



UNIVERSIDADE FEDERAL DO PARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Josivan Rodrigues dos Reis
Número: TD 07/2025

**Metodologia de Previsão de Cenários do Preço
da Curva *Forward* de Energia Elétrica nos
Horizontes de Médio e Longo Prazo**

Belém
2025

Josivan Rodrigues dos Reis
Número: TD 07/2025

**Metodologia de Previsão de Cenários do Preço da Curva
Forward de Energia Elétrica nos Horizontes de Médio e
Longo Prazo**

Metodologia de Previsão de Cenários do Preço da Curva *Forward* de Energia Elétrica nos Horizontes de Médio e Longo prazo apresentado ao Programa de Pós-Graduação em Engenharia Elétrica como requisito parcial para obtenção do grau de Doutorado em Engenharia Elétrica.

Universidade Federal do Pará

Orientador: Maria Emília de Lima Tostes

Coorientador: Flávia Pessoa Monteiro

Belém

2025

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

R375m Reis, Josivan Rodrigues dos.
 Metodologia de previsão de cenários do preço da curva forward
 de energia elétrica nos horizontes de médio e longo prazo / Josivan
 Rodrigues dos Reis, . — 2025.
 CXXXV, 135 f. : il. color.

 Orientador(a): Prof^ª. Dra. Maria Emília de Lima Tostes
 Coorientação: Prof^ª. Dra. Flávia Pessoa Monteiro
 Tese (Doutorado) - Universidade Federal do Pará, Instituto de
 Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica,
 Belém, 2025.

 1. Curva Forward. 2. Preço de Energia Elétrica. 3. Redes
 Bayesiana Dinâmica. 4. Algoritmo Genético. I. Título.

CDD 621.310285



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

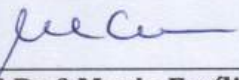
**“METODOLOGIA DE PREVISÃO DE CENÁRIOS DO PREÇO DA CURVA
FORWARD DE ENERGIA NOS HORIZONTES DE MÉDIO E LONGO PRAZO”**

AUTOR: JOSIVAN RODRIGUES DOS REIS

TESE DE DOUTORADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA ELÉTRICA NA ÁREA DE SISTEMAS DE ENERGIA ELÉTRICA.

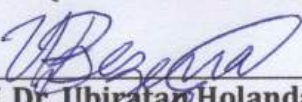
APROVADA EM: 02/04/2025

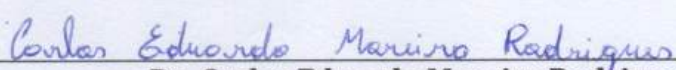
BANCA EXAMINADORA:


Prof.ª Dr.ª Maria Emília de Lima Tostes
(Orientadora - PPGE/ITEC/UFPA)

Documento assinado digitalmente
gov.br FLAVIA PESSOA MONTEIRO
Data: 05/04/2025 12:13:36-0300
Verifique em <https://validar.iti.gov.br>

Prof.ª Dr.ª Flávia Pessoa Monteiro
(Coorientadora - UFOPA)


Prof. Dr. Ubiratan Holanda Bezerra
(Avaliador Interno - PPGE/ITEC/UFPA)


Dr. Carlos Eduardo Moreira Rodrigues
(Avaliador Externo - INPE)

ALVARO AFONSO FURTADO
LEITE:39689603604

Assinado de forma digital por ALVARO
AFONSO FURTADO LEITE:39689603604
Dados: 2025.04.08 11:48:36 -03'00'

Dr. Álvaro Afonso Furtado Leite
(Avaliador Externo - CENERGEL)

Documento assinado digitalmente
gov.br MARCIO JOSE MOUTINHO DA PONTE
Data: 07/04/2025 17:25:46-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Márcio José Moutinho da Ponte
(Avaliador Externo - UFOPA)

VISTO:

Prof. Dr. Diego Lisboa Cardoso
(Coordenador do PPGE/ITEC/UFPA)

Agradecimentos

Chegar até aqui foi uma jornada desafiadora, repleta de aprendizados, obstáculos e conquistas. Este doutorado não é apenas o resultado do meu esforço individual, mas também do apoio e incentivo de muitas pessoas especiais, às quais expresso minha profunda gratidão.

Em primeiro lugar, agradeço a Deus, que me deu força, sabedoria e perseverança para enfrentar os desafios desta caminhada.

À minha orientadora Maria Emília e coorientadora Flávia Monteiro, por sua dedicação, paciência e por compartilharem seus conhecimentos. Suas orientações foram essenciais para o desenvolvimento deste trabalho.

Aos meus pais, pelo amor incondicional, pelo apoio e pelos valores que me transmitiram, fundamentais para que eu chegasse até aqui. Suas palavras de incentivo e sua presença constante foram meu alicerce em todos os momentos.

À minha esposa, pelo companheirismo, compreensão e apoio inabalável. Obrigado por estar ao meu lado nos momentos de desafio, por me incentivar a seguir em frente e por ser minha fonte de equilíbrio e motivação. À minha filha Lívia Reis, que com seu sorriso e sua pureza trouxe leveza e alegria aos meus dias.

Aos familiares e amigos, que sempre acreditaram em mim e ofereceram palavras de encorajamento ao longo dessa jornada.

Aos pesquisadores envolvidos no projeto P&D e colegas de trabalho, pelo compartilhamento de experiências, pelo suporte mútuo, conversas que trouxeram alívio nos momentos de tensão, e pelo ambiente colaborativo que contribuiu para meu crescimento profissional e acadêmico.

A todos que, de alguma forma, contribuíram para que eu chegasse até aqui, meu mais sincero agradecimento.

*A vida é uma peça de teatro que não permite ensaios...
Por isso, cante, ria, dance, chore e viva intensamente cada momento de
sua vida, antes que a cortina se feche e a peça termine sem aplausos...*
(—CHARLES CHAPLIN)

Resumo

No mercado de energia elétrica, a problemática dos preços de negociação dos contratos entre geradores/comercializadores e compradores toma especial relevância, pois uma modelagem assertiva dos contratos leva a um incremento do retorno financeiro e da sustentabilidade do negócio para os vários agentes participantes. No caso brasileiro, os preços de energia são publicados com periodicidade semanal pela Câmara de Comercialização de Energia, por meio da execução dos modelos computacionais NEWAVE, DECOMP E DESSEM. O principal resultado desses modelos é o Preço de Liquidação das Diferenças (PLD), que serve como referência oficial de preços no mercado de energia e é utilizado para a liquidação das sobras e déficits dos agentes nos ambientes de contratação livre e regulada. Porém, as variáveis consideradas no cálculo do PLD, que em sua maioria, buscam refletir o custo da operação do sistema nos horizontes de curto, médio e longo prazo - não refletem de maneira assertiva as expectativas do mercado e nem incorporam o prêmio do risco intrínseco ao processo de comercialização de energia. Esse desalinhamento é mais evidente nos horizontes de médio e longo prazo, que são menos influenciados por variáveis operacionais e mais sensíveis a fatores que modelam o risco associado ao negócio. Diante desse desacoplamento nos horizontes de médio e longo prazo, faz-se necessário o estabelecimento de metodologias que retornem projeções de cenários de curvas de preços que refletem não somente aspectos intrínsecos à operação da rede, mas também o risco associado à comercialização, de forma a otimizar a modelagem dos contratos firmados entre os agentes. Nesse contexto, buscou-se uma abordagem metodológica mais assertiva para proporcionar o dinamismo presente no mercado de energia, aliada a uma modelagem que foca na robustez e na qualidade de experiência do usuário, possibilitando projeções confiáveis do preço de energia nos horizontes de médio e longo prazo (3 a 10 anos). Como proposta metodológica, adotou-se a técnica estatística de Redes Bayesianas Dinâmicas (RBD), para a geração de cenários de previsão, combinada com Algoritmos Genéticos para otimizar automaticamente o desempenho do modelo, definindo as topologias utilizadas pela RBD. Para definir quais as variáveis são relevantes na criação da topologia, realizou-se um estudo detalhado das variáveis para identificar quais dessas impactam na formação de preço de energia. Por fim, os resultados obtidos com a metodologia proposta, foram comparados com resultados gerados pelas técnicas de Regressão Linear, *Support Vector Regression* e *Extreme Gradient Boosting*, utilizando métricas de avaliação de desempenho.

Palavras-chave: Curva Forward, Preço de Energia Elétrica, Redes Bayesiana Dinâmica, Algoritmo Genético.

Abstract

In the electricity market, the issue of contract price negotiation between generators, traders, and buyers is of particular importance, as an accurate contract modeling approach leads to increased financial returns and enhanced business sustainability for the various participating agents. In the Brazilian case, electricity prices are published weekly by the Chamber of Electric Energy Commercialization (CCEE) through the execution of computational models such as NEWAVE, DECOMP, and DESSEM. The primary output of these models is the Preço de Liquidação das Diferenças (PLD), which serves as the official price reference in the electricity market and is used to settle surpluses and deficits of market agents in both the regulated and free contracting environments. However, the variables considered in the PLD calculation—primarily aimed at reflecting the system's operational costs across short-, medium-, and long-term horizons—do not accurately capture market expectations, nor do they incorporate the risk premium intrinsic to the energy trading process. This misalignment is more pronounced in medium- and long-term horizons, where prices are less influenced by operational variables and more sensitive to factors modeling the business-related risks. Given this decoupling in medium- and long-term horizons, it becomes essential to establish methodologies that generate price curve scenario projections that account not only for the intrinsic aspects of grid operation but also for the risks associated with energy trading. This would enable the optimization of contract modeling among market agents. In this context, a more precise methodological approach was sought to incorporate the dynamic nature of the energy market while ensuring a robust and user-experience-oriented modeling process, allowing for reliable electricity price projections over medium- and long-term horizons (3 to 10 years). As a methodological proposal, the statistical technique of Dynamic Bayesian Networks (DBN) was adopted for scenario forecasting, combined with Genetic Algorithms to automatically optimize model performance by defining the topologies used by the DBN. To determine the most relevant variables for topology creation, a detailed study was conducted to identify which factors most significantly impact electricity price formation. Finally, the results obtained from the proposed methodology were compared with those generated by traditional techniques, such as Linear Regression, Support Vector Regression, and Extreme Gradient Boosting, using performance evaluation metrics to assess their effectiveness.

Keywords: Forward Curva, Energy Price, Dynamic Bayesian Network, Genetic Algorithm.

Lista de ilustrações

Figura 1.	Metodologia para estudo de anterioridade baseada na declaração PRISMA.	19
Figura 2.	Publicações envolvendo <i>electricity power forecast</i> nos últimos anos. . .	20
Figura 3.	Distribuição das publicações envolvendo <i>electricity power forecast</i> . . .	21
Figura 4.	Mapa temático de palavras-chave separado por categorias relevantes referente <i>electricity power forecast</i>	22
Figura 5.	Mapa temático de palavras chaves separado por ano de publicação referente a <i>electricity power forecast</i>	23
Figura 6.	Publicações envolvendo <i>electricity forward curve</i> nos últimos anos. . .	24
Figura 7.	Distribuição das publicações envolvendo <i>electricity forward curve</i> . . .	25
Figura 8.	Mapa temático de palavras-chave separado por categorias relevantes referente <i>electricity forward curve</i>	25
Figura 9.	Mapa temático de palavras chaves separado por ano de publicação referente <i>electricity forward curve</i>	26
Figura 10.	Estrutura do Setor Elétrico Brasileiro	39
Figura 11.	Metodologia e horizonte que determina o Preço de Liquidação das Diferenças	44
Figura 12.	Regiões dos Submercados de Energia Elétrica no Brasil	46
Figura 13.	Produtos do mercado de energia elétrica e diferentes tempos de maturação	49
Figura 14.	Sistematização das metodologias de previsão do preço <i>spot</i>	51
Figura 15.	Representação de relação hierárquica da Rede Bayesiana.	52
Figura 16.	Tipo de conexões entre variáveis: a) conexão Linear, b) conexão Divergente e c) conexão Convergente.	53
Figura 17.	Redes Bayesianas como Gráficos Direcionados Acíclicos.	54
Figura 18.	Exemplo da relação temporal entre os nós da rede bayesiana dinâmica.	58
Figura 19.	Operador de <i>crossover</i> entre indivíduo binários.	60
Figura 20.	Operador de mutação binário entre indivíduo binário.	61
Figura 21.	Fluxograma do algoritmo genético.	63
Figura 22.	Exemplo de matriz de correlação aplicada à análise de risco em mercados financeiros.	67
Figura 23.	Variáveis prospectadas pelo estudo de correlação.	68
Figura 24.	Fluxograma da Arquitetura proposta para a previsão da curva <i>forward</i> de Preço de energia elétrica.	71
Figura 25.	Módulo de Aquisição dos dados de diversas fontes.	72
Figura 26.	Módulo de análise e seleção dos dados: avalia as variáveis que mais contribuem na modelagem do problema.	74
Figura 27.	Módulo de previsão de Energia Elétrica	76

Figura 28.	É demonstrado as etapas detalhas do Módulo de previsão de energia. .	77
Figura 29.	Esquema visual da metodologia definida da representação da topologia em cromossomo no Algoritmo Genético.	78
Figura 30.	Fluxograma das etapas do Algoritmo Genético.	80
Figura 31.	Esquema visual da Discretização dos Dados utilizando método <i>K-means</i> .	82
Figura 32.	Esquema visual da modificação na inferência de uma variável, dado a discretização dos dados.	83
Figura 33.	Curvas forward de médio (azul) e longo (vermelho) prazo, obtidas por meio da empresa de consultoria DCide.	86
Figura 34.	Consumo por setor, segundo o Balanço Energético Nacional.	87
Figura 35.	Os principais setores econômicos para a composição PIB brasileiro. . .	88
Figura 36.	Série histórica do Produto Interno Bruto (PIB) entre os anos 2012 a 2023.	89
Figura 37.	Curva <i>forward</i> de médio prazo e o PLD no horizonte de 2012 a 2023. .	91
Figura 38.	Curva <i>forward</i> de Longo prazo e o PLD no horizonte de 2012 a 2023. .	91
Figura 39.	Comparação da Curva <i>forward</i> de Médio e Longo prazo e o Custo Marginal de Operação nos anos de 2012 a 2023.	92
Figura 40.	Carga Total do SIN e nos submercados em MWmed, no horizonte de janeiro de 2012 a dezembro 2023.	93
Figura 41.	Geração de energia do SIN (total) e por fonte de energia, no horizonte de janeiro de 2012 a dezembro 2023.	94
Figura 42.	Geração de energia por submercado (Sudeste/Centro-Oeste, Sul, Nordeste e Norte), no horizonte de janeiro de 2012 a dezembro 2023. . . .	97
Figura 43.	Energia Armazenada em MWmês do SIN e por subsistema, no horizonte de janeiro de 2012 a dezembro 2023.	98
Figura 44.	Capacidade Instalada por tipo de geração no período de 2012 a 2023. .	100
Figura 45.	Crescimento da capacidade instalada das fontes de geração distribuída de 2013 a 2022.	101
Figura 46.	Exemplos de topologias geradas automaticamente pelo Algoritmo Genético	104
Figura 47.	Variáveis disponíveis para os experimentos pelos métodos de regressão.	106
Figura 48.	Resultado da avaliação do fator de importância entre as variáveis de entrada com o preço no longo prazo fornecido pela DCide	108
Figura 49.	Variáveis selecionadas como entrada para os experimentos dos métodos de regressão.	108
Figura 50.	Série histórica do banco de dados: Divisão entre conjunto de treinamento e teste para cenários de projeções de 6 anos.	109
Figura 51.	Série histórica do banco de dados: Divisão entre conjunto de treinamento e teste para cenários de projeções de 3 anos.	109

Figura 52.	Resultados do Cenário 1: (a) RBD sem inferência de informações externa ao modelo comparando com os métodos: XGBoost; (b) <i>Support Vector Regression</i> e (c) Regressão Linear.	110
Figura 53.	Resultados do Cenário 2: (a) RBD com inferência de informações externas na variável Energia Armazenada comparando com os métodos: XGBoost; (b) <i>Support Vector Regression</i> e (c) Regressão Linear. . . .	112
Figura 54.	Resultados do Cenário 3: (a) RBD com inferência de informações externas nas variáveis Energia Armazenada e o Preço da Liquidação das Diferenças comparando com os métodos: XGBoost; (b) <i>Support Vector Regression</i> e (c) Regressão Linear.	114
Figura 55.	Resultados do Cenário 4: (a) RBD sem inferência de informações externas ao modelo comparando com os métodos: XGBoost; (b) <i>Support Vector Regression</i> e (c) Regressão Linear no horizonte de 3 anos. . . .	116
Figura 56.	Resultados do Cenário 5: (a) RBD com inferência de informações externas na variável Energia Armazenada comparando com os métodos: XGBoost; (b) <i>Support Vector Regression</i> e (c) Regressão Linear no horizonte de 3 anos.	117
Figura 57.	Resultados do Cenário 6: (a) RBD com inferência de informações externas nas variáveis Energia Armazenada e o Preço de Liquidação das Diferenças comparando com os método: XGBoost; (b) <i>Support Vector Regression</i> e (c) Regressão Linear no horizonte de 3 anos.	118
Figura 58.	Resultados dos Cenários de Previsão do Preço da Curva <i>Forward</i> no horizonte de 3 anos.	120
Figura 59.	Resultados dos Cenários de Previsão do Preço da Curva <i>Forward</i> no horizonte de 6 anos.	121
Figura 60.	Resultados dos Cenários de Previsão do Preço da Curva <i>Forward</i> no horizonte de 4 anos.	122
Figura 61.	Resultados dos Cenários de Previsão do Preço da Curva <i>Forward</i> no horizonte de 4 anos, com inferências de informações externas na variável Energia Armazenada.	123

Lista de tabelas

Tabela 1.	Resultado da correlação entre as variáveis PIB e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	89
Tabela 2.	Resultado da correlação entre as variáveis Preço da Liquidação das Diferenças e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	92
Tabela 3.	Resultado da correlação entre as variáveis Custo Marginal de Operação e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo. . .	93
Tabela 4.	Resultado da correlação entre as variáveis Carga total do SIN e no submercado Sudeste/Centro-Oeste e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	94
Tabela 5.	Resultado da correlação entre as variáveis Geração total no SIN e por tipo de usinas (Eólica, Hidrelétrica, Nuclear, Solar e Térmica) e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	95
Tabela 6.	Resultado da correlação entre as variáveis capacidade de Geração de energia por submercado e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	97
Tabela 7.	Resultado da correlação entre as variáveis energia armazenada total no SIN e por submercado e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	99
Tabela 8.	Resultado da correlação entre as variáveis capacidade instalada por tipo de geração e o Preço da Curva <i>Forward</i> de energia para Médio e Longo Prazo.	100
Tabela 9.	Parâmetros de configuração do Algoritmo Genético	102
Tabela 10.	Parâmetros de configuração utilizado no método de Regressão Linear .	106
Tabela 11.	Parâmetros de configuração utilizado no método de <i>Support Vector Regression</i>	106
Tabela 12.	Parâmetros de configuração utilizado no método de <i>Extreme Gradient Boosting</i>	107
Tabela 13.	Tabela com a taxa MAPE dos experimentos entre a Rede Bayesiana Dinâmica e os métodos: XGBoost; <i>Support Vector Regression</i> e Regressão Linear nos horizontes de 3 e 6 anos	119
Tabela 14.	Tabela com a taxa MAPE dos Experimentos de Cenários sem inferência e com inferências, na variável Energia Armazenada no horizonte de 4 anos.	124

Lista de abreviaturas e siglas

- (ACL) Ambiente de Contratação Livre
- (ACR) Ambiente de Contratação Regulado
- (AG) Algoritmo Genético
- (API) Interface de Programação de Aplicações
- (ANEEL) Agência Nacional de Energia Elétrica
- (BEN) Balanço Energético Nacional
- (CCEAR) Contratos de Comercialização de Energia no Ambiente Regulado
- (CCEE) Câmara de Comercialização de Energia Elétrica
- (CCEN) Contratos de Cotas de Energia Nuclear
- (CCGF) Contratos de Cota de Garantia Física
- (CEAMAZON) Centro de Excelência em Eficiência Energética da Amazônia
- (CMO) Custo Marginal de Operação
- (CMSE) Conselho de Monitoramento do Setor Elétrico
- (CNPE) Conselho Nacional de Política Energética
- (CPDs) Tabela de Probabilidades Condicionais
- (DAG) Grafo Acíclico Dirigido
- (EPE) Empresa de Pesquisa Energética
- (IBGE) Instituto Brasileiro de Geografia e Estatística
- (IPEA) Instituto de Pesquisa Econômica Aplicada
- (MAPE) Mean Absolute Percentage Error
- (MME) Ministério de Minas e Energia
- (N) Norte

(NE) Nordeste

(ONS) Operador Nacional do Sistema Elétrico

(PCA) Análise de Componentes Principais

(PIB) Produto Interno Bruto

(PLD) Preço de Liquidação das Diferenças

(PRISMA) *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*

(RB) Redes Bayesianas

(RBD) Redes Bayesianas Dinâmicas

(S) Sul

(SE/SO) Sudeste e Centro-Oeste

(SEB) Sistema Elétrico Brasileiro

(SIN) Sistema Interligado Nacional

(SVR) Support Vector Regression

(UFOPA) Universidade Federal do Oeste do Pará

(UFPA) Universidade Federal do Pará

Sumário

1	CAPÍTULO 1	16
1.1	Introdução	16
1.2	Motivação	16
1.3	Objetivos	17
1.4	Organização da Tese	18
1.5	Revisão Bibliográfica	18
1.5.1	Metodologia para Revisão da Literatura	18
1.5.2	Análise Bibliométrica	20
1.5.3	Levantamento Bibliográfico	27
1.6	Contribuições	32
2	CAPÍTULO 2 - MERCADO DE ENERGIA ELÉTRICA	34
2.1	Desenhos de Mercado	34
2.1.1	Monopólio	34
2.1.2	Single Buyer	35
2.1.3	Competição no Atacado	36
2.1.4	Competição no Varejo	37
2.2	Mercado de Energia Elétrica no Brasil	38
2.2.1	Formação de Preços	43
2.2.2	Previsão de Preços: <i>Eletricity Power Forecasting</i> vs <i>Curva forward</i>	49
3	CAPÍTULO 3 - TECNOLOGIAS UTILIZADAS PRA OTIMIZAÇÃO E PREVISÃO	52
3.1	Redes Bayesianas Clássicas	52
3.1.1	Métodos de discretização	56
3.1.1.1	Discretização em quantis	56
3.1.1.2	Discretização por clusterização <i>K-means</i>	57
3.2	Redes Bayesianas Dinâmica	57
3.3	Algoritmo Bioinspirado	59
3.3.1	Algoritmos Genéticos	59
3.4	Aplicação da Teoria dos Grafos na Modelagem do Algoritmo Genético	63
3.5	Método de Correlação: Coeficiente de Correlação de Pearson	65
3.5.1	Levantamento das variáveis	67
4	CAPÍTULO 4 - METODOLOGIA	69
4.1	Arquitetura do Método Proposto	69

4.1.1	Módulo Aquisição dos Dados	72
4.1.2	Módulo Análise e Seleção de Dados	73
4.1.3	Módulo Previsão de Energia Elétrica	75
4.1.3.1	Etapa 1 - Seleção dos Dados e Geração de Topologias	76
4.1.3.2	Etapa 2 - Processamento e transformação dos dados	81
4.1.3.3	Etapa 3 - Redes Bayesianas Dinâmicas	82
4.1.3.4	Etapa 4 - Análise dos Resultados de Projeção de Preço	82
5	CAPÍTULO 5 - RESULTADOS E DISCUSSÕES	85
5.1	Resultados do Estudo das Variáveis	85
5.2	Resultados do Algoritmo Genético para Geração de Topologias . . .	102
5.3	Resultados de Previsão dos Cenários do Preço da Curva <i>Forward</i> de Energia Elétrica	105
6	CONSIDERAÇÕES FINAIS	125
6.1	Publicações	126
	REFERÊNCIAS	129

1 Capítulo 1

1.1 Introdução

Este capítulo introduz o escopo desta tese, demonstra o nível de interesse pelo tema a nível mundial, faz uma revisão na literatura de métodos utilizados para previsão de energia no Brasil, trazendo a motivação para a proposta de pesquisa em questão.

1.2 Motivação

A precificação da energia elétrica desempenha um papel crucial no setor elétrico, influenciando diretamente as estratégias de comercialização, investimentos e decisões operacionais. No Brasil, o Preço de Liquidação das Diferenças, calculado e divulgado pela Câmara de Comercialização de Energia Elétrica, serve como referência para as transações no mercado de curto prazo. No entanto, sua aplicação no médio e longo prazo é limitada, uma vez que sua modelagem é fortemente influenciada por fatores operacionais e hidrológicos de curto prazo, os quais não refletem adequadamente as dinâmicas futuras de precificação.

O mercado de energia elétrica apresenta desafios significativos para a previsão de preços, pois envolve a interação de variáveis econômicas, operacionais, planejamento, ambientais e regulatórias. Em particular, o Brasil possui uma matriz energética predominantemente renovável, onde a geração hídrica representa 64% da capacidade instalada, tornando o sistema altamente dependente de condições hidrológicas. Em períodos de crise hídrica, a volatilidade dos preços se intensifica, exigindo medidas de mitigação de risco por parte dos agentes do setor.

O risco é um componente inerente de todas as atividades econômicas, variando em intensidade e natureza dependendo do setor e da função específica. No setor elétrico, esse risco se manifesta em todos os seus segmentos - geração, comercialização, transmissão, distribuição e consumo — embora em diferentes magnitudes (CABRAL et al., 2015). Assim, tentar tratar a instabilidade desse risco para assegurar a estabilidade e eficiência operacional do setor permite que as empresas não apenas minimizem perdas potenciais, mas também maximizem oportunidades, contribuindo para a sustentabilidade e a competitividade do setor no longo prazo.

As comercializadoras de energia estão entre os agentes mais impactados por essa volatilidade, pois operam na intermediação entre compradores e vendedores, lidando diretamente com as flutuações de preços e demanda. Nesse contexto, a ausência de um modelo preditivo robusto para a curva *forward* de preço da energia elétrica dificulta a

formação de contratos futuros e a gestão eficiente do portfólio de energia.

Diante desses desafios, torna-se essencial o desenvolvimento de metodologias avançadas para previsão de cenários de preços nos horizontes de médio e longo prazo, permitindo a construção de uma curva *forward* confiável. O aprimoramento dessas previsões pode trazer benefícios significativos para o setor, incluindo:

- Redução da incerteza na precificação de contratos futuros;
- Melhoria na gestão de risco para agentes comercializadores e consumidores;
- Apoio à tomada de decisão estratégica para investidores e reguladores.

Neste contexto, esta pesquisa propõe o desenvolvimento de uma metodologia para previsão de cenários da curva *forward* de preços de energia elétrica para o médio e longo prazo no mercado brasileiro, incorporando diferentes fontes de dados e abordagens estatísticas e computacionais avançadas. O objetivo é fornecer um modelo robusto e realista, capaz de capturar a dinâmica do mercado de energia e auxiliar na tomada de decisão dos agentes do setor.

Esta pesquisa é um dos resultados da Chamada Pública realizado pela Eletrobras Furnas de P&D em 2020, intitulado “Sistema Adaptativo de Criação de Cenários de Previsão de Preços de Mercado de Energia de Médio e Longo Prazo: Uma Nova Metodologia utilizando técnicas de *Data Science*, Algoritmos Genéticos Multi-objetivos Aplicados a Redes Bayesianas Clássicas e Dinâmicas” (FURNAS, 2020), onde foi desenvolvido pelo Centro de Excelência em Eficiência Energética da Amazônia (CEAMAZON)/ Universidade Federal do Pará (UFPA) em parcerias com a Universidade Federal do Oeste do Pará (UFOPA) e CENERGEL Sistema Energéticos.

1.3 Objetivos

O objetivo principal deste trabalho é desenvolver uma metodologia para previsão de cenários de preço da curva *forward* de energia elétrica no horizonte de médio e longo prazo para o mercado brasileiro. Como objetivos específicos têm-se:

- Criar um banco de dados com variáveis endógenas à empresa e exógenas;
- Realizar um estudo das possíveis variáveis que mais impactam a formação de preço de energia elétrica no médio e longo prazo;
- Propor um método para criação automática de modelos de topologias utilizadas pelas Redes Bayesianas Dinâmicas;

- Utilizar as Redes Bayesianas Dinâmicas para previsão de cenários de preço da curva *forward* de energia elétrica de médio e longo prazo.

1.4 Organização da Tese

Esta tese está organizado em seis capítulos. O Capítulo 1 é apresenta as considerações iniciais e a revisão da literatura sobre o preço de energia. Neste capítulo é demonstrado a complexidade de prever preço de energia no longo prazo para o mercado brasileiro.

No Capítulo 2 é explanado sobre os desenhos de mercado de energia e a estrutura do mercado brasileiro de energia elétrica.

O Capítulo 3 contextualiza as técnicas utilizadas na metodologia. Apresentam-se os conceitos de Redes Bayesianas Clássicas e Dinâmicas, Algoritmos Genéticos e contextualização de teoria de grafos.

O Capítulo 4 demonstra todas as etapas da arquitetura proposta para previsão de curvas de preço de energia no longo prazo, desde a etapa de aquisição e análise dos dados, modelagem do problema utilizando Algoritmo Genético, a previsão do preço.

No Capítulo 5 concentram-se os resultados e discussões das previsões nos horizontes de 3 e 6 anos, juntamente com a comparação com outros métodos de previsão.

Por fim, o Capítulo 6 dá a ver as considerações finais, resumizando as contribuições e as publicações acadêmicas desenvolvidas no período desta tese.

1.5 Revisão Bibliográfica

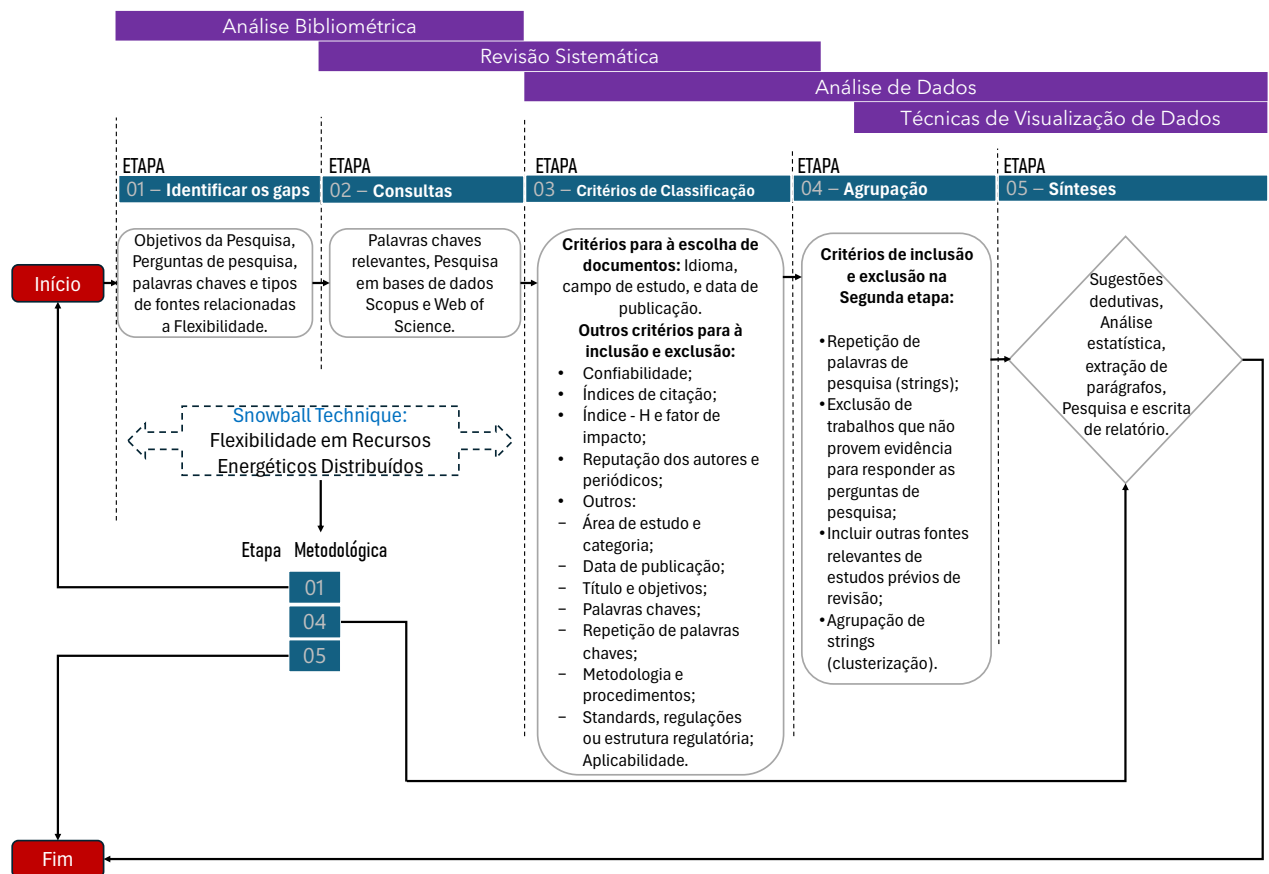
1.5.1 Metodologia para Revisão da Literatura

Visando implementar uma análise sistemática da bibliografia relativa ao tema, foi desenvolvido um estudo de anterioridade baseado na declaração PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), desenvolvido por um grupo de acadêmicos de saúde pública conforme apresentado em (ELSHATER; ABUSAADA, 2022) e (MOHER et al., 2009). A metodologia é dividida em cinco etapas, como apresentada na Figura 1; e inclui uma análise bibliométrica que visa mapear a produção científica, identificar padrões de pesquisa e avaliar a influência de periódicos e instituições na área em questão.

A análise bibliométrica é um método quantitativo para a avaliação da literatura científica e acadêmica em um campo específico ou sobre um tema específico. Esse tipo de análise utiliza dados bibliográficos, como número de publicações, frequência de citações, coautoria, palavras-chave e a distribuição geográfica dos autores, para medir o impacto e

as tendências de pesquisa. Ela permite identificar os principais pesquisadores, instituições e países contribuintes, bem como as redes de colaboração e os tópicos emergentes dentro de uma área de estudo. Além disso, a análise bibliométrica pode ajudar a avaliar o progresso científico e a direção do desenvolvimento de campos específicos ao longo do tempo.

