modelo é capaz de prever, na maior parte das vezes, o aumento e a queda do valor do PLD, sendo falho, no entanto, na assertividade desses valores.

Ou seja, se por um lado o modelo não apresenta um desempenho tão satisfatório, do ponto de vista quantitativo, para a previsão dos valores do PLD para a 12ª semana, ele responde razoavelmente bem na previsão de mudanças de tendências.

## 5 CONCLUSÕES E SUGESTÕES

### 5.1 CONCLUSÕES

O principal objetivo deste trabalho foi comparar o desempenho das metodologias: suavização exponencial ou Holt-Winters, Box-Jenkins e de redes neurais artificiais na previsão de séries temporais, mais precisamente do Preço de Liquidação das Diferenças (PLD) ou preço *spot* da energia no mercado de curto prazo brasileiro.

Foi realizada uma revisão das metodologias utilizadas afim de trazer um embasamento teórico para o estudo de caso da aplicação dos modelos de previsão. Ainda, foi feito um estudo a respeito do Setor Elétrico Brasileiro com o objetivo de conduzir um melhor conhecimento sobre o panorama do mercado de energia do Brasil, as transformações regulatórias do setor e as principais variáveis que afetam na formação dos preços de energia. Por fim, foi realizada uma pesquisa bibliográfica em previsões de séries temporais aplicadas na previsão do preço da energia e especificamente à previsão do PLD.

Para o estudo, foram seguidas as etapas de previsão de séries temporais estabelecidas no capítulo MATERIAIS E MÉTODOS DE PESQUISA. Em suma, as etapas abrangem o tratamento dado às variáveis, as configurações dos hiperparâmetros dos modelos preditivos e as métricas utilizadas para avaliação de desempenho de previsão.

O PLD foi previsto em 12 semanas à frente pelos modelos Holt-Winters, ARIMA, SARIMA e Multilayer Perceptron (MLP) com entrada univariada, apenas o próprio PLD, e multivariado, considerando outras variáveis preditoras para a entrada do modelo. A capacidade de previsão dos modelos foi avaliada e comparada por meio das métricas de desempenho: *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE) e *Mean Absolute Percentage Error* (MAPE), além da avaliação pelo coeficiente U de Theil que examina o desempenho do modelo em relação à previsão ingênua ou trivial.

Os resultados obtidos demonstram superioridade de todos os modelos em relação à previsão ingênua ou trivial, apresentado um valor de coeficiente U de Theil abaixo de uma unidade para a predição da primeira semana. Nessa métrica, o modelo Holter-Winters (ETS) foi o que exibiu o pior desempenho ( $U = 0,996$ ) dentre os modelos considerados e o MLP multivariado o que apresentou o melhor desempenho ( $U = 0,897$ ).

Todos os modelos apresentaram seus melhores desempenhos nas previsões da primeira semana. A medida em que as previsões avançam para semanas mais distantes, mais aumentam os erros de previsão. Isso se deve ao fato dos modelos utilizarem a previsão de uma semana como dado de entrada para a próxima previsão, transferindo também os erros associados.

Nesse aspecto, o modelo MLP multivariado apresentou um comportamento distinto em relação aos demais por evidenciar uma estabilidade nos erros a partir da sétima semana. Essa particularidade é atribuída à seleção otimizada das variáveis de entrada para a previsão de cada uma das semanas. Ou seja, para cada semana prevista, buscou-se as variáveis de entrada ótimas para o modelo, dentre um conjunto especificado, que minimize o erro de previsão.

O modelo multivariado ainda se mostrou mais eficiente que os demais em todas as previsões e métricas consideradas. É evidente a superioridade de desempenho de previsão da rede neural MLP com múltiplas variáveis de entrada em comparação aos demais modelos univariados. Esse melhor desempenho é atribuído às variáveis preditoras utilizadas e à capacidade das redes neurais artificiais de capturar as informações relevantes para a previsão.

Além do próprio PLD, o modelo multivariado demonstrou que variáveis hidrológicas, como energia natural afluyente (ENA) e energia armazenada (EAR) e variáveis de geração hidráulica e térmica (principalmente) são relevantes no processo preditivo, promovendo melhorias de até 50% em relação à modelos univariados. Essas variáveis têm uma significância ainda maior à medida que as previsões são de semanas mais avançadas.

