

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
ENGENHARIA DE ENERGIA

**PREVISÃO DO COMPORTAMENTO DO PREÇO DE
LIQUIDAÇÃO DAS DIFERENÇAS (PLD) COM FERRAMENTAS ESTATÍSTICAS**

por

William Lagasse

Monografia apresentada à Comissão de Graduação do Curso de Engenharia de Energia da Escola de Engenharia da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do diploma de Bacharel em Engenharia de Energia.

Porto Alegre, julho de 2020.



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
ENGENHARIA DE ENERGIA

PREVISÃO DO COMPORTAMENTO DO PREÇO DE
LIQUIDAÇÃO DAS DIFERENÇAS (PLD) COM FERRAMENTAS ESTATÍSTICAS

por

William Lagasse

ESTA MONOGRAFIA FOI JULGADA ADEQUADA COMO PARTE DOS
REQUISITOS PARA A OBTENÇÃO DO TÍTULO DE
BACHAREL EM ENGENHARIA DE ENERGIA.
APROVADA EM SUA FORMA FINAL PELA BANCA EXAMINADORA

Prof. Dr^a Leticia Jenisch Rodrigues
Coordenadora do Curso de Engenharia de Energia

Orientador: Prof. Dr. Paulo Smith Schneider

Coorientadora: Eng.^a Simone Ferigolo Venturini

Banca examinadora:

M.^a Bibiana Maitê Petry Ferraz – LASEP / UFRGS

Dr. Julian David Hunt – IIASA

Prof.^a Dr.^a Mariana Resener – DELAE / UFRGS

Porto Alegre, 16 de julho de 2020.

AGRADECIMENTOS

À Universidade Federal do Rio Grande do Sul, por fornecer ensino público, gratuito e de qualidade durante a minha formação.

Aos meus pais, Leila e Arsildo, que me deram todo suporte necessário e permitiram que eu tivesse acesso à melhor formação possível.

Ao meu orientador, Prof. Dr. Paulo Smith Schneider, que me auxiliou ao longo desse trabalho, sempre disponível quando eu precisava de orientação.

À minha coorientadora, Eng.^a Simone Ferigolo Venturini, por todo suporte, conselhos e contribuições no desenvolvimento desse trabalho.

A todos os professores que me acompanharam na graduação, pelo compartilhamento de seus conhecimentos e pelas lições ensinadas durante minha formação.

À equipe do PLD do LETA, pelas discussões semanais e pelos conhecimentos compartilhados.

A todos meus amigos, que participaram dessa trajetória e contribuíram para que ela fosse mais fácil e divertida.

LAGASSE, W. **Previsão do Comportamento do Preço de Liquidação das Diferenças (PLD) com Ferramentas Estatísticas**. 2020. 23 folhas. Monografia (Trabalho de Conclusão do Curso em Engenharia de Energia) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

RESUMO

A energia elétrica no mercado de curto prazo, ou mercado *spot*, é comercializada com base no cálculo do Preço de Liquidação das Diferenças (PLD), e seu cálculo segue uma metodologia complexa e dispendiosa. Buscando uma opção alternativa, o presente trabalho propõe o emprego de métodos estatísticos de previsão dos valores e tendências do PLD, com o intuito de mitigar a insegurança na tomada de decisão dos agentes de mercado. Os modelos propostos empregam como entrada a energia natural afluenta, a energia armazenada, as gerações hidrelétrica, térmica e eólica, e a demanda de energia elétrica, divulgadas publicamente pelo ONS, além da série histórica do próprio PLD, divulgada pela CCEE, buscando praticidade para utilização do modelo desenvolvido. As previsões são elaboradas por métodos de análise de séries temporais e de regressão para o horizonte de uma semana no submercado Sul para o patamar de carga pesado. Sequências históricas são empregadas para a construção dos modelos e seus resultados de previsão são validados para os períodos excluídos dessas sequências de dados. O modelo que apresenta melhor acurácia na previsão do PLD é o modelo de análise de séries temporais que emprega a suavização exponencial simples com erros médios de 17,91 % (percentual) e de R\$ 28,01/MWh (absoluto). O melhor percentual de acerto de tendência é de 61,36 % obtido a partir de um modelo de regressão linear múltipla.

PALAVRAS-CHAVE: Energia Elétrica, Preço de Liquidação das Diferenças, Modelos de Previsão, Séries Temporais.

LAGASSE, W. **Forecasting the spot price behavior of energy in the Brazilian market with statistical tools**. 2020. 23 pages. Dissertation (Completion work of the Energy Engineering Course) – School of Engineering, Federal University of Rio Grande do Sul, Porto Alegre, 2020.

ABSTRACT

Electricity in the short-term market, or spot market, is sold based on the calculation of the Difference Settlement Price (PLD), and its calculation follows a complex and expensive methodology. Looking for an alternative option, the present work proposes the use of statistical methods for forecasting PLD values and trends, in order to mitigate the insecurity in the decision making of the market agents. The proposed models use as input the affluent natural energy, the stored energy, the hydroelectric, thermal and wind generations, and the demand for electric energy, publicly disclosed by ONS, in addition to the historical series of the PLD itself, disclosed by CCEE, seeking practicality for use of the developed model. Forecasts are elaborated using methods of time series analysis and regression for the one-week horizon in the southern Brazilian submarket for the heavy load level. Historical sequences are used to build the models and their forecast results are validated for the periods excluded from these data sequences. The model that presents the best accuracy in the PLD forecast is the time series analysis model that employs simple exponential smoothing with average errors of 17.91% (percentage) and R\$ 28.01/MWh (absolute). The best percentage of trend hit is 61.36%, obtained from a multiple linear regression model.

KEYWORDS: Electric Power, Short-Term Electricity Price, Forecasting Models, Time Series.

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Revisão Bibliográfica	1
1.2	Objetivos	3
2	FUNDAMENTAÇÃO TEÓRICA	3
2.1	O Setor Elétrico Brasileiro	3
2.1.1	Segmentação do Setor Elétrico.....	3
2.1.2	Sistema Interligado Nacional	4
2.1.3	Mercado de Energia.....	4
2.1.4	Formação do Preço de Liquidação das Diferenças.....	6
2.2	Métodos de Previsão	7
2.2.1	Análise de Regressão.....	7
2.2.2	Análise de Séries Temporais	7
2.2.3	Coefficientes de Determinação	8
2.2.4	Erros	9
3	METODOLOGIA DESENVOLVIDA	9
4	ESTUDO DE CASO E RESULTADOS.....	11
4.1	Modelos de Regressão para o PLD	12
4.2	Modelos de Séries Temporais para o PLD	16
5	DISCUSSÃO SOBRE OS RESULTADOS OBTIDOS.....	19
6	CONCLUSÃO	22

1 INTRODUÇÃO

Uma variável chave do setor elétrico é o preço da energia elétrica no mercado de curto prazo, também chamado de mercado *spot*¹. Emprega-se no Brasil o Preço de Liquidação das Diferenças (PLD) para valorar os montantes liquidados decorrentes das operações de compra e venda de energia elétrica realizadas no Mercado de Curto Prazo (MCP). O PLD promove flexibilidade nas negociações, possibilita ajustes entre a energia contratada e a energia gerada, e é adotado como referência para contratos de longo prazo, constituindo uma ferramenta importante de ajuste entre ofertantes e demandantes (LEITE, CASTRO e TIMPONI, 2013).

O preço *spot* possui uma forma de cálculo complexa e que exige grande esforço computacional, fazendo uso de complexos modelos matemáticos. Os mesmos programas utilizados pelo Operador Nacional do Sistema Elétrico (ONS) para definir a programação da operação do sistema elétrico nacional são utilizados pela Câmara de Comercialização de Energia Elétrica (CCEE) para calcular o PLD. O planejamento da operação de sistemas hidrotérmicos compreende uma vasta gama de atividades, indo do planejamento ao longo e médio prazo até a programação diária da operação dos reservatórios (CCEE, 2020f).

Ao contrário da maioria dos países, o Sistema Elétrico Brasileiro (SEB) é predominantemente hidrelétrico com complementação termelétrica. Conforme o Balanço Energético Nacional de 2019, em 2018 a matriz elétrica brasileira teve 66,6% da geração de energia elétrica no sistema brasileiro sendo proveniente de recursos hídricos (EPE, 2019). Devido à predominância da geração hidrelétrica, os modelos de otimização do cálculo do PLD buscam encontrar um equilíbrio entre o benefício imediato do uso da água dos reservatórios e o benefício futuro de seu armazenamento, medido através da economia esperada dos combustíveis utilizados para operar as usinas termelétricas (CCEE, 2020e).

Significativas variações do PLD podem ocorrer em curtos períodos de tempo, pois seu cálculo considera condições hidrológicas, demanda de energia, preços dos combustíveis das termelétricas, custo de déficit, entrada de novos projetos e disponibilidade de equipamentos de geração e transmissão, de maneira a obter o despacho ótimo da geração para cada submercado (CCEE, 2020e).

A volatilidade do preço *spot* impacta o mercado de energia, sendo uma das maiores preocupações dos agentes do setor elétrico brasileiro, principalmente no contexto do Ambiente de Contratação Livre (ACL). Os agentes possuem diferentes estratégias de comercialização de energia para a maximização do seu lucro e uma variação inesperada do PLD pode acarretar um importante impacto financeiro, o que acaba refletindo em diversos setores da economia brasileira.

Diante desse cenário, é evidenciada a necessidade de uma melhor compreensão da formação dos preços do mercado de curto prazo de maneira a diminuir as incertezas dos agentes do setor elétrico e reduzir os riscos no processo de tomada de decisão e de elaboração de estratégias de comercialização de energia. Modelos alternativos para o cálculo do PLD são úteis para orientar tomadas de decisão por parte dos agentes do setor, preferencialmente quando empregam dados facilmente obtidos e com reduzido esforço computacional.

1.1 Revisão Bibliográfica

A literatura apresenta diversos trabalhos que correlacionam as variáveis do setor elétrico com os valores históricos do PLD para obter modelos de estimativa alternativos.

Rodrigues (2009) desenvolveu um modelo de previsão do preço de energia utilizando Redes Neurais Artificiais (RNA). O autor utilizou dados de carga, geração em Itaipu, geração nas demais hidrelétricas (exceto Itaipu), termelétricas, térmicas nuclear, fonte eólica, preços anteriores, nível de armazenamento, energia natural afluyente, intercâmbios entre submercados e importação e exportação de outros países, sendo todos esses dados disponibilizados pelo ONS. Ele fez projeção dos preços com base semanal utilizando dados das seis semanas anteriores para os quatro submercados brasileiros. O maior erro relativo médio foi de 12% para o patamar de carga pesada. O autor realizou testes de sensibilidade

¹Fulgencio (2007, p. 416) define Mercado *Spot* como “Mercado no qual são negociadas quantidades marginais de produto não cobertos por contratos. O mercado *Spot* considera a oferta e demanda do produto no momento da negociação de compra e venda, para entrega imediata”.

dos resultados retirando as variáveis carga ou energia armazenada e obteve grande imprecisão, o que indica que essas variáveis são de grande importância no desenvolvimento de um modelo alternativo de previsão de preços. Também foi enfatizada pelo autor a volatilidade que se encontra na hidrologia do submercado Sul.

Outros trabalhos também fazem uso de RNAs para realizar previsões de PLD, mas não apresentam resultados com alta confiabilidade (COELHO, 2018) (SILVA, 2018). Coelho (2018) usou redes neurais recorrentes para gerar cenários de PLD em um horizonte de cinco anos e demonstrou a importância de não considerar os valores históricos do PLD que apresentam *outliers*², pois estes podem causar falha na previsão de cenários futuros de preço. As RNAs recorrentes foram consideradas como adequadas para as previsões de PLD, pois possuem uma boa memória para o curto prazo, porém o melhor cenário obtido apresentou um erro percentual de 64,93%. Silva (2018) desenvolveu modelos de previsão do PLD para o submercado Sudeste/Centro-Oeste utilizando como variáveis de entrada a energia armazenada nos reservatórios, energia proveniente de geração hidrelétrica, termelétrica, e carga do submercado. A energia natural afluyente não foi utilizada como dado de entrada. Também foi elaborada uma previsão baseada apenas na série temporal do preço, tendo alcançado sua melhor previsão, apresentando erro percentual de 89,72% e um desvio padrão de R\$ 90,3421. Ambos modelos apresentaram resultados que acompanharam as variações dos preços reais, porém com um desvio bastante significativo.

Com uma abordagem diferente, Sousa (2013) analisou a volatilidade do PLD utilizando modelos estatísticos para prever preços a partir de séries temporais do PLD com o auxílio dos *softwares* Eviews® e R®, que utilizam os modelos ARIMA³ e o modelo GARCH⁴. As análises foram feitas com base em uma série temporal do período de 2005 a 2010. Para a previsão de preços o melhor resultado apresentou erro percentual absoluto médio de 81,28 %, obtido com um modelo GARCH.