Figura 1. Metodologia para estudo de anterioridade baseada na declaração PRISMA.



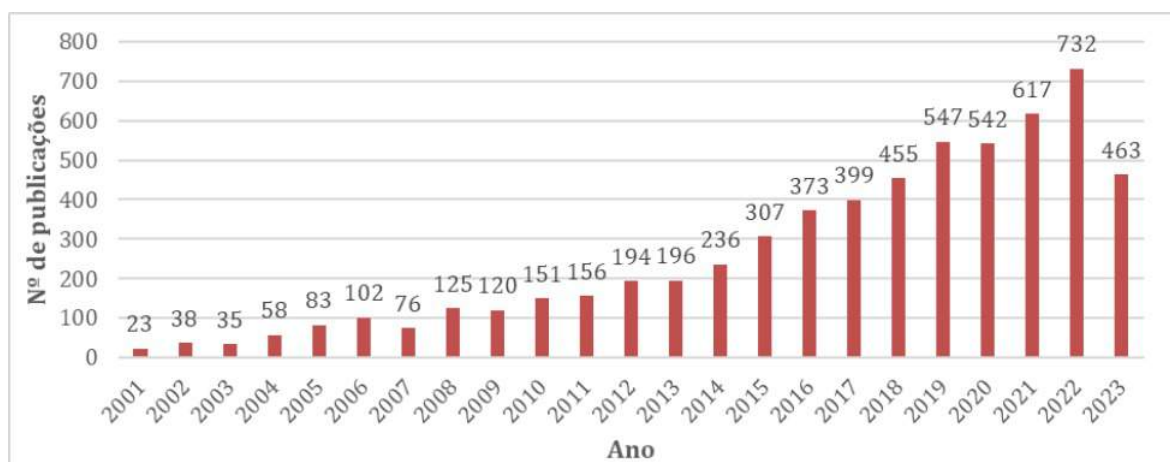
Fonte: Adaptado de (ELSHATER; ABUSAADA, 2022) e (MOHER et al., 2009).

Portanto, visando analisar como está o interesse a nível nacional e mundial sobre os temas *electricity power forecast* e *electricity curve forward*, foi realizado um levantamento na plataforma *Scopus* para seleção dos trabalhos objeto da análise. Trata-se de uma de fonte de busca de maior concentração de base de dados de resumos e citações de literatura técnica-científica revisada por pares, com capacidade de indexação dos mais relevantes repositórios, além de possuir ferramentas bibliométricas para acompanhar, analisar e visualizar a pesquisa.

1.5.2 Análise Bibliométrica

Desta forma, um levantamento das publicações contendo o termo *electricity power forecast* foi realizado considerando os últimos 23 anos, conforme apresentado na Figura 2. Nesta, observa-se que, até 2014, o crescimento das publicações ocorreu de forma gradual, variando entre 23 publicações em 2001 e 236 publicações em 2014. No entanto, a partir de 2014 há um crescimento significativo e consistente no número de trabalhos publicados, atingindo 732 publicações no ano de 2022, o maior número registrado no período analisado, levando em consideração que até o primeiro semestre de 2023 já havia 463 publicações.

Figura 2. Publicações envolvendo *electricity power forecast* nos últimos anos.



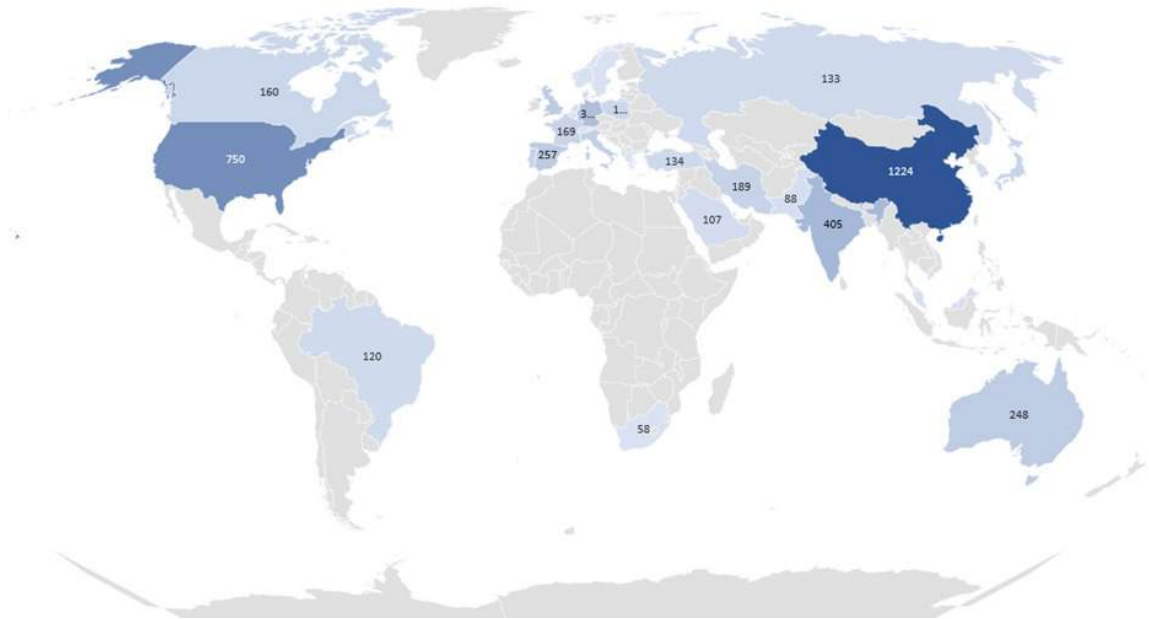
Fonte: Autor

Esse aumento expressivo pode ser atribuído ao avanço de técnicas de aprendizado de máquina, inteligência artificial e análise de séries temporais aplicadas à previsão de preços e demanda de energia elétrica. Além disso, a crescente complexidade dos mercados de energia, impulsionada pela maior penetração de fontes renováveis e a volatilidade dos preços, reforçou a necessidade de abordagens mais sofisticadas para previsões energéticas.

Mundialmente, as pesquisas envolvendo o tema *electricity power forecast* são lideradas pela China (1224), seguida pelos Estados Unidos (750) e Índia (405). O Brasil aparece na 18ª posição, com um total de 120 publicações, conforme demonstrado na Figura 3.

Vale ressaltar que, dadas as palavras-chave inseridas como objeto da pesquisa (*electricity power forecast*) e, selecionando as áreas de interesse (engenharia, energia, ciência da computação e matemática), foram encontrados 6056 publicações entre artigos em periódicos (3166), artigos em conferências (2571) e demais documentos técnicos como notas técnicas, capítulos de livros e demais editoriais. Dada essa extensa base de publicações, foi utilizado o software *VOSViewer* visando a geração de conhecimento com base na geração de um mapa de relações entre as palavras-chave, de acordo com o número de vezes em que

Figura 3. Distribuição das publicações envolvendo *electricity power forecast*.

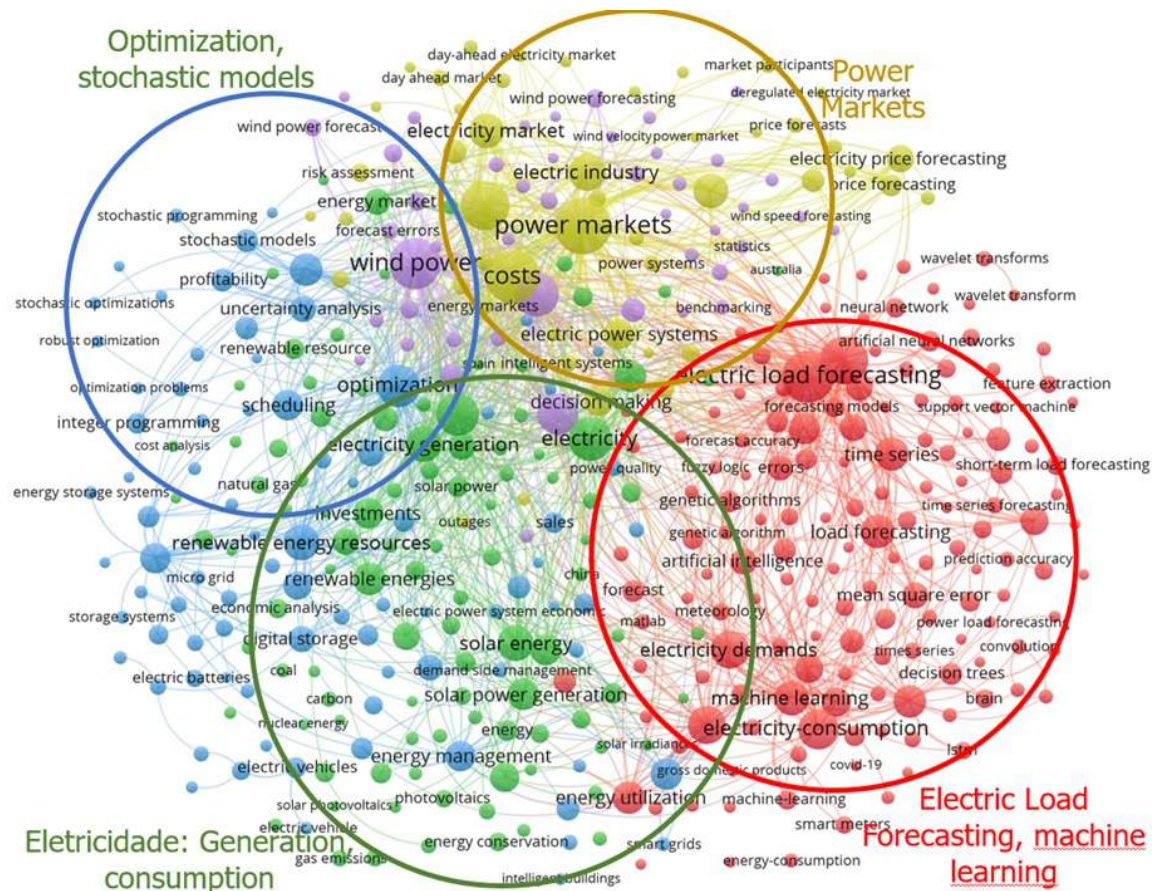


Fonte: Autor

elas são citadas nos estudos, bem como as conexões com outros estudos e palavras chave presentes em outras publicações da base de dados selecionada.

A Figura 4 apresenta o resultado da clusterização da relação das palavras-chave das publicações da base de dados, ressaltando as relações entre estas observadas em cada cluster. Dado o mapa apresentado, cabe ressaltar dois grupos principais de técnicas de determinação do preço *spot*: otimização associado a modelos estocásticos e *machine learning*, representando a relevância dessas duas áreas nas publicações realizadas.

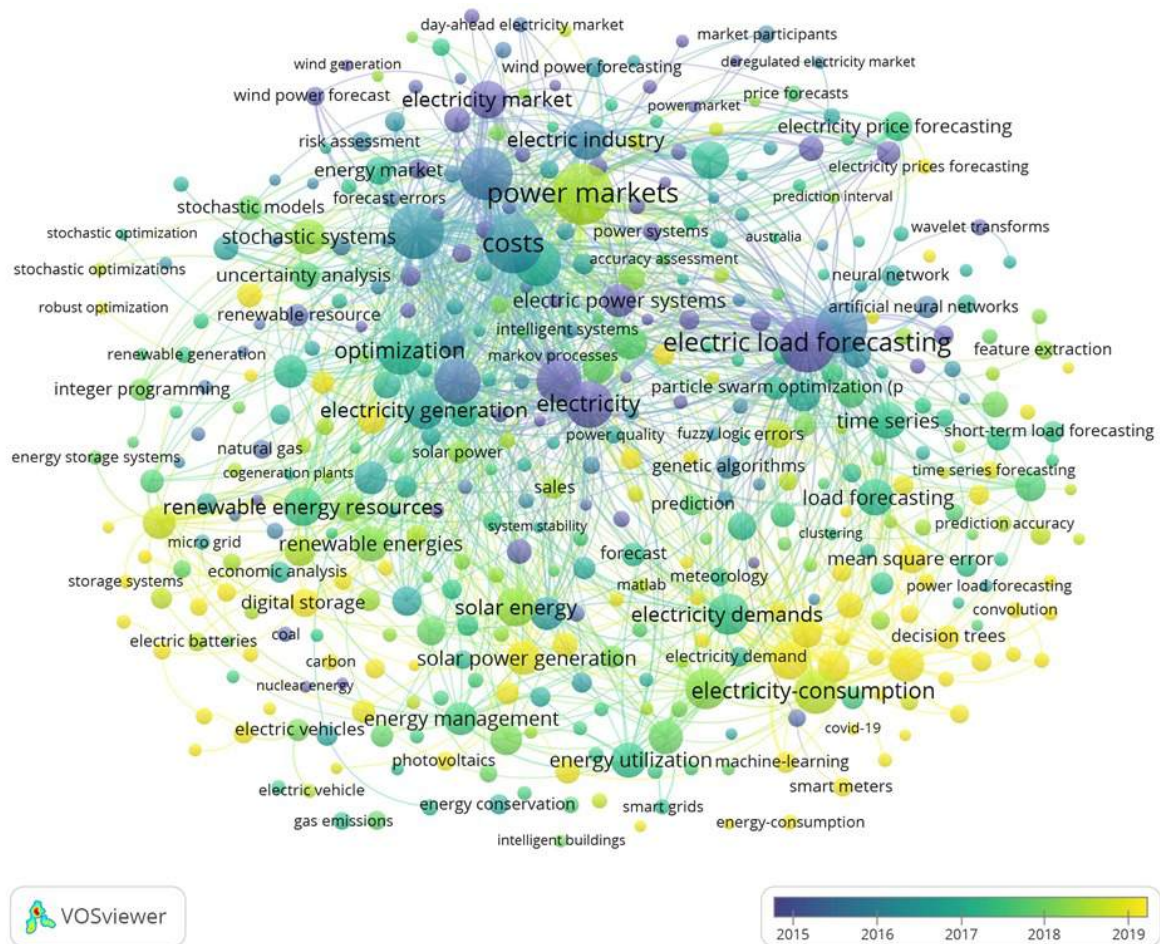
Figura 4. Mapa temático de palavras-chave separado por categorias relevantes referente *electricity power forecast*.



Fonte: Autor

De forma a rastrear os tópicos mais relevantes na atualidade na base de dados selecionada, o software permite realizar o mapeamento das palavras-chave identificadas conforme o ano das publicações em que estão associadas (Figura 5). Com efeito, as publicações atualmente mais relevantes se concentram nas áreas de inteligência computacional, envolvendo técnicas como *machine learning*, *deep learning*, árvores de decisão, *random forests* e redes neurais. A coloração do balão explicar que quanto mais clara, são as técnicas mais utilizadas nos últimos anos. Portanto, em síntese, o tema está em evidência nos últimos anos, juntamente com as técnicas associadas, como *deep learning*, abordagem recente com aplicações em expansão por diversas áreas.

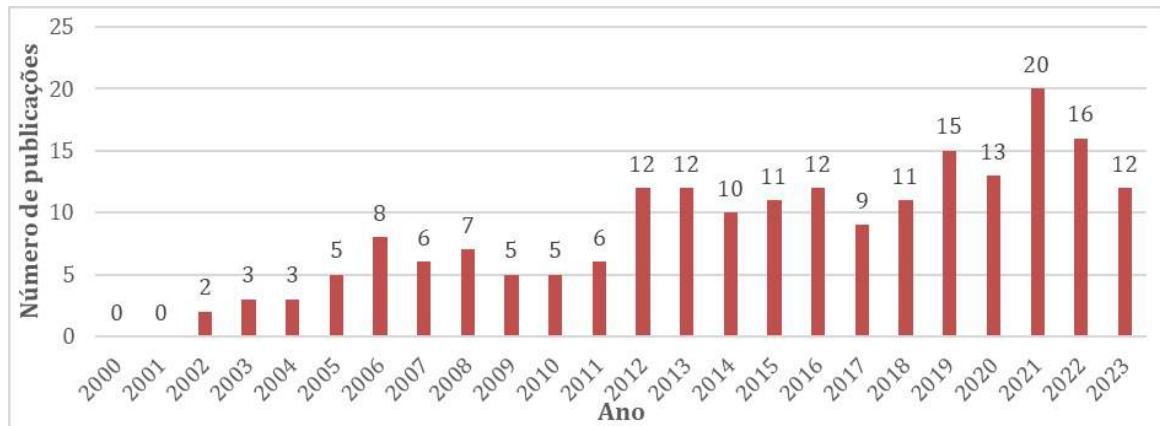
Figura 5. Mapa temático de palavras chaves separado por ano de publicação referente a *electricity power forecast*



Fonte: Autor

Com o objetivo de realizar a mesma análise bibliométrica, foram selecionadas publicações da base *Scopus* com o termo *electricity forward curve*, considerando os últimos 23 anos. A Figura 6 apresenta a evolução do número de publicações, onde os dados indicam um crescimento gradual no interesse acadêmico sobre o tema, especialmente a partir de 2003, quando as primeiras publicações começaram a aparecer. Nos anos seguintes, observa-se um padrão de crescimento moderado e oscilante, com períodos de estabilidade e pequenos aumentos no número de trabalhos publicados.

Figura 6. Publicações envolvendo *electricity forward curve* nos últimos anos.



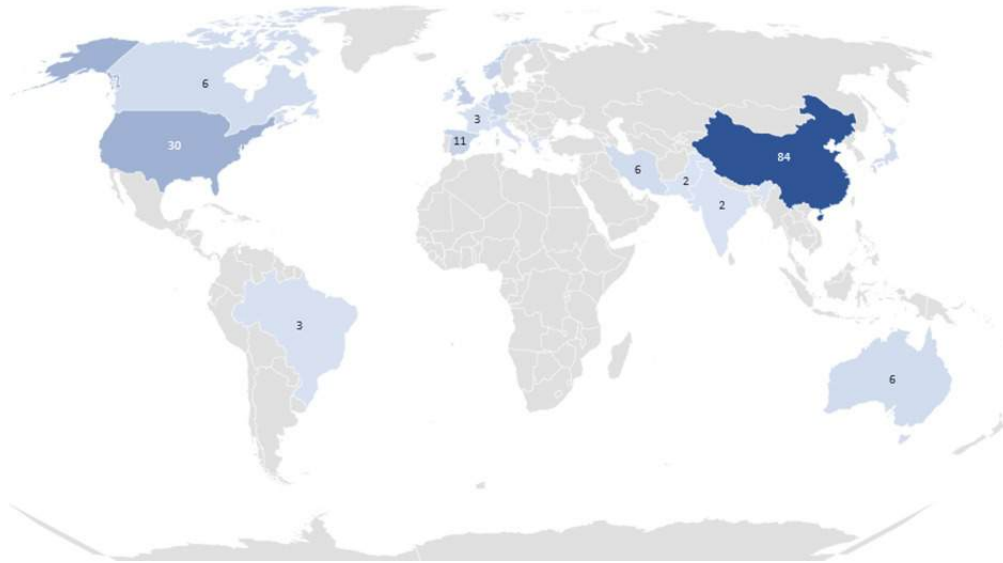
Fonte: Autor

Entre 2006 e 2011, o número de publicações variou entre 5 e 8 por ano, sugerindo um interesse contínuo, mas sem crescimento exponencial. A partir de 2012, nota-se uma leve expansão, com alguns picos de publicações em 2012, 2014 e 2016, mantendo-se sempre acima de 9 publicações anuais. O ano de 2021 destaca-se como o período com o maior número de publicações sobre o tema, totalizando 20 artigos, o que representa o ápice do interesse acadêmico registrado no período analisado. Esse aumento pode estar associado ao crescimento da complexidade dos mercados de energia elétrica, impulsionado por fatores como maior volatilidade dos preços, integração de fontes renováveis, avanço das metodologias de modelagem financeira e também o ano do apogeu da pandemia de COVID-19.

No geral, o comportamento do gráfico indica que o estudo sobre o tema *electricity forward curve* tem recebido atenção significativa, consolidando-se como um campo relevante dentro da literatura acadêmica sobre precificação de energia elétrica.

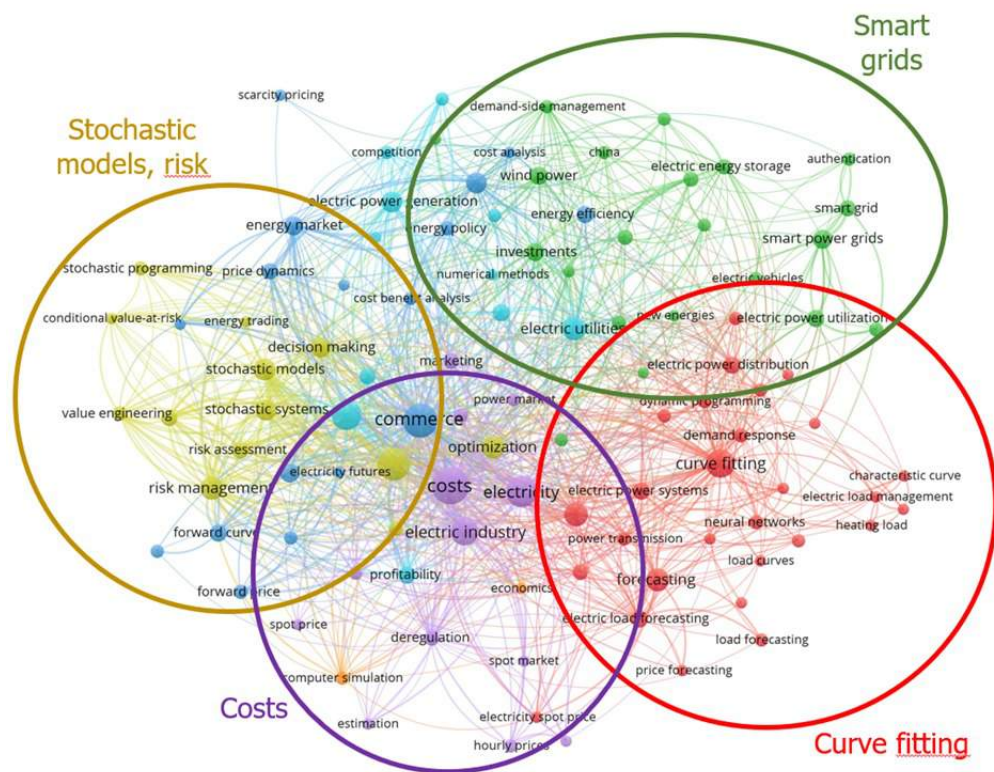
Em âmbito global, as pesquisas envolvendo o *electricity forward curve* têm a China como líder, com 84 publicações, seguida pelos Estados Unidos, com 30, e pelo Reino Unido, com 11. O Brasil ocupa a 17ª posição, contabilizando um total de 3 publicações, conforme demonstrado na Figura 7, referente à distribuição de publicações.

Figura 7. Distribuição das publicações envolvendo *electricity forward curve*.



Fonte: Autor

Figura 8. Mapa temático de palavras-chave separado por categorias relevantes referente *electricity forward curve*



Fonte: Autor

A Figura 8 apresenta o resultado da clusterização da relação das palavras-chave das

tons amarelados destacando assuntos mais recentes (próximos de 2020). Conforme pode ser observado, as publicações, atualmente, mais relevantes se concentram em áreas de inteligência computacional, envolvendo técnicas como *machine learning*, *deep learning*, árvores de decisão (*decision trees*), *random forests* e redes neurais.

Essas metodologias vêm ganhando relevância nos últimos anos devido à crescente complexidade do mercado de energia elétrica e à necessidade de técnicas mais sofisticadas para modelagem, previsão de preços e otimização de estratégias de negociação. Além disso, tópicos relacionados à gestão de energia e à integração de fontes renováveis, como *smart grids*, *electric power distribution*, *demand response* e *electric energy storage*, também aparecem como áreas emergentes. Isso reflete a transição energética global e o aumento da participação de fontes intermitentes, como eólica e solar, que requerem melhores modelos de previsão e gerenciamento de incerteza.

Por outro lado, temas como precificação de energia (*spot price*, *forward price*, *electricity spot price*) e gestão de risco (*risk management*, *stochastic models*) permanecem centrais na literatura, mas já são explorados há mais tempo, conforme indicado pelas cores mais escuras do gráfico.

Em síntese, a bibliometria evidencia que o campo de estudo sobre *electricity forward curve* está se expandindo e incorporando novas abordagens computacionais para lidar com desafios modernos do setor elétrico. O crescimento da aplicação de inteligência artificial e aprendizado de máquina reflete uma tendência de digitalização e automação do setor, tornando esses métodos indispensáveis para análises futuras.

1.5.3 Levantamento Bibliográfico

Vários estudos têm sido realizados ao longo dos últimos anos com o objetivo de compreender e modelar as curvas forward de preço de energia elétrica nos horizontes de médio e longo prazo. Um dos trabalhos seminais na área trata-se do estudo realizado por (BESSEMBINDER et al., 2002), que propõem um modelo de equilíbrio para o mercado de energia, permitindo analisar o sistema através de um número definido de agentes sob análise de oferta. Examinaram o mercado americano de energia e suas taxas *spot* e forward através de um modelo discreto no tempo e identificaram que o prêmio de risco no mercado futuro dependente da variação e assimetria direita dos preços *spot*. Identificaram também um prêmio de risco decrescente na curva forward com a entrada de especuladores no mercado.

Nesse mesmo sentido, foi proposta a construção de uma curva de preço forward da energia no mercado comum do norte europeu utilizando como base os preços de contratos a termo disponíveis para diferentes datas no mesmo mercado (FLETEN; LEMMING, 2003). Os autores relatam que, mesmo para sistemas elétricos com despacho centralizado, como

é o caso do mercado norte europeu, é necessário inserir variáveis que reflitam a aversão ao risco por parte dos agentes de mercado, necessariamente atrelando a construção das curvas forward a modelos financeiros.

Em (DU; YU; HAYES, 2011), avalia fatores que potencialmente influenciam a volatilidade dos preços do petróleo bruto e a possível ligação entre essa volatilidade e os mercados de *commodities* agrícolas. Modelos estocásticos de volatilidade são aplicados aos preços futuros semanais de petróleo bruto, milho e trigo de novembro de 1998 a janeiro de 2009. Os parâmetros do modelo são estimados usando métodos *Bayesian Markov Chain Monte Carlo*. O autor conseguiu encontrar evidências da volatilidade entre os mercados de petróleo bruto, milho e trigo após o outono de 2006. Isso pode ser amplamente explicado pela estreita interdependência entre o petróleo bruto e esses mercados de *commodities* induzida pela produção de etanol. Ambos os materiais são relevantes aos cenários para a precificação da energia.

Considerando o mercado brasileiro, o trabalho realizado em (CABRAL et al., 2015) se propôs a uma análise da situação da gestão de risco de mercado na comercialização de energia elétrica no país por meio de uma pesquisa com os principais agentes atuantes no setor elétrico, incluindo geradores, comercializadores e consumidores. Por meio da pesquisa, foi verificado que as empresas que possuem maior aversão ao risco são aquelas com perfil mais voltado para geração, ao passo que as comercializadoras apresentam, de uma forma geral, maior exposição ao risco. Porém, o setor como um todo tem aprimorado suas metodologias, pois as deficiências na gestão do risco têm sido traduzidas em perdas financeiras. Foi verificado na pesquisa que, ainda que muitas empresas utilizem o PLD como insumo de gestão de risco, tem sido crescente o uso das curvas forward como estratégia de implementação do valor em risco na comercialização.

O trabalho apresentado em (LEME et al., 2015) detalhou as particularidades do mercado brasileiro e utilizou-se de informações de um serviço de fornecimento de preços de fechamento de contratos individuais no ambiente de contratação livre (ACL) com o objetivo de criar uma curva de consenso de preços e analisar suas características. Foi observado que os preços de vencimento mais curtos (dentro do ano corrente) apresentaram características diferentes dos preços dos vencimentos mais longos da curva, com variações até cinco vezes maiores, indicando menor risco associado a posições de longo prazo.

Por sua vez, o trabalho apresentado em (AMADEU, 2020) objetivou a aplicação da simulação de Monte-Carlo aliada ao Movimento Browniano Geométrico e a decomposição de Cholesky, procedimentos estocásticos amplamente utilizados no mercado financeiro e de ações, para simular curvas estocásticas de preços futuros para o mercado brasileiro de energia, com foco no submercado Sudeste/Centro-Oeste, que corresponde à maior parte da comercialização de energia no país. Os resultados obtidos se mostraram robustos mesmo diante dos períodos de alta volatilidade observados nos últimos anos, abrangendo crises

climática, econômica e sanitária, que impactaram sobremaneira a volatilidade dos preços.

Em (ARAUJO, 2020) foi proposto um modelo baseado em dois agentes representativos que administram sua exposição ao risco através de um contrato por diferenças entre o preço futuro esperado da energia elétrica no submercado Sudeste/Centro-Oeste e um preço de referência de forma a abranger todos os participantes do mercado, tais como comercializados e especuladores. O resultado alcançado pelo modelo foi satisfatório, considerando que a única informação utilizada são os preços *spot* esperados, informação amplamente disponível no mercado futuro de eletricidade.

Tomando novas metodologias de abordagem da formação da curva de preços, pode-se citar o trabalho realizado em (YOUSEFI; SIANAKI; SHARAFI, 2019) o qual focou no emprego de técnicas de *machine learning* para a previsão de preços de energia mensalmente, em um horizonte de cinco anos, para o estado da Califórnia. Para isso, foram coletados dados históricos ao longo de dezesseis anos, totalizando setenta e oito variáveis analisadas com o objetivo de determinar a correlação entre as mesmas e a sua influência nos preços de médio e longo prazo.

Em (MONTEIRO et al., 2022) foi implementado um modelo semi-paramétrico visando obter os preços de contratos de duração diária, denominados contratos elementares, cujo portfólio constitui o período de entrega dos contratos observados no mercado. A metodologia proposta consiste em três etapas principais: a primeira consiste em um modelo estrutural semi-paramétrico que obtém as menores peças que compõem o período de entrega de um swap de eletricidade. Em seguida, ocorre a redução da dimensionalidade do conjunto de resíduos via algoritmo PCA (Análise de Componentes Principais) e, posteriormente, é realizada a estimação de um modelo de séries temporais nas componentes resultantes.

Em (BALAN et al., 2022) foi proposto um sistema de decisão com o objetivo de definir a estratégia ótima de comercialização constituindo por módulos integrados de cenarização de geração e preços via clusterização, cálculo do custo marginal de expansão, precificação dos contratos, balanço estático, projeção de tarifa e otimização de portfólio de contratos. Desta forma, a estratégia de decisão ótima de comercialização se baseia em aspectos que vão desde a projeção da geração até a culminação com a definição dos contratos a serem transacionados.

A pesquisa realizada por (MENDONÇA et al., 2019) se concentra em realizar a previsão da curva forward de preço de energia elétrica no Brasil, utilizando quatro modelos de previsão: SARIMAX (*Seasonal Autoregressive Integrated Moving Average*), LSTM (*Long Short Term Memory*), GRU (*Gated Recurrent Unit*) e o modelo híbrido CNN-LSTM (*Convolutional Neural Networks*). Os dados utilizados para a projeção do preço utilizado são do preço *spot* o Preço de Liquidação de Diferenças e a curva *forward* de preço fornecido pela Dcide. Para o treinamento dos modelos foram utilizados dados históricos da série de preço semanais de energia dos anos de 2012 a 2016 e para testes os anos de 2017 e 2018.

A projeção é realizada semanalmente para os dois anos. Desta forma, foram realizadas previsões semanais sobre as séries de preços futuros de energia com maturidades de 0 a 4 semanas ($A+[0 - 4]$). Ao final, verificou-se que os melhores modelos para cada semana de maturidade foram: o modelo SARIMAX para previsões de maturidade ($A+0$ e $A+4$), o LSTM para maturidade ($A+1$), e o CNN-LSTM para maturidade ($A+2$, $A+3$). Observa-se que para grandes variações nos preços futuros nenhum dos modelos conseguiu prever.

Em (PAPPAS; SANTOS, 2019) foi avaliado o poder de mercado na formação de preços via oferta, visando analisar a influência de fatores como a quantidade de agentes e a presença de contratos bilaterais. Tal estudo visou, sobretudo, suscitar a discussão sobre a mudança de paradigma da formação de preços no mercado de energia brasileiro, saindo de uma formação explícita, por meio de modelos, para um novo paradigma que leve mais em consideração o poder das forças de mercado.

Continuando nesse assunto, (BASTOS; CUNHA; BARROSO, 2017) propuseram uma metodologia para a separação da comercialização de energia e lastro no Brasil através da captura do valor econômico da escassez no mercado de eletricidade, buscando determinar um modelo de contratação de lastro no setor elétrico. Tal demanda se deve ao fato de que, no mercado brasileiro, determina-se a comercialização conjunta de dois produtos distintos: energia e lastro. Enquanto o primeiro é a energia elétrica em si, a ser gerada e consumida, o segundo se refere a um valor adicional a ser recebido por geradores pela capacidade de geração de energia em momentos de necessidade, o que também é comumente chamado de produto confiabilidade. Dessa forma, faz-se necessário permitir que os agentes incorporem às suas ofertas às suas expectativas de remuneração no mercado de energia.

Em (CAMARGO; GUARNIER; RAMOS, 2015) foi proposta uma estratégia ótima de comercialização para consumidores livres considerando o *trade-off* entre contratação imediata e postergação de decisão, com base nas incertezas nos preços de curto prazo e na precificação de contratos bilaterais, resultantes de condições estruturais e sistêmicas que impactam de forma direta na formação do preço. Desta forma, foi implementado um modelo de otimização estocástica para a definição da estratégia de contratação para consumidores livres, no qual o controle de risco é realizado pela técnica CVaR (*Conditional Value at Risk*).

Com relação à gestão de risco na comercialização de energia, (GOMES et al., 2022) realizaram uma avaliação de risco de exposição ao mercado de curto prazo na ocasião da aquisição de novos ativos de energia. Desta forma, as alternativas de investimento são avaliadas individualmente e combinadas, de forma a proporcionar uma comparação da receita esperada e do risco associado. A utilização da metodologia proposta permite a avaliação de alternativas, não só quanto aos seus custos, mas também é possível quantificar o risco associado à escolha de determinada alternativa. Os resultados, considerando simulações realizadas para o portfólio de um gerador composto por usinas hidráulicas,

contratos de venda e usinas eólicas mostraram, por exemplo, que seria desejável aumentar os recursos desse portfólio, dado que os cenários analisados apresentaram preços *spot* crescentes.