Na análise da previsão da 12ª semana do modelo MLP multivariado constata-se que esse modelo responde razoavelmente bem para a previsão da tendência média mensal do PLD, apresentando uma acurácia de 65,2%, prevendo, portanto, sua variação na maior parte das vezes. Este resultado é um indicativo que pode ser considerado pelos agentes de mercado na previsão da alteração da tendência do valor dos preços da energia.

A prática demonstrou que o comportamento da série histórica do PLD no subconjunto de testes impacta diretamente nos resultados de desempenho dos modelos. Por exemplo, um mesmo modelo MLP treinado com os dados históricos disponíveis até 2014, apresenta melhores resultados de desempenho nos testes entre o período de 2015 a 2016 (MAPE de 13,14%) do que de 2017 a 2018 (MAPE de 18,84%). Pela Figura 21 é possível perceber que o período entre 2015 e 2016 o PLD apresenta menor volatilidade do que o período seguinte, gerando, por consequência, menos erros de previsão.

Por fim, cabe relatar algumas tentativas frustradas de melhoria do desempenho de previsão não incluídas no capítulo de resultados. Uma dessas tentativas foi a utilização do *Adaptive Boosting (Adaboost)*. O *Adaboost* é um algoritmo meta-heurístico que pode ser

utilizado para melhorar a performance de outros modelos sendo, normalmente, implementado a partir de suas saídas. No caso, ele foi utilizado combinando várias configurações de saídas dos demais modelos, mas nenhuma delas apresentou melhoria de performance.

Outra tentativa foi a utilização de Análise de Componentes Principais ou *Principal Component Analysis* (PCA) para redução da dimensionalidade de entradas defasadas. Em particular, o PCA foi utilizado para reduzir a dimensionalidade das variáveis com 5 atrasos cada no modelo MLP multivariado para apenas 5 vetores de entrada. Mais uma vez os resultados não apresentaram melhorias na performance de previsão.

## 5.2 SUGESTÕES DE TRABALHOS FUTUROS

O assunto abordado nesta dissertação é complexo e muito vasto, e como tal é possível estabelecer um conjunto de direções de futuras investigações que possam vir a melhorar este tema ou ainda a algo correspondente.

Em relação ao modelo MLP multivariado, é possível utilizar outras variáveis de entrada disponibilizadas pelo Operador Nacional do Sistema (ONS). Nessa linha, os resultados deste trabalho demonstram que dados de geração térmica têm uma influência significativa na melhora do desempenho do modelo. Portanto, dados históricos como a geração por tipo de combustível na operação de usinas térmicas podem ser testados para a melhoria no desempenho de previsão.

Outra componente relacionada às usinas térmicas que também poderá trazer melhorias no desempenho de previsão do modelo MLP multivariado é o Custo Variável Unitário (CVU). Esse valor, expresso em reais por megawatt-hora (R\$/MWh), representa o custo necessário para cobrir todas as despesas operacionais de empreendimentos termelétricos. O CVU é um valor atribuído a cada usina térmica e é um dos critérios para classificar a ordem de mérito de despacho das mesmas. Ou seja, usinas térmicas com elevado CVU são despachadas somente em situações hidrológicas extremamente desfavoráveis, contribuindo para o aumento do PLD.

Em relação aos modelos de previsão, outras arquiteturas de redes neurais artificiais podem ser utilizadas, além do *Multilayer Perceptron*, como as Redes Neurais Recorrentes, que já vêm sendo empregadas em modelos preditivos de séries temporais, inclusive na previsão do PLD, principalmente as redes *Long Short Term Memory* (LSTM). Nesse contexto, a sugestão é utilizar pesquisa em grade em uma rede LSTM para a configuração de seus hiperparâmetros e também para a avaliação e seleção otimizada das variáveis de entrada.



## REFERÊNCIAS

AGGARWAL, C. C. **Neural networks and deep learning**. Springer, v. 10, p. 978, 2018.

ALBUQUERQUE, Á. R. **Fluxo de caixa em risco: uma nova abordagem para o setor de distribuição de energia elétrica**. Dissertação (Mestrado em Engenharia Elétrica), Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

ALMEIDA, G. P. S. **Estudo paramétrico para estimação no curto e médio prazo do preço spot de energia elétrica utilizando redes neurais**. Dissertação (Mestrado em Engenharia Elétrica), Universidade de Brasília, Brasília, 2012.