O trabalho desenvolvido por Maia *et. al* (2016), se preocupou em entender os principais fatores que influenciam o valor do PLD e modelar os comportamentos empregando regressão múltipla conjuntamente com séries temporais. O trabalho analisou a relação entre os preços do mercado *spot* e a Energia Natural Afluyente em MW_{med}, a Energia Armazenada em GWh e o Intercâmbio de Energia em GWh entre os submercados, de forma concomitante. Foi destacado o caráter importador de energia do submercado Sul devido a sua baixa capacidade de armazenamento em reservatórios, enquanto o submercado Sudeste/Centro-Oeste possuiu um caráter exportador, devido à sua capacidade de armazenamento. Os autores obtiveram bons resultados de predição para todos submercados, com um coeficiente de determinação⁵, R², superior a 0,74, exceto para a região Sul, que apresentou um R² inferior a 0,12. Salienta-se que um dos possíveis motivos para a não-aderência do modelo ao submercado Sul deveu-se ao caráter importador de energia da região, pois o modelo adotado não considerou dados das demais regiões. Dentre as três variáveis estudadas na região Sul, a energia natural afluyente e o intercâmbio de energia foram as que se mostraram mais relevantes. De acordo com o exposto no artigo, os autores sugerem a adoção de um modelo autorregressivo para obtenção de melhores resultados para a região Sul.

Um modelo de Projeção Dinâmica simplificada utilizando como dados de entrada apenas informações fornecidas pelo ONS e CCEE foi desenvolvido por (CARVALHO e SARDINHA, 2011). Os valores de entrada utilizados foram Energia Natural Afluyente em MW_{med}, Energia Armazenada em MWh, Carga em MW_{med} e a série do PLD em R\$/MWh todos em base mensal média. Os autores não utilizaram dados de geração e nem de intercâmbio entre subsistemas. No trabalho a análise das variáveis e suas inter-relações foram feitas utilizando matrizes de correlações e estatísticas descritivas com o auxílio do *software* Eviews®. Para a projeção dos preços foi utilizado um modelo de Vetores Auto Regressivos com Variáveis Externas (VARx). Os resultados obtidos para estimativa do PLD para os submercados Sudeste/Centro-Oeste e Sul obtiveram R² inferior a 0,70, já para os submercados Norte e Nordeste o melhor R² ficou próximo de 0,75, porém as curvas dos valores previstos não acompanharam a série real na integridade do período analisado.

² Em estatística, um *outlier* é um valor atípico, que apresenta grande afastamento dos demais pontos da série.

³ Modelo auto-regressivo integrado de médias móveis, do inglês *Auto-Regressive Integrated Moving Average*.

⁴ Modelo Auto-Regressivo Generalizado de Heterocedasticidade Condicional, do inglês *Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH)*.

⁵ R² representa a porcentagem de variação na resposta que é explicada pelo modelo.

Contudo, não há um consenso na literatura de quais modelos estatísticos são mais indicados para prever o PLD. Os trabalhos apresentados anteriormente apresentaram previsões com erros significativos, com exceção do trabalho desenvolvido por (Rodrigues, 2009). Entretanto, (Rodrigues, 2009) utilizou dados da série histórica do PLD anteriores a 2009, época em que os limites do PLD apresentavam uma metodologia de cálculo que não é mais praticada e eram utilizadas versões anteriores dos softwares de cálculo do PLD.

Diante do exposto, surge a seguinte questão de pesquisa: é possível identificar um algoritmo estatístico para realizar a previsão dos valores e tendências do PLD para o submercado Sul, em um horizonte de uma semana?

1.2 Objetivos

Neste contexto, o objetivo deste trabalho é desenvolver e comparar diferentes modelos estatísticos, alternativos aos modelos computacionais utilizados pela CCEE, para previsão dos valores do PLD no submercado Sul em um horizonte de uma semana.

Os objetivos específicos são:

- Definir métodos estatísticos, causais e de séries temporais, que tenham potencial aplicabilidade para projeções do PLD com reduzido muito esforço computacional;
- Desenvolver os modelos de previsões para o PLD a partir dos métodos estatísticos definidos;
- Avaliar os erros de previsões dos modelos desenvolvidos, indicando quais obtiveram melhores resultados.
- Estimar as tendências dos preços uma semana à frente.

Salienta-se que o presente trabalho não possui como objetivo refazer os cálculos de PLD conforme as metodologias utilizadas pelo ONS ou CCEE, mas utilizar indiretamente os resultados obtidos por essas metodologias para subsidiar a elaboração de um modelo alternativo que obtenha valores aproximados.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 O Setor Elétrico Brasileiro

O Setor Elétrico Brasileiro (SEB) é reconhecido como uma indústria de base e pode ser relacionado à inclusão social e ao crescimento das atividades econômicas do país, estando relacionado à qualidade de vida da população (CNI, 2017). É organizado para garantir a segurança do fornecimento de energia elétrica, promover a modicidade tarifária e a inserção social a partir da universalização do atendimento (CCEE, 2020c).

2.1.1 Segmentação do Setor Elétrico

O SEB é constituído por quatro segmentos: geração, transmissão, distribuição e comercialização. Os setores de geração e comercialização foram progressivamente desregulados e são caracterizados como segmentos competitivos. Esses setores possuem muitos agentes, sendo a energia elétrica considerada uma *commodity* em consequência da sua homogeneidade. A transmissão e a distribuição constituem monopólios naturais devido a inviabilidade econômica de haver competição entre agentes para uma mesma área de concessão, tendo em vista as suas estruturas físicas o que resultaria em tarifas mais caras de uso dos sistemas de transmissão ou distribuição para os consumidores (ABRADEE, 2020).

Na geração, a energia elétrica é produzida e injetada nos sistemas de transmissão ou distribuição para chegar até os consumidores. De acordo com o Sistema de Informações de Geração (SIGA) da Agência Nacional de Energia Elétrica (ANEEL), a principal fonte de geração de energia elétrica no Brasil é a hidrelétrica (água corrente dos rios), seguida das termelétricas (gás natural, carvão mineral, combustíveis fósseis, biomassa e nuclear) que complementam a geração hidrelétrica em períodos de estiagem e em horários de pico. Em junho de 2020 o Brasil possuía 62,51 % de sua capacidade instalada em operação sendo de origem hidrelétrica e 26,74 % de termelétricas, conforme pode ser observado na Tabela 2.1 (ANEEL, 2020e).

Tabela 2.1: Matriz Elétrica Brasileira – Capacidade Instalada dos Empreendimentos em Operação.

Tipo	Potência Outorgada (kW)	Quantidade	% (Pot. Outorgada)
Usina Hidrelétrica	102.975.428,00	219	58,98%
Usina Termelétrica	44.688.887,89	3.058	25,60%
Central Geradora Eólica	15.822.003,86	644	9,06%
Pequena Central Hidrelétrica	5.359.590,49	417	3,07%
Central Geradora Solar Fotovoltaica	2.940.359,95	3.895	1,68%
Usina Termonuclear	1.990.000,00	2	1,14%
Central Geradora Hidrelétrica	805.145,13	733	0,46%
Central Geradora Unid-Elétrica	50,00	1	0,00%
Total	174.581.465,32	8.969	100%

Fonte: Adaptado de (ANEEL, 2020e).

O sistema de transmissão é responsável por transportar grandes quantidades de energia oriundas das unidades geradoras, operando linhas com tensão igual ou superior a 230 kV. Por sua vez, o sistema de distribuição também é responsável por transportar energia e distribuir aos consumidores e opera com tensões abaixo de 230 kV. Em alguns casos, unidades geradoras com menores potências podem injetar energia diretamente no sistema de distribuição. Os segmentos de distribuição e transmissão têm seus preços regulados pela ANEEL, que é a agência reguladora do setor (ABRADEE, 2020).

O segmento de comercialização de energia possui um papel mais ligado ao contexto econômico e institucional. Seu surgimento está relacionado com a reestruturação do setor elétrico brasileiro, ocorrida na década de 1990, onde se buscou a desestatização de empresas e inclusão gradual de características de livre mercado ao setor. Os agentes de comercialização de energia elétrica no Brasil possuem autorização, concessão ou permissão para realizarem operações de compra e venda de energia elétrica na CCEE, sendo que muitos deles atuam como intermediários entre usinas e consumidores livres (ABRADEE, 2020; ANEEL, 2020d; CNI, 2017).

2.1.2 Sistema Interligado Nacional

O Sistema Interligado Nacional (SIN) é um sistema hidro-termo-eólico de grande porte, que engloba os sistemas de produção e transmissão de energia elétrica das regiões do Brasil e que possibilita o suprimento de energia elétrica nas regiões do país. O SIN é constituído pelos subsistemas Sudeste/Centro-Oeste, Sul, Nordeste e Norte, e permite o intercâmbio da energia produzida em todas as regiões do Brasil, exceto nos sistemas isolados que são localizados principalmente na região Norte (ONS, 2020).

2.1.3 Mercado de Energia

O modelo atual do mercado de energia nacional entrou em vigor em 2004 com as Leis nº 10.847 e 10.848, de 15 de março de 2004, e pelo Decreto nº 5.163, de 30 de julho de 2004, que dentre outras providências, regulamentou a comercialização de energia elétrica, estabelecendo que o preço seja calculado com periodicidade máxima semanal (CCEE, 2020g).

O novo modelo definiu a criação de três órgãos: a CCEE, cuja função é dar continuidade às atividades do Mercado Atacadista de Energia (MAE), relativas à comercialização de energia elétrica no sistema interligado; a Empresa de Pesquisa Energética (EPE), que é responsável pelo planejamento do setor elétrico em longo prazo; o Comitê de Monitoramento do Setor Elétrico (CMSE), que é uma instituição com a função de avaliar permanentemente a segurança do suprimento de energia elétrica. O exercício do Poder Concedente foi delegado ao Ministério de Minas e Energia (MME). O setor tem como órgão regulador a ANEEL, e o ONS como responsável pela operação das instalações de geração e transmissão nos sistemas interligados brasileiros (CCEE, 2020g; MME, 2020).

No que concerne a comercialização de energia, foram estabelecidos dois ambientes para a realização de contratos de compra e venda de energia, o Ambiente de Contratação Regulada (ACR), do

qual participam agentes de geração e distribuição, e o Ambiente de Contratação Livre (ACL), com geradores, distribuidores, comercializadores, importadores e exportadores, além dos consumidores livres e especiais. Todos os contratos acordados nesses dois ambientes de contratação são registrados na CCEE (CCEE, 2020g).

No ACR a compra e venda de energia é firmada a partir de contratos que possuem regulação específica, tais como o preço da energia, o submercado de registro do contrato e a vigência de suprimento, não sendo passíveis de alterações bilaterais pelos agentes. A contratação é realizada por meio de leilões de energia promovidos pela CCEE, sob delegação da ANEEL (CCEE, 2020a; CCEE, 2020c). Esses leilões buscam contratar energia pelo menor preço possível, buscando a modicidade tarifária, além de atrair investidores para construção de novas usinas para expandir a geração e reter a geração existente (ANEEL, 2020c).

No ACL, também chamado de Mercado Livre, geradoras, comercializadoras, consumidores livres e especiais têm liberdade para negociar o preço e o volume de compra e venda da energia a ser determinado nos contratos (CCEE, 2020c). De acordo com a portaria nº 514 de 27 de novembro de 2018, os consumidores livres devem possuir, no mínimo, 2000 kW de demanda contratada para poder contratar energia proveniente de qualquer fonte de geração, enquanto que os consumidores especiais devem possuir demanda contratada igual ou maior que 500 kW e menor que 2000 kW. Os consumidores especiais podem contratar energia proveniente apenas de usinas eólicas, solares, a biomassa, pequenas centrais hidrelétricas (PCH) ou hidráulicas de empreendimentos com potência inferior ou igual a 50.000 kW (BRASIL, 2009).

Como todos os contratos celebrados nestes ambientes de contratação são registrados na CCEE, ela realiza a apuração dos montantes de energia efetivamente consumidos ou produzidos por cada agente. As diferenças averiguadas, sejam elas negativas ou positivas, são contabilizadas e posteriormente liquidadas financeiramente no Mercado de Curto Prazo, sendo valoradas pelo Preço de Liquidação das Diferenças (PLD). O MCP, também conhecido como Mercado das Diferenças, é um segmento da CCEE que promove o ajuste das diferenças entre os montantes contratados e os montantes de geração e consumo de energia elétrica efetivamente verificados e atribuídos aos respectivos agentes (CCEE, 2020b).

Assim, o PLD é um indicador calculado antecipadamente (*ex-ante*) pela CCEE, com periodicidade máxima semanal, definido por submercado (Sul, Sudeste/Centro-Oeste, Nordeste e Norte), por patamar de carga (leve, médio e pesado) e limitado entre um valor mínimo e máximo, definidos anualmente pela ANEEL.

No artigo 57 do Decreto nº 5.163, de 30 de julho de 2004, cuja redação é dada pelo Decreto nº 9.143, de 22 de agosto de 2017, são definidas as características de divulgação e cálculo para o PLD, conforme extrato de texto a seguir.

§ 1o O PLD, a ser publicado pela CCEE, será calculado antecipadamente, com periodicidade máxima semanal e terá como base o custo marginal de operação, limitado por preços mínimo e máximo, e deverá observar o seguinte:

- I - a otimização do uso dos recursos eletro-energéticos para o atendimento aos requisitos da carga, considerando as condições técnicas e econômicas para o despacho das usinas;
- II - as necessidades de energia elétrica dos agentes;
- III - os mecanismos de segurança operativa, podendo incluir curvas de aversão ao risco de déficit de energia;
- IV - o custo do déficit de energia elétrica;
- V - as restrições de transmissão entre submercados;
- VI - as interligações internacionais; e
- VII - os intervalos de tempo e escalas de preços previamente estabelecidos que deverão refletir as variações do valor econômico da energia elétrica. (BRASIL, 2004a).