Nesse mesmo sentido, (MARQUES et al., 2022) propuseram um modelo de apoio à tomada de decisão sobre operações de hedge na comercialização de energia elétrica. Em geral, as operações de hedge funcionam como um mecanismo contra a flutuação do preço da energia, contribuindo para a gestão de riscos e, conseqüentemente, trazendo maior previsibilidade de resultados para as companhias. Para isso, os autores propuseram um modelo baseado em uma função de preferência que permite modelar a variação do nível de aversão ao risco de um agente considerando diferentes bandas de preferência, mensurando o risco por meio das medidas VaR (*Value at Risk*), CVaR ou ES (*Expected Shortfall*). Os resultados do modelo indicam, então, o preço ótimo de compra além do montante ideal de energia.

Ainda com relação à gestão de risco, em (CARVALHO; VIANA; MELLO, 2022) foi proposta uma metodologia para precificar contratos de derivativos do tipo por quantidade e collar (de parâmetros piso, teto e ágil), indexados ao PLD, de forma que sejam traçadas curvas de mesma preferência para as possíveis combinações dos parâmetros, permitindo, desta forma, o estabelecimento de estratégias de negociação de comercializadores frente aos consumidores, levando em consideração o perfil de risco de ambas as partes envolvidas. Para análise dos resultados, foram obtidas três curvas de mesma preferência para as combinações desses parâmetros que possibilitassem ao agente comercializador ter uma ferramenta de apoio para o auxílio na negociação dos contratos.

Em (IBRAHIM, 2024) é proposto um modelo híbrido utilizando técnicas de memória de longo prazo (LSTM) baseado em um algoritmo aprimorado de busca de grade (IGS) para prever os preços de energia elétrica. O algoritmo (IGS) aumenta a eficiência do ajuste de hiperparâmetros, permitindo uma seleção sistemática e precisa dos parâmetros ideais para o modelo LSTM. Foi realizada a integração de fatores exógenos como feriados e eventos significativos no modelo de previsão, e com isso obteve-se uma melhora significativamente na confiabilidade e precisão na previsão de preços no mercado atacadista. Foi realizado a previsão do preço de energia de curtíssimo prazo (1 dia) utilizando LSTM-IGS, que em comparação com métodos tradicionais como o LSTM, FFNN e suas variações, se mostrou superior. Logo, variáveis do tipo exógenas afetam diretamente os preços de mercado de energia.

Na pesquisa realizada em (HE; LI; QI, 2024) estuda-se a previsão do preço da eletricidade na província de Shanxi. Para realização da previsão foi utilizado modelagem multi-modelo que incorpora a regressão, redes neurais e somada a *grey theory modeling*. O modelo prevê o preço *spot* mensal de energia que auxilia novas empresas (que não tem tanto influência no setor) na tomada de decisão das transações. Para a previsão de curto

prazo foi utilizado regressão e algoritmos de rede neural e para melhorar a taxa de acerto foi adicionado a *grey theory modeling* e a suavização exponencial tripla que trouxe uma melhor assertividade no resultado da previsão, pois eles permitem o tratamento dos preços extremos da energia elétrica que as redes neurais não conseguem tratar.

Por fim, (REIS et al., 2025) apresentam uma revisão de literatura sobre os métodos e técnicas utilizados na previsão de preços e demanda de energia elétrica, abrangendo desde abordagens estatísticas até técnicas de aprendizado de máquina. Os autores identificam os métodos mais empregados em diferentes horizontes de previsão curto, médio e longo prazo e detalha as variáveis mais relevantes para cada abordagem. Além disso, discutem as limitações de cada método, apontam possíveis melhorias e analisam criticamente os pontos fracos dos estudos revisados.

1.6 Contribuições

Diante do exposto na seção 1.5.3 de Levantamento Bibliográfico, projetar preço da energia no mercado brasileiro é desafiador e complexo. Existe o Preço da Liquidação das Diferenças que é periodicamente disponibilizado semanalmente pela Câmara de Comercialização de Energia Elétrica, mas, como demonstrado, para períodos longos é impraticável.

Várias abordagens objetivam prever o preço da energia, empregando métodos estocásticos, redes neurais, entre outros. A maioria dessas abordagens foi direcionada para períodos de até dois anos, enquanto outras abordaram até quatro anos. Todas essas metodologias utilizaram dados das séries históricas do Preço de Liquidação das Diferenças.

O estudo realiza uma análise detalhada das variáveis que influenciam diretamente a formação do preço de energia, tanto no nível agregado pelo Sistema Integrado Nacional quanto por submercado. Apesar de os preços do PLD serem informados por submercado, observa-se que esses valores ficam bem parecidos entre eles. Será possível verificar se outras variáveis, além das operacionais utilizadas pelos modelos computacionais do Operador Nacional do Sistema Elétrico, influenciam também na formação do preço de energia de longo prazo. Para investigar a correlação entre as variáveis prospectadas, foi utilizado o coeficiente de correlação de Pearson.

Depois de prospectadas e analisadas, as variáveis serviram de entrada para a metodologia proposta de previsão de cenários de preço da curva *forward* de energia elétrica.

Essa metodologia utiliza as Redes Bayesianas Dinâmicas para as projeções dos cenários futuros. No entanto, as entradas para RBDs consistem de uma topológica de Grafo Acíclico Dirigido. Assim, para criar essa topologia é essencial possuir um conhecimento profundo sobre o problema, logo é uma atividade complexa. Entretanto, o Algoritmo

Genético é utilizado para criar estas topologias que representem o problema definido, por se um método que trata problemas de otimização.

Esta metodologia de previsão de cenários, utilizando Redes Bayesianas Dinâmicas, destaca-se pela sua flexibilidade, permitindo que o usuário faça apontamento/inferências para qualquer período de tempo dentro dos anos de projeção. Por exemplo, ao realizar uma projeção para seis anos, é possível fazer apontamentos em qualquer variável presente no modelo e em qualquer mês ou semestre do horizonte de projeção. Assim, são projetados diversos cenários do preço de energia. A robustez da metodologia também é evidente, pois ela pode ser adaptada às mudanças no cenário do problema. Basta inserir no modelo as novas variáveis que se deseja considerar.

A metodologia proposta busca auxiliar os participantes do mercado de energia elétrica, oferecendo a capacidade de gerar múltiplos cenários de preços para o médio e longo prazo. Assim, proporcionando a esse agente autonomia para fazer apontamentos na geração desses cenários.

2 Capítulo 2 - Mercado de Energia Elétrica

Nesta seção será contextualizado: (i) os desenhos do mercado de energia; (ii) a contextualização do mercado de energia brasileiro, juntamente com sua regulamentação e; (iii), a formação do preço de energia no Brasil juntamente com suas limitações.

2.1 Desenhos de Mercado

Tomando o cenário mundial em perspectiva, observa-se que diferentes países adotam distintos desenhos para seus mercados de energia que refletem intrinsecamente questões históricas, culturais, regulatórias e de disponibilidade das fontes primárias. Entretanto, um ponto comum a ser destacado é o impacto dos desenhos de mercado nos preços verificados ao consumidor, e na complexidade e sucesso de ferramentas para previsão dos preços da energia elétrica.

Tal compreensão é de vital importância, pois a previsão de preços da energia elétrica para um determinado mercado está intrinsecamente ligada ao desenho do mercado de energia para o ambiente em que será realizada, refletindo o impacto regulador do Estado e as forças de mercado de oferta e demanda entre os comercializadores e consumidores.

De uma forma geral, os modelos de desenho de mercado observados a nível mundial se relacionam diretamente com o grau de participação do consumidor na definição do suprimento da energia elétrica, tendo como base a conversão de que os serviços de transporte, mesmo em mercados altamente competitivos, permanecem atrelados ao modelo de monopólios naturais. Desta forma, os modelos de mercado atualmente observados se dividem em (VIANA, 2018).

2.1.1 Monopólio

Trata-se do modelo predominante em muitos países em boa parte do século XX, até a implementação das reformas estruturantes verificadas nos mercados de energia elétrica de diversos países.

Neste modelo, uma única empresa ou um número limitado de empresas controla todas as etapas da cadeia de valor da eletricidade, desde a geração, transmissão até a distribuição ao consumidor final, visando a obtenção de ganhos de sinergia e facilitação do controle do produto estratégico a nível nacional. Geralmente, estas empresas operam sob regulação estatal rigorosa para evitar práticas monopolistas e garantir o fornecimento de energia. Nesse modelo facilita o planejamento e a coordenação do sistema elétrico, e a empresa tem um controle total sobre o investimento e a operação em todas as etapas, mas,

em contra partida, pode levar a eficiências operacionais reduzidas e falta de incentivo para reduzir custos ou inovar devido à ausência de concorrência.

No cenário brasileiro, tal modelo foi observado até meados dos anos 2000, quando da implementação de reformas de modernização do setor elétrico após cenários de escassez e apagões ocorridos no início da década. Até então, prevaleciam dois modelos de monopólio: uma variante na qual uma única estatal atuava realizando todas as funções da indústria e outra na qual uma empresa federal atuava na geração e transmissão, enquanto uma distribuidora estadual atuava no setor de distribuição de energia (VIANA, 2018).

2.1.2 Single Buyer

O modelo *Single Buyer* ou comprador único surgiu nos Estados Unidos na década de 1970 visando incrementar a eficiência alocativa dos recursos diante do cenário de crise energética resultante do choque do petróleo. Neste modelo, surge a figura do IPP (*Independent Power Producer*) que se constituem em agentes privados que competem para fornecer energia elétrica para os centros de carga já existentes por meio da participação em leilões ou chamadas públicas reguladas por uma agência governamental.

O objetivo desse modelo é combinar a eficiência da produção e a confiabilidade do fornecimento com algum nível de competição no setor de geração, mantendo o controle sobre a transmissão e a distribuição.

A nível de geração, diversos produtores de energia, que podem ser empresas estatais ou privadas, geram eletricidade e vendem sua produção para a entidade compradora única. Geralmente, essas vendas são realizadas através de contratos de longo prazo, que especificam quantidades, preços dentre outras condições.

A entidade compradora única, que opera sob regulação estatal, tem a responsabilidade de garantir que haja suprimento suficiente de eletricidade para atender à demanda atual e futura a preços razoáveis. Ela compra energia dos geradores e a vende para as empresas de distribuição ou diretamente para grandes consumidores. Essa entidade também pode ter a responsabilidade de promover o investimento em novas capacidades de geração, especialmente em fontes renováveis de energia, para garantir a segurança e a sustentabilidade do suprimento a longo prazo.

Nesse modelo as redes de transmissão e distribuição geralmente permanecem como monopólios naturais regulados, devido à sua natureza de infraestrutura essencial e à necessidade de evitar duplicação desnecessária de redes (não seria viável pelos custos em infraestrutura). Essas redes podem ser operadas pela entidade compradora única ou por outras empresas designadas, sob estrita regulação para garantir acesso não discriminatório e tarifas justas.

No entanto, esse modelo também enfrenta críticas, especialmente em termos de

eficiência e inovação. A falta de competição direta no mercado pode levar a ineficiências operacionais e reduzir o incentivo para inovação e redução de custos. Além disso, a dependência de uma única entidade para a compra de energia pode introduzir riscos sistêmicos e desafios regulatórios.

Tal modelo trata-se de um primeiro estágio no sentido da implementação de um mercado competitivo, pois, apesar de haver competição dos agentes no momento da ocorrência dos leilões, ainda assim os consumidores não dispõem de escolha do seu supridor de energia (VIANA, 2018).

2.1.3 Competição no Atacado

O modelo competição no atacado, também conhecido como “mercado atacadista de energia”, teve início em sua implementação a nível mundial na década de 1980 e posterior consolidação na década de 1990. O modelo é uma estrutura de mercado onde a geração de eletricidade é desagregada das atividades de transmissão e distribuição, permitindo que diversas empresas geradoras compitam para vender energia em um mercado aberto. Onde diferentes empresas podem possuir e operar instalações de geração de energia. E essa geração pode ser partir de uma variedade de fontes, incluindo combustíveis fósseis, nuclear, hidroelétrica, solar, eólica, entre outras. A diversidade de geradores promove uma competição baseada tanto no preço quanto na sustentabilidade da energia produzida.

A eletricidade gerada é vendida em um mercado atacadista, onde os preços são determinados principalmente pela oferta e demanda. Este mercado pode operar em diferentes horizontes de tempo, incluindo mercados *spot* (para entrega imediata de energia) e mercados de futuros (para entrega de energia em uma data futura). Operadores de mercado ou bolsas de energia facilitam as transações, garantindo a transparência e a eficiência na formação de preços.

Neste modelo surgem dois novos atores que desempenham um papel central na coordenação do sistema: o Operador de Mercado, responsável pela contabilização e liquidação financeira das posições; e, o Operador do Sistema, responsável pelo equilíbrio e manutenção da estabilidade física do consumo e geração, ou seja, garantindo que a demanda de eletricidade seja atendida e mantendo a estabilidade e a confiabilidade da rede. Este operador também é responsável por gerenciar o mercado de energia, incluindo o despacho de geração, que determina quais usinas geradoras serão utilizadas para atender à demanda com base no custo e na eficiência.

A infraestrutura de transmissão, que é regulada e muitas vezes de propriedade do estado ou operada sob concessões, transporta a eletricidade do ponto de geração para os distribuidores ou grandes consumidores. As tarifas de transmissão são reguladas para garantir que o acesso à rede seja justo e não discriminatório. Embora a competição no

atacado foque na geração e venda de energia em grande escala, a distribuição até os consumidores finais pode também ser aberta à competição, dependendo do modelo de mercado específico do país ou região. Em alguns casos, os consumidores finais podem escolher entre diferentes fornecedores de energia no mercado de varejo, que compram energia no mercado atacadista e a vendem para indústrias, empresas e residências.

A competição no atacado tem a vantagem de promover a eficiência e a inovação entre as geradoras, potencialmente reduzindo os custos de energia para os consumidores. Também incentiva o investimento em novas tecnologias e fontes de energia renováveis, à medida que as geradoras buscam vantagens competitivas. Contudo, este modelo exige uma regulação e supervisão rigorosas para evitar manipulação de mercado e garantir a confiabilidade do fornecimento. A complexidade do sistema de mercado e a necessidade de investimentos significativos em infraestrutura de transmissão são também desafios importantes. Tal modelo apresenta predominância em sua adoção pelo setor elétrico brasileiro.

2.1.4 Competição no Varejo

Este modelo é caracterizado por uma participação ativa dos consumidores sejam esses residenciais, comerciais ou industriais nas operações do mercado, independentemente do seu tamanho ou aspectos técnicos tais como nível de tensão e potência instalada para ingresso. Entretanto, ainda podem existir agentes como comercializadores varejistas ou agregadores que negociam diretamente com uma base de clientes, mediante a compra de energia no mercado atacadista, assumindo riscos de mercado pela sua atuação.

Para que este modelo funcione adequadamente, é necessário um grande número de comercializadores e consumidores e que estes detenham ferramentas de comparação de preços e realização de transações, conferindo liquidez ao mercado (VIANA, 2018).

Nesse modelo os consumidores têm o poder de escolher entre diferentes fornecedores de energia com base em critérios como preço, qualidade do serviço, fontes de energia (como por exemplo, renováveis ou não renováveis), dentre outras.

Os fornecedores de energia no mercado de varejo compram eletricidade no mercado atacadista ou diretamente de geradores e a vendem aos consumidores finais.

Apesar da competição no fornecimento de energia, a distribuição continua sendo um monopólio regulado, operado por uma ou mais empresas designadas na área de serviço. Estas empresas são responsáveis pela manutenção da rede de distribuição e pela entrega física da eletricidade aos consumidores, independentemente de qual fornecedor eles escolheram. Os distribuidores cobram uma tarifa regulada pelo uso da rede, que é separada do custo da energia fornecida pelos varejistas.

Neste cenário, o poder estatal centralizador na formação dos preços é bastante

reduzido, fazendo com que os mecanismos de mercado tais como a oferta e demanda sejam mais determinantes na definição dos preços da energia elétrica. Neste ambiente, semelhantemente a uma bolsa de valores, o preço da energia elétrica estará atrelado mais diretamente à percepção do valor do produto e aos riscos do que necessariamente à estrutura de custo de geração e disponibilização.

Esse modelo estimula a concorrência diretamente no ponto de venda ao consumidor, com o objetivo de melhorar o serviço e reduzir os preços da energia, com isso tem-se a vantagem em melhorar o atendimento ao cliente e maior inovação, à medida que os fornecedores buscam diferenciar suas ofertas. No entanto, também pode apresentar desafios, como a complexidade da escolha para os consumidores, o risco de práticas de venda enganosas e a necessidade de uma regulação eficaz para evitar a concentração de mercado e proteger os interesses dos consumidores.

Os desenhos básicos dos mercados de energia elétrica são estruturados para atender às necessidades específicas de cada sistema elétrico, refletindo as políticas energéticas, os objetivos ambientais e socioeconômicos, e as características técnicas da geração, transmissão, e distribuição de eletricidade. Estes desenhos são fundamentais para garantir a eficiência operacional e a sustentabilidade do setor elétrico.

2.2 Mercado de Energia Elétrica no Brasil

O mercado brasileiro de energia elétrica passou por diversas reestruturações desde seus primórdios e, hoje, possibilita a alguns consumidores opções de escolha em relação ao custo de consumo. Nesse âmbito, (BRAGA, 2018) resume todo o funcionamento do mercado de energia elétrica, começando pelo processo histórico do setor. Este inicia-se na época da Proclamação da República, época que o país começou a ser urbanizado, aumentando o consumo de energia elétrica com o crescimento da indústria do café. De acordo com o autor entretanto, a crise de 1929 foi crucial para a percepção da necessidade de delinear os papéis do Estado na política das nações. Sendo assim, em 1934, o aproveitamento do potencial hidroelétrico passou a ser patrimônio estatal com a publicação do Código das Águas.

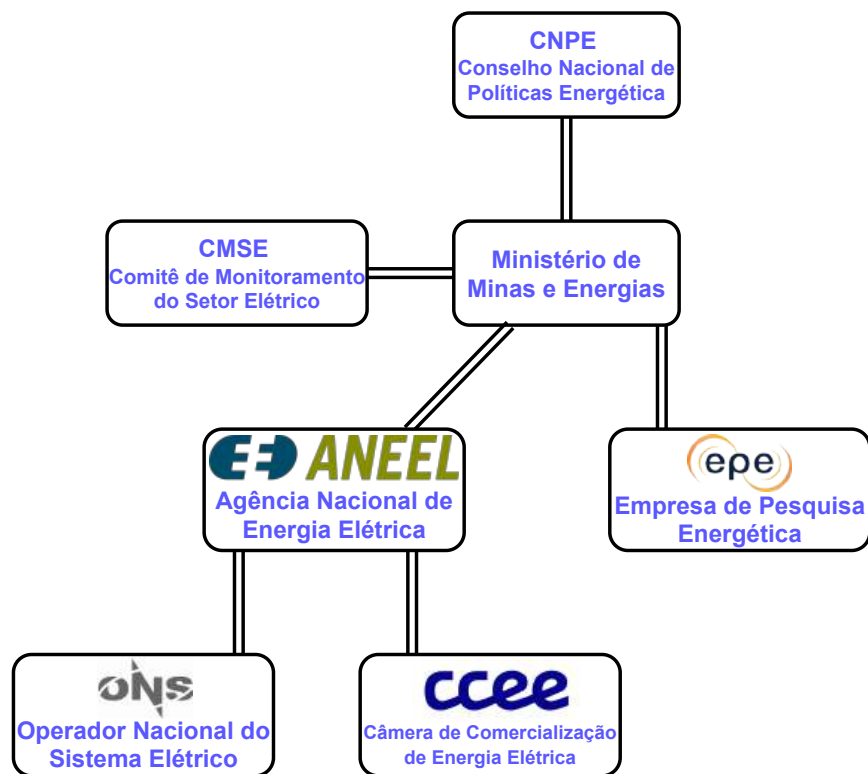
O modelo do mercado de energia do Brasil conhecido nos dias atuais começou seu desenvolvimento em 1990, a partir de uma reestruturação do setor elétrico (RESEB), baseado na política que o Estado não deveria mais ser o executor direto das atividades relacionadas à geração, transmissão e distribuição de energia, mas sim atuar como um regulador e promotor do desenvolvimento setorial, o que iniciou um processo de privatização das empresas e criação de autarquias públicas.

O modelo proposto não conseguiu expandir a oferta de energia de forma suficiente, assim, alguns novos ajustes foram realizados a partir de 2004, porém muito semelhantes às

mudanças anteriores. Dessa forma, apenas em 2012 o setor elétrico ganhou as características conhecidas atualmente, a partir da Medida Provisória 579 (REPÚBLICA, 2012), a qual foi convertida, posteriormente, na Lei 12.783/2013 (REPÚBLICA, 2013). O atual setor elétrico brasileiro possui como características principais: segregação das atividades de geração, transmissão e distribuição, existência simultânea de empresas públicas e privadas, coexistência de consumidores livres e cativos, negociações de caráter livre entre geradores consumidores e comercializadores, e leilões para contratação de energia.

A estrutura regulatória do setor elétrico brasileiro é complexa e multidimensional, refletindo a necessidade de conciliar objetivos econômicos, técnicos, ambientais e sociais. Ela é composta por diversos órgãos, entidades reguladoras e mecanismos de mercado, cada um com responsabilidades específicas (MME, 2021). Esta estrutura tem como objetivo não apenas assegurar a oferta de energia elétrica de forma eficiente a preços razoáveis, mas também promover a expansão sustentável do setor, na Figura 10 é demonstrando essa estrutura.

Figura 10. Estrutura do Setor Elétrico Brasileiro



Fonte: Autor

De acordo com esta estrutura, cada órgão do setor desempenha funções particularizadas e específicas, a saber:

1. O Conselho Nacional de Política Energética (CNPE) é de fato o principal órgão formador de políticas públicas para o suprimento de energia no Brasil e tem como função propor ao Presidente da República políticas destinadas a: Promover o aproveitamento racional dos recursos de energia do Brasil; assegurar o suprimento de insumos de energia para os locais mais distantes e de difícil acesso do País, submetendo ao Congresso Nacional as medidas que implicarem na criação de subsídios; estabelecer diretrizes para programas relacionados ao uso do gás natural, do carvão, da energia termonuclear, dos biocombustíveis, da energia solar, da energia eólica e de energia proveniente de outras fontes alternativas, por exemplo. Além disso, o Conselho Nacional de Política Energética estabelecer diretrizes para importação e exportação, para atender às necessidades de consumo interno de petróleo, biocombustíveis, gás natural de modo a garantir o funcionamento adequado do Sistema Nacional de Estoques de Combustíveis e o cumprimento do Plano Anual de Estoques Estratégicos de Combustíveis dentre outras (CIVIL, 1997).
2. O Ministério de Minas e Energia (MME) tem a responsabilidade de formulação e implementação das políticas nacionais de energia desempenhando um papel central na definição das diretrizes para o desenvolvimento do setor elétrico.
3. O Conselho de Monitoramento do Setor Elétrico (CMSE), possui o papel de avaliar e acompanhar a continuidade e a segurança do suprimento eletroenergético do país, por meio do acompanhamento das atividades de geração, transmissão, distribuição, comercialização, exportação e importação de energia elétrica, gás natural, petróleo e seus derivados; da avaliação das condições de abastecimento e de atendimento; da análise integrada de segurança de abastecimento e atendimento ao mercado de energia elétrica, gás natural e petróleo, levando em consideração a demanda, oferta e qualidade de insumos energéticos, o funcionamento do sistema de oferta e produção e a configuração dos sistemas de transporte locais, regionais e internacionais (CMSE, 2024).
4. A Agência Nacional de Energia Elétrica (ANEEL) é um autarquia federal sob o regime especial, vinculada ao MME, com a função de regular e fiscalizar a geração, transmissão, distribuição e comercialização de energia elétrica, garantindo a estabilidade e a qualidade do fornecimento; a fiscalização das concessões, das permissões e dos serviços de energia elétrica; A implementação de políticas do governo federal, com relação à exploração da energia elétrica e ao aproveitamento dos potenciais hidráulicos; o estabelecimento de tarifas dentre outras (ANEEL, 2022).
5. Empresa de Pesquisa Energética (EPE) tem por função prestar serviços ao MME relativos às áreas de estudo e pesquisa relacionadas ao subsidiamento do planejamento do setor de energia, levando em consideração a energia elétrica, o petróleo e o gás

natural e seus derivados e biocombustíveis. Ademais, a EPE possui o objetivo de resgatar a responsabilidade constitucional do Brasil em assegurar as bases para o desenvolvimento sustentável da infraestrutura energética nacional (EPE, 2024b).

6. O Operador Nacional do Sistema Elétrico (ONS) é uma associação civil sem fins lucrativos, que opera o sistema nacional de energia elétrica e tem como atribuições ampliações e reforços das redes, portanto durante a elaboração do Plano da Operação Elétrica do Sistema Interligado Nacional (SIN), foram estabelecidas condições para garantir a segurança do sistema para identificação das necessidades de expansão da Rede Básica, o reforço de infraestruturas de transmissão e distribuição que possam afetar o funcionamento sistêmico, bem como a avaliação das melhorias associadas à substituição de equipamentos obsoletos e deteriorados; supervisão e o controle dos centros de operações; a previsão de carga e o planejamento da operação dos sistemas isolado, dentre outras (ONS, 2024a). E por fim,
7. A Câmara de Comercialização de Energia Elétrica (CCEE), é uma associação civil sem fins lucrativos, que opera o mercado de energia elétrica do país e tem como atribuições: o registro de contratos de compra e venda dos agentes dos mercados regulado e livre; o cálculo do preço do PLD para o dia seguinte, diariamente; a gestão de contas setoriais; a divulgação de resultados das operações do mercado; o monitoramento do mercado, com a identificação e análise de ações dos agentes que não estão de acordo com a legislação ou com as boas práticas comerciais, além da aplicação de penalidades em caso do descumprimento das regras de mercado; o desenvolvimento de plataformas tecnológicas para as operações; a operacionalização de negociações para a venda de energia excedente das distribuidoras para agentes do mercado livre interessados na compra; elaborar regras de comercialização; responsável por realizar leilões para a contratação de energia no SIN; contabilizar Liquidação dos valores dos contratos e dos dados de medição, dentre outras (CCEE, 2023).

No mercado brasileiro a partir da publicação da Lei Nº 10.848/2004, as transações bilaterais de compra e venda de energia são realizadas em dois ambientes: Ambiente de Contratação Livre (ACL), cujos contratos são de livre negociação entre os agentes envolvidos, respeitada a regulamentação vigente emitida pela Câmara de Comercialização de Energia; e, Ambiente de Contratação Regulada (ACR), subdivididos nos Contratos de Comercialização de Energia no Ambiente Regulado (CCEAR), Contratos de Cota de Garantia Física (CCGF), Contratos de Cotas de Energia Nuclear (CCEN), Contratos de Itaipu, Contratos do PROINFA, Contratos de Leilão de Ajuste e Contratos Bilaterais Regulados (CCEE, 2023).

De uma forma geral, o ACR visa atender a demanda dos consumidores cativos, que são os consumidores residenciais, de serviço e indústrias. No ACR as distribuidoras

de energia elétrica atendem os consumidores cativos através de uma contratação em um ambiente *spot*, realizada por meio de leilões de energia elétrica, nos quais os geradores que oferecem os menores preços são os ganhadores. Por meio dos CCEARs, as distribuidoras de energia são obrigadas a contratar 100% de sua carga de energia a partir da contratação de lastro de energia e potência. Nesse sentido, os termos e as condições dos CCEAR's e o mecanismo de garantias bilaterais servem para assegurar o cumprimento das obrigações contratuais que são definidos pela ANEEL, no âmbito de sua competência para elaboração dos editais dos leilões de energia, conforme previsto na Lei nº 10.848/2004 e pelo Decreto nº 5.163/2004 (CASTRO; BRANDÃO, 2021).

Este ambiente é caracterizado por uma sistemática de leilões organizados pelo governo, nos quais são firmados contratos bilaterais de longo prazo entre geradores e distribuidoras de energia, em que estas arcam com o risco de crédito dos consumidores cativos. Há leilões de vários tipos, dentre os quais se destacam os Leilões de Energia Existente, para compra de energia de usinas já em operação comercial, e os Leilões de Energia Nova, destinados a viabilizar a construção de novos empreendimentos de geração, sendo estes realizados majoritariamente no atual cenário, com contratos de longo prazo, de até 30 anos (CASTRO; BRANDÃO, 2021).

Por outro, o ACL é composto de consumidores de maior volume de consumo, com a comercialização de energia realizada de forma bilateral entre geradores, comercializadores, consumidores livres e especiais. Neste ambiente, os termos de contratação como preço, quantidade, condições de registro dos contratos, garantias, dentre outros, são pactuados livremente entre os agentes. Desta forma, a CCEE se limita a registrar os prazos e as quantidades destes contratos, sendo seus valores restritos às partes contratantes (CASTRO et al., 2017a).

O ACL está estruturado como um balcão não organizado, sem uma entidade central para administrar os contratos bilaterais firmados. Apesar disso, ocorreram medidas para a criação de sistemas de negociação de contratos padronizados por meio de entidades privadas, sendo a mais bem sucedida o Balcão Brasileiro de Comercialização de Energia (BBCE) (CASTRO et al., 2017b). Entretanto, o volume de contratos firmados nesta plataforma ainda se mostra incipiente se comparado aos contratos firmados diretamente entre comercializadores e consumidores e que não possuem seus preços divulgados o que, consequentemente, prejudica o conhecimento dos agentes envolvidos dos preços usualmente praticados no mercado, representando uma oportunidade para o desenvolvimento de metodologias e ferramentas para a previsão de preços de comercialização.

As tendências para a comercialização de energia para o mercado brasileiro nos próximos anos é uma acentuação do crescimento do ACL e retração do ACR, com migração definitiva da maioria das empresas (BRANDÃO, 2019). Segundo o Anuário Estatístico de Energia Elétrica (EPE, 2023), com dados relativos ao ano de 2022, o percentual de

consumo no ACL já representa 39,66% do total de energia consumida no Brasil, o que se traduz na entrada contínua de novos agentes e mais liquidez nos contratos firmados.

2.2.1 Formação de Preços

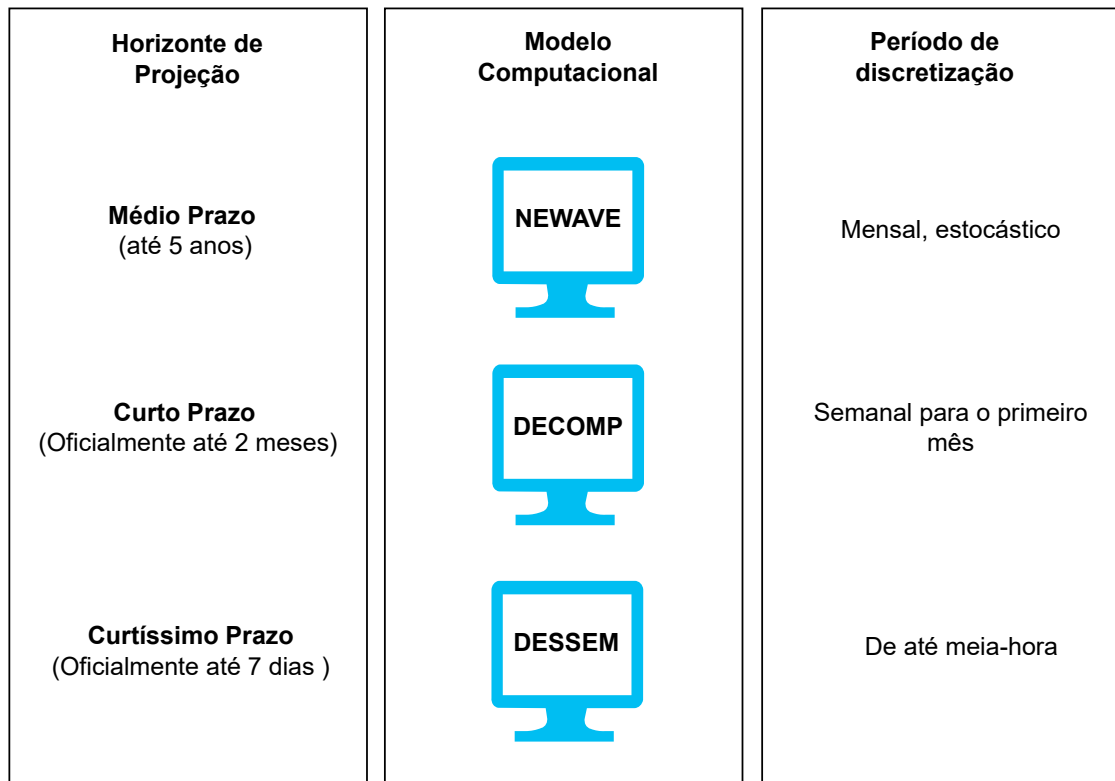
No mercado de energia elétrica, a problemática dos preços de negociação dos contratos entre geradores/comercializadores e compradores toma especial relevância pois uma modelagem assertiva dos contratos leva a um incremento do retorno financeiro e da sustentabilidade do negócio para os vários agentes participantes, levando-os à realização de investimentos em setores especializados em previsão de preços e análises de riscos.

O mercado brasileiro de energia elétrica apresenta características únicas se comparado aos demais mercados internacionais em função do desenho contratual com mecanismos que garantem a expansão equilibrada e dinâmica entre a oferta e a demanda em um mercado competitivo com baixa participação da geração térmica (CASTRO; BRANDÃO; DANTAS, 2014). Desta forma, o mercado de energia elétrica considera-se um mercado de contratos e não de energia, de longo prazo em que os agentes consumidores necessitam garantir 100% de contratação do seu consumo, sob pena de sujeição a penalidades.