ANEEL. **Comitê Coordenador do Planejamento da Expansão dos Sistemas Elétricos - CCPE**, 1999. Disponível em: <<https://www.aneel.gov.br/>>. Acesso em: 21 jun. 2021.

\_\_\_\_\_. **A ANEEL**, 2019. Disponível em: <<https://www.aneel.gov.br/a-aneel>>. Acesso em: 20 jun. 2021.

BEZERRA, J. A. M. **Um histórico das políticas públicas no setor elétrico brasileiro no período de 1879-2018**. Monografia (Bacharel em Ciência e Tecnologia), Universidade Federal Rural do Semi-Árido, Angicos, 2020.

BOWDEN, N.; PAYNE, J. E. **Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models**. Energy Economics, v. 30, n. 6, p. 3186–3197, 2008.

BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**. John Wiley & Sons, 2015.

BRASIL. **Comitê de Monitoramento do Setor Elétrico - CMSE**, 2004. Disponível em: <<https://www.gov.br/mme/pt-br/assuntos/conselhos-e-comites/cmse>>. Acesso em: 26 jun. 2021.

\_\_\_\_\_. **Portal da Câmara dos Deputados Emenda Constitucional nº 95 de 2016.**, 2016a. Disponível em: <<https://www2.camara.leg.br/legin/fed/decret/1930-1939/decreto-24643-10-julho-1934-498122-publicacaooriginal-1-pe.html>>. Acesso em: 12 jun. 2021

\_\_\_\_\_. **PAC - Ministério do Planejamento**, 2016b. Disponível em: <<http://www.pac.gov.br/noticia/68777baf>>. Acesso em: 22 jun. 2021.

BRESEGHELLO, F. N. **Estudo comparativo de métodos de previsão de demanda: uma aplicação ao caso dos aeroportos com tráfego aéreo regular administrados pelo DAESP**. Dissertação (Mestrado em Ciências), Instituto Tecnológico da Aeronáutica, São José dos Camps, 2005.

CCEE. **Capacita CCEE**, 2013. Disponível em: <<https://capacita.ccee.org.br/>>. Acesso em: 12 jul. 2021.

\_\_\_\_\_. **História da CCEE e do setor elétrico brasileiro**, 2015. Disponível em: <[https://www.ccee.org.br/portal/faces/pages\\_publico/quem-somos/historia](https://www.ccee.org.br/portal/faces/pages_publico/quem-somos/historia)>. Acesso em: 21 jun. 2021.

\_\_\_\_\_. **Média Semanal**, 2020. Disponível em: <[https://www.ccee.org.br/portal/faces/pages\\_publico/o-que-fazemos/como\\_ccee\\_atua/precos/](https://www.ccee.org.br/portal/faces/pages_publico/o-que-fazemos/como_ccee_atua/precos/)>. Acesso em: 28 dez. 2020.

\_\_\_\_\_. **CCEE - Câmara de Comercialização de Energia Elétrica**, 2021. Disponível em: <<https://www.ccee.org.br/portal/>>. Acesso em: 12 jul. 2021.

CEPEL. **NEWAVE - Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Longo e Médio Prazo**, 2018. Disponível em: <[http://www.cepel.br/pt\\_br/produtos/newave-modelo-de-planejamento-da-operacao-de-sistemas-hidrotermicos-interligados-de-longo-e-medio-prazo.htm](http://www.cepel.br/pt_br/produtos/newave-modelo-de-planejamento-da-operacao-de-sistemas-hidrotermicos-interligados-de-longo-e-medio-prazo.htm)>. Acesso em: 11 jul. 2021.

CONTRERAS, J.; ESPINOLA, R.; NOGALES, F. J.; CONEJO, A. J. **ARIMA models to predict next-day electricity prices**. IEEE transactions on power systems, v. 18, n. 3, p. 1014–1020, 2003.

DA SILVA, R. G.; DE OLIVEIRA, A. B.; DA SILVA, I. C.; DE OLIVERIA, T. F. **Application of a demand forecasting model in a rental company of billiard tables**. ITEGAM-JETIA, v. 4, n. 15, p. 53–58, 2018.