Porém, a partir de janeiro de 2021 o PLD será calculado em base horária, conforme definido no cronograma da Comissão Permanente para Análise de Metodologias e programas Computacionais do Setor Elétrico (CPAMP). De abril de 2018 até a implementação do preço horário, a CCEE realiza e divulga o cálculo de Preço Horário Sombra em paralelo ao PLD em base semanal. A divulgação do preço sombra tem objetivo antecipar os eventuais impactos da adoção do preço horário, preparando os agentes para a nova base de cálculo de vigorará em janeiro de 2021 (CCEE, 2020d).

2.1.4 Formação do Preço de Liquidação das Diferenças

O PLD é o valor de referência utilizado para efetuar a contabilização dos montantes de energia elétrica comercializados no SIN e para a liquidação financeira dos valores decorrentes das operações de compra e venda de energia elétrica realizadas no MCP (BRASIL, 2004b). Cabe a CCEE a apuração desse valor, por meio de modelos de otimização desenvolvidos pelo Centro de Pesquisas de Energia Elétrica (CEPEL) (CEPEL, 2020a). A CCEE determina semanalmente esses valores e os divulga nas sextas-feiras, por submercado (Sul-S, Sudeste-SE/Centro-Oeste-CO, Nordeste-NE e Norte-N) e por patamar de carga, limitado a um preço mínimo e máximo estabelecido anualmente pela ANEEL (CCEE, 2013; BRASIL, 2004a). O PLD é calculado antecipadamente (*ex-ante*) pela CCEE e tem por base de cálculo o Custo Marginal da Operação (CMO). Semanalmente, a CCEE assume por base o CMO, nos patamares de carga pesada, média e leve, e enquadra-o entre um valor mínimo (PLD_{min}) e um valor máximo (PLD_{max}). Esses valores são determinados anualmente pela ANEEL de acordo com a legislação vigente, e são adotados nos quatro submercados que formam o sistema elétrico brasileiro (AMARANTE *et al.*, 2018).

O PLD é calculado pelos modelos de otimização NEWAVE (Modelo Estratégico de Geração Hidrotérmica a Subsistemas Equivalentes Interligados) e DECOMP (Determinação da Coordenação da Operação a Curto Prazo), ambos desenvolvidos pelo CEPEL (CEPEL, 2020a). Esses modelos são utilizados pelo ONS para o Planejamento da Operação hidrotérmica e o despacho do SIN, e pela CCEE, para o estabelecimento do preço de energia no mercado de curto prazo. A EPE utiliza os modelos no cálculo da garantia física das usinas, para instituir leilões de energia e na formulação do Planejamento da Expansão (CEPEL, 2020b).

Os preços máximos e mínimos definidos anualmente pela ANEEL vêm sofrendo alterações nas suas metodologias de cálculo. O Decreto nº 5.163, de 30 de julho de 2004 atribui à ANEEL a competência de determinar os limites máximos e mínimos do PLD. O valor máximo deve considerar os custos variáveis das usinas termelétricas, enquanto os valores mínimos devem considerar os custos de operação e manutenção das usinas hidrelétricas (BRASIL, 2004a).

Em 2004 o valor máximo do PLD foi determinado pela Resolução ANEEL 682/2003, cuja atualização deveria ocorrer anualmente e considerar o menor valor entre o preço da usina termoeletrica mais cara, com capacidade instalada maior que 65 MW, na época, a UTE Alegrete (ANEEL, 2020b), e a atualização do valor máximo daquele ano pela variação do Índice Geral de Preços – Disponibilidade Interna (IGP-DI). Entre 2004 a 2014 o critério de atualização dos preços prevaleceu sendo pelo IGP-DI, atingindo em 2014 o valor máximo de R\$ 822,83/MWh (ANEEL, 2020a).

Entre 2015 e 2019 a metodologia de definição dos valores máximos e mínimos do PLD foi definida de acordo Resolução Normativa ANEEL nº 663/2014, com validade entre a primeira e a última semana operativa⁶ de cada ano. O valor máximo passou a ser definido em dezembro do ano anterior com base no Custo Variável Unitário (CVU⁷) da usina termelétrica em operação comercial, a gás natural, mais cara contratada no ACR (ANEEL, 2020a). Já o PLD mínimo era calculado no mês de dezembro de cada ano com base no maior valor entre: i) o calculado com base na Receita Anual de Geração (RAG) das usinas hidrelétricas em regime de cotas, nos termos da Lei nº 12.783/2013, excluídos os valores relacionados à remuneração e reintegração de investimentos, e adicionada a estimativa de Compensação Financeira pelo Uso dos Recursos Hídricos (CFURH); e ii) as estimativas dos custos de geração da usina de Itaipu para o ano seguinte, fornecidas pela Itaipu Binacional para fins de reajustes e/ou revisões tarifárias.

Para 2020, ocorreu uma nova mudança nas metodologias de cálculo dos limites do PLD conforme a Resolução Normativa ANEEL nº 858/2019. Foram estabelecidos dois novos limites para o PLD máximo: o limite máximo estrutural; e o limite máximo horário, calculado com base média ponderada, pela potência instalada, dos Custos Variáveis Unitários das usinas termelétricas a óleo diesel. Os limites máximos do PLD serão atualizados pela ANEEL anualmente pela variação do Índice Nacional de Preços ao Consumidor Amplo (IPCA) em dezembro de cada ano para utilização em janeiro do ano seguinte a partir do valor de R\$ 556,58/MWh, e de R\$ 1.141,85/MWh, para o PLD máximo estrutural, sendo esses preços de setembro de 2019. Se a média diária dos PLDs horários for superior ao PLD máximo estrutural,

⁶ Período que inicia às 0h de sábado e termina às 24h da sexta-feira da semana subsequente.

⁷ Custo Variável Unitário, valor expresso em Reais por Megawatt-hora (R\$/MWh), necessário para cobrir todos os custos operacionais do empreendimento termoeletrico.

a CCEE deve ajustar a série de PLDs horários até que a média de seus valores seja igual ao PLD máximo estrutural. Por sua vez, o valor mínimo do PLD é calculado como base no maior valor entre a Tarifa de Energia de Otimização (TEO⁸) da UHE Itaipu e TEO das outras usinas hidrelétricas do Sistema Interligado Nacional – SIN.

2.2 Métodos de Previsão

Os métodos de previsão podem ser classificados como qualitativos e quantitativos. Os métodos qualitativos são subjetivos e dependem da experiência acumulada de especialistas ou grupo de pessoas para predizer as probabilidades dos resultados de eventos. No presente trabalho serão abordados métodos quantitativos, que são estruturados e podem ser aplicados por outros analistas. Eles são classificados em métodos causais, ou seja, baseados na análise estatística de realizações passadas que se correlacionam com a variável de interesse, e de séries temporais, que envolvem apenas análises estatísticas dos dados passados da variável a ser prevista. Dentre os métodos causais estão os métodos de análise de regressão (ARCHER, 1980; ARMSTRONG, 1993).

2.2.1 Análise de Regressão

Este método estatístico estuda a relação das variáveis independentes, também chamadas de preditoras ou variáveis de entrada, com a variável dependente, também chamada de variável de saída ou de interesse. Essa relação é representada por uma equação que associa essas variáveis, a chamada equação de regressão. O modelo pode ser de regressão simples, com uma única variável independente, ou regressão múltipla, quando incorpora mais variáveis independentes. A determinação dos coeficientes das equações pode ser feita pelo método dos mínimos quadrados, entre outros (ARCHER, 1980).

2.2.2 Análise de Séries Temporais

Uma série temporal é um conjunto de observações ordenadas no tempo, sendo a sequência dos dados fundamental para o estudo. A abordagem utiliza o comportamento passado da variável de interesse para fazer previsões sobre o seu futuro. As séries temporais podem ser decompostas em até quatro elementos: tendência, sazonalidade, variações cíclicas e variações irregulares. A tendência é o comportamento da série no longo prazo. A sazonalidade representa variações que ocorrem com periodicidade constante, tipicamente em função das estações do ano. As variações cíclicas são variações com duração superior a um ano que se repetem com certa periodicidade, usualmente resultado de expansão ou recessão da economia, enquanto a variação irregular representa as oscilações aleatórias (MORETTIN e TOLOI, 2006).

A técnica de análise de séries temporais por decomposição não é recomendada quando a série de dados a ser analisada não apresenta comportamento bem definido. Padrões e tendências não sazonais podem ser extrapolados usando um modelo de média móvel ou de suavização. A suposição básica por trás dos métodos de média móvel e suavização é que a série temporal é localmente estacionária com uma média que varia lentamente (NAU, 2020).

A média móvel simples (MMS) e o método de suavização exponencial simples (SES) calculam uma média local de dados no final da série, supondo que essa seja a melhor estimativa do valor médio atual em torno do qual os dados estão flutuando. Esses métodos assumem que a média da série varia lenta e aleatoriamente sem tendências persistentes. Normalmente é preferível a utilização da suavização exponencial simples ao invés da média móvel simples, porque sua média ponderada exponencialmente faz um trabalho mais cuidadoso de descontar os dados mais antigos, e porque o parâmetro de suavização (α) é contínuo e pode ser facilmente otimizado (NAU, 2020).

Os métodos MMS e SES assumem que não há tendência de qualquer tipo nos dados, o que é geralmente bom para previsões com um passo à frente de dados que possuem muita aleatoriedade. Porém, para gerar previsões de curto prazo, quando os dados apresentam alguma tendência e não possuem uma componente sazonal, é indicada a utilização do método de suavização exponencial dupla (SED), também chamado de suavização exponencial linear de Holt. O método SED calcula estimativas locais de nível e tendência suavizando os dados recentes, incluindo duas constantes de suavização, uma para o nível (α) e

⁸ Tarifa de Energia de Otimização (TEO), destinada à cobertura dos custos incrementais de operação e manutenção (O&M) das usinas hidrelétricas e ao pagamento da Compensação Financeira pelo Uso dos Recursos Hídricos (CFURH).

outra para a tendência (β). A qualquer momento i , existe uma estimativa L_i do nível local e uma estimativa T_i da tendência local. Os valores das previsões um passo a frente (\hat{Y}_{i+1}) são computados recursivamente, a partir do valor de Y_i (valor da série original observado no tempo i), e das estimativas anteriores do nível e tendência, por duas equações que aplicam suavização exponencial a eles separadamente (MINITAB, 2020; NAU, 2020).

Na Tabela 2.2 estão listadas as equações empregadas no desenvolvimento da análise de séries temporais utilizados nesse trabalho.

Tabela 2.2: Métodos de análise de séries temporais e suas respectivas equações.

Métodos	Sigla	Equações
Média Móvel Simples	MMS	$\hat{Y}_{i+1} = \frac{Y_i + Y_{i-1} + Y_{i-m+1}}{m} \quad (2.1)$
Suavização Exponencial Simples	SES	$L_i = \alpha Y_i + (1 - \alpha)L_{i-1} \quad (2.2)$
		$\hat{Y}_{i+1} = L_i \quad (2.3)$
Suavização Exponencial Dupla	SED	$L_i = \alpha Y_i + (1 - \alpha)(L_{i-1} + T_{i-1}) \quad (2.4)$
		$T_i = \beta(L_i - L_{i-1}) + (1 - \beta)T_{i-1} \quad (2.5)$
		$\hat{Y}_{i+k} = L_i + k \times T_i \quad (2.6)$

Fonte: Adaptado de (MINITAB, 2020; NAU, 2020).

Com \hat{Y}_i o valor previsto no instante i , Y_i o valor real da série observado no instante i , m o tamanho da média móvel simples, L_i a estimativa de nível no tempo i , α a constante de suavização de nível, T_i a estimativa de tendência, β a constante de suavização de tendência e k o número de períodos à frente de i em que se deseja fazer a previsão.

Quanto maior o valor de α mais rápida é a resposta do modelo em assumir mudanças no comportamento da série, e de forma oposta, quanto menor for seu valor, maior a demora do modelo a assumir mudanças. A interpretação da constante de suavização de tendência β é análoga à da constante de suavização de nível α . Modelos com valores pequenos de β assumem que a tendência muda muito lentamente com o tempo, enquanto modelos com β maior assumem que ela está mudando mais rapidamente (MINITAB, 2020; NAU, 2020).

2.2.3 Coeficientes de Determinação

O coeficiente de determinação, R^2 , mede o ajuste de um modelo estatístico linear generalizado, como a regressão linear, em relação aos valores observados. O R^2 representa a porcentagem de variação na resposta que é explicada pelo modelo, ou seja, quanto mais próximo de 100%, melhor o modelo ajusta os dados utilizadas para fazer a regressão. (MAIA, 2017; MINITAB, 2020).

O valor de R^2 sempre aumenta conforme novas variáveis preditoras são adicionadas ao modelo, mesmo quando não existe uma verdadeira melhoria estatística. Assim, para comparar modelos com diferente número de variáveis de entrada, usa-se o R^2 ajustado, denotado por $R^2 (aj)$, que é a porcentagem de variação na resposta que é explicada pelo modelo, ajustada para o número de preditoras do modelo em relação ao número de observações. Com a incorporação do número de variáveis utilizadas no modelo, o $R^2 (aj)$ subsidia a escolha do modelo mais adequado entre modelos com quantidades diferentes de preditoras (MAIA, 2017; MINITAB, 2020). Na Tabela 2.3 estão representadas as equações para R^2 e $R^2 (aj)$.

Tabela 2.3: Coeficientes de determinação.