Adicionalmente, ao contrário dos mercados europeus e estadunidense, o despacho é realizado de forma centralizada, com base em um modelo computacional de otimização de custos e não a partir de ofertas de preços e quantidades dos geradores e consumidores. Tal desenho se deve, sobretudo, às características específicas do mercado brasileiro, marcado pela predominância da geração hídrica. Desta forma, o preço de curto prazo da energia, definido com base no custo marginal de operação calculado pelos modelos computacionais NEWAVE, DECOMP e DESSEM, tende a ser demasiado baixo sempre que ocorrerem hidrologias normais e demasiado elevado em caso de hidrologias desfavoráveis.

Os modelos computacionais tem como objetivo otimizar curva futura dos custos de energia com base em uma série de informações sobre o Sistema Elétrico Brasileiro (SEB). A partir dessa otimização, determinar uma política de operação que minimize os custos, mantendo um nível aceitável de risco de não atender à demanda, e assim definir o custo marginal de operação (FELIZATTI et al., 2019), ou seja, são ferramentas de simulação utilizadas pelo Operador Nacional do Sistema Elétrico e por outros agentes do setor elétrico brasileiro para o planejamento e a operação do sistema elétrico. Cada modelo desempenha função específica dentro do processo de gestão da geração e distribuição de energia elétrica no Brasil. A Figura 11 demonstra o horizonte de projeção, modelo (produto) computacional para cada horizonte de tempo e o período de discretização.

Figura 11. Metodologia e horizonte que determina o Preço de Liquidação das Diferenças



Fonte: Autor, informação da (CCEE, 2023)

O NEWAVE é um modelo de planejamento da operação de médio prazo que tem um horizonte de projeção até 5 anos, utilizado principalmente para a otimização da operação do sistema interligado e para a avaliação da segurança do suprimento de energia elétrica (considerando a produção das usinas hidrelétricas e termelétricas). Um dos objetivos é conseguir definir estratégias ideais de geração hidráulica e térmica em cada estágio e retornar o valor mínimo esperado do custo de operação para o período do planejamento e tem como um dos seus principais resultados as funções de custo futuro, que é essencial para o planejamento energético, permitindo a análise de cenários e a tomada de decisões estratégicas sobre o uso dos recursos energéticos disponíveis. E essa função é essencial para o modelo de curto prazo, pois representa o impacto da utilização da água armazenada nos reservatórios. As variáveis gerais de entrada desse modelo são: dados históricos e previstos sobre energias afluentes, características das usinas (hidrelétricas e termelétricas), dados de vazões, energia armazenada dos reservatórios, cargas agregadas, custo operacional térmica e etc (CCEE, 2023).

O modelo considerado de curto prazo é o DECOMP que tem com objetivo definir as metas de geração de cada usina de cada sistema hidrotérmico sujeito a afluência estocástica de forma a minimizar o valor do custo de operação ao longo do período de

planejamento. O modelo analisa as restrições físicas e operacionais associados aos problemas de armazenamento, atendimento a demanda e etc. O modelo detalha a operação semanal do sistema, levando em consideração algumas informações como: previsões de cargas, vazões, disponibilidades, limites de transmissão entre submercados, função de custo futuro do modelo NEWAVE.

E por último tem-se o modelo de curtíssimo prazo o DESSEM, que tem como horizonte até 7 dias, e discretização de até meia-hora, com representação detalhada das características e restrições operativas das usinas hidrelétricas, usinas termelétricas e por unidades geradoras de forma individualizada. Este modelo permite uma análise mais detalhada do despacho de geração, incluindo a gestão de restrições operativas do sistema e a otimização do despacho das usinas termelétricas e hidrelétricas, levando em conta a variabilidade da demanda ao longo do dia e as condições hidrológicas.

Como definição o Preço de Liquidação das Diferenças é um mecanismo financeiro fundamental no mercado de energia elétrica, particularmente em sistemas que adotam o modelo de comercialização de energia em ambiente de contratação livre. O PLD possui como base de cálculo o Custo Marginal de Operação (CMO), e tem seus valores limitados por um preço mínimo e máximo.

No cálculo do PLD um dos fatos que mais impactam são o volume de água nos reservatórios e a previsão de vazão das bacias hidrográficas, pois a matriz energética brasileira é predominantemente hídrica. Porém leva também em consideração fatos como as demandas de energias dos agentes; implicações de transmissão entre os submercados; o custo de déficit de energia dentre outros (CASTRO; BRANDÃO, 2021).

Determinar o PLD envolve uma complexidade significativa, especialmente considerando que uma grande parcela da produção de energia elétrica do Brasil provém de usinas hidrelétricas. Assim, é fundamental que o cálculo incorpore tanto o valor imediato da utilização da água quanto o valor futuro associado ao seu armazenamento, bem como a economia gerada pela redução do uso de combustíveis em usinas termelétricas. O verdadeiro desafio reside em conceber uma solução ideal que harmonize adequadamente todos esses elementos (CCEE, 2023).

O PLD é determinado para cada subsistema de energia elétrica. Os subsistemas são divididos em 4 regiões: Sudeste e Centro-Oeste (SE/CO), Sul (S), Nordeste (NE) e a maior parte da região Norte (N), conforme demonstrado na Figura 12. O único estado que ainda não está interligado é Roraima.

Figura 12. Regiões dos Submercados de Energia Elétrica no Brasil



Fonte: Adaptado de (PREX, 2024)

Essas regiões são interconectadas pelo Sistema Interligado Nacional. A malha de transmissão que interliga os sistemas elétricos facilita a movimentação de energia entre diferentes subsistemas, que permite a obtenção de ganhos sinérgicos e aproveitamento da variação nos padrões hidrológicos das bacias. Essa conexão entre os meios de geração e transmissão assegura que o mercado seja suprido de maneira segura e econômica. A capacidade instalada de geração do SIN é composta, principalmente, por usinas hidrelétricas distribuídas em dezesseis bacias hidrográficas nas diferentes regiões do país.

As regiões Nordeste e Sul apresentaram nos últimos anos um forte crescimento na instalação de usinas eólicas, aumentando a importância dessa geração para o atendimento do mercado. As usinas térmicas, geralmente localizadas nas proximidades dos principais centros de carga, desempenham papel estratégico relevante, pois contribuem para a segurança do SIN, pois são acionadas em momentos de crise hídrica, que ocasionam a redução no nível dos reservatórios de água que abastecem as usinas hidrelétricas para

complementar a geração de energia e assegurar o atendimento futuro. O sistema de transmissão integra as diferentes fontes de produção de energia elétrica e possibilita o suprimento do mercado consumidor.

Os preços de energia elétrica brasileira são publicados com periodicidade semanal por meio da Câmara de Comercialização de Energia. No Brasil o PLD é o único preço oficial do mercado de energia elétrica, servindo para liquidar as sobras e déficits e energia entre os agentes, tanto no ambiente de comercialização livre quanto no regulado, valora as penalidades e é utilizado como balizador de preços em contratos de balcão.

Em essência, portanto, os valores assumidos pelo Preço da Liquidação das Diferenças são resultantes da hidrologia e não da interação entre preferências dos consumidores e estruturas de custos dos geradores (CASTRO; BRANDÃO; DANTAS, 2014). Tal característica se torna evidente em momentos de crise hídrica com a observada no ano de 2021, em que os reservatórios do Sudeste e Centro-Oeste, responsáveis por 70% do armazenamento de energia do país, estavam no patamar mínimo histórico de 26% (SANTOS; FIGER, 2022).

Devido a tal volatilidade, o sistema brasileiro não possui uma das principais características de um mercado competitivo, que é a capacidade de autorregulação através de ações dos agentes participantes induzidos pelos preços de mercado. Isto se deve ao fato de o preço de curto prazo não fornecer uma sinalização adequada para o aumento ou diminuição da oferta, para novos investimentos ou para desinvestimentos no setor: um preço de curto prazo elevado não é por si só indicador de necessidade de investimentos e um preço muito baixo não é indicação clara de que há excesso de capacidade ociosa (CASTRO; BRANDÃO; DANTAS, 2014). Desta forma, de acordo com os autores, há uma discrepância acentuada entre o PLD e os custos de produção da geração de energia elétrica. E esta discrepância não dá sinais econômicos para orientar as decisões de investimento e desinvestimento por parte dos geradores (CASTRO et al., 2017b).

Portanto, o PLD não fornece sinal econômico para investimentos e desinvestimentos para os agentes do mercado de energia elétrica. Como consequência direta, ele não pode servir de base para a expansão do sistema, pois é imprevisível no curto, médio e no longo prazo e, assim, constitui base inadequada para a tomada de decisões de investimento e financiamento (CASTRO et al., 2017b), bem como transações de blocos de energia livremente negociados entre os agentes do mercado.

Como evidência dessa disparidade, no mercado brasileiro as atividades comerciais e físicas são realizadas por agentes distintos: ao passo que a Câmara de Comercialização de Energia Elétrica é o operador comercial que registra as transações de compra e venda de energia e é responsável pela liquidação financeira do mercado de curto prazo de energia elétrica, o Operador Nacional do Sistema Elétrico é responsável pela coordenação e controle da operação das instalações de geração e transmissão de energia elétrica (PORTO; SACCHI,

2019).

Desta forma, a CCEE calcula o PLD com capilaridade horária para cada patamar de carga e submercado, com limitações em um preço máximo e um preço mínimo, vigentes para cada período de apuração. O cálculo do PLD é baseado no despacho ex-ante, ou seja, é apurado com as informações previstas, anteriores à operação real do sistema, considerando os valores de disponibilidades declaradas de geração e a demanda prevista para cada submercado. Na prática, adota-se uma abordagem *tight pool*, onde esse valor é definido por uma instituição independente, no caso brasileiro (ONS) e a base do preço é determinado pelo preço por custo, sem a realização dos leilões do tipo para o dia seguinte (*day-ahead* ou dentro do dia (*intra-day*) para a operação do sistema (VIANA, 2018).

Diante deste paradigma, tanto os agentes participantes do ACL quanto do ACR dependem de metodologias e ferramentas para a precificação da energia sob os diversos horizontes de curto, médio e longo prazos. Tais metodologias devem ser capazes de refletir as interações entre oferta e demanda e o prêmio do risco envolvido na comercialização da energia, trazendo menos volatilidade ao setor e permitindo a viabilidade econômica de novos empreendimentos apesar de períodos extremos como os observados em casos de crise hídrica.

Entretanto, as variáveis utilizadas no cálculo do PLD, cujo propósito principal é representar os custos operacionais do sistema nos horizontes de curto, médio e longo prazo, não conseguem refletir de maneira precisa as expectativas do mercado e não incorporam o prêmio pelo risco inerente ao processo de comercialização de energia. Esse desacoplamento é particularmente evidente nos horizontes de médio e longo prazo, onde as variáveis tradicionalmente operacionais têm menor impacto, enquanto aquelas que modelam o risco associado ao negócio tornam-se mais significativas (BASTOS; CUNHA; BARROSO, 2017).

Desta forma, mediante tal desacoplamento nos horizontes de médio e longo prazo, faz-se necessário o estabelecimento de metodologias que retornem projeções de curvas de preços que reflitam não somente aspectos intrínsecos à operação da rede, mas também o risco associado à comercialização, de forma a otimizar a modelagem dos contratos firmados entre os agentes. Nesse contexto, introduz-se o conceito de curvas forward como ferramentas capazes de capturar as expectativas mercado tais como prêmio de risco, custo de oportunidade, liquidez, concentração de mercado e volatilidade de forma a refletir com maior acurácia o preço de comercialização da energia (PILIPOVIC, 2007).

O pool de preços trata-se de uma base de dados na qual os agentes do mercado inserem suas expectativas de preço em determinados horizontes de tempo. Tal expectativa de preço reflete os riscos de mercado para a comercialização de energia e, devido a se tratar de um consentimento dos agentes participantes do setor, torna-se então um parâmetro significativo na modulação dos contratos. Vale ressaltar que, dado o setor elétrico do qual o agente é participante (geração, comercialização e consumo), da assimetria de

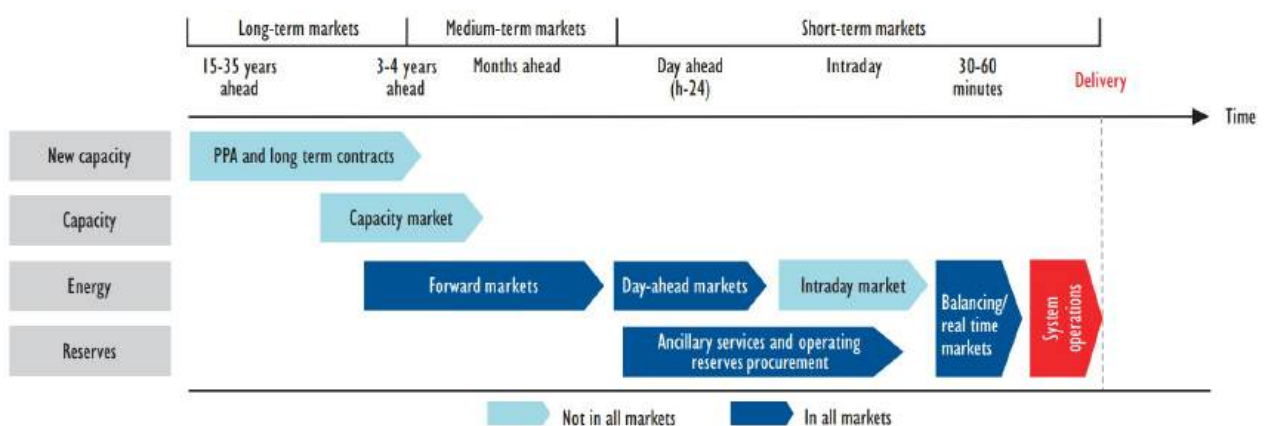
informações entre eles e das diferentes visões macro do setor elétrico, as expectativas de preço tendem a ser distintas, cabendo, desta forma, uma ponderação para a determinação de curvas razoáveis de preços que realmente reflitam um consenso do mercado (FELIZATTI; HANSEN, 2019).

2.2.2 Previsão de Preços: *Electricity Power Forecasting* vs *Curva forward*

Ao tratar-se da previsão de preços no mercado de energia elétrica, usualmente dois termos são evidenciados pela literatura especializada: o *electricity power forecasting* e a modelagem da curva *forward*. Apesar de ambos os termos guardarem uma similaridade entre si, já que tratam de mecanismos de previsão de preços da energia elétrica, diferem entre si substancialmente devido a alguns pontos: *i* – a existência de prêmios de risco que variam com o tempo de maturidade do contrato; e *ii* – a elevada incerteza na realização de previsões de preço *spot* para horizontes de tempo muito longos (LAMPS, 2021)

Neste aspecto, vale ressaltar os diferentes produtos e suas respectivas escalas de tempo no ambiente do mercado de eletricidade, conforme apresentado na Figura 13. Portanto, no longo prazo (*long-term markets*), estão usualmente presentes as estruturas relacionadas aos contratos de disponibilidade de novas unidades de geração (*PPAs – Power Purchase Agreements*). No médio prazo são firmados contratos de capacidade para as usinas já existentes, além de contratos de entrega de energia futura (contratos *forward*). No curto prazo, por sua vez, estão localizadas as operações do mercado de energia e de serviços ancilares nos horizontes *Day-Ahead Market* e o *Intraday/Real Time Market*. Portanto, considerando os diferentes produtos e diferentes horizonte de tempo de maturação, tem-se o estabelecimento das diferenças fundamentais entre os dois conceitos envolvidos na previsão de preços: *Electricity Power Forecasting* e *Curvas Forward*

Figura 13. Produtos do mercado de energia elétrica e diferentes tempos de maturação



A Curvas *forward* trata-se de curvas que registram o preço de contratos em que a negociação da energia elétrica inicia-se à frente da entrega e do consumo físico realizados por meio de balcões organizados, bolsas de energia ou mesmo contratos bilaterais entre as partes, com foco, geralmente, nos mercados de médio e longo prazo. Devido a esta característica, incorporam em sua construção aspectos inerentes do mercado, tais como o risco associado na compra ou venda de um determinado montante de energia em um determinado horizonte de tempo (VIANA, 2018).

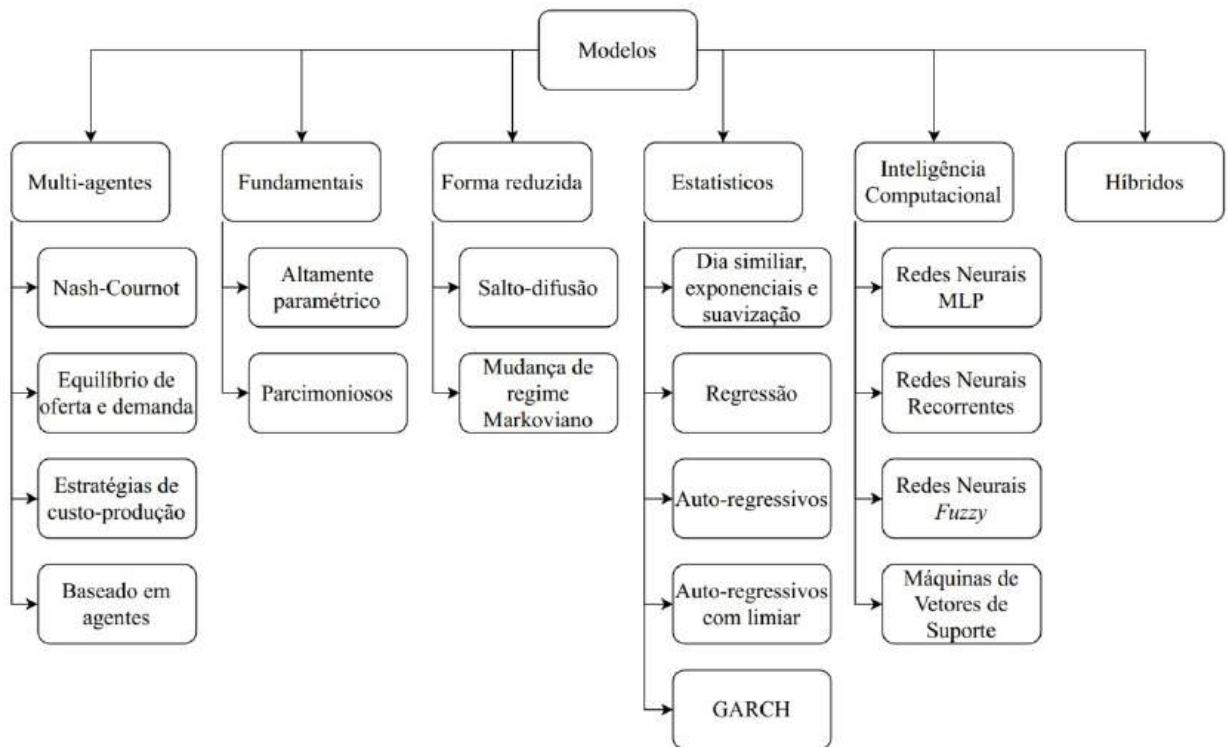
Conforme (LEME et al., 2015), a construção de uma curva forward depende das especificidades do mercado. Para commodities não estocáveis, não se pode utilizar o método padrão de precificação de commodities empregando custos de armazenamento e valores ajustados ao *spot* usando prêmios de risco. No caso de mercados de eletricidade com liquidez suficiente em bolsas de valores abertas, é possível usar os preços desses contratos padrão para a construção da curva, entretanto onde a liquidez é insuficiente, os agentes geralmente dependem da alternativa da curva de consenso.

Quando se trata de mercado de curto prazo, a previsão de energia elétrica trata-se essencialmente da previsão do preço *spot*, da energia em um determinado instante de tempo. No mercado brasileiro, como já mencionado é resultado da cadeia de resultados dos modelos computacionais utilizados pela CCEE, que estão intrinsecamente atrelados aos fundamentos operacionais e técnicos do sistema.

Quando se pesquisa sobre métodos para previsão de preço de energia elétrica no curto, médio e longo prazo, depare-se com diversas modelos. A nível mundial, a literatura reporta uma diversidade de metodologias e técnicas para definição do preço. Em (NAMETALA, 2023) é apresentado alguns métodos de previsão de preço *spot* de energia elétrica, mas alguns desse modelos também estão presentes em previsões nos horizontes de médio e longo prazo.

A Figura 14, apresenta uma sistematização dessas metodologias de previsão, com destaque para os métodos: multi-agentes; fundamentais; forma reduzida; estatísticos; inteligência computacional; híbridos. Destacando-se um crescimento nos últimos anos conforme demonstrado em seção 1.5.2, na utilização de métodos híbridos fazendo combinações entre métodos computacionais e também estatísticos, para previsões.

Figura 14. Sistematização das metodologias de previsão do preço *spot*



Fonte: (NAMETALA, 2023), pag 167.

3 Capítulo 3 - Tecnologias Utilizadas pra Otimização e Previsão

Nesse capítulo, é realizado uma revisão dos conceitos sobre (i) Redes Bayesianas Clássicas, Dinâmicas e descrição dos métodos de inferências; (ii) do Algoritmo genético que será utilizado para definir a topologia utilizada pelas Redes Bayesianas Dinâmicas para projeção; (iii) aplicação da Teoria dos Grafos na Modelagem do Algoritmo Genético e; (iv) o método de correlação aplicada nas variáveis, que serão utilizadas para previsão da curva *forward* de energia.

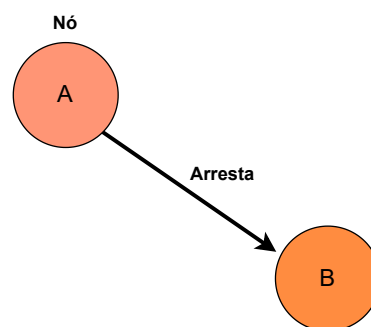
3.1 Redes Bayesianas Clássicas

A Rede Bayesiana (RB) é representado por um grafo acíclico dirigido para estruturar um conjunto de variáveis aleatórias $X = X_1, X_2, \dots, X_p$ e suas relações entre si. Os arcos direcionados representam as relações de dependência entre as suas variáveis (nós). Os principais componentes utilizados para representar uma rede Bayesiana são:

- **Vértices ou Nós:** Representam variáveis aleatórias que podem assumir um conjunto de estados distintos;
- **Arestas ou Arcos:** São direcionados e representam as relações de dependências entre os nós alvo e origem. Essa relação é dada conforme a orientação do arco;

As relações de dependência estão diretamente relacionadas à hierarquia da rede: o nó pai corresponde ao vértice do qual parte uma aresta direcionada a outro nó, denominado nó filho. Um exemplo é demonstrado na Figura 15.

Figura 15. Representação de relação hierárquica da Rede Bayesiana.



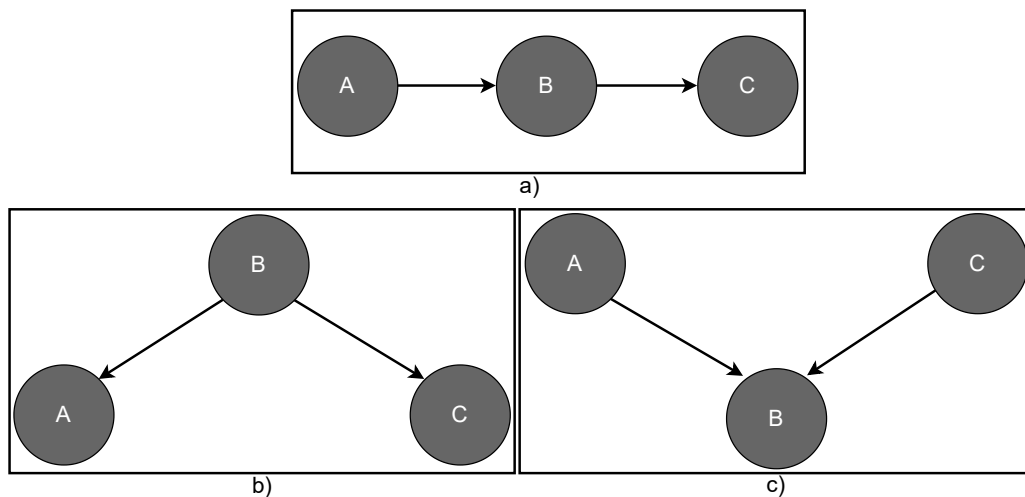
Fonte: Autor

As Redes Bayesianas destacam-se pelo uso do raciocínio probabilístico, o qual possibilita a tomada de decisão mesmo quando não existem informações suficientes para provar o acontecimento de um determinado evento (CHARNIAK, 1991).

Formalmente, uma Rede Bayesiana é um Grafo Acíclico Dirigido ou do inglês *Directed Acyclic Graph* - DAG, que é definido por um par $G = (V, E)$, onde V é um conjunto de vértices e, E é um conjunto de arestas dirigidas, no qual cada aresta é composta por um par ordenado (u, v) , tal que $u, v \in V$ e $u \neq v$, logo não existe caminho que parta de u e retorne para u , garantindo a ausência de ciclos, ou seja, o que significa que não é possível começar em um vértice e seguir uma sequência de arestas dirigidas que eventualmente retornam ao vértice inicial. Esta característica assegura que exista uma “ordem” que pode ser seguida dentro do grafo, de modo que, começando de um vértice pode percorrer o grafo sem nunca entrar em um *loop* (DUDA; HART; STORK, 2000).

A estrutura de um DAG em uma Rede Bayesiana permite inferir sobre as variáveis, dadas certas evidências. As conexões entre os vértices (variáveis) podem ser classificadas com base na direção das relações de dependência, e três tipos de conexões são particularmente importantes para a compreensão e análise destas redes que são: conexão linear, divergente e convergente, conforme demonstrado na Figura 16.

Figura 16. Tipo de conexões entre variáveis: a) conexão Linear, b) conexão Divergente e c) conexão Convergente.



Fonte: Autor

A Figura 16a) apresenta uma conexão Linear, onde é visto o fluxo de influencia entre as variáveis A , B e C . Nesse caso A influencia em B , e B por sua vez em C , ou seja, qualquer influencia em A , irá afetar B e C .

Na Figura 16b) é demonstrada a conexão divergente que ocorre quando uma variável

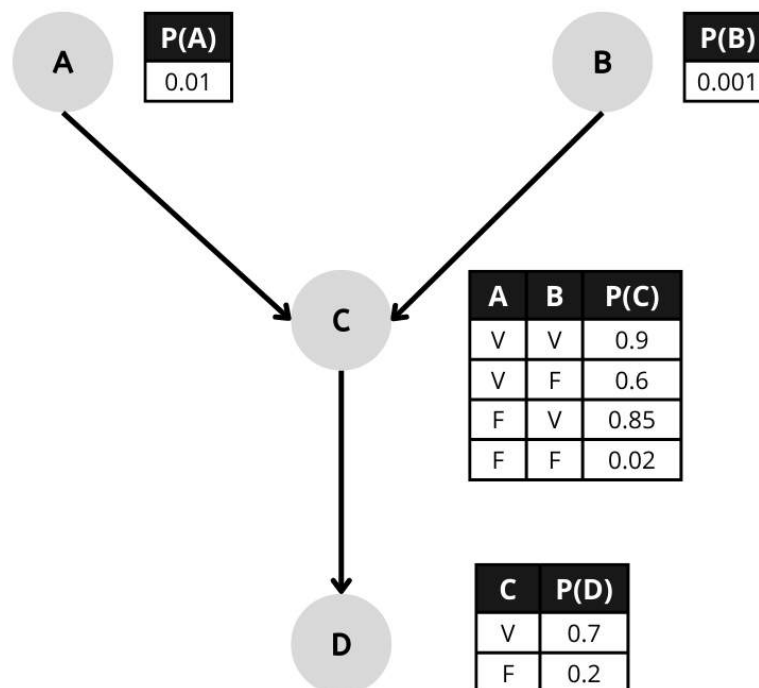
influencia duas ou mais variáveis que não tem outra conexão direta entre si. As variáveis A e C divergem de B , ou seja, a influência ocorrida em B será transmitida para todos os seus filhos.

E por fim, na Figura 16c) é demonstrado a conexão Convergente, onde as variáveis A e C convergem para B , ou seja, a variável B depende de A e C . Logo quando se sabe o estado de B , A e C não são d-separados, pois são condicionalmente dependentes dado A .

Os DAGs são especialmente úteis em situações que envolvem a modelagem de processos ou sistemas onde é essencial representar a precedência ou a dependência entre tarefas, eventos, ou estados.

O modelo de Rede Bayesiana é representado por um DAG, tal como na Figura 17, pelo qual seus vértices ou nós representam variáveis aleatórias que podem assumir um determinado conjunto de estados distintos e os relacionamento entre essas variáveis são representados através de arcos direcionados. Sendo importante ressaltar que cada nó possui uma tabela de probabilidades condicionais, a qual representa a probabilidade de tal nó assumir estados específicos com base nos estados de seus antecessores.

Figura 17. Redes Bayesianas como Gráficos Direcionados Acíclicos.



Fonte: Autor

De acordo com a Regra de Bayes, a probabilidade que dois eventos A e B

ocorram simultaneamente é representada pela Equação 3.1.

$$P(A, B) = P(A|B).P(B) = P(B|A).P(A) \quad (3.1)$$

O Teorema de Bayes consiste em considerar o evento A como sendo uma “causa” do evento B , sendo atribuídas probabilidades deste evento acontecer na ocorrência de B , conforme equação 3.2.

$$P(A, B) = \frac{P(B|A).P(A)}{P(B)} \quad (3.2)$$

onde: $P(A|B)$ é a probabilidade a posteriori (probabilidade condicionada) de A condicional a B ; $P(B|A)$ é a probabilidade a posteriori (probabilidade condicionada) de B condicional a A e $P(A)$ e $P(B)$ são as probabilidade a priori de A e B .

Cada nó possui uma Tabela de Probabilidades Condicionais (CPDs) contendo as probabilidades de tal nó assumir seus estados específicos. A tabela de CPDs é baseada nos nós, cujos arcos apontam para o nó em questão, também chamados de nós pais. Assim a distribuição conjunta de todas as variáveis de uma Rede Bayesiana é definida pela Equação 3.3.

$$P(X) = P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | Pais(X_i)) \quad (3.3)$$

Este modelo de Rede Bayesiana, que representa as relações de dependências entre um conjunto de variáveis, permite a realização de análises, perguntas e hipóteses a partir do processo denominado inferência. O Teorema de Bayes e suas fórmulas de probabilidade condicional e probabilidade condicional conjunta são utilizados para realizar o cálculo dessas inferências. A Equação 3.4 representa a regra generalizada de Bayes.

$$P(A = v_i | B) = \frac{P(B|A = v_i).P(A = v_i)}{\sum_{j=1}^n P(B|A = v_k).P(A = v_k)} \quad (3.4)$$

Realizar uma inferência a partir de um modelo corresponde ao processo de encontrar a distribuição de probabilidades condicionais sobre um conjunto de variáveis. Durante este processo também é possível aplicar conhecimentos preexistentes relacionados ao domínio do modelo com a utilização de evidências, assumindo que as probabilidades de certos eventos ocorrerem são de 100%. Com o uso de evidências, as redes Bayesianas tornam-se uma ferramenta poderosa de análise de cenários, aliando métodos estatísticos e o conhecimento de especialistas.

Inicialmente, o processo de construção de redes Bayesianas era realizado de maneira empírica por especialistas de determinado contexto. Esse processo é composto por duas

etapas distintas, definição de estrutura e estimação de parâmetros. Atualmente, ambos contam com algoritmos que permitem a automatização de todos os processos.

A Rede Bayesiana é de natureza qualitativa, pois mesmo que utilize das probabilidades condicionais para sua modelagem, estas podem ser estimadas subjetivamente diante a expectativa de ocorrência de um evento, de forma a invalidar as regras de cálculo de probabilidades (JENSEN, 1996). Logo, técnicas de discretização são necessárias para o processamento de certos dados quantitativos.

Uma série de métodos meta-heurísticos pode ser utilizada para realizar uma busca automática pela topologia ideal de Rede Bayesiana. O método de *Hill Climb Search* é possível realizar uma busca local e, dado uma função de pontuação, como a *K2 Score* que ranqueia cada topologia analisada com o objetivo de encontrar o máximo global no espaço de busca que contém as possíveis combinações de nós e seus respectivos arcos.

Para realizar o aprendizado de Redes Bayesianas é necessário utilizar uma função de ranqueamento do método *BD Score* proposta por (COOPER; HERSKOVITS, 1991). Neste método é utilizada a função *K2 Score*. A função de ranqueamento é utilizada para calcular $P(B|T)$, a distribuição de probabilidades posterior, dado uma distribuição das possíveis redes B e sendo condicionadas ao conjunto de dados T . A função *K2 Score* é representada na equação 3.5.

$$K2(B, T) = \log(P(B)) + \sum_{i=1}^n \sum_{j=1}^{q_i} \left(\log\left(\frac{(r_i - 1)!}{(N_{ij} + r_1 - 1)}\right) \right) + \sum_{k=1}^{r_i} \log(N_{ijk}!) \quad (3.5)$$

Na Equação 3.5 tem-se r_i sendo o número finito de estados para a variável X_i , q_i o número de possíveis combinações de pais para X_i e N o total de instâncias.

As Redes Bayesianas Clássicas não possuem habilidades de representar seus nós de maneira temporal. Para superar essa limitação, foram introduzidas as Redes Bayesianas Dinâmicas, que incorporam a dimensão temporal ao modelo probabilístico.

3.1.1 Métodos de discretização

Devido à natureza das redes Bayesianas, tanto clássicas como dinâmicas, é necessário trabalhar apenas com dados discretos. A seguir serão apresentadas algumas técnicas de discretização distintas, onde cada técnica possui características distintas e representa a distribuição dos dados de maneira diferente.

3.1.1.1 Discretização em quantis

A discretização em quantis permite que o conjunto de dados sejam discretizados em duas formas. Lembrando que esse método necessita que os dados sejam inicialmente ordenados.

1. Divide os dados de maneira que o intervalo que cada faixa seja igual para todos os n intervalos selecionados, ou seja, se o número n for igual a 25% dos dados, logo o número máximo de intervalo será 4. Este método é ideal para dados que possuem distribuição uniforme.
2. Divide os dados de maneira que cada grupo tenha, aproximadamente, a mesma quantidade de amostras, ou seja, dividir as amostras de forma que se interprete a sua variação (diferença entre a amostra atual e a anterior quando estas estão ordenadas no tempo).