DE CAMPOS, L. M. L.; PEREIRA, J. H. A.; DUARTE, D. S.; DE OLIVEIRA, R. C. L. **Bio-Inspired System for Electricity Price Forecast in the Brazilian Market**. Proceedings of the International Joint Conference on Neural Networks, p. 0–7, 2020

DE OLIVEIRA, A. F. **Setor Elétrico, Perspectivas e Desafios para a Contabilidade de Custos**. Anais do Congresso Brasileiro de Custos-ABC. Recife, 2000

EHLERS, R. S. **Análise de séries temporais**. p. 1-118, Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná, Curitiba, 2007

ELETROBRÁS. **Memória da Eletricidade**, 2015. Disponível em: <<http://www.eletrabras.gov.br>>. Acesso em: 26 jun. 2021.

EPE. **A Empresa**, 2017. Disponível em: <<https://www.epe.gov.br/sites-pt/a-epe>>. Acesso em: 22 jun. 2021.

FERREIRA, J. C. **Medida provisória 579/2012 e as ações do risco hidrológico: intervencionismo judiciário ou ativismo judicial? Uma análise crítica do caso GSF**. Monografia (Bacharel em Direito), Centro Universitário de Brasília, Brasília, 2020.

FILHO, J. C. R. **Previsão multi-passos a frente do preço de energia elétrica de curto prazo no mercado brasileiro**. Tese (Doutorado em Engenharia Elétrica). Universidade Federal do Pará, Belém, 2014.

FLORES, J. H. F. **Comparação de modelos MLP/RNA e modelos Box-Jenkins em séries temporais não lineares**. Dissertação (Mestrado em Engenharia de Produção), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.

FOUNDATION, P. S. **Python**, 2020. Disponível em: <<https://www.python.org/>>. Acesso em: 31 ago. 2020.

GOMES, D. T. **Modelos de Redes Neurais recorrentes para previsão de Séries Temporais de memórias curta e longa**. Dissertação (Mestrado), Universidade Estadual de Campinas, Campinas, 2005.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. Pearson Education India, 2010.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. OTexts, 2018.

JANNUZZI, A. C. **Regulação da qualidade de energia elétrica sob o foco do consumidor**. Dissertação (Mestrado em Engenharia Elétrica), Universidade de Brasília, Brasília,, 2007.

JENSEN, A. F.; CHERCHI, E.; MABIT, S. L.; ORTÚZAR, J. D. D. **Predicting the potential market for electric vehicles**. *Transportation Science*, v. 51, n. 2, p. 427–440, 2017.

JUNIOR, F. R. P. **Redes neurais diretas e recorrentes na previsão do preço de energia elétrica de curto prazo no mercado brasileiro**. Dissertação (Mestrado em Engenharia Elétrica), Universidade Federal do Pará, Belém, 2016.

JUPYTER, P. **Project Jupyter**, 2020. Disponível em: <<https://jupyter.org/>>. Acesso em: 31 ago. 2020.

KAPUSTIN, N. O.; GRUSHEVENKO, D. A. **Long-term electric vehicles outlook and their potential impact on electric grid**. *Energy Policy*, v. 137, p. 111103, 2020.

LAGO, J.; DE RIDDER, F.; DE SCHUTTER, B. **Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms**. *Applied Energy*, v. 221, p. 386–405, 2018.

LORENZO, H. C. O setor elétrico brasileiro: passado e futuro. **Perspectivas: Revista de Ciências Sociais**, 2001.

LUZ, C. P. DE M. S. **O mercado a termo de energia elétrica no Brasil: evidências sobre sua dinâmica a partir de um estudo exploratório**. Dissertação (Mestrado em Administração), Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2011.

MEDEIROS, L. **Previsão do Preço Spot no Mercado de Energia Elétrica**. Tese (Doutorado em Engenharia Elétrica), Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2003.

MME. **Consultas Públicas - Ministério de Minas e Energia**, 2017. Disponível em: <<http://antigo.mme.gov.br/web/guest/servicos/consultas-publicas>>. Acesso em: 24 jun. 2021.

MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to time series analysis and forecasting**. John Wiley & Sons, 2015.

MORETTIN, P. A.; TOLOI, C. **Análise de séries temporais**. p. 538, Edgard Blücher, São Paulo, 2006.

NASCIMENTO, R. A.; FERREIRA, R. P. M. **Location of charging stations for electric vehicles**. CILAMCE, Foz do Iguaçu, 2020

ONS. **O que é o SIN**, 2017. Disponível em: <<http://www.ons.org.br/paginas/sobre-o-sin/o-que-e-o-sin>>. Acesso em: 26 jun. 2021.