Significado	Equação
Coeficiente de determinação	$R^2 = \left(1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \right) \times 100\% \quad (2.7)$
Coeficiente de determinação ajustado	$R^2(aj) = \left(1 - \frac{n-1}{n-(k+1)} \times (1 - R^2) \right) \times 100\% \quad (2.8)$

Fonte: Elaborado pelo Autor com base em (MAIA, 2017).

Com Y_i o valor original da série no período i , \hat{Y}_i o valor estimado pelo modelo no período i , \bar{Y} o valor médio das amostras da série original, k é o número de variáveis independentes utilizadas no modelo de regressão e n é o número de períodos considerados pelo cálculo da medida de erro.

O parâmetro R^2 predito, $R^2(pred)$, estima o quão bem o modelo de regressão prediz as respostas para novas observações. Para o cálculo de $R^2(pred)$ são removidas de forma sistemática cada observação do conjunto de dados, estimada a equação de regressão e então determinando o quão bem o modelo prediz a resposta para a observação removida. Se o valor de $R^2(pred)$ for significativamente inferior ao valor de R^2 isso pode significar que o modelo está com excesso de ajuste (MINITAB, 2020). Assim como para R^2 , os valores de $R^2(aj)$ e $R^2(pred)$ variam entre 0 a 100 %.

2.2.4 Erros

Medidas de erro são utilizadas para identificar a exatidão de um método de previsão ou como critério de seleção de métodos de previsão. A Tabela 2.4 apresenta as medidas de erros selecionadas nesse trabalho, entre aqueles disponíveis na literatura (LEMOS, 2006).

Tabela 2.4: Medidas de erro para os modelos de previsão.

Sigla	Significado	Equação
MAE ⁹	Erro absoluto médio	$MAE = \frac{1}{n} \times \sum_{i=1}^n Y_i - \hat{Y}_i \quad (2.9)$
MAPE ¹⁰	Média dos erros percentuais absolutos	$MAPE = \left(\frac{1}{n} \times \sum_{i=1}^n \left \frac{Y_i - \hat{Y}_i}{Y_i} \right \right) \times 100\% \quad (2.10)$
MSE ¹¹	Erro quadrático médio	$MSE = \frac{1}{n} \times \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (2.11)$

Fonte: Adaptado de Lemos (2006).

As medidas de erro MAE e MAPE são medidas importantes para analisar a acurácia média das previsões em termos absolutos. O MSE é uma medida importante para identificar se há valores que destoam dos erros absolutos médios das previsões.

3 METODOLOGIA DESENVOLVIDA

O presente trabalho usa como hipótese a existência de um sistema causal que exerce influência sobre os valores futuros do PLD. A metodologia de sete etapas apresentadas na Figura 3.1 foi empregada para avaliar as projeções dos valores do PLD para o patamar de carga pesado no submercado Sul. A sigla

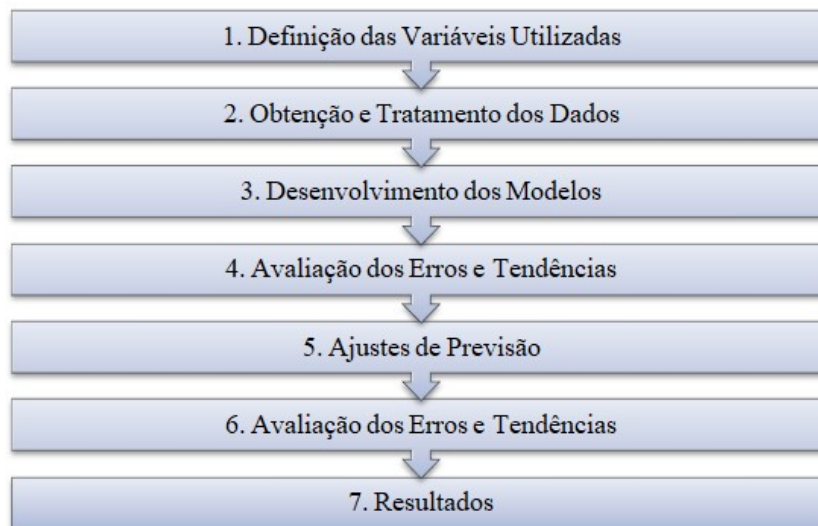
⁹ Do inglês *Mean Absolute Error*.

¹⁰ Do inglês *Mean Absolute Percentage Error*.

¹¹ Do inglês *Mean Squared Error*.

PLD foi usada aqui para referenciar o patamar de carga pesado no submercado Sul com periodicidade semanal.

Figura 3.1: Fluxograma simplificado da metodologia desenvolvida.



Fonte: Autor.

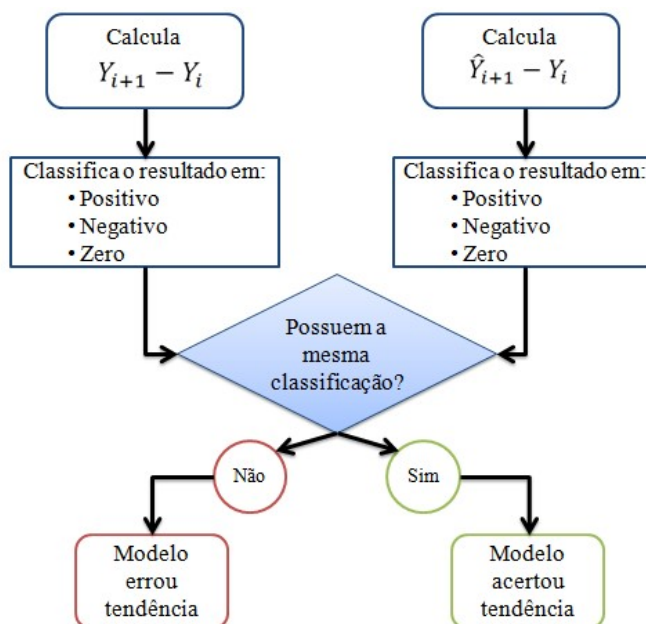
Passo 1 - As variáveis utilizadas pelos programas NEWAVE e DECOMP são escolhidas e analisadas para o desenvolvimento dos modelos. É dada prioridade aos dados disponíveis de maneira pública e que sejam facilmente extraídos para uso, buscando praticidade para a utilização nos modelos desenvolvidos.

Passo 2 – As variáveis são obtidas a partir dos bancos de dados do ONS e da CCEE e é conduzido um pré-processamento dos dados, visando compensar dados faltantes e realizar um ajuste de base temporal. É importante que todas as séries de dados históricos estejam na mesma base temporal e com o mesmo número de elementos, para que não haja erros nas previsões devido a algum eventual descompasso temporal entre as variáveis.

Passo 3 - São desenvolvidos modelos de previsão dos preços a partir das técnicas de regressão linear múltipla e modelos para os métodos de séries temporais apresentados na Tabela 2.2.

Passo 4 - Para cada modelo são calculados os erros apresentados na Tabela 2.4, comparando os valores previstos com os preços divulgados pela CCEE, e o nível de acerto de tendências do modelo. A análise de tendências também é usada por ser um indicador antecipado da direção em que os preços tendem a seguir e assim subsidiar a tomada de decisão e os processos de negociação. O algoritmo de análise de tendência está representado na Figura 3.2.

Figura 3.2: Fluxograma do algoritmo de análise do acerto de tendência.



Fonte: Autor.

Primeiramente calculam-se as diferenças dos valores da série original do PLD da semana $i + 1$ com o valor da semana i , ou seja, é calculado $Y_{i+1} - Y_i$. O resultado dessa diferença irá indicar se o valor do PLD aumentou, diminuiu ou se manteve constante de uma semana para a outra. Esse resultado é então classificado em positivo, negativo ou zero, para aumento ($Y_{i+1} - Y_i > 0$), diminuição ($Y_{i+1} - Y_i < 0$), ou estagnação ($Y_{i+1} - Y_i = 0$) dos valores do PLD de uma semana para outra, respectivamente.

Em paralelo é realizado o mesmo cálculo para a diferença entre os valores previstos pelos modelos para o PLD para a semana $i + 1$ com o valor real do PLD da semana i , ou seja, é calculado $\hat{Y}_{i+1} - Y_i$. A mesma classificação em positivo, negativo ou zero, é feita para esse cálculo.

As classificações dos cálculos das diferenças $\hat{Y}_{i+1} - Y_i$ e $Y_{i+1} - Y_i$, para uma mesma semana são então comparadas. Caso possuam a mesma classificação, o modelo desenvolvido para a previsão acertou a tendência. Caso contrário, considera-se que o modelo errou a tendência de preços. Essa mesma métrica é realizada para todos os resultados da previsão de cada modelo, com i variando de 1 até n . Após realizar essa análise para todos os valores estimados pelo modelo, calcula-se o percentual de acertos.

Passo 5 - As previsões obtidas no Passo 3 são ajustadas para que os valores previstos do PLD não extrapolem os valores teto e piso definidos anualmente pela ANEEL.

Passo 6 - Repete-se o Passo 4 para os modelos ajustados obtidos no Passo 5.

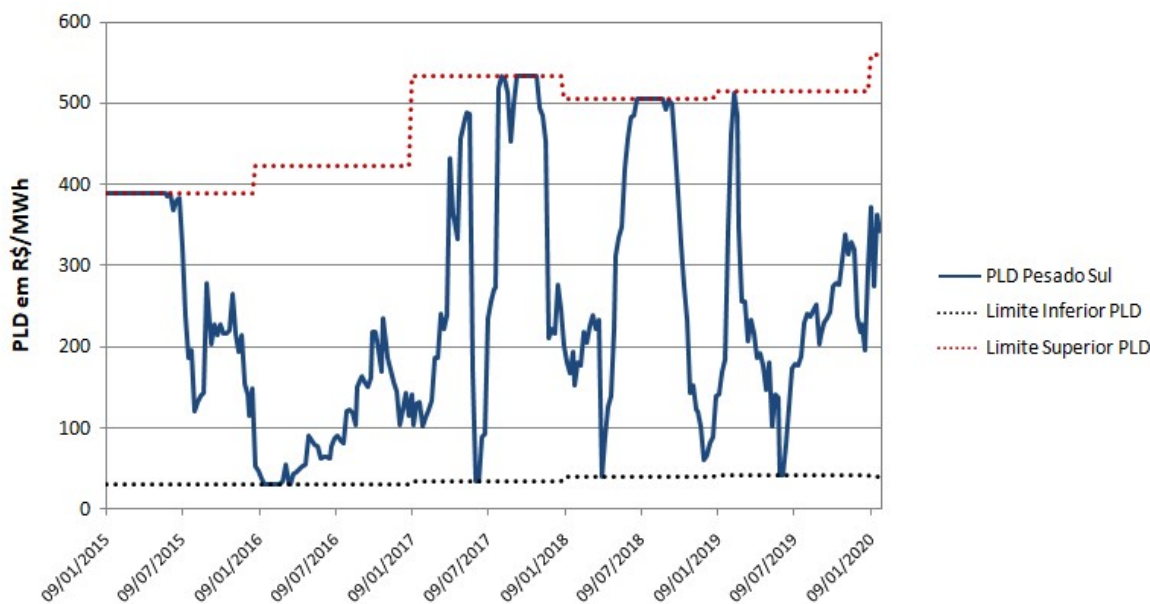
Passo 7 - A partir das medidas de erro, identifica-se qual dentre os modelos desenvolvidos foi o mais eficiente para previsão do PLD no submercado Sul em um horizonte de uma semana.

4 ESTUDO DE CASO E RESULTADOS

Para desenvolver os modelos de previsão dos preços e tendências do PLD foram utilizados dados a partir da segunda semana operativa de 2015 até a última semana operativa de janeiro de 2020. O intervalo de tempo utilizado no presente trabalho é função da mudança significativa na metodologia e nos valores dos limites máximos e mínimos do PLD estipulados anualmente pela ANEEL, conforme a Resolução Normativa 663 de 2014 da ANEEL. Houve nova alteração nesses limites do PLD, que entrou em vigor em 2020, entretanto, os valores dos limites em vigor em 2020 apresentaram menos de 9% de diferença absoluta relativa aos limites praticados em 2019. Devido a pandemia da COVID-19, que resultou em uma redução significativa da carga e, por consequência, em valores atípicos e extremamente baixos do PLD, optou-se por não incluir na análise os dados a partir de fevereiro de 2020, que poderiam mascarar a identificação de alguma característica sazonal, cíclica ou até mesmo de tendência nas séries históricas. A série de dados tratados do PLD usada no presente trabalho está representada na Figura 4.1,

bem como os valores de piso e teto estipulados pela ANEEL, onde as datas representadas no eixo das abscissas representam o último dia da semana operativa na qual vigoraram os valores representados.

Figura 4.1: Série histórica do PLD para o patamar de carga pesado no submercado Sul.



Fonte: Elaborado pelo Autor.

Pode-se observar que o PLD não possui um comportamento bem definido, apresentando grande volatilidade. Em alguns períodos da série os valores acabam apresentando valores iguais de uma semana para a outra devido aos limites inferior e superior definidos pela ANEEL.