3.1.1.2 Discretização por clusterização *K-means*

É um método de aprendizagem de máquina não supervisionado, que extrai padrões de dados não rotulados. Particiona os dados em k grupos (a definição do número de grupos é baseada em conhecimento do domínio ou critérios estatísticos) distintos com base na similaridade dos valores.

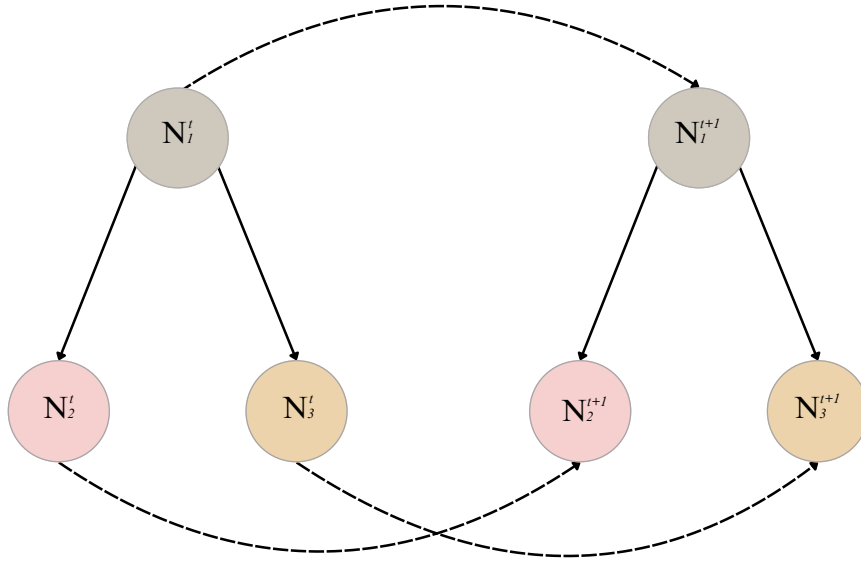
A ideia básica por trás do uso do algoritmo *k-means* para discretização é agrupar valores contínuos em k categorias, em que cada categoria representa um *cluster* formado pelo algoritmo. Cada *cluster* é definido de maneira a minimizar a variância dos dados, tornando os valores dentro de cada grupo os mais similares possível, ao mesmo tempo em que maximiza a diferença entre os grupos.

3.2 Redes Bayesianas Dinâmica

Devido a limitações nos modelos de Redes Bayesianas clássicas, que também são conhecidos como modelos baseados em variáveis, diversos modelos especializados foram sendo desenvolvidos ao longo do tempo. As Redes Bayesianas Dinâmicas surgem devido à necessidade de introduzir um fator temporal aos modelos, permitindo a realização de inferências em um tempo t , em que algumas variáveis possam depender dos valores que ocorreram em um tempo $t - 1$.

As Redes de Bayesianas Dinâmica (DBN) são formadas por uma sequência de Redes Bayesianas Clássicas interligadas, onde cada instante de tempo t é representado por uma dessas instâncias, correspondendo a um estado do sistema. A dependência temporal entre esses estados é modelada por arcos temporais, que conectam os nós ao longo do tempo, conforme demonstrado da Figura 18, em um exemplo de rede com três nós, onde o i -ésimo nó N_i^t no tempo t estabelece uma relação com seu correspondente no instante futuro $N_i^{(t+1)}$, capturando assim a evolução do sistema ao longo do tempo.

Figura 18. Exemplo da relação temporal entre os nós da rede bayesiana dinâmica.



Fonte: Autor

Para a modelagem de modelos com fatores temporais, observa-se como as variáveis se comportam ao longo de diferentes períodos de tempo, assumindo que um evento no passado pode influenciar a ocorrência de um evento futuro. Cada variável recebe um tempo t associado, assim, tendo-se o conjunto de n variáveis $\{X_0, X_1, \dots, X_n\}$, será usado X_t para prever o valor de X_{t+1} .

É possível descrever uma DBN pela função de probabilidades da sequência de T variáveis $X = \{X_0, X_1, \dots, X_n\}$ de estado oculto, as quais são resultado da função de probabilidade e não diretamente observadas, e de T variáveis $Y = \{Y_0, Y_1, \dots, Y_n\}$ de observação, em que T representa o limite de tempo para o evento em questão, e pode ser representado pela Equação 3.6.

$$Pr(X, Y) = \prod_{i=1}^{T-1} Pr(X_t | X_{t-1}) \prod_{i=0}^{T-1} Pr(Y_t | X_t) Pr(X_0) \quad (3.6)$$

De acordo com (MIHAJLOVIĆ; PETKOVIC, 2001), para a composição de uma DBN são necessários três parâmetros principais:

- Estados de transição $Pr(X_t | X_{t-1})$, os quais especificam as dependências temporais entre os estados.
- Observações $Pr(Y_t | X_t)$, as quais especificam as dependências dos nós observados em relação aos demais nós do instante t .

- Distribuição do estado iniciais $Pr(X_0)$, o qual representa a distribuição de probabilidade inicial no começo do processo.

Como já mencionado anteriormente, a RBD é amplamente utilizada para prever a probabilidade de eventos futuros ocorrerem, analisando fatores do passado. No entanto, seu processo de inferência é mais complexo se comparado com as RBs clássicas. Será mencionado três métodos de inferência:

1. O método da Filtragem, onde calcula-se a distribuição a posteriori do estado atual, dada toda evidência observada até o momento.
2. O método de Predição, que é calculada a distribuição a posteriori do estado futuro, dada toda evidência observada até o momento e.
3. O método de Suavização que calcula a distribuição a posteriori do estado passado, dado toda a evidência observada até o momento.

3.3 Algoritmo Bioinspirado

Algoritmos bioinspirados são técnicas de resolução de problemas que se baseiam em processos ou fenômenos naturais observados na biologia e na natureza. Essas técnicas procuram replicar ou se inspirar em mecanismos encontrados em sistemas biológicos, como evolução, comportamento animal, sistemas imunológicos entre outros, para resolver problemas complexos de diferentes áreas.

A ideia por trás dos algoritmos bioinspirados é aproveitar as soluções eficazes que a evolução e a adaptação natural desenvolveram ao longo de milhões de anos e aplicá-las a problemas humanos. Essas técnicas frequentemente se destacam em problemas de difícil resolução por abordagens tradicionais.

3.3.1 Algoritmos Genéticos

O Algoritmo Genético (AG) é inspirado pela teoria da evolução de Darwin, e trabalham com populações de soluções candidatas, ou seja, o AG é um algoritmo de otimização probabilística baseado na teoria da seleção natural, onde os indivíduos mais aptos sobrevivem e suas características são propagadas para às próximas gerações, produzindo soluções melhoradas.

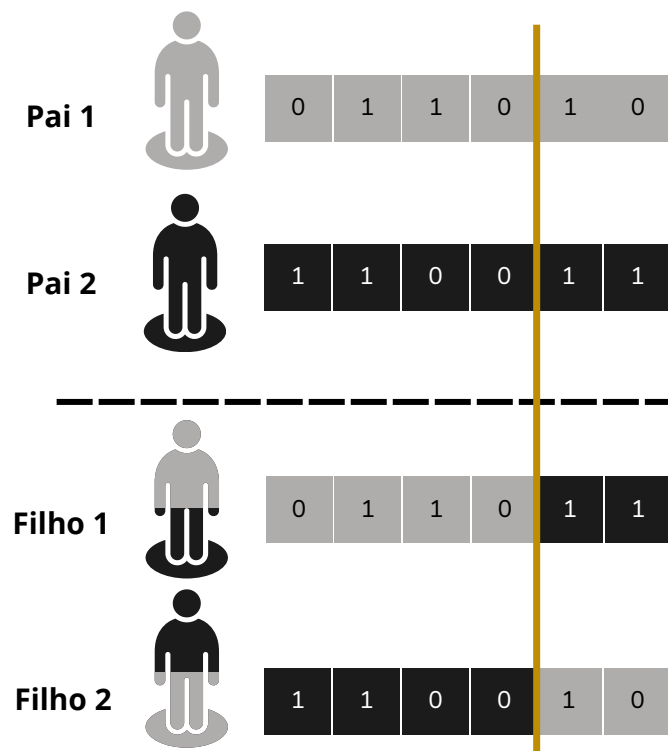
O AG resolve problemas de otimização introduzindo múltiplas soluções prováveis no espaço de busca e tem como objetivo encontrar a melhor solução aplicando operadores genéticos como seleção, *crossover* e mutação, que evoluem ao longo de várias gerações para encontrar melhores soluções (BRAGA; CARVALHO; LUDEMIR, 2000).

Cada solução candidata é representada por um cromossomo, geralmente formada por um vetor ou sequência de valores que codificam a solução. A forma como o cromossomo é codificado depende do problema específico. O processo começa com a criação de uma população inicial de indivíduos gerados aleatoriamente, ou usando algum conhecimento prévio do problema.

Cada indivíduo na população é avaliado quanto à sua aptidão (*fitness*), com base em uma função de avaliação que quantifica o quão boa é a solução representada por esse indivíduo. Indivíduos com maior aptidão têm uma maior probabilidade de serem selecionados para reprodução, seguindo o princípio de “sobrevivência dos mais aptos”. Diferentes estratégias de seleção, como a roleta, ranking ou torneio, podem ser utilizadas.

Os indivíduos selecionados são combinados por meio do *crossover*, que envolve a troca de características entre dois pais que serão herdados pelas gerações futuras. O objetivo é explorar diferentes combinações de características, e esse cruzamento é demonstrado na Figura 19.

Figura 19. Operador de *crossover* entre indivíduo binários.



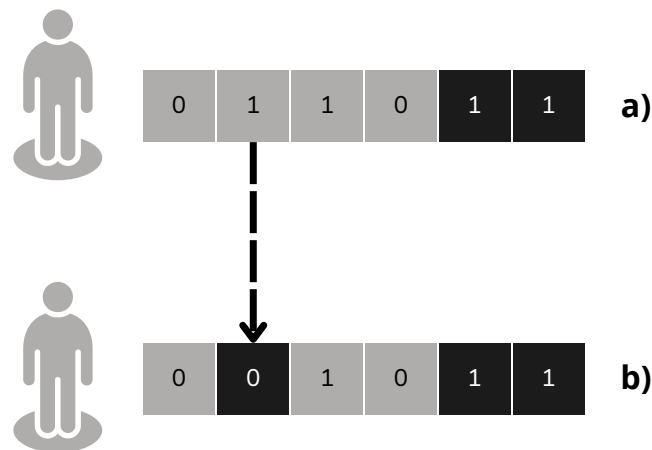
Fonte: Autor

O operador de mutação é aplicado aos descendentes após o processo de *crossover*, promovendo alterações em um indivíduo sem a necessidade de interação com um parceiro. A taxa de mutação é definida para cada indivíduo com base em uma probabilidade

preestabelecida. No entanto, é crucial que essa taxa seja adequadamente ajustada, pois uma taxa muito alta pode tornar a busca aleatória, levando à perda das características herdadas dos pais; por outro lado, uma taxa muito baixa pode restringir a busca a uma vizinhança limitada, aumentando o risco de convergência para um mínimo local (RUSSELL; NORVIG, 2013).

Logo um dos objetivos do operador de mutação é evitar que o algoritmo fique preso no problema de mínimo local, pois o operador introduz uma aleatoriedade no indivíduo. Na Figura 20, o indivíduo (a) representa um indivíduo binário antes da aplicação do operador de mutação, enquanto o indivíduo (b) ilustra o novo indivíduo resultante após a mutação. No caso de indivíduos binários, a operação de mutação consiste simplesmente em inverter o valor de um gene, alterando '1' para '0' ou vice-versa.

Figura 20. Operador de mutação binário entre indivíduo binário.



Fonte: Autor

A nova população é composta por indivíduos resultantes do *crossover* e da mutação. Essa nova população substitui a geração anterior. O processo de seleção determina quais serão os indivíduos da população que poderão participar da fase de reprodução, e esses indivíduos são selecionados de acordo com seu índice de aptidão. Existem alguns métodos de seleção como: Método da Roleta, Torneio e Ranking.

Um dos métodos de seleção mais amplamente utilizados é o método da Roleta, cujo funcionamento pode ser descrito da seguinte maneira: cada indivíduo tem associado um valor índice de aptidão, e esse indivíduo ocupa uma parte da roleta condizente com o valor de seu índice. Dessa maneira o indivíduo com maior aptidão (ocupa uma maior área na roleta) tem uma probabilidade maior de ser selecionado, ao contrário do indivíduo com menor valor de aptidão. Uma das características desse método é que qualquer indivíduo poderá ser selecionado.

O método de ranking organiza toda a população em ordem decrescente com base na nota de aptidão de cada indivíduo e, em seguida, seleciona os n indivíduos mais aptos. Essa técnica tende a excluir os indivíduos menos aptos, reduzindo suas chances de serem selecionados.

O método de seleção por torneio percorre todos os indivíduos da população em pares, comparando seus índices de aptidão e selecionando sempre o mais apto. Nesse método, o pior indivíduo nunca será escolhido, pois a seleção ocorre sempre entre dois indivíduos, privilegiando aquele com maior aptidão. Uma das principais vantagens desse método é a facilidade em percorrer toda a população, o que reduz a probabilidade de uma convergência prematura e mantém a diversidade da população por mais gerações, evitando a estagnação em ótimos locais.

O número de gerações ou evoluções em um algoritmo genético é determinado pelo critério de parada. Esse critério pode ser baseado na perda de diversidade, ou seja, quando os indivíduos gerados após um determinado número de gerações apresentam características similares ou por um número fixo de gerações preestabelecido (BRAGA; CARVALHO; LUDEMIR, 2000).

O algoritmo continua a evoluir a população por um número específico de gerações ou até que um critério de parada seja atingido, como a obtenção de uma solução satisfatória ou o alcance de um número máximo de iterações. Os Algoritmos Genéticos são amplamente utilizados em problemas de otimização, reconhecimento de padrões, aprendizado de máquina e diversas outras áreas. Eles se destacam especialmente em problemas onde o espaço de soluções é complexo e não linear, dificultando a busca por soluções ótimas devido à presença de múltiplos ótimos locais.

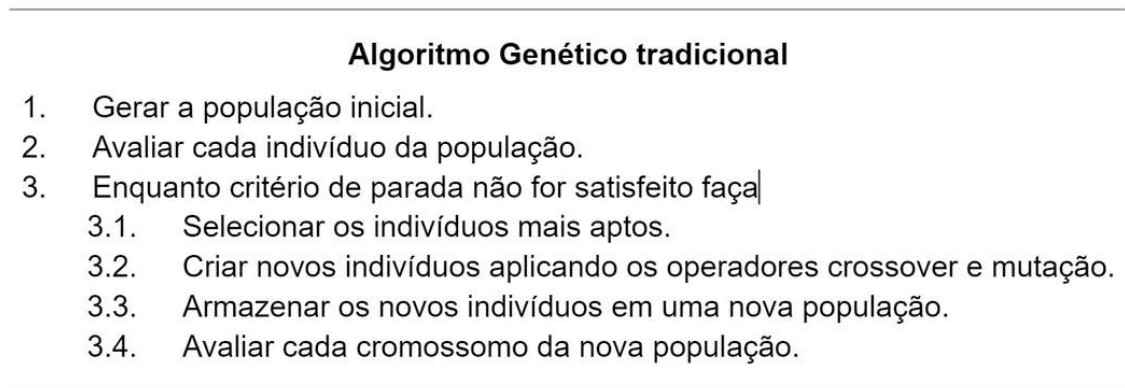
Vale ressaltar que os AG's não garantem encontrar a solução globalmente ótima, mas são capazes de fornecer boas aproximações em um tempo computacional viável, tornando-se uma abordagem eficiente para a resolução de problemas complexos e de alta dimensionalidade. Além disso, sua capacidade de explorar amplamente o espaço de soluções os torna uma alternativa poderosa para problemas onde métodos determinísticos tradicionais podem falhar ou serem computacionalmente inviável.

Os elementos básicos que compõem a estrutura de algoritmos genéticos são: o cromossomo, indivíduos, população, geração e função de aptidão. Os cromossomos são sequências de genes que representam as características dos indivíduos. Os indivíduos são soluções potenciais que são avaliadas pela função de aptidão. A população é um conjunto de indivíduos que evolui ao longo das gerações, por meio do cruzamento e mutação, em busca de soluções cada vez melhores.

O fluxo de um algoritmo genético é composto por diversas etapas, conforme demonstrado na Figura 21. Na etapa 1, ocorre a geração inicial da população, que pode

ser criada de forma aleatória ou baseada em dados reais do conjunto de treinamento. Na etapa 2, cada indivíduo da população é avaliado, e seu índice de aptidão é calculado.

Figura 21. Fluxograma do algoritmo genético.



Fonte: (BRAGA; CARVALHO; LUDEMIR, 2000).

Em seguida, nas etapas 3.1 e 3.2, os indivíduos mais aptos são selecionados, de acordo com o método de seleção definido para o problema. Esses indivíduos serão utilizados para gerar uma nova população por meio dos operadores de cruzamento e mutação. Após a criação dos novos indivíduos, na etapa 3.3, eles formam uma nova população, e cada um é novamente avaliado para calcular seu índice de aptidão, conforme descrito na etapa 3.4. Por fim, o critério de parada é avaliado. Se esse critério não for satisfeito, o algoritmo retorna à etapa 3. O critério de parada pode ser baseado em um número pré-definido de iterações ou na perda de diversidade da população.

3.4 Aplicação da Teoria dos Grafos na Modelagem do Algoritmo Genético

O algoritmo genético é uma classe de algoritmos de busca e otimização inspirados pelo processo de evolução natural. Eles operam sobre uma população de soluções candidatas (indivíduos) e aplicam operadores genéticos, como seleção, cruzamento e mutação, para evoluir a população em direção a soluções melhores. Grafos desempenham um papel fundamental nesse tipo de otimização, uma vez que podem ser empregados para modelar e representar problemas complexos de otimização.

Um grafo é uma estrutura de dados que consiste em nós (também chamados de vértices) e arestas (também chamadas de ligações). Os nós representam objetos, enquanto as arestas representam as relações entre os objetos.

Um grafo G é uma tripla (V, E, I) , onde V e E são conjuntos finitos e $I \subset V \times E$ satisfaz $1 \leq N(\{v \in V; (v, e) \in I\}) \leq 2$, para qualquer $e \in E$. Chama-se $V(G)$, $E(G)$ e I de vértices, arestas e relação de incidência do grafo G , respectivamente. Nota-se, pelo número de elementos do conjunto $I_e = \{v \in V; (v, e) \in I\}$, que existem dois tipos de arestas. Se $N(I_e) = 1$, isto é, se existe apenas um vértice incidente à aresta e , diz-se que é um laço. Quando o grafo não possui laços nem arestas múltiplas, isto é, dois vértices ligados por mais de uma aresta, o grafo é dito simples. Vértices que não se ligam a nenhum outro são ditos isolados. Convenciona-se desenhar um grafo representando cada vértice por um círculo e as arestas por linhas ligando estes círculos.

Os grafos podem ser usados para representar uma variedade de relacionamentos, incluindo relacionamentos entre pessoas, relacionamentos entre cidades e relacionamentos entre genes.

Um dos aspectos cruciais em algoritmos genéticos é a representação das soluções. Grafos podem ser usados para codificar soluções de problemas complexos. Cada nó do grafo pode representar um componente da solução, e as arestas podem indicar relações, restrições ou conexões entre esses componentes. Isso permite que o espaço de busca seja mapeado para um espaço de grafos, simplificando a manipulação das soluções e a aplicação de operadores genéticos.

No contexto de algoritmos genéticos, o crossover é uma operação que combina características de dois indivíduos para criar descendentes. Em problemas que envolvem estruturas de rede, como redes de distribuição, redes de comunicação ou rotas de veículos, os grafos podem ser usados para realizar crossovers. Por exemplo, os nós e as arestas de dois grafos parentais podem ser combinados para criar um descendente que herda as características de ambos os pais.

O cruzamento é um operador genético que é usado para criar novas soluções a partir de soluções existentes. O objetivo do cruzamento é criar soluções que sejam melhores do que as soluções existentes. Existem muitos tipos diferentes de cruzamento. Alguns dos tipos de cruzamento mais comuns são:

- Cruzamento ponto único: O cruzamento ponto único é o tipo de cruzamento mais simples. O cruzamento ponto único escolhe um ponto aleatório no cromossomo e troca os genes à esquerda e à direita do ponto.
- Cruzamento ponto múltiplo: O cruzamento ponto múltiplo é semelhante ao cruzamento ponto único, mas escolhe vários pontos aleatórios no cromossomo.
- Cruzamento uniforme: O cruzamento uniforme escolhe cada gene do cromossomo com uma probabilidade igual.

- Cruzamento por faixa: O cruzamento por faixa escolhe uma faixa aleatória do cromossomo e troca os genes dentro da faixa.
- Cruzamento por restrição: O cruzamento por restrição aplica restrições ao cruzamento. Por exemplo, o cruzamento por restrição pode impedir que o cruzamento troque genes que são importantes para a solução.

O tipo de cruzamento mais adequado para um problema específico dependerá das características do problema. Por exemplo, o cruzamento ponto único é geralmente uma boa escolha para problemas com cromossomos curtos, enquanto o cruzamento ponto múltiplo é geralmente uma boa escolha para problemas com cromossomos longos.

A avaliação da aptidão das soluções candidatas é uma parte central dos algoritmos genéticos. Grafos podem ser utilizados para modelar relações de custo, distância, tempo, etc., entre os elementos da solução. Ao calcular a aptidão, é possível analisar a estrutura do grafo para determinar o desempenho da solução em relação ao objetivo de otimização.

Problemas de otimização combinatória, como o problema do caixeiro-viajante, coloração de grafos e programação de horários, podem ser abordados utilizando algoritmos genéticos que operam sobre representações de grafos. A estrutura de grafos permite explorar a combinatoriedade dos problemas de maneira eficaz, auxiliando na busca de soluções ótimas ou aproximadas.

Os grafos oferecem uma estrutura poderosa para a resolução de problemas de otimização em várias áreas. Ao modelar problemas complexos com grafos, é possível explorar propriedades específicas da estrutura para desenvolver algoritmos eficientes e encontrar soluções aproximadas ou mesmo ótimas. Com a combinação certa de técnicas de modelagem e algoritmos, os grafos se tornam ferramentas essenciais na resolução de problemas complexos de otimização.

3.5 Método de Correlação: Coeficiente de Correlação de Pearson

O método de correlação é empregado para identificar as variáveis mais relevantes na previsão dos preços de mercado da energia elétrica, permitindo a seleção daqueles fatores que possuem maior impacto na formação dos preços. Com base na listagem das variáveis mais significativas, foi possível estruturar a metodologia utilizada para analisar e estabelecer as relações entre elas, garantindo uma abordagem mais robusta e fundamentada na modelagem preditiva.

Essencialmente, a análise de correlação busca identificar a existência de uma associação entre duas variáveis, ou seja, o grau de interdependência entre uma variável dependente e uma independente. No entanto, essa análise deve ser conduzida com rigor,

pois a correlação apenas evidencia uma variação simultânea entre as variáveis, sem necessariamente indicar uma relação de causalidade. Isso ocorre porque outros fatores não considerados na análise podem influenciar o comportamento da variável dependente, tornando essencial uma interpretação criteriosa dos resultados obtidos.

Nesse aspecto, será utilizado o coeficiente de correlação de Pearson como medida da correlação entre as variáveis, pois trata-se de um índice adimensional com valores situados entre -1.0 e 1.0 que reflete a intensidade de uma relação linear entre dois conjuntos de dados. Portanto, valores iguais a 1.0 representam uma correlação perfeita positiva entre as duas variáveis. Por outro lado, valores iguais a -1.0 significam uma correlação perfeita negativa. Adicionalmente, um valor igual a 0 indica não haver uma relação linear entre as variáveis, embora possa existir outra dependência não linear entre as mesmas (MORETTIN; SINGER, 2022).

O cálculo do coeficiente de Pearson é apresentado na equação 3.7.

$$Corr(X, Y) = \frac{Cov(X, Y)}{Var(X)Var(Y)} \quad (3.7)$$

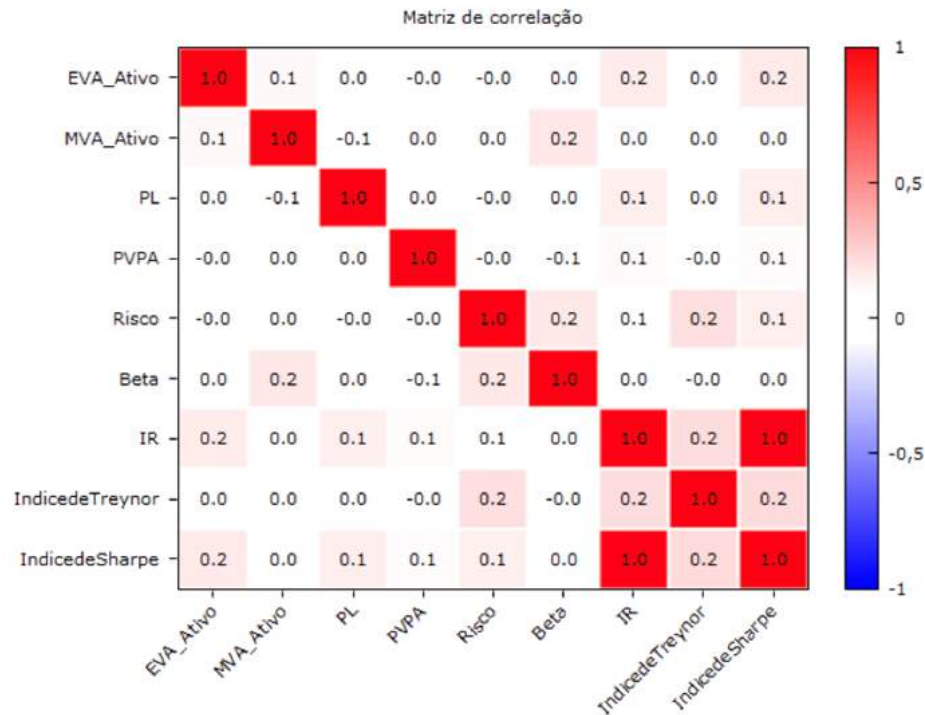
Em que:

- Covariância entre as variáveis X e Y ;
- Variância de X ;
- Variância de Y .

Após o cálculo dos coeficientes de Pearson, é construída a matriz de correlação, uma estrutura bidimensional que apresenta a relação entre as diversas variáveis analisadas. Essa matriz é uma ferramenta valiosa para identificar possíveis interdependências entre as variáveis envolvidas. Essa abordagem permite visualizar padrões de correlação, facilitando a seleção de variáveis relevantes para a modelagem preditiva, conforme apresentado um exemplo na Figura 22. Além disso, a matriz de correlação serve como um ponto de partida para a análise estatística dos dados históricos, proporcionando uma compreensão mais clara das inter-relações entre os fatores considerados (SCHMULLER; BATISTA, 2019).

A análise de correlação, será aplicada na Seção 5, para avaliar o grau de dependência entre uma lista de variáveis significativas para a projeção de preços e os valores históricos da energia no mesmo período. Esse procedimento permitirá identificar quais variáveis apresentam maior influência na formação dos preços. Com os coeficientes de correlação calculados, será construída a matriz de correlação, que servirá como base para a seleção das variáveis mais relevantes. Essa matriz fornecerá subsídios essenciais para a implementação do módulo de Redes Bayesianas Dinâmicas, garantindo uma modelagem mais estruturada e fundamentada.

Figura 22. Exemplo de matriz de correlação aplicada à análise de risco em mercados financeiros.



Fonte: Autor

3.5.1 Levantamento das variáveis

Nesta seção, será realizado o levantamento das variáveis que possuem relação com a precificação da energia nos horizontes de médio e longo prazo. A seleção dessas variáveis será baseada em sua relevância e impacto potencial nos preços da energia, compondo a base de dados utilizada para as análises subsequentes.

O processo de seleção envolve uma revisão da literatura e a análise de dados históricos, garantindo que as variáveis escolhidas sejam pertinentes, representativas e adequadas aos objetivos da modelagem preditiva. Que essas variáveis consiga capturar os principais fatores que influenciam a formação dos preços de energia ao longo do tempo.

Levando em consideração a importância da seleção de variáveis que apresentem uma influência direta na previsão do preço da energia, faz-se necessário a adoção de uma métrica que possa captar de forma assertiva a relação intrínseca entre as variáveis e o preço verificado da energia em um determinado horizonte de tempo. Portanto, conforme definido na seção 3.5, foi determinada a utilização do coeficiente de Pearson como ferramenta para explorar o nível de correlação entre variáveis prospectadas e o preço da energia de forma a subsidiar os algoritmos de previsão.

Na Seção 5.1, será apresentado o estudo de correlação entre as variáveis, inicialmente

focado no submercado de energia Sudeste/Centro-Oeste, com possibilidade de expansão para os demais submercados. Será analisado o comportamento das séries históricas das variáveis selecionadas e, posteriormente, será demonstrada a correlação dessas variáveis com a série histórica do preço de energia fornecido pela empresa de consultoria Dcide (DCIDE, 2024) (projeções de médio e longo prazo), considerando sua ampla aceitação entre os agentes do mercado, o que a torna uma referência no segmento.

Na Figura 23, são apresentadas as variáveis que, que foram definidas como objeto de estudo de correlação, além das respectivas fontes de extração dos dados.

Figura 23. Variáveis prospectadas pelo estudo de correlação.



Fonte: Autor

4 Capítulo 4 - Metodologia

Neste capítulo, serão apresentados e explicados os seguintes módulos: (i) Aquisição dos Dados, (ii) Análise e Seleção dos Dados e (iii) Previsão de Energia. Esses módulos compreendem as etapas da metodologia proposta para a geração de cenários da curva de preço de energia.

4.1 Arquitetura do Método Proposto

A previsão de preço de energia elétrica no mercado brasileiro enfrenta diversos desafios, principalmente devido às características particulares do sistema elétrico do país e à volatilidade inerente a esse mercado. A matriz energética brasileira é predominantemente hidrelétrica, sendo fortemente influenciada por variáveis meteorológicas. Além disso, o crescimento das fontes renováveis, como solar, fotovoltaica e eólica, adiciona novas dinâmicas ao setor, tornando a modelagem ainda mais complexa. Esses fatores impactam diretamente a formação do preço da energia, tornando sua previsão um grande desafio, especialmente quando se considera a influência de variáveis naturais.

O Preço de Liquidação das Diferenças é a principal referência para precificação no mercado brasileiro de energia elétrica. No entanto, sua aplicação para previsões de médio e longo prazo torna-se impraticável devido à complexidade dos fatores envolvidos e à necessidade de modelos mais robustos para capturar a incerteza inerente ao setor.

Diante do exposto, propõem-se uma metodologia de previsão da curva *forward* de energia para o horizonte de médio e longo prazo, que são até 4 e 10 anos respectivamente. Logo a curva *forward* representa o preço de energia em um determinado tempo no futuro. O preço de energia é influenciado por diversas variáveis como climáticas, Econômicas, Mercadológicas, Operacionais e Planejamento, ou seja, determinar o preço de energia elétrica a médio e longo prazo demanda o uso de diversas variáveis, tanto endógenas quanto exógenas.

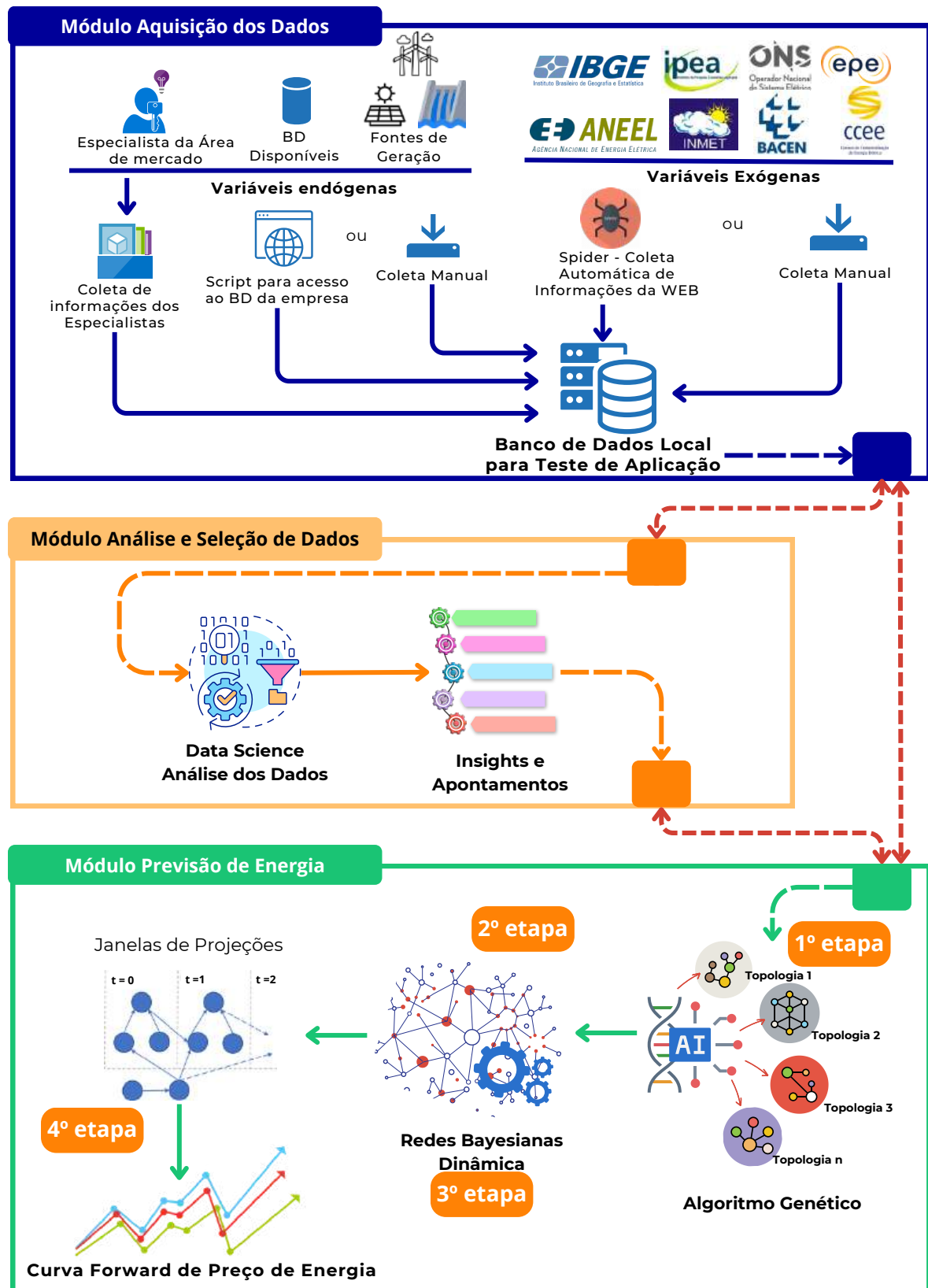
O preço da energia elétrica é influenciado por diversas variáveis, conforme apresentado na Seção 3.5.1, onde é analisada a correlação entre o preço e fatores climáticos, econômicos, mercadológicos, operacionais e de planejamento. Dessa forma, a determinação do preço da energia elétrica no médio e longo prazo requer a consideração de múltiplas variáveis, tanto endógenas quanto exógenas. Um dos principais desafios deste trabalho foi definir um modelo capaz de representar a relação e o impacto das variáveis no preço da energia elétrica. Durante o estudo, observou-se que a formação do preço é baseada em diferentes cenários, os quais são influenciados por diversas variáveis, algumas das quais

podem ou não apresentar correlação ao longo do tempo.