\_\_\_\_\_. **Sobre a ONS - O que é a ONS**, 2018. Disponível em: <<http://www.ons.org.br/paginas/sobre-o-ons/o-que-e-ons>>. Acesso em: 21 jun. 2021.

\_\_\_\_\_. **O Sistema em Números**, 2021a. Disponível em: <<http://www.ons.org.br/paginas/sobre-o-sin/o-sistema-em-numeros>>. Acesso em: 30 jun. 2021.

\_\_\_\_\_. **Histórico da Operação**, 2021b. Disponível em: <<http://www.ons.org.br/paginas/resultados-da-operacao/historico-da-operacao>>. Acesso em: 19 jan. 2021b.

PANAPAKIDIS, I. P.; DAGOUMAS, A. S. **Day-ahead electricity price forecasting via the application of artificial neural network based models**. Applied Energy, v. 172, p. 132–151, 2016.

PORTO, V. G. B. A.; OLIVI, L. R. **Prediction of Brazilian Electric Energy Price Using Recurrent Artificial Neural Networks and Correction Filter**. Congresso Brasileiro de Automática-CBA. Anais 2020

RODRIGUES, A. L. **Redes neurais artificiais aplicadas na previsão de preços do mercado spot de energia elétrica**. Dissertação (Mestrado em Energia), Universidade de São Paulo, São Paulo, 2009.

SANTOS, G. **Uma aplicação de redes neurais recorrentes do tipo LSTM à previsão dos preços de curto prazo do mercado de energia elétrica brasileiro**. Sustainability (Switzerland), v. 11, n. 1, p. 1–14, 2019.

SENADO. **PLS 232/2016 - Senado Federal**, 2020. Disponível em: <<https://www25.senado.leg.br/web/atividade/materias/-/materia/126049>>. Acesso em: 25 jun. 2021.

SHMUELI, G.; LICHTENDAHL JR, K. C. **Practical time series forecasting with r: A hands-on guide**. [s.l.] Axelrod Schnall Publishers, 2016.

SHUMWAY, R. H.; STOFFER, D. S.; STOFFER, D. S. **Time series analysis and its applications**. v3, Springer, 2000.

SIGA. **Sistema de Informações de Geração da ANEEL**, 2021. Disponível em: <<https://bit.ly/2IGf4Q0>>. Acesso em: 29 jun. 2021.

SILVA, I. V. G. **Previsão de Vento para Geração de Energia Elétrica**. Dissertação (Mestrado Engenharia Elétrica) Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2010.

SINGHAL, D.; SWARUP, K. S. **Electricity price forecasting using artificial neural networks**. International Journal of Electrical Power & Energy Systems, v. 33, n. 3, p. 550–555, 2011.

SOARES, R. A. R.. **Abordagem à Previsão do Preço de Energia Elétrica Via Métodos de Suavização Exponencial**. Dissertação (Mestrado em Engenharia Eletrotécnica), Instituto Superior de Engenharia de Lisboa, Lisboa, 2013.

TEODORO, D. M. **A reestruturação do setor elétrico brasileiro e os reflexos em uma empresa estatal: um estudo de caso na Centrais Elétricas de Santa Catarina-CELESC**. Dissertação (Mestrado em Economia) , Universidade Federal de Santa Catarina, Florianópolis, 2006.

TOSHIOKA, F. **Previsão de preço semanal de energia elétrica através de redes neurais artificiais com limites de saturação**. Dissertação (Mestrado em Desenvolvimento de Tecnologia), Instituto de Tecnologia para o Desenvolvimento, Curitiba, 2017.

WADY, A. F. **Os sinais regulatórios necessários para o desenvolvimento da mobilidade elétrica na cidade: uma abordagem com foco na infraestrutura de recarga**. Dissertação (Mestrado em Política Científica e Tecnológica) , Universidade Estadual de Campinas, Campinas, 2021.

WALVIS, A.; GONÇALVES, E. D. L. **Avaliação das reformas recentes no setor elétrico brasileiro e sua relação com o desenvolvimento do mercado livre de energia**. p 1-63, FGV-CERI, 2014

WANG, D.; LUO, H.; GRUNDER, O.; LIN, Y.; GUO, H. **Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm**. Applied Energy, v. 190, p. 390–407, 2017.

ZIEL, F.; STEINERT, R. **Probabilistic mid-and long-term electricity price forecasting**. Renewable and Sustainable Energy Reviews, v. 94, p. 251–266, 2018.