As variáveis de entrada selecionadas para o presente estudo de caso foram a Energia Natural Afluente (ENA), Energia Armazenada (EAR), Geração Hidroelétrica (GH), Geração Termelétrica (GT), Geração Eólica (GE), Intercâmbio de Energia entre os submercados Sul e Sudeste (IE) e Demanda de Energia (D), obtidas por meio do ONS¹², e as próprias séries históricas dos valores do PLD, obtidas por meio da CCEE¹³. A influência do Índice Geral de Preços do Mercado (IGPM) na composição do PLD foi investigada por se tratar de um indicador macroeconômico, calculado mensalmente pelo Instituto Brasileiro de Economia da Fundação Getúlio Vargas (FGV IBRE)¹⁴. Adotou-se o uso do IGPM visto que ele é utilizado para ajustar as tarifas da energia elétrica, além de acompanhar as mudanças tanto de valor da moeda quanto da alteração nos preços no cenário nacional.

Conforme o Passo 2 da metodologia, as variáveis de entradas foram compiladas em um mesmo arquivo, em base semanal. O índice econômico IGPM é disponibilizado com frequência mensal, portanto, no presente trabalho foi considerado que para cada semana operativa, dentro de um mesmo mês e ano, esse índice é igual. Também foram feitas exclusões de dados semanais que estavam duplicados ou com a base temporal diferente da semanal. Dados faltantes foram adotados como sendo a média simples entre o dado imediatamente anterior e o dado imediatamente posterior da série em ordem cronológica.

4.1 Modelos de Regressão para o PLD

No Passo 3, os modelos de regressão linear múltipla para realizar previsões dos valores do PLD foram gerados com o auxílio do *software* Minitab®. Antes de desenvolver os modelos de regressão foram analisados os melhores subconjuntos dentre as variáveis de entrada escolhidas. Na Tabela 4.1 estão as siglas e unidades de cada variável utilizada nas regressões.

¹² <http://www.ons.org.br/>, acessado em abril de 2020.

¹³ <https://www.ccee.org.br/>, acessado em abril de 2020.

¹⁴ <https://portalibre.fgv.br/>, acessado em maio de 2020.

Tabela 4.1: Variáveis utilizadas nas ferramentas de regressão e suas respectivas siglas e unidades.

Variável	Sigla	Unidade
Demanda de Energia Elétrica	<i>D</i>	MW _{med}
Energia Armazenada	<i>EA</i>	GWh
Energia Natural Afluente	<i>ENA</i>	MW _{med}
Geração Eólica	<i>GE</i>	MW _{med}
Geração Hidrelétrica	<i>GH</i>	MW _{med}
Geração Térmica	<i>GT</i>	MW _{med}
Índice Geral de Preços do Mercado	<i>I_{IGPM}</i>	Índice acumulado
Intercâmbio de Energia	<i>IE</i>	MW _{med}
Preço de Liquidação das Diferenças	<i>PLD</i>	R\$/MWh

Fonte: Autor.

O indicador econômico IGPM foi utilizado como variável de entrada, considerando que este índice era 100 em janeiro de 1993.

O coeficiente de determinação de predição $R^2(pred)$ foi adotado para determinar quais subconjuntos de variáveis possuem maior efeito explicativo sobre o preço da energia elétrica no mercado *spot* e que resultam no modelo de regressão com melhor efeito preditivo.

Na Tabela 4.2 estão representados os dez melhores subconjuntos de variáveis preditoras, dentre as variáveis analisadas, para o modelo de regressão linear múltipla para o PLD pesado Sul.

Tabela 4.2: Melhores subconjuntos de regressão para o PLD pesado Sul.

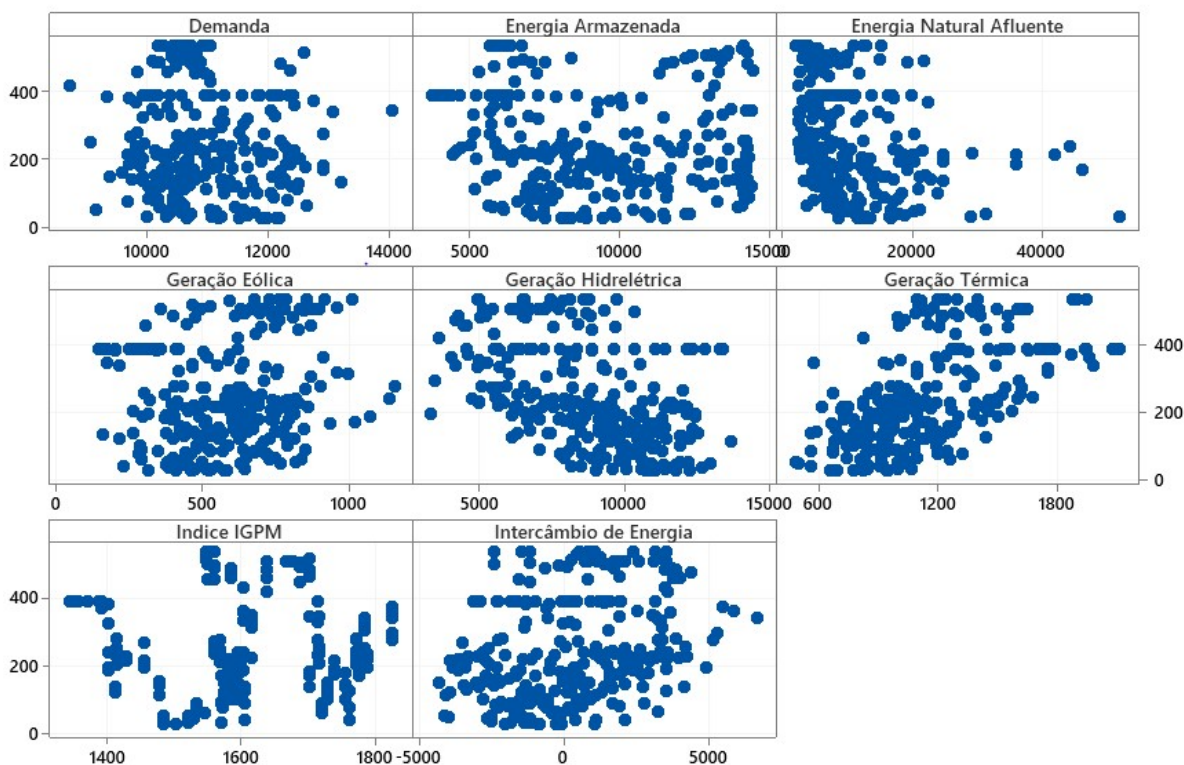
Variáveis de Entrada								Nº de variáveis utilizadas	$R^2(pred)$	$R^2(aj)$	R^2
<i>D</i>	<i>EA</i>	<i>ENA</i>	<i>GE</i>	<i>GH</i>	<i>GT</i>	<i>I_{IGPM}</i>	<i>IE</i>				
<i>X</i>	<i>X</i>		<i>X</i>	<i>X</i>		<i>X</i>	<i>X</i>	6	53,0 %	54,5 %	55,5 %
<i>X</i>	<i>X</i>			<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	6	53,0 %	54,5 %	55,5 %
<i>X</i>	<i>X</i>			<i>X</i>		<i>X</i>	<i>X</i>	5	53,0 %	54,3 %	55,2 %
<i>X</i>	<i>X</i>	<i>X</i>		<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	7	52,8 %	54,5 %	55,7 %
<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>		<i>X</i>	<i>X</i>	7	52,8 %	54,5 %	55,7 %
<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	8	52,5 %	54,4 %	55,8 %
<i>X</i>			<i>X</i>	<i>X</i>		<i>X</i>	<i>X</i>	5	50,3 %	51,5 %	52,4 %
<i>X</i>				<i>X</i>		<i>X</i>	<i>X</i>	4	50,3 %	51,5 %	52,3 %
	<i>X</i>			<i>X</i>	<i>X</i>	<i>X</i>		4	46,5 %	47,6 %	48,4 %
<i>X</i>				<i>X</i>			<i>X</i>	3	43,3 %	44,5 %	45,1 %

Fonte: Autor.

Em cada linha está representado um subconjunto. Quando a coluna correspondente a determinada variável de entrada estiver marcada com um *X* significa que a variável foi utilizada no subconjunto. Os subconjuntos estão representados em ordem decrescente de $R^2(pred)$, ou seja, os conjuntos das linhas superiores possuem melhor capacidade de predição para novas observações com relação aos conjuntos que estão representados nas linhas inferiores. Também estão representados os valores de $R^2(aj)$, que determina o ajuste dos modelos independente do número de variáveis utilizadas, bem como os valores de R^2 e do número de variáveis utilizadas em cada subconjunto.

O maior $R^2(pred)$ obtido foi de 53%, sendo que três subconjuntos apresentaram esse mesmo valor para o coeficiente em questão, sendo dois subconjuntos com seis variáveis e um subconjunto com cinco variáveis preditoras. Entretanto, esse valor de $R^2(pred)$ de 53% não representa um bom efeito de predição para o modelo, assim como os valores de $R^2(aj)$ na faixa de 54% também não representam um bom efeito explicativo sobre a amostra de dados.

Para compreender os motivos de baixos coeficientes de correlação para os modelos de regressão foram analisadas as variáveis de entrada. Na Figura 4.2 estão representados os gráficos do PLD pesado Sul no eixo das ordenadas e as respectivas variáveis preditoras nos eixos das abscissas.

Figura 4.2: Gráficos de dispersão do PLD pesado Sul *versus* variáveis de entradas.

Fonte: Elaborado pelo Autor.

Observa-se uma correlação entre geração hidrelétrica (GH) e geração térmica (GT) com o PLD, sendo a GH uma curva decrescente (valor do PLD diminui com uma maior geração hidrelétrica) e a GT com uma curva crescente (valor do PLD aumenta com o aumento de despacho térmico). As demais variáveis não apresentam uma considerável correlação, pois há dispersão significativa. Todas as variáveis foram representadas na Figura 4.2 com as suas respectivas unidades, conforme indicado na Tabela 4.1.

Na Tabela 4.3 estão representadas as equações de regressão linear múltipla dos três modelos que apresentaram o maior valor de $R^2(pred)$, conforme da Tabela 4.2.

Tabela 4.3: Equações de regressão que correlacionam os melhores subconjuntos de variáveis de entrada com o PLD.

Modelo	Equação
<i>i</i>	$PLD = 619 + 0,1946 \times D + 0,0097 \times EA - 0,0527 \times GE - 0,2079 \times GH - 0,4309 \times I_{IGPM} - 0,1795 \times IE$ (5.1)
<i>ii</i>	$PLD = 598 + 0,1624 \times D + 0,01009 \times EA - 0,1738 \times GH + 0,0453 \times GT - 0,4438 \times I_{IGPM} - 0,1460 \times IE$ (5.2)
<i>iii</i>	$PLD = 652 + 0,1886 \times D + 0,00946 \times EA - 0,1999 \times GH - 0,4728 \times I_{IGPM} - 0,1708 \times IE$ (5.3)

Fonte: Elaborado pelo autor através do *software* Minitab.

Foram utilizados 80% dos dados das séries históricas para montar o modelo de regressão e 20% dos dados visando obter valores de forma não-tendenciosa. Para avaliar as previsões foram usados os valores históricos das variáveis de entrada e então foram comparados os valores obtidos pela regressão com os valores reais do PLD, divulgados pela CCEE. O presente trabalho não tem como objetivo realizar a

previsão das variáveis de entrada utilizadas nas regressões, sendo o foco do trabalho desenvolver um modelo com potencial de aplicabilidade para realizar previsões.

Os modelos de regressão linear múltipla da Tabela 4.3 foram usados para estimar valores do PLD e compará-los com os valores originais disponibilizados pela CCEE, conforme o Passo 4 da metodologia. A Tabela 4.4 apresenta as métricas de erro dos modelos de regressão empregados.

Tabela 4.4: Erros e acertos de tendência para os modelos de regressão.

Modelo	Tipo	Erros das previsões dos modelos			Acerto de tendência
		MAE	MAPE	MSE	
<i>i</i>	Regressão	79,79	58,28 %	9705,20	56,06 %
<i>ii</i>	Regressão	79,62	57,94 %	9704,20	55,68 %
<i>iii</i>	Regressão	79,63	58,63 %	9773,06	56,44 %

Fonte: Autor.

Dentre os modelos de regressão, o modelo *ii* apresentou os menores erros para as previsões dos valores futuros do PLD, enquanto o modelo *iii* apresentou melhor acurácia para o acerto de tendências. No entanto, os três modelos apresentaram valores próximos para todos os parâmetros analisados, com erros calculados elevados e acertos de tendência baixos.

Assim, no Passo 5, foram aplicados limitadores superiores e inferiores, baseados nos valores máximos e mínimos estipulados anualmente pela ANEEL, aos resultados dos modelos. Se o valor previsto pelo modelo para a semana *i* do ano *a* estava acima do teto estabelecido pela ANEEL para esse mesmo ano *a*, então esse valor previsto pelo modelo foi corrigido para o valor teto estipulado para o PLD no ano *a*, e analogamente para o caso de valor inferior ao piso. Essa correção fez com que os modelos finais tenham todos os seus valores previstos situados entre os limites máximos e mínimos estipulados anualmente pela ANEEL.

Os erros de previsão e percentuais de acerto de tendência para os três modelos de regressão com a aplicação de limites inferior e superior aos resultados previstos estão indicados na Tabela 4.5.

Tabela 4.5: Erros e tendências para os modelos de regressão com aplicação de limites superior e inferior.