Diante disso, foi proposta uma metodologia que, a partir da análise dos dados, identifica e modela a relação entre essas variáveis, gerando um modelo relacional que serve como entrada para o método estatístico de previsão. Esse método se destaca pela sua flexibilidade, permitindo a incorporação de novas variáveis e ajustes baseados em informações adicionais, o que contribui para aprimorar a precisão das previsões.

A arquitetura da metodologia proposta é apresentada na Figura 24. Em um nível macro, a metodologia é estruturada em três módulos principais: Módulo de Aquisição de Dados, Módulo de Análise e Seleção de Dados e Módulo de Previsão de Energia.

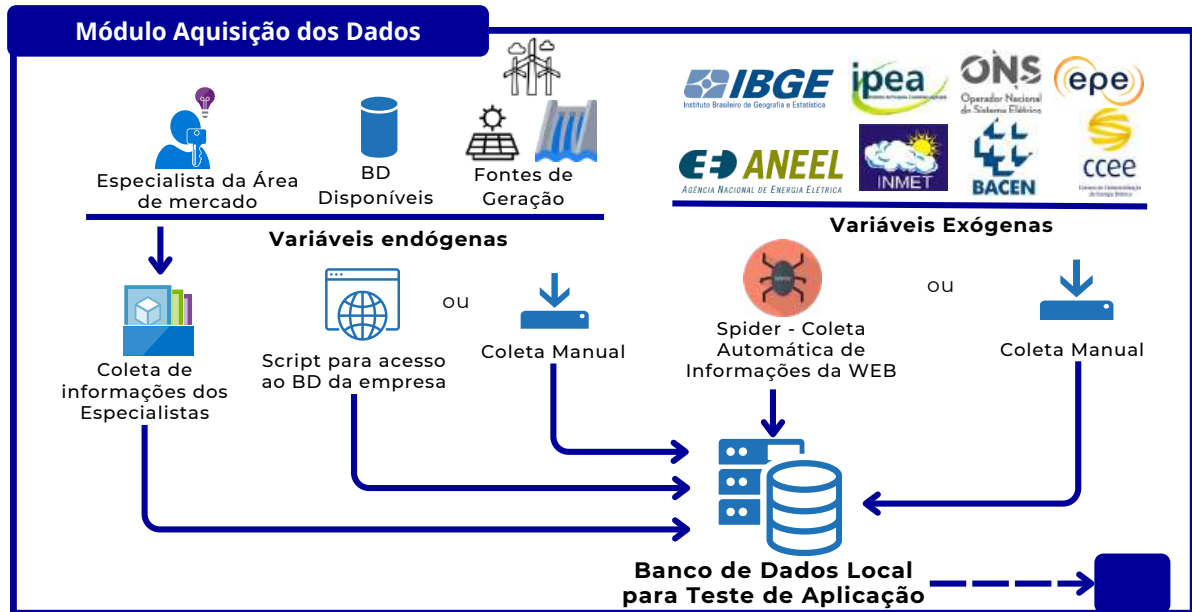
Figura 24. Fluxograma da Arquitetura proposta para a previsão da curva *forward* de Preço de energia elétrica.



4.1.1 Módulo Aquisição dos Dados

A fase de coleta de dados, demonstrado na Figura 25, é um componente essencial da metodologia, pois a qualidade e a disponibilidade dessas informações impacta diretamente nas previsões. Todos os dados utilizados foram obtidos a partir de bases de dados abertas e acessíveis online.

Figura 25. Módulo de Aquisição dos dados de diversas fontes.



Fonte: Autor

As variáveis utilizadas na criação do banco de dados incluem tanto endógenas ¹ à empresa quanto exógenas ². Essa abordagem é abrangente na coleta de dados, pois permite uma análise mais detalhada e robusta, proporcionando um entendimento mais profundo dos fatores que determinam a dinâmica de mercado e a formação dos preços no setor elétrico.

As fontes de dados que foram utilizadas são: Operador Nacional do Sistema Elétrico - ONS, Instituto Brasileiro de Geografia e Estatística - IBGE, Agência Nacional de Energia Elétrica - ANEEL, Câmara de Comercialização de Energia Elétrica - CCEE, Instituto Nacional de Meteorologia - INMET, Empresa de Pesquisa e Energia - EPE, Instituto de Pesquisa Econômica Aplicada - IPEA, Banco Central do Brasil - BACEN e série histórica

¹ Variável Endógena é aquela cujo valor é determinado dentro do modelo em questão e é influenciada por outras variáveis no modelo e, em contextos de análise, representa tipicamente a variável dependente (WOOLDRIDGE, 2017).

² Variável Exógena é aquela cujo valor é determinado fora do modelo e que afeta as variáveis dentro do modelo, mas não é afetada por elas no contexto da análise em questão. Essas variáveis são consideradas variáveis independentes que influenciam, mas não são influenciadas por outras variáveis dentro do modelo (WOOLDRIDGE, 2017).

da DCIDE. Esse processo de extração dos dados pode ser realizado de forma manual, com o usuário baixando direto dos sites, ou automaticamente, por meio de *API* (Interface de Programação de Aplicações) que foram desenvolvidas. Na Figura 23 é demonstrado com detalhes as variáveis que estão sendo prospectada de cada fonte.

Esses dados, uma vez extraídos, são submetidos a um processo de tratamento devido às diferentes frequências em que são disponibilizados, necessitando assim de padronização. Observou-se que alguns dados eram retornados em frequências diversas, como horário, quinzenal, mensal ou trimestral. Portanto, para a criação da base de dados, foi realizada uma normalização dessas frequências. Além disso, algumas variáveis apresentaram lacunas de dados para determinados períodos da série histórica. A série histórica dos dados utilizados compreende o intervalo dos anos de 2012 a 2023. Esta abrangência temporal permite uma análise detalhada das tendências e padrões ao longo dos anos.

A aquisição de dados no setor elétrico, em especial a partir do Portal de Dados da ONS, que é uma plataforma de acesso a diversas informações relacionadas ao sistema elétrico brasileiro. O portal fornece uma gama de dados, relatórios e estudos que são essenciais para análises, planejamentos e tomadas de decisões no setor elétrico. Quanto às diversas frequências encontradas nas bases de dados do ONS, elas refletem a natureza complexa e dinâmica do setor elétrico. As informações podem ser coletadas e registradas em diferentes intervalos de tempo, para atender às necessidades específicas de monitoramento e operação do sistema. Algumas das frequências comuns nas bases de dados incluem:

- **Frequência Horária:** Dados registrados a cada hora, fornecendo uma granularidade mais fina para análises detalhadas das variações no consumo e na geração de energia ao longo do dia.
- **Frequência Diária:** Agregação de dados ao longo do dia, oferecendo uma visão mais ampla dos padrões de consumo e geração em uma escala diária.
- **Frequência Mensal:** Dados consolidados em intervalos mensais, permitindo análises de tendências ao longo de períodos mais extensos e sazonalidades mensais.
- **Frequência Anual:** Sumarização dos dados em base anual, útil para avaliações de desempenho e planejamento a longo prazo.
- **Frequência Semanal:** Dados consolidados em base semanal oferecem uma visão intermédia entre as escalas mais granulares (horária e diária) e as mais amplas (mensal e anual).

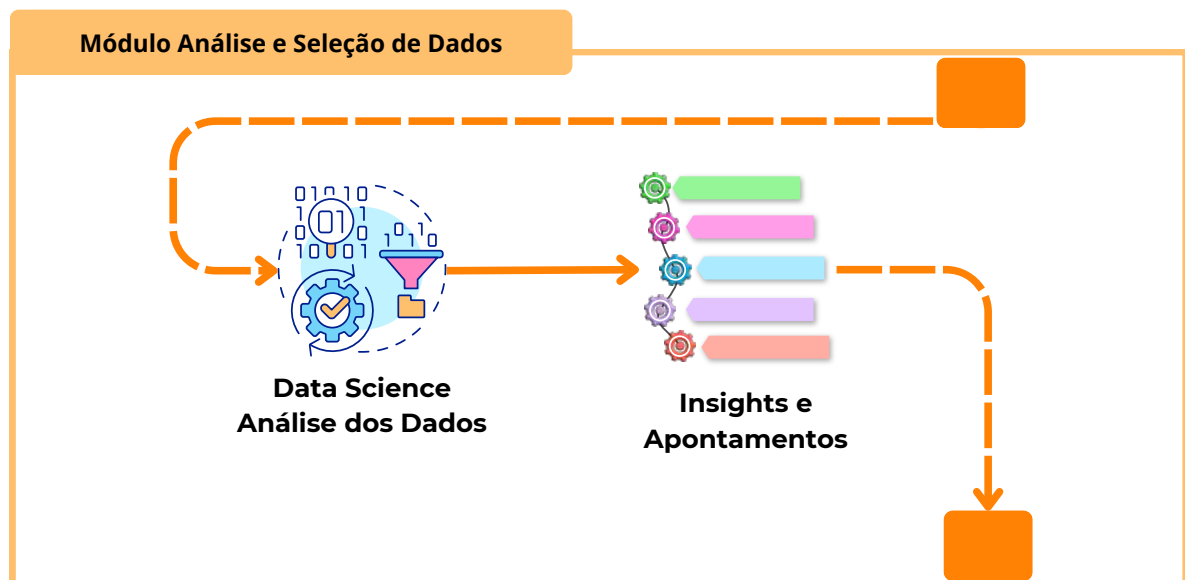
4.1.2 Módulo Análise e Seleção de Dados

A análise exploratória dos dados é um procedimento estatístico essencial para a compreensão das informações coletadas, pois permite identificar padrões, detectar

anomalias e detectar variáveis relevantes para o modelo. As técnicas utilizadas incluem análises gráficas — como histogramas e gráficos de dispersão — e métodos quantitativos, como a análise de correlação, possibilitando a validação de suposições tanto por meio de representações visuais quanto por métricas estatísticas (BRUCE; BRUCE, 2019).

O módulo de análise e seleção de dados, ilustrado na Figura 26, realiza um estudo detalhado de todas as variáveis selecionadas, ou seja, do banco de dados construído na fase de aquisição. Esse processo permite uma triagem criteriosa dos fatores mais relevantes para a modelagem da previsão de preços.

Figura 26. Módulo de análise e seleção dos dados: avalia as variáveis que mais contribuem na modelagem do problema.



Fonte: Autor

Para determinar quais variáveis impactam o preço da energia no longo prazo, foi realizado uma análise de correlação entre o valor do preço de referência — representado pela curva *forward* de energia elétrica fornecida pela Dcide (DCIDE, 2024) — e as variáveis prospectadas. Esse procedimento auxilia na identificação dos principais fatores que influenciam a precificação da energia ao longo do tempo.

A Dcide é uma empresa dedicada ao desenvolvimento de soluções de informação, processamento e modelagem quantitativa para o setor de energia elétrica. Ela fornece um boletim da Curva *Forward* semanal, com os valores do preço de energia a nível mensal e anual até 4 anos para os quatro submercados. Possui uma plataforma colaborativa de pool de preço de energia, onde os agentes mais comercialmente ativos inserem suas referências de preços individuais para um horizonte de cinco anos em troca de métricas agregadas de todos os participantes do processo. Após ter os valores de cada agente, a

Dcide, para garantir a representatividade e qualidade da informação, realiza um processo de validação das curvas individuais, para minimizar a influência de valores atípicas (valores que não representativos ao mercado atual, ou seja, valores que poderiam tendenciar a curva forward tanto positivamente quanto negativamente) utilizando um algoritmo estatístico (FELIZATTI et al., 2019).

Os valores de preço fornecido pela Dcide serão utilizados como referência para os preços praticados no mercado, no cenário de longo prazo. Dessa forma, esse valor é utilizado como benchmark no estudo, permitindo a identificação das variáveis que impactam a formação do preço da energia elétrica. Para isso, será aplicada a análise de correlação entre as variáveis, auxiliando na compreensão das relações que influenciam a dinâmica do mercado.

Para essas análises, foi selecionado o coeficiente de correlação de Pearson, conforme descrito na Seção 3.5, como a técnica a ser utilizada na metodologia. Esse coeficiente fornece uma medida quantitativa da dependência entre as variáveis, permitindo identificar o grau de correlação entre elas e contribuindo para uma análise mais precisa da influência de cada fator na formação dos preços de energia.

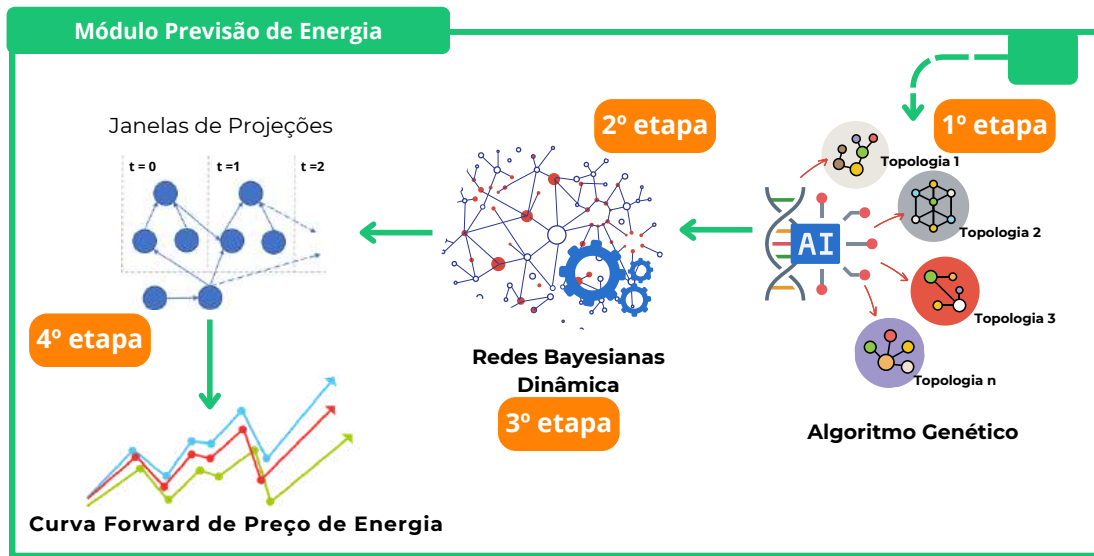
Esse módulo apresenta alta flexibilidade, permitindo a inserção de novas variáveis a qualquer momento para análise. Essa capacidade de adaptação possibilita a atualização e o aprimoramento contínuos do processo, garantindo sua relevância diante de mudanças no ambiente de dados ou nos requisitos do estudo.

A saída desse módulo consiste em *insights* e apontamentos estratégicos sobre os dados, identificando as variáveis mais significativas para a definição dos cenários de projeção. Essas informações serão utilizadas como entrada para o módulo de previsão de energia, assegurando que o modelo esteja embasado em uma análise robusta e dinâmica.

4.1.3 Módulo Previsão de Energia Elétrica

O módulo de previsão de energia elétrica, onde tem como entradas os dados direto dos portais abertos que são coletados por API, como também os que veem do modulo de análise, onde são normalizados e analisadas suas correlações com a variável algo definida pelo problema. Este módulo conforme demonstrado na Figura 27, a nível macro é composto pelas etapas de definição de topologias que são os modelos de entrada para a Rede Bayesiana Dinâmicas, a técnica adotada para realizar previsões nos horizontes de médio e longo prazo. Esse módulo é responsável pela geração das curvas *forward* de energia elétrica.

Figura 27. Módulo de previsão de Energia Elétrica



Fonte: Autor

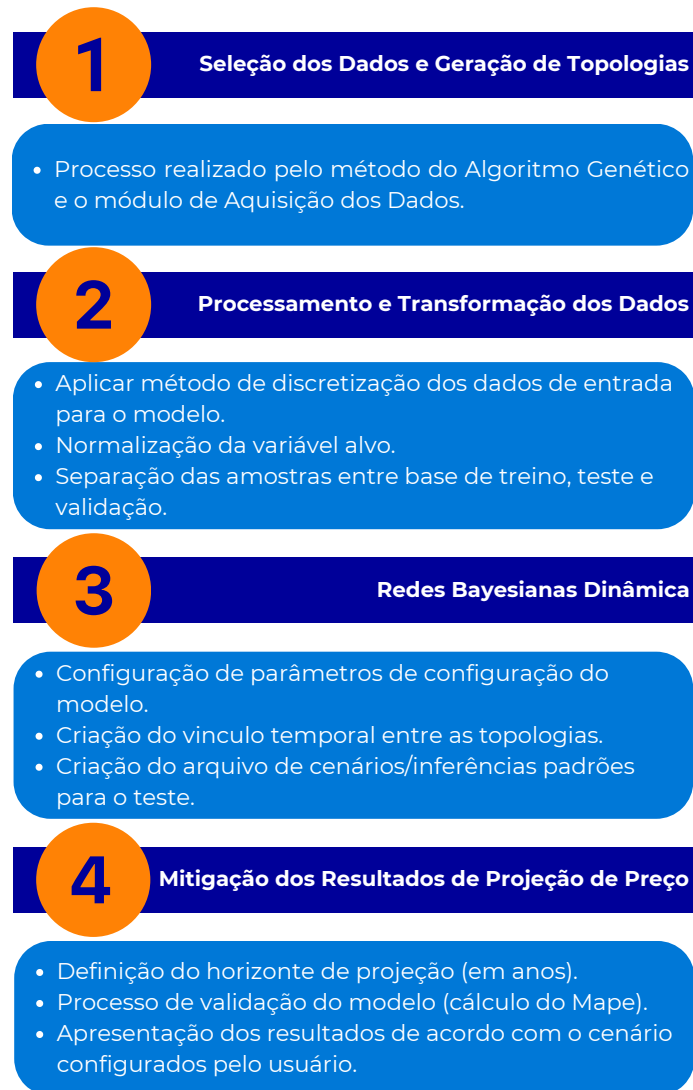
A Figura 28 apresenta cada uma das 4 etapas e serão explanadas detalhadamente, proporcionando uma compreensão abrangente de cada fase do processo de previsão.

4.1.3.1 Etapa 1 - Seleção dos Dados e Geração de Topologias

Esta etapa utilizar como entrada tanto os dados provenientes diretamente do método de aquisição quanto os dados e os apontamentos oriundos do módulo de análise de dados. Com essas informações, o Algoritmo Genético é empregado para selecionar as variáveis (atributos) que irão compor a topologia. Esta seleção é crucial para estruturar as conexões e relações entre os atributos na modelagem da rede.

Desta forma, determinar a melhor topologia é um problema não-polinomial e, portanto, não é viável obter uma solução ótima em tempo computacionalmente viável, por definição de problemas não-polinomiais. Serão formadas topologias de cenários possíveis com o auxílio de Redes Bayesianas Clássicas. Então, pode-se utilizar de heurísticas para obtenção de um resultado viável, que não tem garantias matemáticas de ser a melhor solução, mas sendo caracterizada como uma solução viável. Diante disso a heurística adotada foi o Algoritmo Genético.

Figura 28. É demonstrado as etapas detalhas do Módulo de previsão de energia.

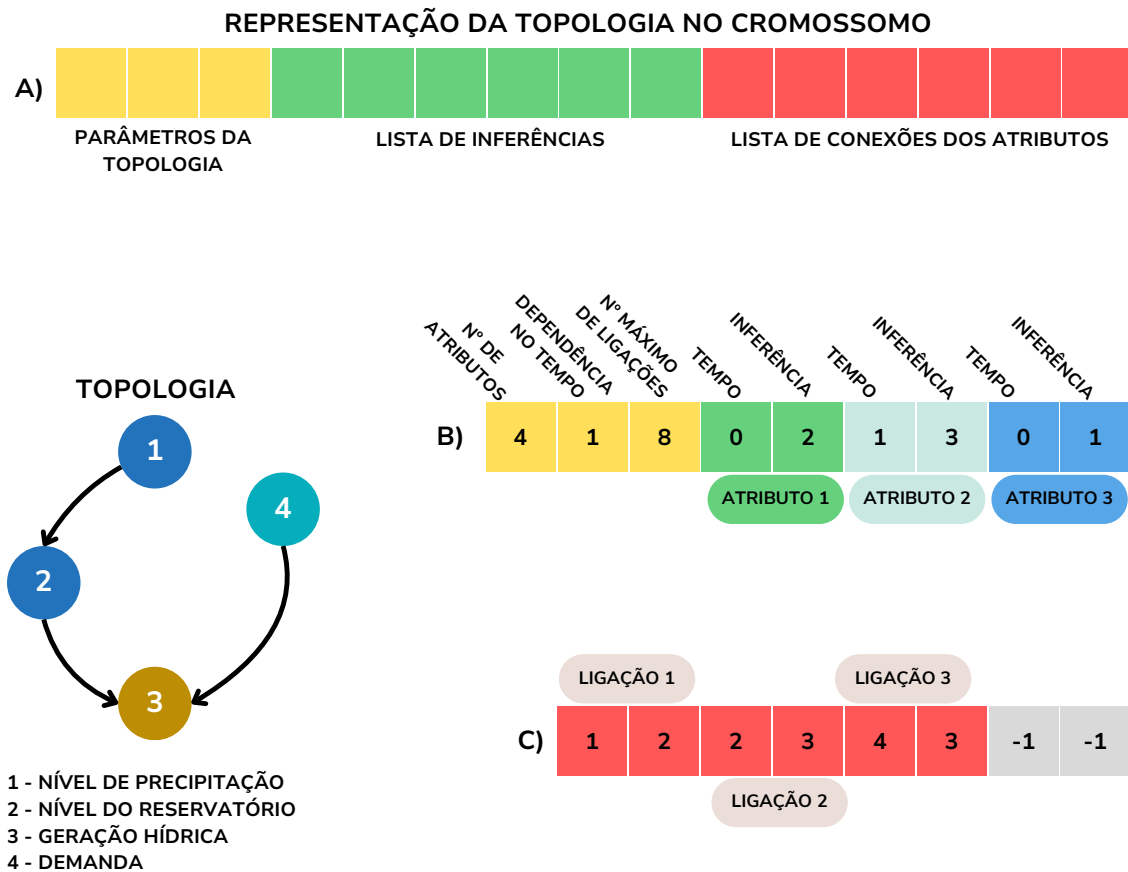


Fonte: Autor

Como o algoritmo genético foi empregado para gerar os modelos de topologias, tornou-se essencial desenvolver um esquema para representar-las em cromossomos, que são os indivíduos usados no processo de evolução durante o treinamento. A Figura 29 ilustra como essa representação foi realizada, detalhando a metodologia aplicada para codificar as estruturas da topologia em formato genético, facilitando assim a seleção e a evolução dos modelos ao longo das gerações do algoritmo.

Essa representação, denominada cromossomo, foi dividida em três partes, cada uma correspondendo a uma combinação distinta de genes: Parâmetros da Topologia, Lista de Inferências e Lista de Conexões dos Atributos, conforme ilustrado na Figura 29A. Essa estruturação facilita a organização e a manipulação dos dados genéticos necessários para a

Figura 29. Esquema visual da metodologia definida da representação da topologia em cromossomo no Algoritmo Genético.



Fonte: Autor

modelagem e análise subsequente.

Nos parâmetros da topologia, conforme ilustrado na Figura 29B, a configuração é composta por três genes. O primeiro gene define a quantidade de atributos (variáveis) disponíveis. O segundo gene determina o número de atributos que se relacionam temporalmente, ou seja, o atributo no tempo (t) que influenciará no tempo ($t + 1$), refletindo o conceito das Redes Bayesianas Dinâmicas onde um atributo pode afetar o mesmo ou outros atributos no tempo. O terceiro gene estabelece o número máximo de ligações possíveis na topologia, dado o número de atributos definidos, indicando quantas arestas seriam necessárias para conectar todos os atributos.

Dando continuidade à Figura 29B, na etapa de lista de inferências, define-se quais atributos poderão ou não receber inferências ³ e, se aplicável, qual será o estante do tempo

³ A inferência em Redes Bayesianas Dinâmicas refere-se ao processo de calcular a distribuição de probabilidade de um conjunto de variáveis de interesse, dadas algumas evidências ou observações. Este

para esses. Assim, a parte do cromossomo que representa a lista de inferências é composta por dois genes - tempo e inferência — para cada atributo, estabelecendo uma estrutura genética detalhada para a modelagem das inferências.

Na sequência da representação da topologia, a etapa de Lista de Conexões dos Atributos define quais atributos são conectados, sendo cada ligação representada por 2 genes, ou seja, nesse momento é definido qual variável influencia ou é influenciada por outras. Essa representação é ilustrada na Figura 29C. Como exemplo, é apresentado um modelo da topologia com 4 atributos, que são: 1 - Nível de precipitação, 2 - Nível do Reservatório, 3 - Geração Hídrica e 4 - Demanda. Esse é um exemplo de uma topologia simples que poderia ser empregada para realizar a previsão de geração de energia por uma determinada hidrelétrica, onde a variável alvo (a variável para a qual será realizada a previsão) é a Geração Hídrica.

Ao analisar a topologia, observa-se que o nó 1 se conecta ao nó 2, e a variável alvo nó 3 (variável que será projetada) gera conexões dos nós 2 e 4, ou seja, essas duas variáveis possuem ligações direta (maior influência) com a geração hídrica. Com base nessa configuração, as ligações são representadas no cromossomo por pares: a ligação 1 é representada por 1, 2, a ligação 2 por 2, 3 e a ligação 3 por 4, 3. Ao encontrar uma ligação -1, -1 no cromossomo, isso indica que não existem mais ligações a serem consideradas, conforme demonstrada na Figura 29 C).

A configuração do cromossomo no processo de evolução do indivíduo foi definido da seguinte forma:

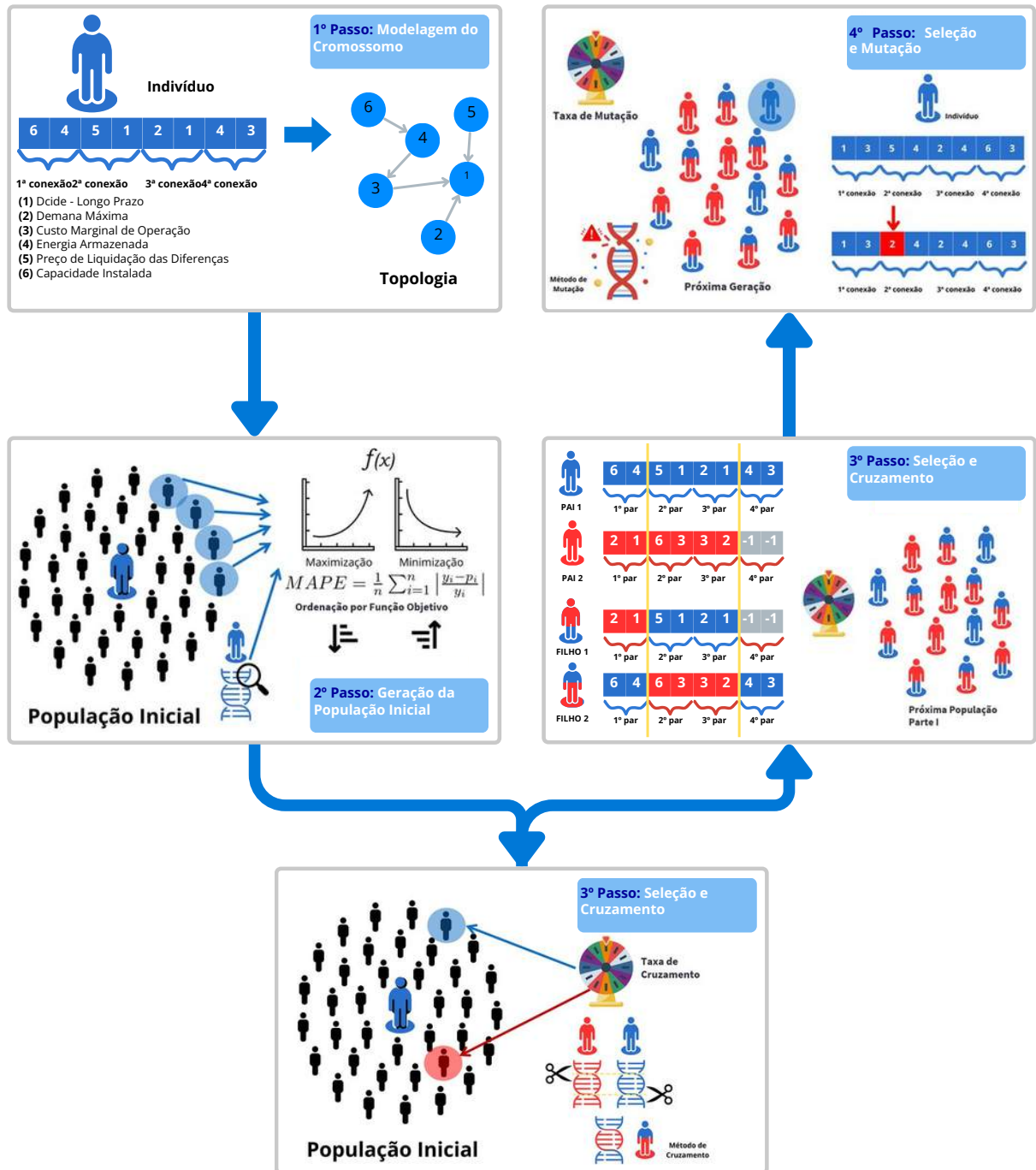
- Taxa de Mutação: é escolhido um gene do cromossomo.
- Cruzamento: O cromossomo é dividido em 3 partes.
- Dependência no Tempo: 0 ou 1.

O método de seleção dos indivíduos adotado foi o da roleta, cujo funcionamento é detalhado na seção 3.3.1. Este método seleciona os indivíduos para reprodução com base na proporção de sua aptidão relativa ao conjunto total, favorecendo aqueles com maior aptidão, mas ainda preservando a diversidade genética ao possibilitar a seleção de indivíduos menos aptos.

Para elucidar como foi desenvolvido o processo envolvendo a criação da população, seleção dos indivíduos, cruzamento e mutação do Algoritmos Genéticos, a sequência completa dessas operações é detalhada na Figura 30.

processo é fundamental em DBNs para entender a dinâmica do sistema modelado, especialmente para prever estados futuros, filtrar informações presentes ou suavizar estimativas passadas (DUDA; HART; STORK, 2000).

Figura 30. Fluxograma das etapas do Algoritmo Genético.



Fonte: Autor

No passo 1, refere-se à modelagem do cromossomo, onde a escolha dos nós e as conexões (arestas) entre eles são realizadas de maneira aleatória à medida que esse indivíduo vai sendo gerado. Durante esse processo, é verificado se o cromossomo representa um Grafo Acíclico Dirigido [3.1]. Caso não represente, outro indivíduo é gerado. Após a geração dos indivíduos, forma-se então uma população inicial. No passo 2, cada indivíduo

da população é avaliado usando uma função de aptidão específica, que no caso a utilizada é o Erro Absoluto Percentual Médio (MAPE). Quanto menor for o valor de MAPE, mais apto é considerado o indivíduo. No passo 3, ocorre o cruzamento entre os indivíduos da população. Para realizar o cruzamento, pares de indivíduos são selecionados através do método da Roleta. Neste método, indivíduos com maior aptidão ocupam uma proporção maior na “roleta” e, conseqüentemente, têm uma probabilidade maior de serem escolhidos para o cruzamento. Esse processo de seleção favorece a evolução da população, pois tende a aumentar a presença de indivíduos mais aptos à medida que o cruzamento acontece.

O quarto e último passo é o método de mutação, fundamental para preservar a diversidade genética da população ao longo das gerações. Esse processo desempenha um papel crucial na eficiência do algoritmo, permitindo uma exploração mais abrangente do espaço de soluções e reduzindo o risco de convergência prematura para ótimos locais. Foi aplicada uma taxa de mutação de 10%, o que implica em uma probabilidade significativa de alteração aleatória dos genes nos cromossomos. Essa variação contribui para a descoberta de novas soluções potencialmente superiores, enriquecendo o processo evolutivo do algoritmo.

Após a execução do Algoritmo Genético, são geradas várias topologias. A criação destas topologias com base no problema é complexa, especialmente na ausência de conhecimento especializado sobre o cenário, ou seja, não se sabe quais atributos se conectam ou quais exercem maior influência sobre a variável alvo. Em alguns casos, pode ser necessário realizar entrevistas, elaborar formulários, entre outras abordagens, para tentar construir essa topologia. Com a metodologia proposta, é possível gerar essas topologias sem a necessidade de possuir conhecimento prévio do cenário ou das relações entre as variáveis, facilitando significativamente o processo de modelagem.

4.1.3.2 Etapa 2 - Processamento e transformação dos dados

Após a definição e seleção das variáveis que compõem a topologia, inicia-se a etapa de pré-processamento dos dados. A topologia gerada pode não abranger todas as variáveis disponíveis no banco de dados. Em seguida, ocorre a categorização dos dados, um procedimento conhecido como discretização. Esse processo consiste na conversão de variáveis contínuas em variáveis categóricas, segmentando o espectro de valores contínuos em intervalos pré-definidos.