Modelo	Tipo	Erros das previsões dos modelos			Acerto de tendência
		MAE	MAPE	MSE	
<i>i</i>	Regressão	75,20	56,58 %	9277,70	60,98 %
<i>ii</i>	Regressão	75,06	56,28 %	9278,82	60,61 %
<i>iii</i>	Regressão	75,53	56,98 %	9411,94	61,36 %

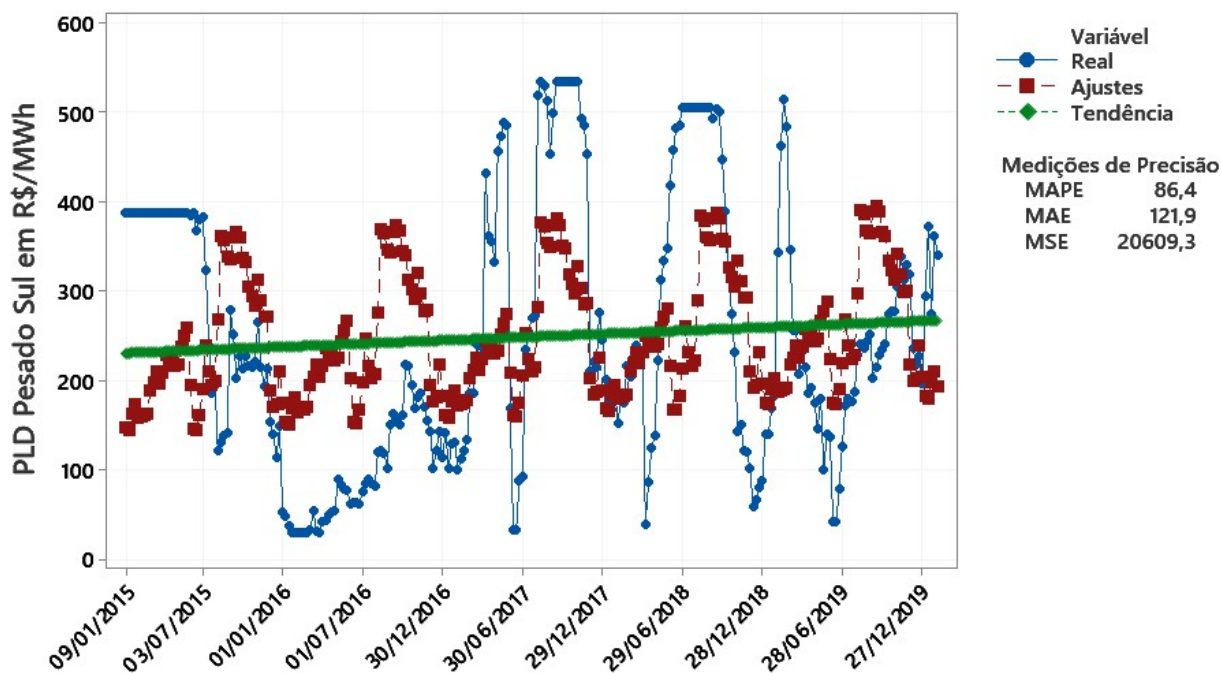
Fonte: Autor.

Após a aplicação dos limites o modelo *ii* novamente apresentou os menores erros (MAE e MAPE) para as previsões dos valores futuros do PLD, assim como o modelo *iii* apresentou 61,36% de acerto de tendências, sendo este o maior percentual de acerto entre os três modelos. Outra vez os três modelos apresentaram valores próximos para todos os parâmetros analisados, sendo que o percentual de acerto das tendências aumentou em torno de 5%, devido ao fato de que em determinadas semanas valores que eventualmente se mantinham constantes e iguais ao limite máximo ou mínimo do PLD estavam sendo estimados pelas regressões de forma a extrapolar esses limites, o que acabava causando erro nas tendências quando comparado com a série original. Os erros percentuais diminuíram um pouco menos de 2% com a aplicação dos limites. Os erros quadráticos médios continuaram com uma ordem de grandeza significativa, o que indica que alguns resultados das previsões diferem de forma significativa dos valores reais.

4.2 Modelos de Séries Temporais para o PLD

O Passo 3 da metodologia foi novamente aplicado à análise de séries temporais. A primeira etapa consistiu em identificar padrões não-aleatórios nos dados históricos do PLD, tais como tendência global, sazonalidade ou flutuações de período variáveis, chamadas de ciclicidade. Para o período analisado, antes da pandemia do COVID-19 e da consequente baixa dos preços do PLD, a tendência global dos preços era levemente crescente, como é possível observar na Figura 4.3.

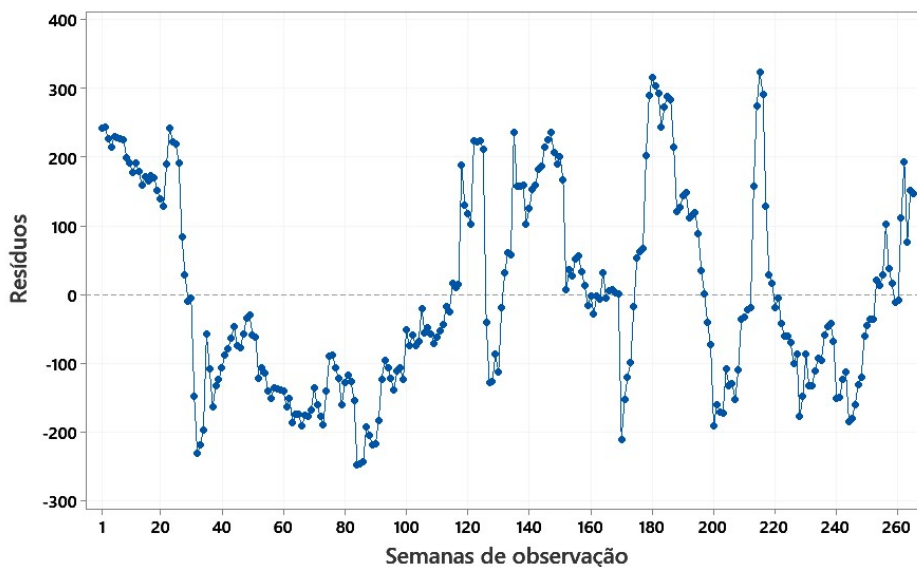
Figura 4.3: Série histórica do PLD pesado Sul e ajuste considerando sazonalidade e de tendência da série.



Fonte: Elaborado pelo Autor.

Ademais, pode ser observada na Figura 4.3 a importante oscilação dos valores do PLD, não apresentando uma característica sazonal importante. Na figura estão apresentados em vermelho os ajustes da série temporal, somando as decomposições de tendência global e de sazonalidade, onde é possível observar que o modelo clássico de séries temporais por decomposição possui diferenças significativas entre os ajustes e os valores reais da série. A sazonalidade foi estimada para um período de 52 semanas, o que representa aproximadamente um ano. Os resíduos entre os valores reais da série e os ajustes por decomposição estão representados no gráfico da Figura 4.4, em R\$/MWh por semana operativa do período analisado em ordem cronológica.

Figura 4.4: Série histórica do PLD pesado Sul e ajuste considerando sazonalidade e de tendência da série.



Fonte: Elaborado pelo Autor.

Pode-se observar um comportamento aleatório dos resíduos ao longo do tempo, o que dificulta uma nova decomposição com base nos resíduos.

Como a série do PLD não apresenta um comportamento bem definido, isso faz com que um método de análise de séries temporais por decomposição não seja indicado para elaborar um modelo de previsão. Desta forma, buscaram-se métodos de análise de séries temporais em que as componentes sazonais e de tendência não tenham pesos importantes no desenvolvimento dos modelos.

Os primeiros modelos de séries temporais desenvolvidos foram modelos de média móvel para suavizar a série e gerar previsões uma semana à frente ($k = 1$). Foram desenvolvidos três modelos usando média móvel: modelo de média móvel simples de tamanho 4 (MMS4), que faz a previsão do valor do PLD para a semana i ser igual a média aritmética das quatro semanas imediatamente anteriores; modelo de média móvel simples de tamanho 3 (MMS3), que faz a previsão do valor do PLD para a semana i ser igual a média aritmética das três semanas imediatamente anteriores; e média móvel simples de tamanho 2 (MMS2), que faz a previsão do valor do PLD para a semana i ser igual a média aritmética das duas semanas imediatamente anteriores.

Apesar de médias móveis mais longas filtrarem de forma mais satisfatória os ruídos, elas são menos sensíveis a mudanças e mais indicadas para séries com bom comportamento sazonal (NAU, 2020). No presente trabalho foram empregadas médias móveis curtas, que dão mais importância ao nível recente da série e menos importância ao fator sazonal, tendo em vista que a série histórica do PLD possui mudanças abruptas sem um comportamento bem definido.

Na Tabela 4.6 estão os erros de previsão e percentuais de acerto de tendência para os três modelos de média móvel.

Tabela 4.6: Erros e acertos de tendências para os modelos de média móvel.

Modelo	Tipo	Erros das previsões dos modelos			Acerto de tendência
		MAE	MAPE	MSE	
<i>iv</i>	MMS4	47,88	33,34 %	5943,29	49,81 %
<i>v</i>	MMS3	41,39	28,85 %	4732,63	50,00 %
<i>vi</i>	MMS2	34,60	23,95 %	3570,35	50,19 %

Fonte: Autor.

Conforme o Passo 5 da metodologia, os resultados obtidos com a aplicação de limites inferior e superior do PLD aos resultados previstos são iguais aos resultados sem aplicação dos limites, o que já era esperado, visto que as médias móveis fazem a média aritmética de resultados da própria série. Eventuais diferenças poderiam ocorrer nos períodos próximos a mudança de ano operativo, momento em que é feito o ajuste dos limites do PLD.

Os menores erros de previsão foram do modelo *vi* que usa média móvel de tamanho 2. Para as estimativas de tendência os três modelos de média móvel apresentaram resultado em torno de 50 % de acertos, o que não pode ser considerado um resultado satisfatório.

Na sequência foi desenvolvido um modelo usando suavização exponencial simples (SES), que suaviza os dados calculando as médias exponencialmente ponderadas das semanas anteriores para realizar a previsão dos valores uma semana à frente. Esse método foi utilizado por ser ideal para obter previsões de curto prazo com dados que não possuem tendência ou comportamento sazonal (NAU, 2020). Para aplicar o algoritmo da suavização exponencial simples foi necessário escolher o valor da constante de suavização de nível, denotada α . O nível é semelhante a uma média móvel das observações. A magnitude da constante α determina a velocidade de resposta do modelo frente a mudanças dos valores da série histórica.

O *software* Minitab possui uma ferramenta que calcula o valor ótimo de α . O valor de α é obtido de forma iterativa, buscando minimizar os erros quadráticos médios ao rodar o modelo para cada valor diferente de α , e, ao final, é obtido um valor ótimo de α , com o qual se estabelece o modelo final. Para a série histórica do PLD pesado Sul o valor ótimo de α obtido foi de 1,12211.

O último modelo desenvolvido usou a técnica de suavização exponencial dupla (SED), que é indicada para dados que possuem uma tendência, mesmo que ela seja pouco significativa, mas que não possuem sazonalidade. Esse modelo leva em conta duas constantes de suavização, a constante de nível (α) e a constante de tendência (β), sendo uma independente da outra. Nesse contexto a tendência é semelhante a uma diferença móvel entre observações consecutivas. Valores mais baixos de β dão menos peso aos dados recentes, de modo que as previsões seguem a tendência geral dos dados, enquanto que valores maiores de β dão mais peso a dados recentes, de modo que as previsões sigam a tendência dos dados mais recente. Foi utilizado o *software* Minitab para rodar esse modelo, o qual calcula os pesos das constantes de maneira iterativa, de forma a minimizar o erro quadrático médio das previsões. Para a série histórica do PLD pesado Sul os pesos ótimos para o modelo de SED foram $\alpha = 1,20731$ e $\beta = 0,01109$.

Os erros de predição e os percentuais de acerto de tendência para os modelos de suavização exponencial simples e de suavização exponencial dupla estão indicados na Tabela 4.7.

Tabela 4.7: Erros e acertos de tendência para os modelos de suavização exponencial.

Modelo	Tipo	Erros das previsões dos modelos			Acerto de tendência
		MAE	MAPE	MSE	
<i>vii</i>	SES	28,30	18,30 %	2524,03	54,54 %
<i>viii</i>	SED	29,85	19,44 %	2574,97	41,67 %

Fonte: Autor.

Os dois modelos de suavização exponencial apresentaram erros de previsões relativamente próximos, tendo o modelo *vii* apresentado erros ligeiramente menores. Com relação às tendências houve uma diferença de 12,87% entre eles, sendo novamente o modelo *vii* o que apresentou melhor resultado.

De acordo com o Passo 5 da metodologia, os valores previstos pelos modelos de suavização exponencial foram então limitados pelos valores máximos e mínimos anuais do PLD. Os resultados da aplicação dos limites aos modelos podem ser conferidos na Tabela 4.8.

Tabela 4.8: Erros e tendências para os modelos de suavização exponencial com aplicação de limites superiores e inferiores.

Modelo	Tipo	Erros das previsões dos modelos			Acerto de tendência
		MAE	MAPE	MSE	
<i>vii</i>	SES	28,01	17,91 %	2510,15	59,85 %
<i>viii</i>	SED	29,01	18,13 %	2536,74	50,38 %

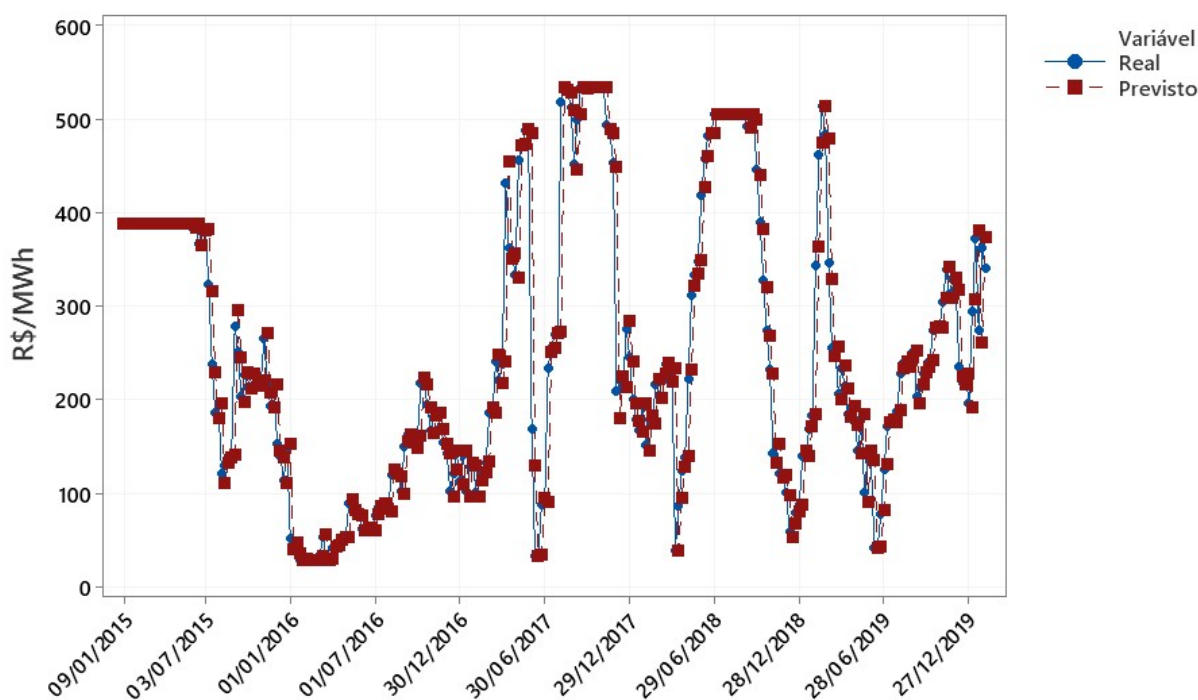
Fonte: Autor.

Os erros de previsão diminuíram levemente para os dois modelos com essa limitação. O acerto de tendências do modelo *vii* aumentou em 5,31% e o acerto de tendências do modelo *viii* aumentou 8,71%. Após aplicação dos limites o modelo *vii* continuou apresentando resultados com maior acurácia em comparação ao modelo *viii*.

5 DISCUSSÃO SOBRE OS RESULTADOS OBTIDOS

Conforme os Passos 4 e 6 da metodologia, a avaliação de erros foi feita para selecionar o método mais apropriado de previsão do valor do PLD uma semana a frente. Foram aplicadas as medidas de acuracidade MAE, MAPE e MSE para mensurar as diferenças entre as estimativas e os dados originais. Dentre os modelos desenvolvidos, o modelo *vii*, que utiliza suavização exponencial simples com aplicação de limites superior e inferior, apresentou os menores erros de previsão para as três medidas avaliadas, com MAE de R\$ 28,01/MWh, MAPE de 17,91 % e MSE de 2510,15. A Figura 5.1 apresenta as previsões do modelo *vii* confrontado com os valores reais do PLD.

Figura 5.1: Valores previstos pelo modelo *vii* com ajuste de limites sobreposto aos valores reais do PLD.



Fonte: Elaborado pelo Autor.

Mesmo apresentando os menores erros de previsão, o modelo *vii* não apresentou os melhores acertos de tendência. Isso ocorre pois os modelos criados a partir de análise de séries temporais acabam sendo mais suscetíveis a errar as tendências de uma semana a frente nos momentos em que a curva do PLD sofre uma mudança no sinal da sua derivada, ou seja, nos momentos em que os valores do PLD apresentam mínimos ou máximos relativos, ou ainda, quando começam ou deixam de apresentar

resultados constantes nos seus valores ao longo das semanas. Isto se deve ao fato de usarem como entrada para as projeções apenas valores passados da própria série do PLD e assim não reconhecerem mudanças de tendências dos preços.

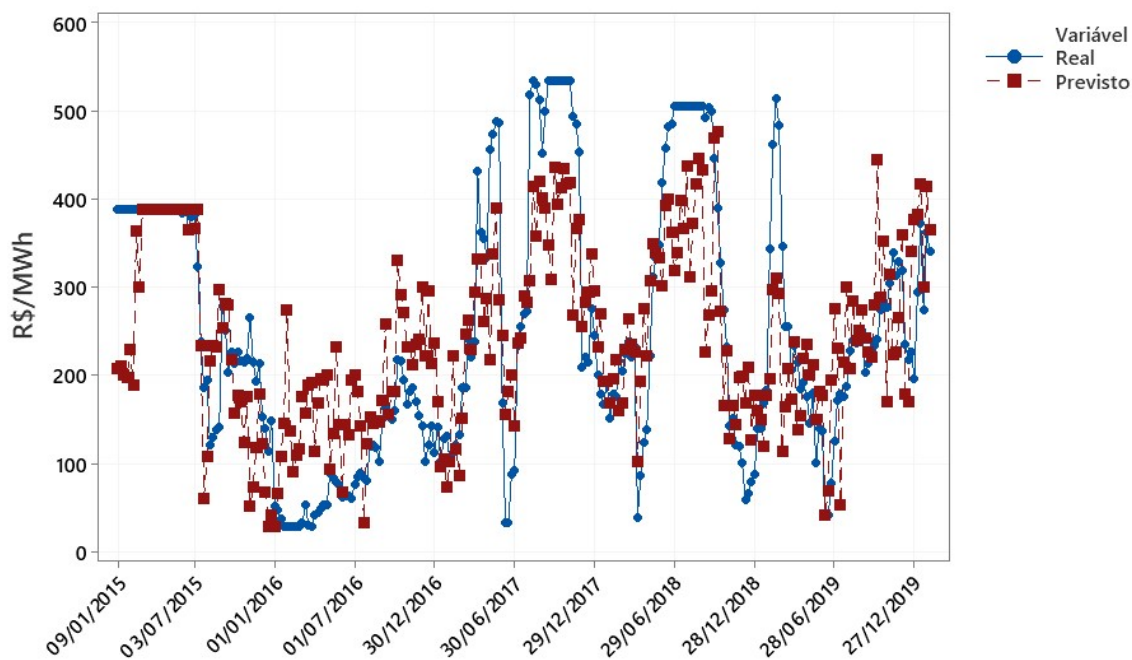
De um modo geral, séries temporais que buscam dar pesos maiores para os valores mais recentes acabam gerando previsões que seguem a tendência global dos últimos valores da série real. No caso analisado nesse trabalho, os modelos de séries históricas apresentaram previsões com medidas de erros menores em comparação aos modelos de regressão, porém erros menores não necessariamente correspondem ao acerto da tendência. Ao limitar os valores previstos pelos modelos de séries temporais com os limites superior e inferior do PLD estipulados pela ANEEL (Passo 5 da metodologia) foi possível melhorar sutilmente o acerto de tendência, devido ao fato de manter constantes eventuais valores que iriam extrapolar os limites. O melhor nível de acerto de tendência entre os modelos de séries temporais também foi do modelo *vii*, com um percentual de 59,85 %.

A tendência global dos valores da série histórica do PLD não foi significativa para previsões, visto que o modelo *viii*, que levou em consideração a tendência global do período, apresentou resultados com erros médios superiores ao modelo *vii*, que não fez uso da tendência global. Deste modo, é reforçada a discussão feita no capítulo 4.2, de que métodos de previsão a partir da decomposição da série histórica não são adequados para realizar previsões do PLD.

Por sua vez, os modelos de regressão desenvolvidos apresentaram erros superiores aos modelos de análise de séries temporais para as previsões uma semana à frente. Regressões lineares simples ou múltiplas não se mostraram eficientes para gerar previsões dos preços futuros devido à complexidade da modelagem da operação do sistema hidro-termo-eólico nacional, ao grande número de variáveis de entrada utilizadas no cálculo do PLD, e principalmente aos diferentes graus de importância e pesos que são atribuídos as variáveis nos programas utilizados pelo CEPTEL que dão origem aos valores do PLD.

Dentre os modelos desenvolvidos e analisados, aqueles baseados em regressão foram os que obtiveram melhores percentuais de acerto de tendências, justamente por considerarem variáveis de entrada relacionadas à operação do SIN. A demanda elétrica do sistema, a geração hídrica e o intercâmbio de energia foram as variáveis que se encontraram nos modelos de regressão com maior potencial preditivo. Dentre os subconjuntos analisados pode-se perceber que mesmo apresentando grandes erros percentuais e absolutos de previsões, os modelos de regressão conseguiram estimar melhor a direção global que os preços devem seguir a partir das variações das variáveis de entrada. O modelo com maior índice de acertos de tendência foi o modelo *iii* com aplicação dos limites, com 61,36 % de acerto das tendências. Esse modelo foi formado a partir da regressão linear múltipla com as variáveis demanda de energia elétrica, energia armazenada, geração hidrelétrica, intercâmbio de energia e com o indicador macroeconômico IGPM. Na Figura 5.2 está o gráfico das previsões do modelo *iii* com aplicação de limites, em conjunto com o gráfico dos valores reais do PLD.

Figura 5.2: Gráfico com os valores previstos pelo modelo *iii* com ajuste de limites sobreposto aos valores reais do PLD.



Fonte: Elaborado pelo Autor.

Ressalta-se também que todos os modelos apresentaram erros quadráticos médios elevados. Mesmo que os erros absolutos e percentuais de um modelo possam ser considerados relativamente baixos, valores de MSE altos é um indicio de que as previsões possuem valores que distinguem de forma significativa dos valores originais, principalmente em períodos em que os preços apresentam valores atípicos em relação às semanas anteriores.

Analisando a série histórica, percebe-se que os preços não tendem a apresentar mudanças muito bruscas de uma semana para a outra. A diferença absoluta dos preços da série real do PLD entre duas semanas consecutivas possui média de R\$ 28,29/MWh e um desvio padrão de R\$ 48,21/MWh. Apesar do desvio padrão ser grande com relação a média absoluta, se em uma determinada semana os valores estiverem próximos do valor piso, pode-se afirmar que é pouco provável que na semana seguinte os valores estarão próximos do valor teto do PLD.

Entretanto, ao calcular as medidas de erros estimando que as previsões dos preços da semana $i + 1$ sejam iguais aos preços da semana i , que é o equivalente a calcular uma média móvel de tamanho 1, obtém-se $MAE = R\$ 28,18/MWh$, $MAPE = R\$ 18,21 \%$ e $MSE = 2565,06$. Observou-se que o erro absoluto médio da projeção é muito próximo da média da diferença absoluta dos preços de duas semanas consecutivas da série real (R\$ 28,29/MWh), considerando que os preços uma semana a frente sejam iguais aos preços da semana anterior (R\$ 28,18/MWh). As medidas de erro para essa média móvel de tamanho 1 foram próximas às medidas de erro para os modelos de suavização exponencial, que foram os dois modelos que apresentaram menores erros de previsão. O modelo desenvolvido a partir de SES com aplicação de limites apresenta erros ligeiramente inferiores do que a simples projeção dos preços da semana seguinte como iguais aos da semana anterior.

Para melhorar o subsídio aos agentes de mercado na avaliação das previsões de tendência do PLD uma semana à frente, foi feita uma nova classificação para as tendências, similar à classificação de tendências apresentada no Passo 4 da metodologia. Na nova classificação os resultados da diferenças dos valores do PLD entre duas semanas consecutivas foram classificados em positivo para $Y_{i+1} - Y_i \geq 25$, negativo para $Y_{i+1} - Y_i \leq -25$, ou zero para $25 > Y_{i+1} - Y_i > -25$. Esses novos limites da classificação também foram aplicados para os valores de $\hat{Y}_{i+1} - Y_i$. Assim, os valores que variaram entre R\$ 25,00/MWh e R\$ - 25,00/MWh entre duas semanas foram considerados com tendência constante, de modo a filtrar, com a análise de tendência, apenas grandes variações entre duas semanas consecutivas. A comparação entre as classificações continuou seguindo a métrica descrita no Passo 4 da metodologia. Os

resultados estão apresentados na Tabela 5.1 e foram feitos apenas com os valores dos modelos no qual foram aplicados os limites inferiores e superiores (após o Passo 5 da metodologia).

Tabela 5.1: Acertos de tendência para os modelos com aplicação de limites superiores e inferiores considerando tendência constante para variações inferiores ao valor absoluto de R\$ 25/MWh.

Modelo	Tipo	Acerto de tendência
<i>i</i>	Regressão	50,76 %
<i>ii</i>	Regressão	52,27 %
<i>iii</i>	Regressão	50,38 %
<i>iv</i>	Média Móvel Simples de Tamanho 4	50,57 %
<i>v</i>	Média Móvel Simples de Tamanho 3	52,67 %
<i>vi</i>	Média Móvel Simples de Tamanho 2	59,70 %
<i>vii</i>	Suavização Exponencial Simples	64,77 %
<i>viii</i>	Suavização Exponencial Dupla	64,77 %

Fonte: Autor.