A Figura 31 apresenta um exemplo de discretização. Para isso, foi selecionada a variável climática de precipitação de chuva, que foi categorizada em três níveis: baixo, médio e alto. A classificação foi realizada com base em limites previamente definidos para cada categoria, e o método empregado rotulou os valores conforme esses limites. Observa-se que, na amostra analisada, a maioria dos dados pertence à categoria do tipo ‘alto’, indicando uma predominância de valores elevados de precipitação.

Figura 31. Esquema visual da Discretização dos Dados utilizando método *K-means*.



Fonte: Autor

O método de discretização utilizado foi aprendizagem não-supervisionada ⁴ chamado *k-means*.

4.1.3.3 Etapa 3 - Redes Bayesianas Dinâmicas

As Redes Bayesianas Dinâmicas são adequadas para modelar sequências temporais de variáveis devido à sua capacidade de integrar informações ao longo do tempo e gerenciar incerteza, e essa noção de tempo é feito através da introdução de estados que se repetem ao longo do tempo, onde esse estado tem uma dependência do anterior. Na etapa 3, após a criação de várias topologias pelo Algoritmo Genético, é selecionada uma para gerar os cenários de projeções.

Após a seleção de uma topologia apropriada, é possível analisar as dependências entre as variáveis no instante de tempo t e suas conexões para o instante subsequente $t + 1$, conforme discutido na Seção 3.2. As conexões temporais apresentam flexibilidade e podem ser ajustadas pelo usuário conforme necessário. Em seguida, são gerados os cenários de previsão, os quais podem ser modificados com base nas inferências temporais realizadas sobre as variáveis, possibilitando ajustes dinâmicos que refletem mudanças nas condições ou a incorporação de novas informações.

4.1.3.4 Etapa 4 - Análise dos Resultados de Projeção de Preço

Após as configurações dos atributos da topologia, é possível gerar diversos cenários de previsão de preço de energia. Esses cenários podem ser modificações de acordo com as inferências e conexões temporais no modelo.

⁴ Método de aprendizagem não-supervisionada é um tipo de aprendizagem de máquina em que o agente recebe um conjunto de dados de treinamento não-rotulado (desconhece a que classe ele pertence) e precisa aprender para definir a classe dos dados

Tais inferências são fundamentais para auxiliar nas previsões dos cenários de curva *forward* de energia. Esses inputs são normalmente extraídos de séries históricas, de empresas de consultoria de previsões (como previsões de períodos de chuva mais intensos, temperatura, tendências de expansão no mercado, entre outras) ou até mesmo da expertise do analista. Esta abordagem integrada melhora significativamente a precisão e a relevância das previsões no setor de energia elétrica.

As RBDs permitem realizar inferências, que são flexíveis no tempo, podendo ser aplicadas a qualquer variável e em qualquer horizonte de tempo. A Figura 32 ilustra como a variável “nível de precipitação”, que foi discretizada em três categorias, pode ser manipulada pelo usuário. Por exemplo, com base em conhecimento prévio fornecido por consultorias meteorológicas sobre o nível de chuvas, o usuário pode informar que, em um determinado período do horizonte de tempo, essa variável estará na categoria 3, correspondente a um nível alto de chuva. Após receber as entradas, a RBD processa os dados para gerar cenários de previsão, adaptando-se dinamicamente às novas informações para aprimorar continuamente as projeções.

Figura 32. Esquema visual da modificação na inferência de uma variável, dado a discretização dos dados.



Fonte: Autor

A flexibilidade dessa abordagem permite realizar diversas simulações de previsões de cenários de preço de energia elétrica, pois cada informação inserida na rede gera modificações nos valores ao longo da curva de valores.

Nesta etapa do processo, define-se o horizonte de previsão, que pode variar de 1 a 10 anos. Com essas configurações iniciais, um cenário é gerado. Conforme mencionado anteriormente, o usuário pode modificar esse cenário e solicitar uma nova projeção, ajustando-a conforme suas necessidades ou eventuais mudanças no ambiente analisado.

As previsões geradas podem ter saídas em diferentes intervalos de tempo, como mensal, trimestral, semestral ou anual, dependendo da configuração do período definida na

seleção das variáveis. Isso significa que, se os dados das variáveis forem organizados para representar um intervalo específico, como o semestral, as Redes Bayesianas Dinâmicas produzirão previsões seguindo essa mesma periodicidade, estendendo-se pelo número total de anos estabelecido na configuração do modelo.

Conforme demonstrado, a metodologia proposta nesta pesquisa apresenta características de robustez e adaptabilidade. Sua robustez decorre da capacidade de gerar previsões a partir de séries temporais, além de permitir a interação do usuário com a ferramenta, possibilitando a exploração de diferentes cenários e a mitigação de incertezas. Essa abordagem também atribui ao usuário um papel ativo na obtenção dos resultados. Além disso, a metodologia é adaptativa, pois permite a incorporação flexível de novas variáveis ao modelo sempre que necessário, possibilitando que o sistema se ajuste dinamicamente às mudanças nos dados e às novas condições do ambiente analisado.

No caso da projeção da curva *forward* de energia, como será demonstrado no capítulo 5 de resultados, observa-se uma influência significativa das variáveis meteorológicas, refletindo a composição da matriz energética brasileira. No entanto, caso ocorresse uma mudança nessa matriz, bastaria prospectar novas variáveis relevantes para o modelo. A metodologia aplicada possui a capacidade de gerar automaticamente as novas topologias necessárias para a construção de diferentes cenários. Essa flexibilidade evidencia a adaptabilidade do modelo a mudanças estruturais no setor energético ou em suas condições operacionais.

5 Capítulo 5 - Resultados e Discussões

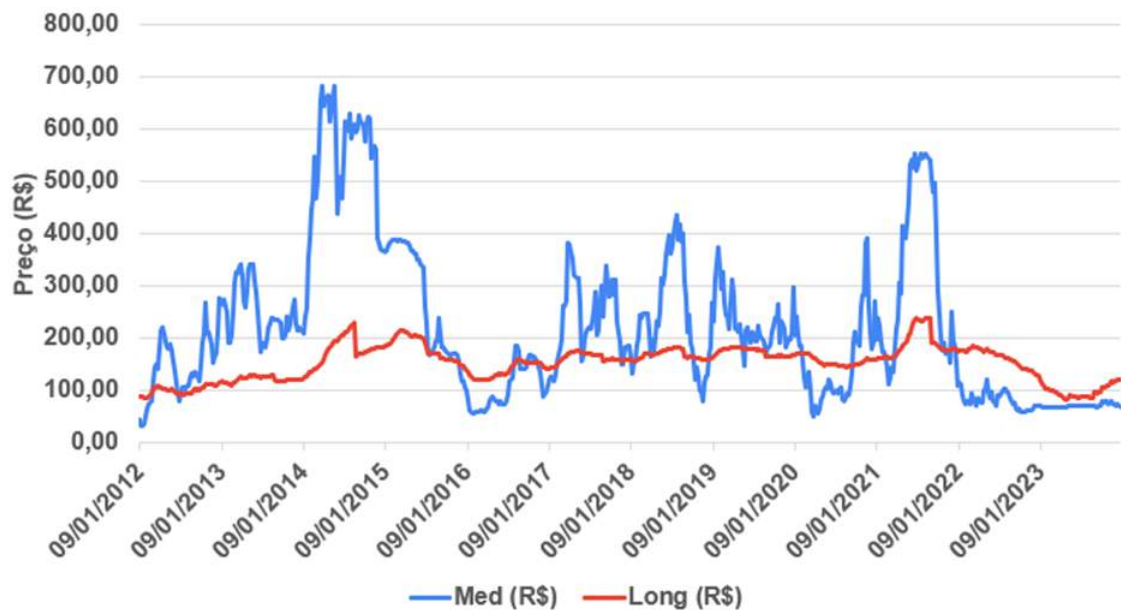
Neste capítulo apresenta-se os resultados preliminares obtidos: (i) estudo das variáveis que foi realizado com o objetivo de identificar o que mais impacta na formação do preço de energia no médio e longo prazo; (ii) resultados do Algoritmo Genético para a criação automática dos modelos de topologias que representam a modelam de problema e, (iii) os resultados das previsões de cenários de preço da curva *forward* no horizonte de até 6 anos.

5.1 Resultados do Estudo das Variáveis

Para previsão de energia de longo prazo, fatores de menor volatilidade tendem a influenciar significativamente o preço da energia, diferentemente do que ocorre no curto prazo. Como se sabe, no mercado brasileiro as fontes principais de geração de energias são renováveis com predominância na geração hidrelétrica, com isso, fatos climáticos e operacional tendem a ter um impacto direto na formação do preço de energia elétrica para longo prazo. Diante disso foi realizado um estudo com objetivo de analisar a correlação entres as 25 variáveis que foram selecionadas para a criação do banco de dados, onde estão distribuídas em econômica, operacional, climáticas e planejamento. Essa correlação foi realizada aplicando o método de correlação de Pearson das 25 variáveis com a variável do preço de energia elétrica de longo prazo fornecido pela Dcide. As variáveis do banco de dados possuem uma série histórica entre os anos de 2012 até 2023. Inicialmente o estudo está focado para o submercado de energia Sudeste/Centro-Oeste.

Como já mencionado que de acordo com o horizonte de tempo os valores de preço de energia tendem a ter comportamento diferentes, isso dever-se pelo comportamento de volatilidade das variáveis de que impactam na formação do preço, esse comportamento pode ser observado na Figura 33, que demonstra a série histórica do preço de energia elétrica dos horizontes e médio e longo prazo fornecido pela Dcide. Deve-se levar em consideração que o horizonte de médio prazo definido pela Dcide é trimestral.

Figura 33. Curvas forward de médio (azul) e longo (vermelho) prazo, obtidas por meio da empresa de consultoria DCide.

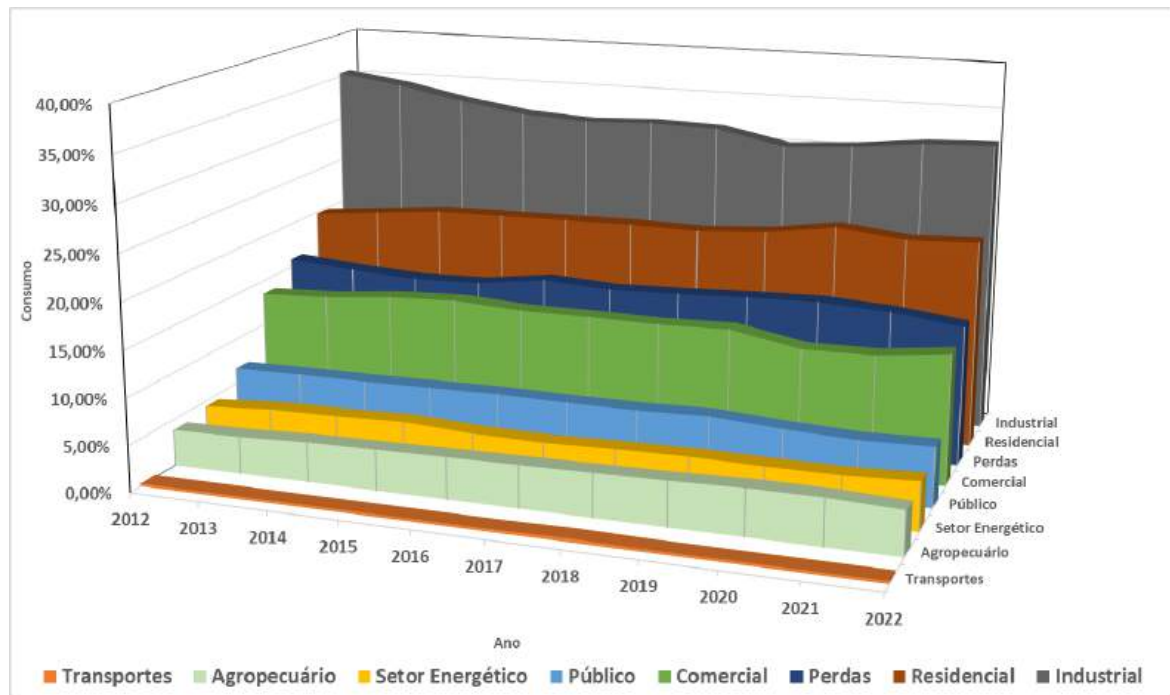


Fonte: Autor

Inicialmente, são consideradas variáveis de natureza econômica. Assim, como primeira análise, é essencial verificar os usos finais da energia elétrica no mercado nacional, a fim de identificar as atividades que mais contribuem para a carga. Essa verificação pode ser realizada por meio da análise do Balanço Energético Nacional (BEN), publicado anualmente pela Empresa de Pesquisa Energética (EPE, 2024a).

Com o objetivo de analisar esse comportamento de consumo por atividade é apresentado na Figura 34 entre os anos de 2012 a 2022, os setores que mais contribuíram para o consumo de energia elétrica. Logo, os setores industrial, residencial, comercial e as perdas técnicas foram os que mais consumiram energia elétrica. No entanto, ao longo deste período, observa-se um aumento na representatividade do consumo residencial e do setor agropecuário, enquanto o setor industrial apresenta uma participação relativa menor. Essa mudança reflete as variações observadas na composição do Produto Interno Bruto (PIB) do país.

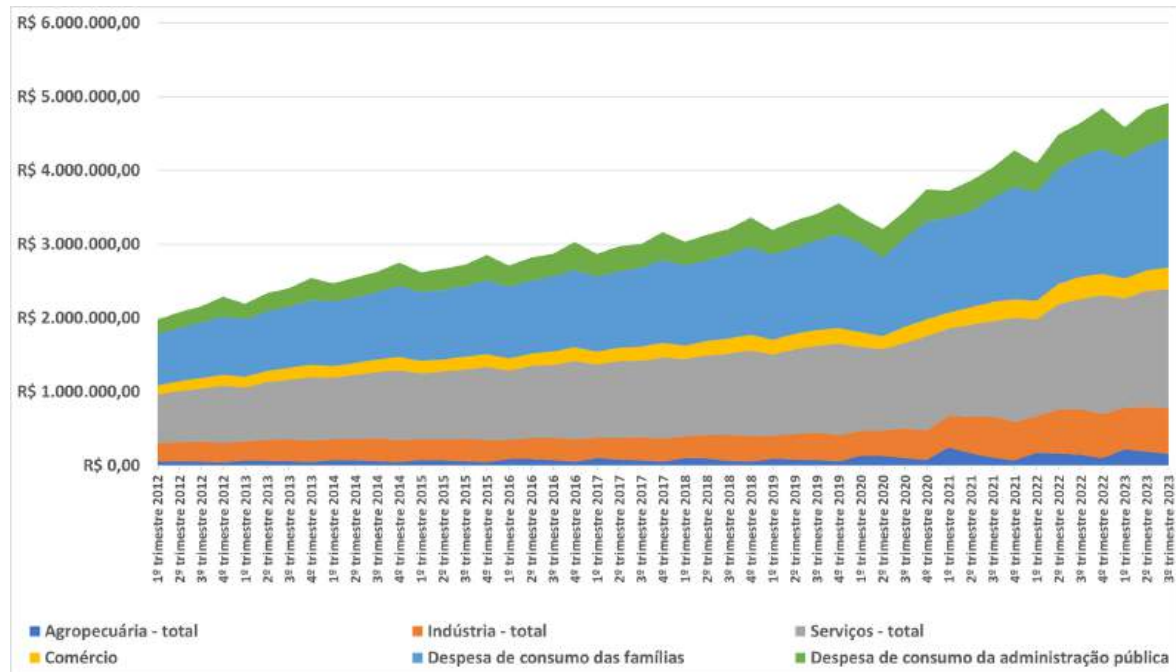
Figura 34. Consumo por setor, segundo o Balanço Energético Nacional.



Fonte: Autor com dados da EPE

Com o objetivo de avaliar a composição do PIB e identificar os setores mais impactantes para a carga, foram analisadas as bases de dados do Instituto Brasileiro de Geografia e Estatística (IBGE) para determinar quais setores da economia brasileira mais contribuem para a geração de riqueza no país (IBGE, 2024). Portanto, conforme ilustrado na Figura 35, ao longo do período observado, os principais setores responsáveis pela composição do PIB foram, respectivamente, o consumo das famílias, o setor de serviços e o setor industrial.

Figura 35. Os principais setores econômicos para a composição PIB brasileiro.

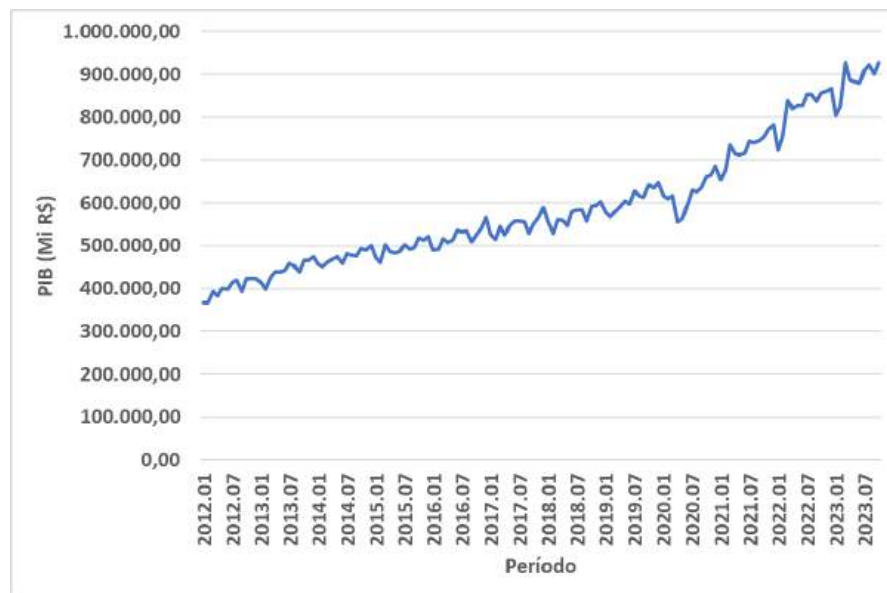


Fonte: Autor com dados do IBGE

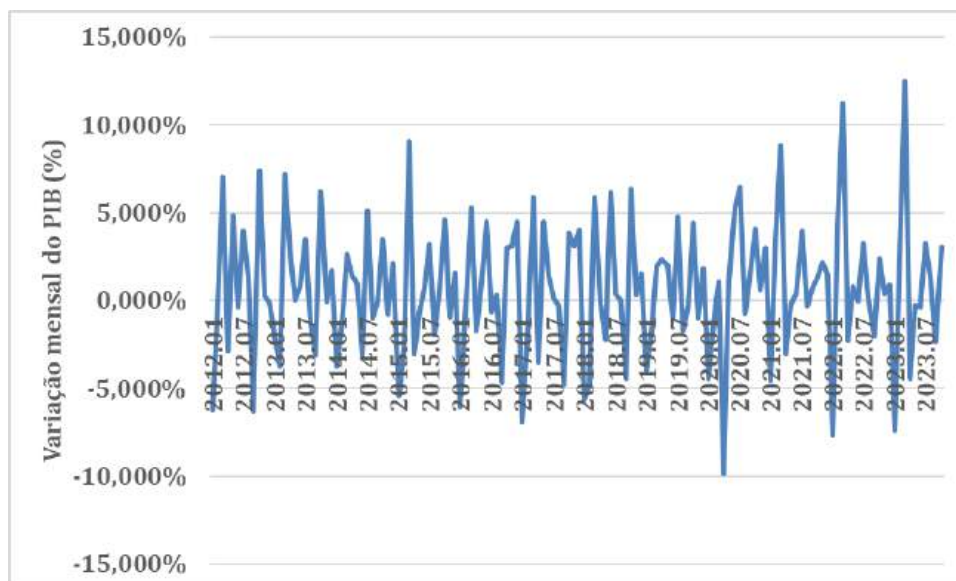
Com o objetivo de avaliar a influência do PIB na previsão dos preços de energia elétrica no Brasil, foram analisadas as correlações entre as curvas *forward* de médio e longo prazo. Para esse estudo, foram considerados os dados mensais do PIB, calculados pelo IBGE e disponibilizados pelo Instituto de Pesquisa Econômica Aplicada (IPEA), abrangendo a série histórica dos anos de 2012 a 2023, conforme apresentado na Figura 36(a). Além disso, para ilustrar a variação mensal do PIB em relação ao mês anterior, a Figura 36(b) apresenta essa evolução ao longo do período analisado e é possível verificar que não apresenta um comportamento sazonal.

Portanto, com o objetivo de analisar essa correlação, foram considerados os dados mensais do PIB. O coeficiente de correlação obtido foi de -0.33036397 para a curva *forward* de médio prazo e de 0.08215320 para a curva *forward* de longo prazo, conforme apresentado na Tabela 1. Entretanto, diante da baixa correlação observada entre a riqueza produzida no país e o preço da energia elétrica, torna-se necessária uma análise mais detalhada dos dados para melhor compreender essa relação.

Figura 36. Série histórica do Produto Interno Bruno (PIB) entre os anos 2012 a 2023.



(a) PIB mensal



(b) Variação mensal do PIB

Fonte: Autor com dados do IBGE

Tabela 1. Resultado da correlação entre as variáveis PIB e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coefficiente de Correlação
PIB x Curva Forward Médio Prazo	-0.33036397
PIB x Curva Forward Longo Prazo	0.08215320

Fonte: Autor

A primeira análise refere-se à base temporal da análise do Produto Interno Bruno

devido ao fato de que, usualmente, os dados são amplamente divulgados pelos principais veículos de comunicações em bases trimestrais, fazendo com que haja um atraso entre a efetiva produção de riqueza e a percepção dos consumidores e investidores com relação ao nível de atividade da economia brasileira. Portanto, se faz necessário averiguar a base temporal que se mostre mais adequada para a correlação com os preços de energia.

A segunda análise a ser realizada refere-se à desagregação das parcelas componentes do PIB, abrangendo a atividade industrial, o setor de serviços, a agropecuária, as exportações e importações, o consumo das famílias e os gastos públicos. Essa análise é fundamental, pois cada um desses componentes tende a apresentar ciclos anuais distintos e a responder de maneira diferenciada às tendências macroeconômicas no Brasil. Consequentemente, essas variações refletem diferentes percepções de consumidores e investidores em relação à projeção dos preços da energia elétrica.

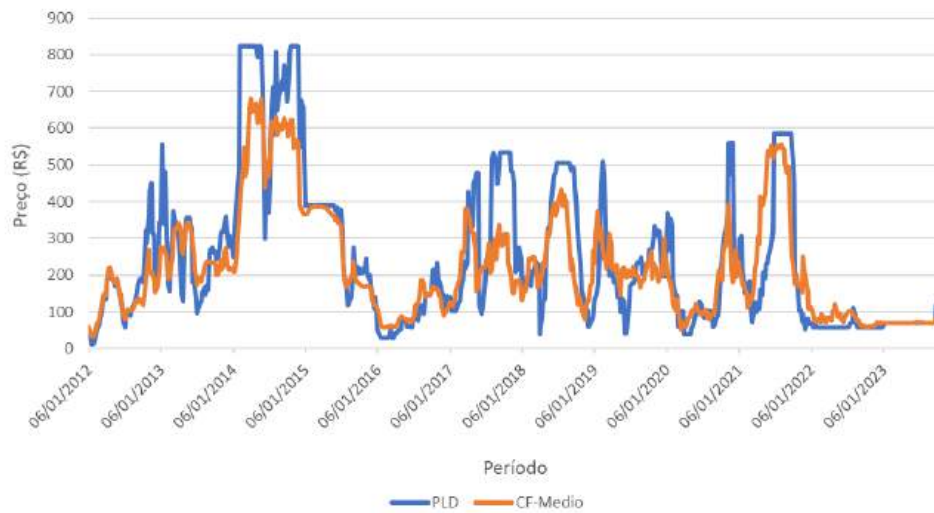
Dando continuidade ao estudo das variáveis, serão analisadas as variáveis operacionais, cujos dados são integralmente extraídos do Operador Nacional do Sistema Elétrico e da Câmara de Comercialização de Energia Elétrica. O Preço de Liquidação das Diferenças é resultado da cadeia de modelos computacionais NEWAVE, DECOMP e DESSEM, que é considerado o preço oficial do mercado de energia elétrica. Seu principal objetivo é a liquidação das sobras e déficits dos agentes nos ambientes de contratação livre e regulada. No entanto, por refletir essencialmente o custo operacional do sistema elétrico, o PLD não capta as expectativas do mercado nem o prêmio associado ao processo de comercialização.

Adicionalmente, o Preço de Liquidação das Diferenças, cujos valores são calculados pelos modelos computacionais NEWAVE, DECOMP e DESSEM, apresenta forte correlação com os regimes hidrológicos nacionais. Como resultado, o PLD tende a ser excessivamente baixo em períodos de hidrologia normal e elevado em cenários de hidrologia desfavorável. Em essência, os valores do PLD são determinados predominantemente pelas condições hidrológicas, e não pela interação entre as preferências dos consumidores e a estrutura de custos dos geradores.

No entanto, por ser o único padrão oficial de preços observado no setor elétrico, ele é inevitavelmente considerado na modulação dos contratos de energia elétrica. Essa influência é ainda mais evidente nos horizontes de curto e médio prazo, nos quais os preços de custo da operação e os preços de mercado da energia elétrica apresentam uma maior correlação.

A Figura 37 apresenta o comportamento do PLD no mercado brasileiro entre os anos de 2012 e 2023, representando o preço *spot* da energia elétrica. Além disso, é exibido o valor da Curva *forward* de médio prazo. Observa-se que o PLD de médio prazo tende a acompanhar a precificação da DCide.

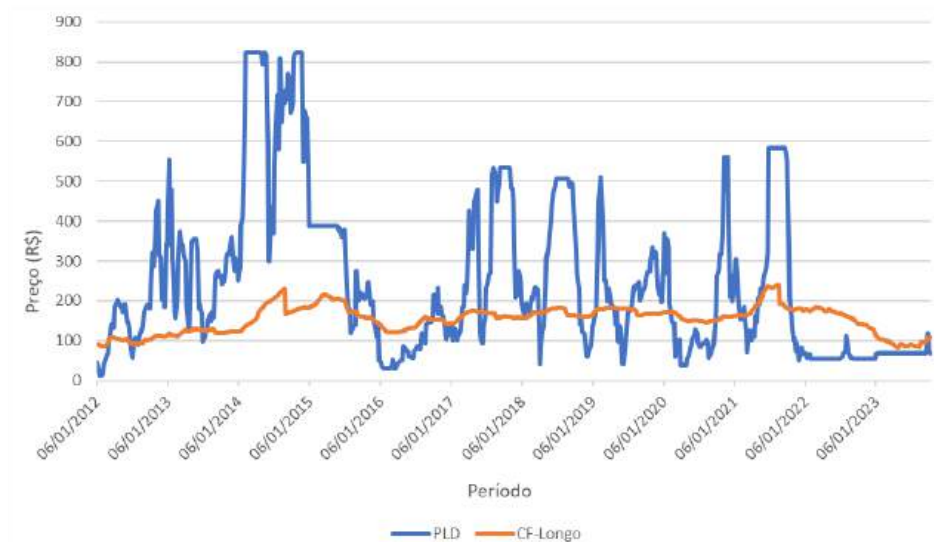
Figura 37. Curva *forward* de médio prazo e o PLD no horizonte de 2012 a 2023.



Fonte: Autor

Por outro lado, ao analisar o PLD em conjunto com a Curva *forward* de longo prazo, conforme ilustrado na Figura 38, nota-se que o preço da DCide apresenta um comportamento mais estável em comparação ao PLD. Isso sugere que, no médio prazo, o PLD pode ser utilizado como referência no mercado de energia. No entanto, no horizonte de longo prazo, há uma grande disparidade entre os valores.

Figura 38. Curva *forward* de Longo prazo e o PLD no horizonte de 2012 a 2023.



Fonte: Autor

Para validar essa análise, foi calculado o coeficiente de correlação entre os preços. Os resultados obtidos indicam um coeficiente de 0.86557953 para o médio prazo e 0.424605885 para o longo prazo, conforme apresentado na Tabela 2.

Tabela 2. Resultado da correlação entre as variáveis Preço da Liquidação das Diferenças e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coeficiente de Correlação
PLD x Curva Forward Médio Prazo	0.86557953
PLD x Curva Forward Longo Prazo	0.424605885

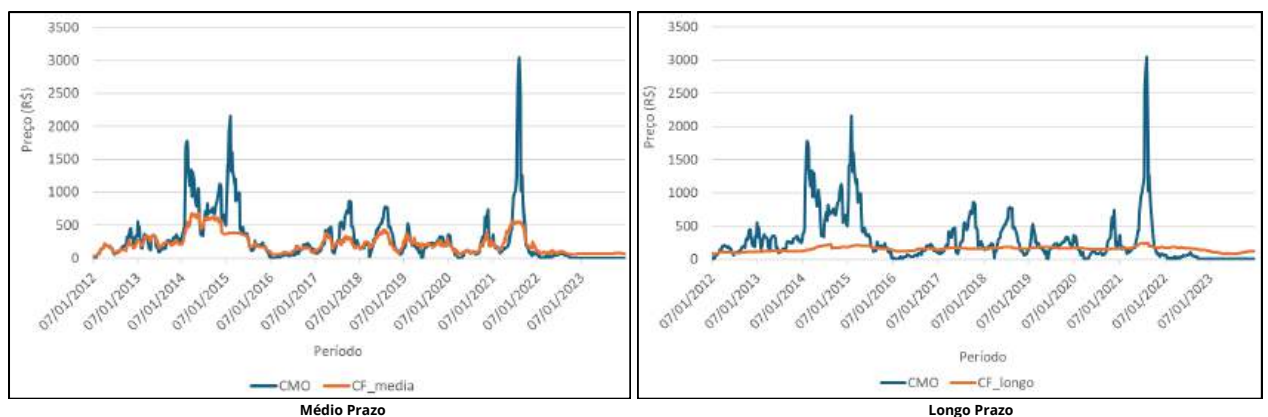
Fonte: Autor

A saída dos modelos NEWAVE, DECOMP e DESSEM é denominada Custo Marginal de Operação. A partir dos valores de saída, é estabelecido um limite mínimo e máximo, e os valores resultante desses limites, são definidos como valores do Preço de Liquidação das Diferenças.

Diante disso, analisou-se o impacto do CMO na formação do preço de energia, a fim de compará-lo com o PLD. O objetivo foi verificar se os dados brutos extraídos dos modelos poderiam exercer uma influência mais significativa na formação dos preços.

Na Figura 39, é apresentada a comparação entre o CMO e os preços de médio e longo prazo, respectivamente. Conforme mencionado anteriormente, a análise da série histórica do CMO revela a ocorrência de picos em determinados períodos, sendo especialmente notável o aumento registrado em 2021, decorrente da crise hídrica que afetou o setor elétrico naquele ano.

Figura 39. Comparação da Curva *forward* de Médio e Longo prazo e o Custo Marginal de Operação nos anos de 2012 a 2023.



Fonte: Autor

Tabela 3 apresenta a correlação entre o Custo Marginal de Operação e as curvas de preços. No médio prazo, observa-se uma correlação significativa de 0.777114144, enquanto

no longo prazo essa relação se reduz pela metade com 0.486578387. Em contraponto ao Preço de Liquidação das Diferenças, conclui-se que o preço do PLD para o médio prazo pode ser praticado no mercado, uma vez que não apresenta os picos de valores que tornariam inviável sua aplicação comercial.

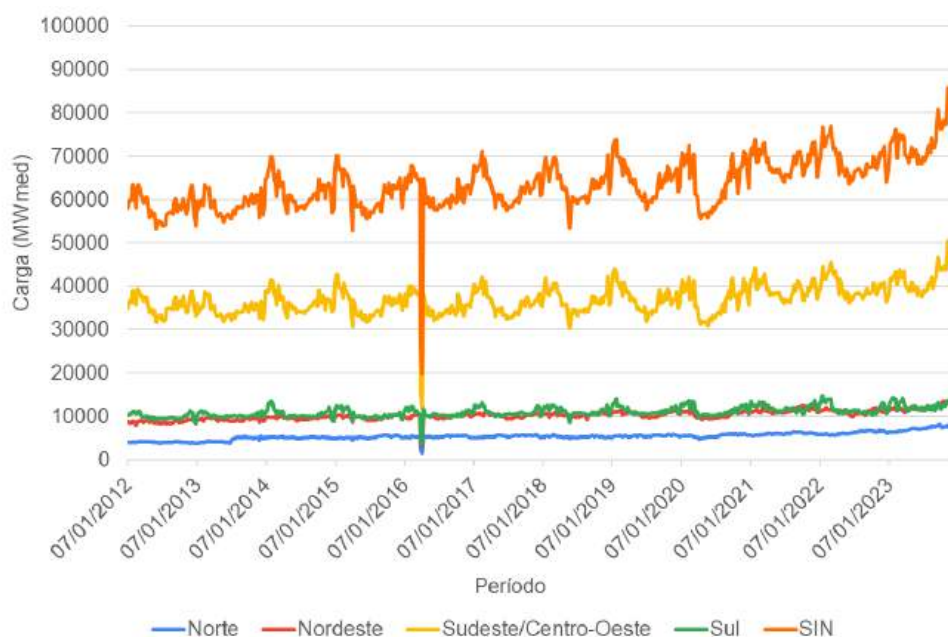
Tabela 3. Resultado da correlação entre as variáveis Custo Marginal de Operação e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coefficiente de Correlação
CMO x Curva Forward Médio Prazo	0.777114144
CMO x Curva Forward Longo Prazo	0.486578387

Fonte: Autor

Dando sequência aos estudos sobre variáveis operacionais, foram prospectados dados da carga total (MWmed) do Sistema Interligado Nacional, além das medições realizadas por submercado (Sudeste/Centro-Oeste, Sul, Nordeste e Norte), com discretização temporal por semana operativa, conforme ilustrado na Figura 40.

Figura 40. Carga Total do SIN e nos submercados em MWmed, no horizonte de janeiro de 2012 a dezembro 2023.



Fonte: Autor

É notável que o submercado Sudeste/Centro-Oeste exerce maior influência sobre a carga total do SIN, representando mais de 55% da demanda total do sistema segundo o boletim mensal de outubro de 2024 do Balanço energético. Analisando esse submercado no final de 2023 houve um aumento confederável em relação ao ano anterior.