Os modelos de séries temporais *vii* e *viii* apresentaram sutil melhora nos seus percentuais de acerto de tendência, ambos atingindo 64,77 %, sendo este o melhor percentual entre todos os modelos. Os modelos de regressão tiveram os percentuais de acerto de tendência reduzidos. Essa redução dos percentuais de acerto foi devida aos significativos erros dos modelos de regressão, que antes acertavam o sinal da tendência, mas apresentavam grandes erros nas previsões. Ao restringir uma tendência constante para uma variação de mais ou menos R\$ 25/MWh as previsões desses modelos de regressão acabam não estando dentro desses limites e, portanto, valores que esses modelos deveriam agora considerar constantes, continuam sendo considerados crescentes ou decrescentes.

6 CONCLUSÃO

O presente estudo apresentou o desenvolvimento e a comparação entre diferentes métodos para a previsão dos valores e tendências do preço de liquidação das diferenças semanal para o patamar de carga pesado no submercado Sul para uma semana à frente, através de métodos de análise de regressão linear múltipla e de análise de séries temporais. Para os modelos de regressão foram analisadas oito variáveis independentes: demanda, energia armazenada, energia natural afluenta, geração eólica, geração hidrelétrica, geração térmica, índice macroeconômico IGPM e o intercâmbio de energia entre os submercados Sul e Sudeste.

A seleção dos melhores subconjuntos de regressão para o desenvolvimento de modelos foi feita com base no coeficiente de determinação de predição. Os modelos de previsão desenvolvidos foram comparados através das medidas de erro absoluto médio, média dos erros percentuais absolutos e erro quadrático médio. Dois modelos obtiveram erros absolutos percentuais médios inferiores a 20%. O modelo de previsão que obteve os melhores resultados para os valores futuros do PLD foi o modelo utilizando o método de suavização exponencial simples com a aplicação dos limites máximo e mínimo do PLD definidos anualmente pela ANEEL. Esse modelo apresentou erro médio absoluto de R\$ 28,01/MWh e média dos erros percentuais absolutos de 17,91 % para as previsões uma semana à frente. Esse resultado pode dar um subsídio inicial à tomada de decisão dos agentes de mercado, embora apresente uma margem de erro que pode ser considerada significativa quando analisados grandes montantes de energia. Esses erros são ligeiramente inferiores aos erros obtidos a partir da simples projeção dos preços da semana futura sendo iguais aos preços da semana anterior.

A avaliação das tendências busca compreender o comportamento dos preços de forma ampla, compreendendo qual direção eles tendem a seguir na semana seguinte. Quando a tendência constante considera apenas valores iguais de uma semana para outra, os modelos de análise de regressão obtiveram os melhores resultados, sendo o melhor percentual de acerto de tendências de 61,36 %. Isso mostra que, apesar dos grandes erros de previsão, modelos causais têm potencial para modelar situações futuras baseados no estudo de condições passadas, e podem ser explorados em trabalhos futuros. Ao considerar a tendência constante para uma variação compreendida entre mais ou menos R\$ 25,00/MWh de uma semana para a seguinte, os melhores resultados foram para os modelos de suavização exponencial, ambos obtendo um percentual de acerto de tendências de 64,77 %.

O estudo proposto no presente trabalho comprovou a complexidade de se buscar um modelo alternativo para cálculo do PLD, mostrando que os modelos de séries temporais aqui explorados apresentam resultados com erros significativos para as previsões dos preços uma semana à frente, bem como os modelos de regressão linear múltipla com as variáveis selecionadas. Assim, para trabalhos futuros abrem-se portas para explorar modelos causais que utilizem outros métodos de regressão, que possuam maior capacidade de aprendizado, e a exploração de outros conjuntos de variáveis independentes. Como sugestão para trabalhos complementares também fica o emprego de outras bases temporais, as quais sejam mais intuitivas para a busca de um comportamento sazonal.

REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. **ANEEL estabelece limites do PLD para 2015.** <<http://www.aneel.gov.br/>>. Acesso em: maio 2020a.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. **BIP-Boletim InformativoPF/ANEEL.** <<http://www2.aneel.gov.br/>>. Acesso em: maio 2020b.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. **Editais de Geração.** Disponível em: <<https://www.aneel.gov.br/geracao4>>. Acesso em: 04 abril 2020c.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. **Resolução Normativa nº 109 de 26 de outubro de 2004. Seção: Informações Técnicas.** Disponível em: <<http://www2.aneel.gov.br/cedoc/ren2004109.pdf>>. Acesso em: abril 2020d.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. **Sistema de Informações de Geração da ANEEL (SIGA).** Disponível em: <<http://www2.aneel.gov.br/>>. Acesso em: 23 junho 2020e.

AMARANTE, A. *et al.* Um Comparativo entre o Modelo Autorregressivo Vetorial e o Modelo Computacional Newave-Decomp na Gestão da Previsão do Preço Spot de Energia Elétrica no Brasil. In: **Encontro de Gestão e Negócios**, Uberlândia, MG, 2018, p. 321 – 336.

ARCHER, B. H. Forecasting demand: Quantitative and intuitive techniques. **International Journal of Tourism Management**, v. 1, n. 1, p. 5-12, Mar. 1980.

ARMSTRONG, J. S. Strategic Planning and Forecasting Fundamentals. In: ALBERT, K. **The Strategic Management Handbook**, New York: McGraw-Hill, 1983.

ASSOCIAÇÃO BRASILEIRA DE DISTRIBUIDORES DE ENERGIA ELÉTRICA - ABRADÉE. **Setor Elétrico: Visão Geral do Setor.** Disponível em: <<https://www.abradee.org.br/setor-eletrico/visao-geral-do-setor/>>. Acesso em: 01 abril 2020.

BRASIL. Decreto Nº 5.163, de 30 de julho de 2004. Regulamenta a comercialização de energia elétrica, o processo de outorga de concessões e de autorizações de geração de energia elétrica, e dá outras providências. **Diário Oficial da República Federativa do Brasil.** Poder Executivo, Brasília, DF, 30 de julho de 2004. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2004/Decreto/D5163.htm>. Acesso em: abril 2020.

BRASIL. Decreto Nº 5.177, de 12 de agosto de 2004. Regulamenta os arts. 4º e 5º da Lei Nº 10.848, de 15 de março de 2004, e dispõe sobre a organização, as atribuições e o funcionamento da Câmara de Comercialização de Energia Elétrica - CCEE. **Diário Oficial da República Federativa do Brasil.** Poder Executivo, Brasília, DF, 12 de agosto de 2004. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2004/Decreto/D5177.htm>. Acesso em: abril 2020.

BRASIL. Lei Nº 11.943, de 28 de maio de 2009. Autoriza a União a participar de Fundo de Garantia a Empreendimentos de Energia Elétrica - FGEE; altera o § 4º do art. 1º da Lei no 11.805, de 6 de novembro de 2008; dispõe sobre a utilização do excesso de arrecadação e do superávit financeiro das fontes de recursos existentes no Tesouro Nacional; altera o art. 1º da Lei no 10.841, de 18 de fevereiro de 2004, as Leis nos 9.074, de 7 de julho de 1995, 9.427, de 26 de dezembro de 1996, 10.848, de 15 de março de 2004, 3.890-A, de 25 de abril de 1961, 10.847, de 15 de março de 2004, e 10.438, de 26 de abril de 2002; e autoriza a União a repassar ao Banco Nacional de Desenvolvimento Econômico e Social - BNDES recursos captados junto ao Banco Internacional para a Reconstrução e o Desenvolvimento - BIRD. **Diário Oficial da República Federativa do Brasil.** Poder Executivo, Brasília, DF, 28 de maio de

2009. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Lei/L11943.htm>. Acesso em: abril 2020.

BRASIL. Portaria Nº 514, de 27 de dezembro de 2018. Regulamentar o disposto no art. 15, § 3º, da Lei nº 9.074, de 7 de julho de 1995, com o objetivo de diminuir os limites de carga para contratação de energia elétrica por parte dos consumidores. **Diário Oficial da República Federativa do Brasil**. Poder Executivo, Brasília, DF, 28 de dezembro de 2018. Disponível em: <http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/57219064/do1-2018-12-28-portaria-n-514-de-27-de-dezembro-de-2018-57218754>. Acesso em: abril 2020.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA - CCEE. **Ambiente livre e ambiente regulado**. Disponível em: <<https://www.ccee.org.br/>>. Acesso em: 01 abril 2020a.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA - CCEE. **Comercialização**. Disponível em: <<https://www.ccee.org.br/>>. Acesso em: 07 abril 2020b.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA - CCEE. **Entenda o mercado e a CCEE**. Disponível em: <<https://www.ccee.org.br/>>. Acesso em: 07 abril 2020c.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA - CCEE. **Preço Horário Sombra**. Disponível em: <<https://www.ccee.org.br/>>. Acesso em: 09 abril 2020d.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA (CCEE). **Preços**. Disponível em: <<http://www.ccee.org.br/>>. Acesso em: março 2020e.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA (CCEE). **Regras de Comercialização – Preço de Liquidação das Diferenças - versão 2020.2.0**. Disponível em: <<https://www.ccee.org.br/>>. Acesso em: 29março de 2020f.

CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA - CCEE. **Setor elétrico**. Disponível em: <<https://www.ccee.org.br/>>. Acesso em: 07 abril 2020g.

CARVALHO, F. M. J.; SARDINHA, H. I. **Modelo de Regressão Dinâmica para Projeções de Preço da Energia no Brasil**. 2011. Projeto de Graduação (Engenharia Elétrica) – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2011.

CENTRO DE PESQUISAS DE ENERGIA ELÉTRICA (CEPEL). **DECOMP - Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Curto Prazo**. Seção: Produtos. Disponível em: <<http://www.cepel.br/>>. Acesso em: 08 abril 2020a.

CENTRO DE PESQUISAS DE ENERGIA ELÉTRICA (CEPEL). **Documentação Técnica das metodologias e modelos de otimização energética do CEPEL**. Seção: Notícias. Disponível em: <<http://www.cepel.br/>>. Acesso em: 07 abril 2020b.

CONFEDERAÇÃO NACIONAL DA INDÚSTRIA - CNI. **A evolução do Setor Elétrico Brasileiro rumo à Sustentabilidade**. In: Fórum de Meio Ambiente do Setor Elétrico, 2017. Disponível em: <https://bucket-gw-cni-static-cms-si.s3.amazonaws.com/media/filer_public/09/de/09de36cb-4f51-4da3-82be-cb9f292269b0/fmase.pdf>. Acesso em: 04 abril 2020.

COELHO, I. **Estudo sobre a aplicação de Redes Neurais Recorrentes para previsão da geração eólica e do Preço de Liquidação das Diferenças**. 2018. Trabalho de Conclusão de Curso (Engenharia Elétrica) – Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2018.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). **Balanco Energético Nacional**. 2019. Disponível em: <<http://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2019>>. Acesso em: março de 2020.

FULGENCIO, P.C. **Glossário VadeMecum: administração pública, ciências contábeis, direito, economia, meio ambiente: 14.000 termos e definições**. Mauad X, 2007.

LEITE, A. L. S.; CASTRO, N. J.; TIMPONI, R. **Causas da volatilidade do preço spot de eletricidade no Brasil**. Ensaio FEE (Impresso) , v. 34, 2013, p. 647-668-668.

LEMOS, F. O. **Metodologia para Seleção de Métodos de Previsão De Demanda**. 2006. Dissertação de Mestrado (Engenharia de Produção) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2006.

MAIA, V. M. *et. al.* Água: Único Fator a Influenciar o Preço da Energia no Mercado Spot? **Sociedade, Contabilidade e Gestão**, v. 11, n. 1, p. 29 – 45, Jan.2016.

MAIA, A. G. **Econometria: Conceitos e Aplicações**. Saint Paul, 1ª edição, 2017.

MINISTÉRIO DE MINAS E ENERGIA - MME. **Implantação do DESSEM nas atividades de programação da operação e formação de preço horário – METODOLOGIA**. Disponível em: <<http://www.mme.gov.br/>>. Acesso em: 07 abril 2020.

MINITAB. **Suporte ao Minitab 19**. Disponível em: <<https://support.minitab.com/pt-br/minitab/19/>> Acesso em: 15 maio 2020.

MORETTIN, P. A., TOLOI, C. M. C. **Análise de séries temporais**. Edgard Blucher, 2ª edição, 2006.

NAU, R. **Statistical forecasting: notes on regression and time series analysis**. Duke University. Disponível em: <<http://people.duke.edu/~rnau/411home.htm>>. Acesso em: mai. 2020.

OPERADOR NACIONAL DO SISTEMA ELÉTRICO - ONS. **O Sistema Interligado Nacional**. Disponível em: <<http://www.ons.org.br/paginas/sobre-o-sin/o-que-e-o-sin>>. Acesso em: 07 abril 2020.

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

SILVA, A. P. S. **Previsão do preço de liquidação das diferenças por meio de redes neurais artificiais**. 2018. Trabalho de Conclusão de Curso (Engenharia Elétrica) – Instituto de Ciências Exatas e Aplicadas, Universidade Federal de Ouro Preto, João Monlevade – MG, 2018.

SOUSA, A. J. S. **Análise e Previsão da Volatilidade do Preço de Liquidação das Diferenças no Mercado Brasileiro Utilizando o Modelo GARCH**. 2013. Dissertação – (Mestrado em Engenharia). Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal da Bahia, Salvador, 2013.