Dessa forma, foram calculadas as correlações entre as curvas *forward* de médio e longo prazo e a carga verificada no mesmo período, tanto no Sistema Interligado Nacional quanto no submercado Sudeste/Centro-Oeste, conforme apresentado na Tabela 4. Observa-se que as correlações entre as variáveis foram negativas, indicando que os valores evoluem em direções opostas, e próximas de zero, sugerindo uma baixa correlação entre elas.

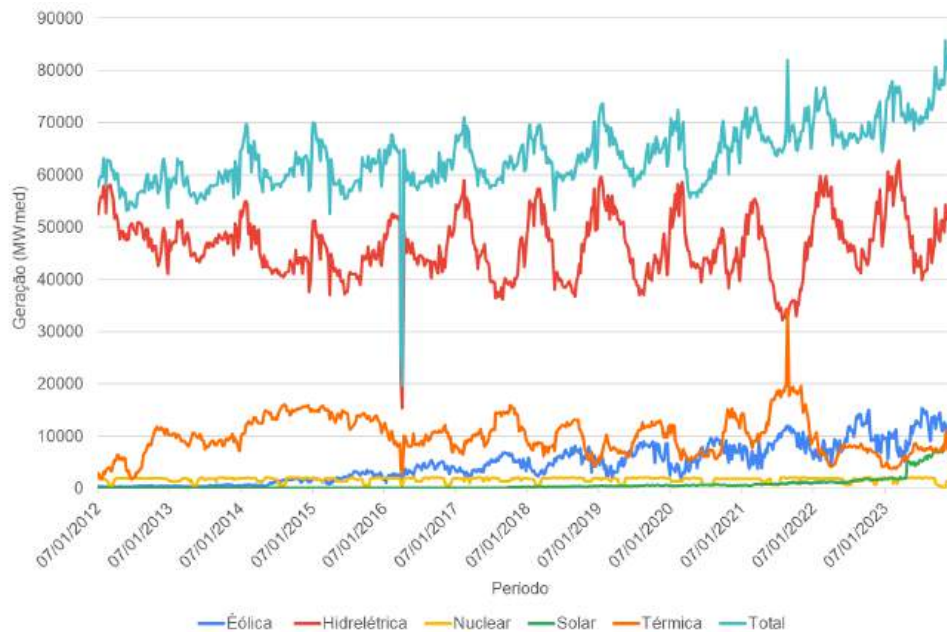
Tabela 4. Resultado da correlação entre as variáveis Carga total do SIN e no submercado Sudeste/Centro-Oeste e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coefficiente de Correlação
Carga SIN x Curva Forward Médio Prazo	-0.196253795
Carga SIN x Curva Forward Longo Prazo	-0.008335536
Carga Sudeste/Centro-Oeste x Curva Forward Médio Prazo	-0.17113102
Carga Sudeste/Centro-Oeste x Curva Forward Longo Prazo	-0.072700166

Fonte: Autor

Ainda em relação às variáveis operacionais, foram utilizados os dados de geração total de energia no SIN e por tipo de usina, agregados a partir das informações individuais de usinas de geração distribuídas pelo país e por submercado. Como ilustrado na Figura 41, a maior parte da geração de energia no Brasil provém de usinas hidrelétricas, seguidas por usinas térmicas e eólicas.

Figura 41. Geração de energia do SIN (total) e por fonte de energia, no horizonte de janeiro de 2012 a dezembro 2023.



Fonte: Autor

As usinas térmicas, que utilizam combustíveis fósseis para geração de energia, são

acionadas com maior sensibilidade às condições de precipitação e aos níveis dos reservatórios das hidrelétricas conectadas ao SIN, ou seja, as usinas técnicas dão uma segurança para sistema de energia elétrica do país a nível de geração. Por outro lado, a geração por fontes eólica e solar tem apresentado crescimento contínuo desde o segundo semestre de 2020 e do segundo semestre de 2023, respectivamente. Em contrapartida, a geração por usinas térmicas tem se mantido estável desde 2022.

Para mensurar o nível de correlação entre a geração total do SIN e os diferentes tipos de geração em relação à curva *forward* de preço de energia nos horizontes de médio e longo prazo, os resultados são apresentados na Tabela 5.

Tabela 5. Resultado da correlação entre as variáveis Geração total do SIN e por tipo de usinas (Eólica, Hidrelétrica, Nuclear, Solar e Térmica) e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coefficiente de Correlação
Geração Total SIN x Curva Forward Médio Prazo	-0.214430188
Geração Total SIN x Curva Forward Longo Prazo	-0.025967674
Geração Eólica x Curva Forward Médio Prazo	-0.298437249
Geração Eólica x Curva Forward Longo Prazo	0.072316758
Geração Nuclear x Curva Forward Médio Prazo	-0.017589382
Geração Nuclear x Curva Forward Longo Prazo	-0.074700718
Geração Hidrelétrica x Curva Forward Médio Prazo	-0.384959774
Geração Hidrelétrica x Curva Forward Longo Prazo	-0.373222203
Geração Solar x Curva Forward Médio Prazo	-0.327047562
Geração Solar x Curva Forward Longo Prazo	-0.331567146
Geração Térmica x Curva Forward Médio Prazo	0.689879482
Geração Térmica x Curva Forward Longo Prazo	0.621401596

Fonte: Autor

Observa-se que as usinas hidrelétricas e térmicas apresentam comportamentos opostos em relação ao impacto no preço da energia. As hidrelétricas possuem correlação negativa com a curva *forward* de preços, indicando que um aumento na geração hidrelétrica tende a reduzir os preços no mercado futuro. Em contrapartida, as térmicas apresentam forte correlação positiva, sugerindo que um aumento na geração térmica está associado a preços mais elevados. Esse efeito pode ser explicado pelo maior custo operacional das usinas térmicas, que utilizam fontes primárias mais dispendiosas.

Fazendo uma análise mais detalhadas das demais variáveis, como a geração total de energia do SIN, no médio prazo, um aumento na geração total está associado a uma leve redução nos preços futuros da energia. No longo prazo, esse efeito praticamente desaparece, sugerindo que a oferta total de energia não influencia significativamente a precificação futura nesse horizonte. Com relação a geração Eólica, no médio prazo, há uma correlação negativa moderada, indicando que o aumento da geração eólica pode reduzir os preços de energia. Isso ocorre porque a energia eólica é uma fonte renovável de baixo custo marginal. No entanto, no longo prazo, essa relação se enfraquece, possivelmente devido à

variabilidade da geração eólica e sua intermitência.

E por fim, em relação a geração Solar, a correlação negativa sugere que um aumento na geração solar tende a reduzir os preços futuros da energia. Isso é esperado, pois a energia solar tem custo operacional praticamente nulo e, quando disponível, reduz a necessidade de despacho de fontes mais caras, como térmicas.

Ao analisar a Figura 41, observa-se que a geração eólica é superior à geração solar. No entanto, ao realizar a mesma análise na Tabela 5 de correlação, verifica-se que a geração solar apresenta uma correlação mais elevada. Esse resultado pode ser explicado pelo fato de que os valores da curva *forward* de preço de energia de médio e longo prazo utilizados no estudo referem-se ao submercado Sudeste/Centro-Oeste, onde se concentra a maior geração de energia solar entre os submercados.

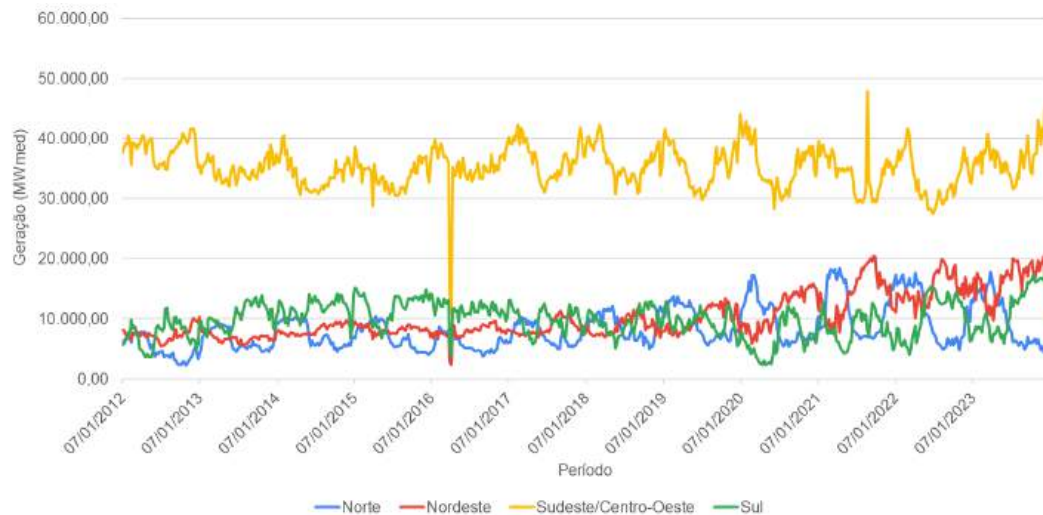
Dessa forma, os resultados reforçam que a composição da matriz energética impacta diretamente a formação dos preços da energia elétrica, sendo a geração hidrelétrica um fator de atenuação de custos, enquanto a geração térmica pode contribuir para o aumento dos preços.

Ainda analisando a variável capacidade de geração, busca-se identificar qual submercado possui a maior capacidade de geração, considerando que cada um apresenta características distintas em termos de infraestrutura, tipos de usinas e outros fatores. A Figura 42 mostra que o submercado Sudeste/Centro-Oeste possui a maior capacidade de geração, o que está diretamente relacionado à predominância maior de usinas hidrelétricas nessa região. Neste contexto, nota-se uma oscilação periódica na geração, possivelmente associada a fatores sazonais, como períodos de seca e chuvas, além de variações na demanda.

Ainda analisando a geração por submercados é verificado que o Nordeste tem uma tendência de aumento na geração nos últimos anos, provavelmente devido à expansão da geração eólica e solar na região. E o Norte apresenta oscilações mais pronunciadas, que podem estar relacionadas à variabilidade da geração hidrelétrica e às condições dos reservatórios.

A Tabela 6 apresenta os resultados da correlação entre a capacidade de geração de energia por submercado e o valor da curva *forward* de médio e longo prazo. O valor da curva *forward*, fornecido pela Dcide, é segmentado por submercado, o que justifica a realização dessa correlação.

Figura 42. Geração de energia por submercado (Sudeste/Centro-Oeste, Sul, Nordeste e Norte), no horizonte de janeiro de 2012 a dezembro 2023.



Fonte: Autor

Tabela 6. Resultado da correlação entre as variáveis capacidade de Geração de energia por submercado e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

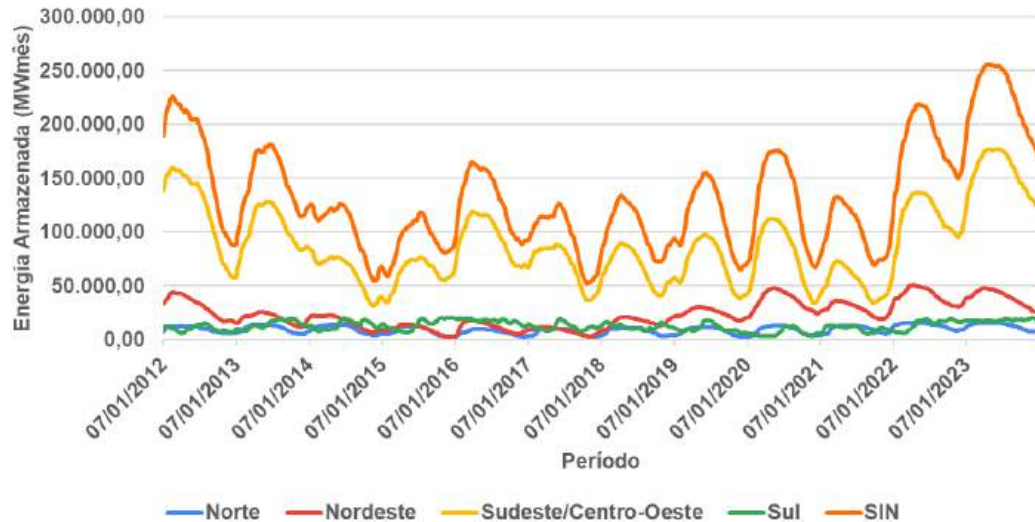
Variáveis	Coefficiente de Correlação
Norte x Curva Forward Médio Prazo	-0.057532759
Norte x Curva Forward Longo Prazo	0.220449051
Sudeste/Centro-Oeste x Curva Forward Médio Prazo	-0.17120011
Sudeste/Centro-Oeste x Curva Forward Longo Prazo	-0.294966835
Sul x Curva Forward Médio Prazo	0.032806668
Sul x Curva Forward Longo Prazo	-0.006609478
Nordeste x Curva Forward Médio Prazo	-0.17369021
Nordeste x Curva Forward Longo Prazo	0.02898716

Fonte: Autor

Os submercados Norte e Sudeste/Centro-Oeste exibiram correlações mais significativas na composição do preço da energia: o Norte apresentou uma correlação positiva, enquanto o Sudeste/Centro-Oeste mostrou uma correlação negativa, indicando que a oferta de energia nesse submercado pode ter um impacto redutor nos preços futuros.

Dando continuidade à análise dos dados operacionais, a energia armazenada representa a capacidade total de geração de energia disponível nos reservatórios coordenados pelo ONS em um determinado instante. A Figura 43 ilustra essa informação tanto para o Sistema Interligado Nacional como para cada submercado. Observa-se que a maior capacidade de armazenamento está concentrada nos reservatórios do submercado Sudeste/Centro-Oeste.

Figura 43. Energia Armazenada em MWmês do SIN e por subsistema, no horizonte de janeiro de 2012 a dezembro 2023.



Fonte: Autor

O Sudeste/Centro-Oeste apresenta os maiores volumes de energia armazenada ao longo do período. Sua curva exhibe oscilações sazonais marcantes, sugerindo um forte impacto da variabilidade hidrológica. E há um pico significativo em 2022-2023, possivelmente associado a um período de alta afluência hídrica e recuperação dos reservatórios, ou seja, a variação sazonal da energia armazenada indica uma dependência significativa das condições hidrológicas.

A Tabela 7 evidencia uma forte correlação inversa entre a energia armazenada nos reservatórios e os preços *forward* da energia no médio e longo prazo. Esse comportamento é esperado, pois um aumento na energia armazenada reduz a necessidade de geração térmica, que, por sua vez, possui um custo operacional mais elevado. Como consequência, há uma tendência de queda nos preços dos contratos negociados no mercado. Além disso, observa-se que o impacto mais significativo nos preços *forward* ocorreu a partir da correlação com a energia armazenada no submercado Sudeste/Centro-Oeste, com valores de coeficiente de correlação de -0.5162029 para médio prazo e -0.667641277 para longo prazo. A maior correlação com este submercado deve-se à sua maior capacidade de armazenamento em reservatórios hidrelétricos, o que ressalta sua importância na formação do preço da energia elétrica no país.

Tabela 7. Resultado da correlação entre as variáveis de energia armazenada total do SIN e por submercado e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coefficiente de Correlação
SIN x Curva Forward Médio Prazo	-0.490126674
SIN x Curva Forward Longo Prazo	-0.587548366
Sudeste/Centro-Oeste x Curva Forward Médio Prazo	-0.516202029
Sudeste/Centro-Oeste x Curva Forward Longo Prazo	-0.667641277
Norte x Curva Forward Médio Prazo	0.011467829
Norte x Curva Forward Longo Prazo	-0.067357365
Sul x Curva Forward Médio Prazo	-0.260256022
Sul x Curva Forward Longo Prazo	-0.289043352
Nordeste x Curva Forward Médio Prazo	-0.357451382
Nordeste x Curva Forward Longo Prazo	-0.272620787

Fonte: Autor

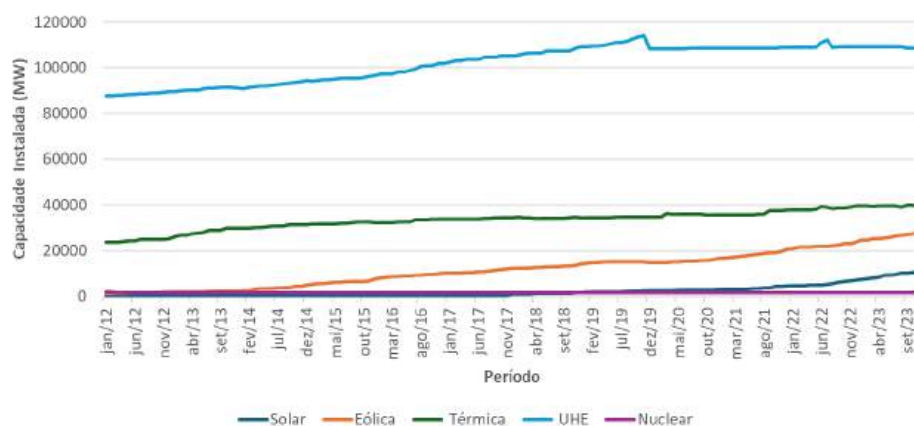
Dando sequência ao estudo das variáveis, serão abordadas aquelas relacionadas ao planejamento, com foco na capacidade instalada por tipo de geração e na geração distribuída. A capacidade instalada é uma métrica fundamental para o planejamento energético, pois permite avaliar se o sistema pode atender à demanda presente e futura, garantindo a segurança energética e a estabilidade do fornecimento. Além disso, desempenha um papel estratégico na promoção de fontes de energia renováveis, à medida que governos e entidades reguladoras buscam diversificar a matriz energética e reduzir a dependência de combustíveis fósseis.

Foram coletados dados da capacidade instalada por tipo de geração, abrangendo as fontes solar fotovoltaica, eólica, térmica, hidrelétrica e nuclear desde 2012 a 2023, conforme ilustrado na Figura 44. No Brasil, observa-se que a matriz energética é predominantemente hidrelétrica, seguida pela geração térmica. As fontes eólica e solar vêm apresentando um crescimento significativo, a eólica especialmente a partir de 2016, refletindo os investimentos na diversificação da matriz energética e na expansão dos parques eólicos e a Solar a partir de 2018, indicando avanço de energia fotovoltaica no país. Atualmente, as duas são responsáveis por 22% da matriz energética brasileira (ONS, 2024b).

A estabilidade da capacidade instalada de geração térmica e nuclear indica que essas fontes continuam sendo utilizadas como suporte à matriz energética, especialmente em períodos de baixa hidrologia.

Com os dados da capacidade instalada, foram realizadas correlações com a curva *forward* de preço de médio e longo prazo, com o objetivo de verificar se a capacidade instalada impacta diretamente a formação do preço da energia elétrica. Os resultados dessa análise estão apresentados na Tabela 8.

Figura 44. Capacidade Instalada por tipo de geração no período de 2012 a 2023.



Fonte: Autor adaptado da ONS

Tabela 8. Resultado da correlação entre as variáveis capacidade instalada por tipo de geração e o Preço da Curva *Forward* de energia para Médio e Longo Prazo.

Variáveis	Coefficiente de Correlação
Solar fotovoltaico x Curva Forward Médio Prazo	-0.381186234
Solar fotovoltaico x Curva Forward Longo Prazo	-0.186891418
Usina Eólica x Curva Forward Médio Prazo	-0.378190313
Usina Eólica x Curva Forward Longo Prazo	0.09615745
Térmicas x Curva Forward Médio Prazo	-0.245960126
Térmicas x Curva Forward Longo Prazo	0.279605511
Hidrelétrica x Curva Forward Médio Prazo	-0.278491761
Hidrelétrica x Curva Forward Longo Prazo	0.313352904
Nuclear x Curva Forward Médio Prazo	-0.13637078
Nuclear x Curva Forward Longo Prazo	-0.231503773

Fonte: Autor

Embora uma usina possua uma determinada capacidade instalada, a quantidade de energia efetivamente gerada pode variar devido a diversos fatores, como manutenção, disponibilidade de combustível (no caso de termelétricas), condições climáticas (para hidrelétricas, eólicas e solares), demanda do sistema e decisões operacionais baseadas em prioridades por tipo de fonte.

Ao analisar a correlação de longo prazo entre a geração das usinas hidrelétricas e a curva *forward*, foi verificado que os dados de potência instalada não impactam diretamente a formação do preço de energia elétrica como esperado. O coeficiente de correlação obtido foi de 0,313352904, indicando uma relação mais fraca do que se antecipava. Como a energia gerada por essas fontes apresenta um custo menor, era esperado que a correlação fosse negativa.

Ao analisar o coeficiente de correlação entre a potência gerada e o preço da curva

forward de longo prazo, obtém-se um valor de $-0,373222203$, conforme apresentado na Tabela 5. Esse resultado indica que, em condições meteorológicas desfavoráveis, a geração das hidrelétricas tende a diminuir, enquanto a geração térmica aumenta, impactando diretamente a oferta de energia.

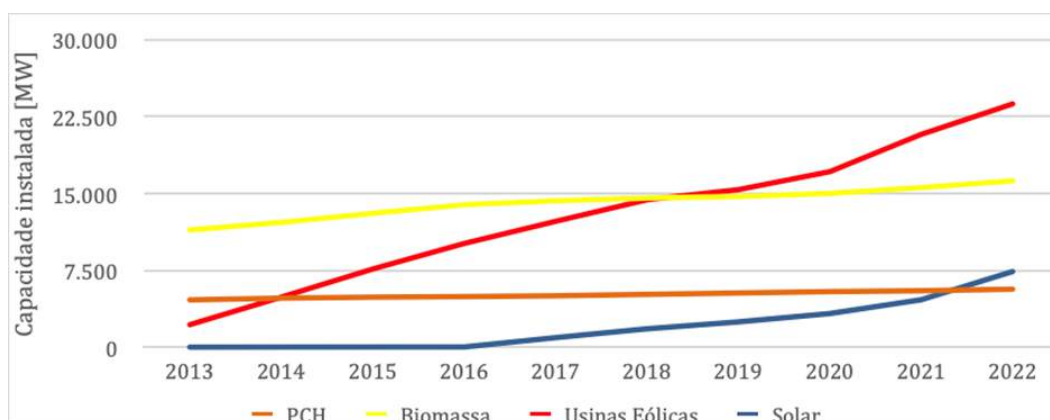
Consequentemente, esperava-se que a potência instalada apresentasse correlações no mesmo sentido. No entanto, como já mencionado, a potência instalada de uma fonte de energia não reflete necessariamente a potência efetivamente gerada, pois depende de fatores como demanda, políticas operacionais e outras variáveis externas.

Outra abordagem em análise é a correlação com a geração distribuída, uma vez que essa modalidade, predominantemente representada pela geração solar fotovoltaica, teve um crescimento significativo na última década, impulsionado pela regulamentação da ANEEL por meio da Resolução Normativa N° 482/2012.

Entretanto, o aumento da penetração de unidades de geração distribuída se traduz em uma redução da demanda observada nos pontos de entrega de energia por parte dos distribuidores de energia. Desta forma, o aumento da presença da geração distribuída representa uma redução da demanda observada pelos agentes geradores, comercializadores e distribuidores do sistema.

A variação da capacidade instalada das fontes de geração distribuída no período de 2013 a 2022 está apresentada na Figura 45. Dando destaque para as fontes geradoras Solar fotovoltaicas e Eólicas.

Figura 45. Crescimento da capacidade instalada das fontes de geração distribuída de 2013 a 2022.



Fonte: Autor adaptado da ONS

Os dados indicam que as fontes renováveis variáveis (solar e eólica) têm sido os principais impulsionadores do crescimento da geração distribuída no Brasil. A expansão da

energia solar, em particular, destaca-se como um segmento emergente, enquanto a energia eólica já se consolidou como uma das principais fontes. Em contraste, PCHs e biomassa apresentam crescimento mais modesto.

Nesta seção, foi realizada uma análise para identificar as variáveis que influenciam, direta e indiretamente, a formação do preço da curva *forward* de energia elétrica no Ambiente de Contratação Livre. Observou-se que algumas variáveis possuem baixa relevância direta; no entanto, isso não significa que serão excluídas da análise, pois podem impactar outras variáveis que afetam diretamente o preço. Por outro lado, variáveis que demonstraram potencial de comprometer a precisão do modelo de previsão de preços foram devidamente descartadas.

5.2 Resultados do Algoritmo Genético para Geração de Topologias

Nesta seção, delineia-se os resultados das topologias gráficas que serão utilizadas pelas Redes Bayesianas Dinâmicas, que foram geradas por meio da aplicação de Algoritmo Genético. A definição dessas topologias é um passo fundamental no processo de modelagem, pois determina a estrutura das dependências probabilísticas entre as variáveis.

A abordagem baseada em Algoritmos Genéticos foi escolhida devido à sua capacidade de explorar eficientemente o espaço de soluções, identificando configurações de redes que maximizam a capacidade preditiva do modelo. A avaliação das topologias geradas foi realizada com base em métricas específicas de desempenho, garantindo que a estrutura final seja representativa e eficaz na captura das relações entre as variáveis analisadas.

Como mencionado anteriormente, a definição dessa modelagem não é trivial, logo foi necessário o uso de um método de otimização. As variáveis ou atributos utilizados foram aqueles analisados e selecionados na Seção 5.1, referente ao estudo das variáveis.

Na execução do Algoritmo Genético para identificar a topologia mais adequada ou aquela que melhor representa o problema, foi necessário considerar diversos parâmetros de configuração, os quais estão apresentados na Tabela 9.

Tabela 9. Parâmetros de configuração do Algoritmo Genético

Parâmetro	Valor
Taxa de cruzamento	10%
Taxa de mutação	2%
Taxa de permanência dos melhores indivíduos	10%
Número máximo de ligações	$2 \times [(\text{número de nó}) - 1]$
Número mínimo de ligações	$(\text{Número de nó} - 1)$
Número de Nós	12

Fonte: Autor

A taxa de cruzamento entre os pares de cromossomos foi definida em 10%. Adicio-

nalmente, a taxa de mutação é estabelecida em 2%, o que ajuda o algoritmo a evitar a convergência para mínimos locais. Além disso, foi adicionado uma taxa de permanência dos melhores indivíduos, onde 10% dos indivíduos mais aptos são selecionados para a próxima geração a cada rodada, assegurando a presença contínua de indivíduos de alta qualidade na população.

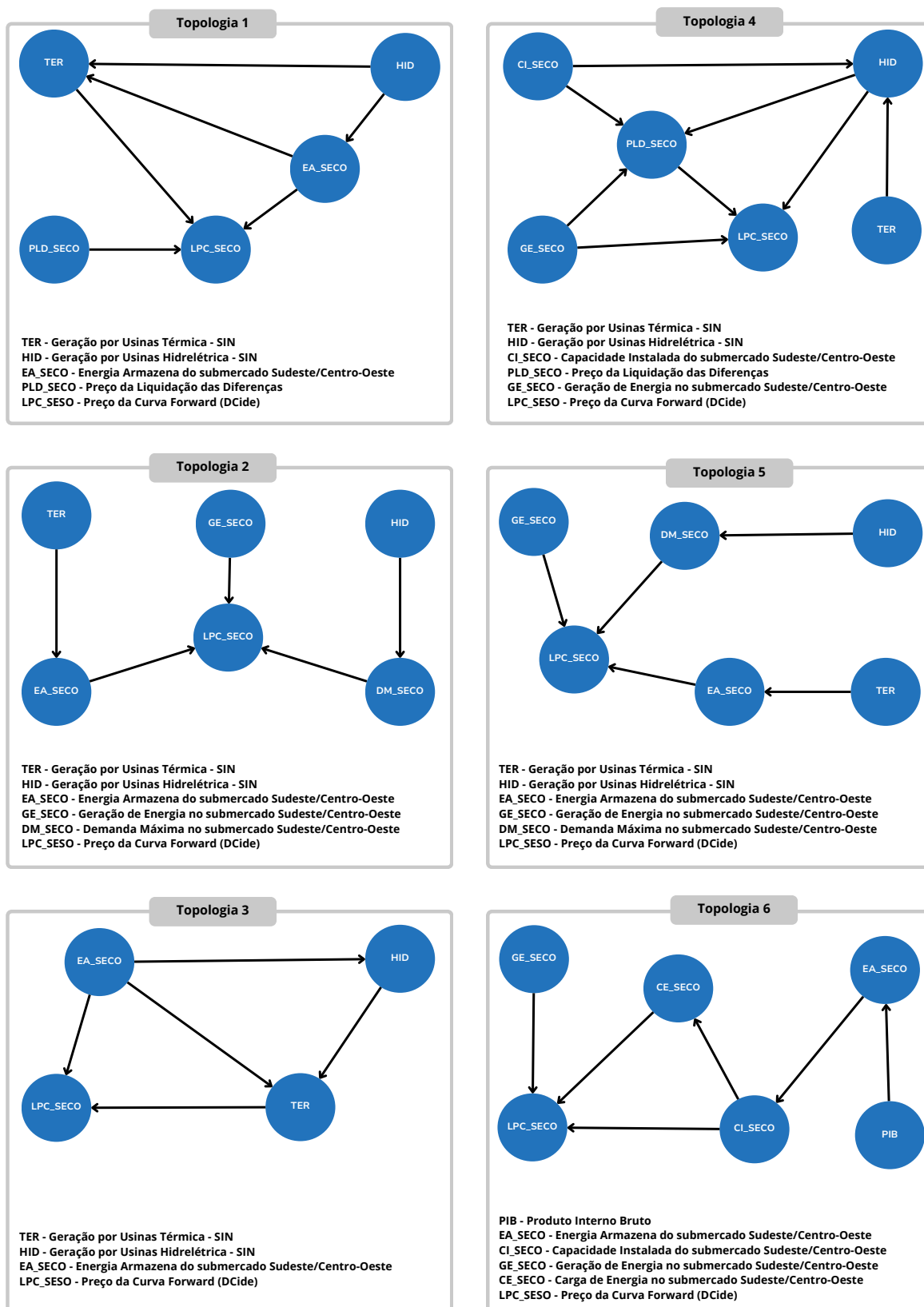
É também essencial definir o número máximo e mínimo de ligações entre os nós. Para isso, foi estabelecido que o número máximo de ligações é duas vezes a quantidade de nós menos um, e o número mínimo é a quantidade de nós menos um, garantindo que cada nó tenha pelo menos uma ligação. Finalmente, a quantidade máxima de nós permitida na topologia é fixada em 11, devido ao número de atributos utilizados, isso pode ser alterado a medida que se incluir ou retira os atributos.

Para os experimentos, foi definido 12 nós, correspondentes ao número de variáveis utilizadas, incluindo a variável alvo. As variáveis selecionadas foram: Geração por Usinas Hidrelétricas do SIN, Geração por Usinas Térmicas do SIN, Carga de Energia do submercado Sudeste/Centro-Oeste, Carga de Energia Total do SIN, Energia Armazenada Total do SIN, Energia Armazenada no Subsistema Sudeste/Centro-Oeste, Demanda Máxima do submercado Sudeste/Centro-Oeste, Preço da Liquidação das Diferenças, Capacidade Instalada da geração hidrelétrica, Produto Interno Bruto, Custo Marginal de Operação e a variável objetivo, que é a Curva de Preço no longo prazo fornecida pela Decide para o submercado Sudeste/Centro-Oeste.

Após a configuração desses parâmetros, foram geradas diversas topologias. A avaliação das melhores topologias foi realizada com base no Erro Percentual Absoluto Médio (MAPE). A Figura 46 apresenta seis topologias geradas pelo Algoritmo Genético.

Conforme mencionado anteriormente sobre a taxa de retenção dos melhores indivíduos, observa-se que, na maioria das topologias, as variáveis Energia Armazenada no Subsistema Sudeste/Centro-Oeste (EA_SECO), Geração por Usinas Hidrelétricas do SIN (HID) e Geração por Usinas Térmicas do SIN (TER) estão frequentemente presentes. Isso indica que essas variáveis têm uma contribuição significativa para a formação do preço, conforme demonstrado na Seção 5.1, que trata do estudo das variáveis.

Figura 46. Exemplos de topologias geradas automaticamente pelo Algoritmo Genético



5.3 Resultados de Previsão dos Cenários do Preço da Curva *Forward* de Energia Elétrica

A previsão da curva *forward* de preço de energia elétrica desempenha um papel fundamental na tomada de decisões estratégicas no setor elétrico, fornecendo subsídios para a gestão de contratos, análise de risco e planejamento de longo prazo. Nesta seção, são apresentados os resultados das projeções realizadas para um horizonte de até seis anos, utilizando a metodologia desenvolvida ao longo deste estudo.

Inicialmente, são apresentadas outras técnicas para comparação com o método proposto, em seguida é demonstrado os resultados dos cenários gerados e por fim os resultados são analisados em termos de desempenho preditivo, considerando métricas estatísticas para avaliar a acurácia do modelo.

Para avaliar o desempenho da técnica proposta, que são as Redes Bayesianas Dinâmicas, foram criados experimentos comparativos com outros algoritmos de regressão. Esses métodos são: Linear Regression (WILEY; SONS, 2013); Support Vector Regression (SVR) (DRUCKER et al., 1996) e (ARA; MAIA, 2023) e Extreme Gradient Boosting (XGBoost) (CHEN; GUESTRIN, 2016) e (GOMES, 2019), todos são utilizados em projeção que utilizam dados de séries temporais.

Nos experimentos realizados com as três técnicas diferentes de regressão, avaliou-se o impacto das variáveis de entrada utilizando o cálculo do fator de importância, um método intrínseco às técnicas de árvore de regressão. Antes de proceder com essa análise, foi necessário normalizar as variáveis.

Para o pré-processamento dos dados, explorou-se o uso de diversas técnicas, incluindo *StandardScaler* (distribuição normal), *MinMaxScaler*, *RobustScaler* e *QuantileTransformer*. Entre estas, o método *MinMaxScaler* foi selecionada para os experimentos por ter apresentado os melhores resultados.

Ao utilizar as bibliotecas de cada método é necessário fazer alguns ajuste nas configurações, em busca de obter resultados melhores. As configurações realizadas nos métodos de Regressão Linear, *Support Vector Regression* e *Extreme Gradient Boosting*, são demonstradas e descritas nas Tabelas 10, 11 e 12, respectivamente.

Tabela 10. Parâmetros de configuração utilizado no método de Regressão Linear

Parâmetro	Valor	Descrição
Fit_intercept	True	Se deve calcular a interceptação para este modelo. Se definido como False, nenhuma interceptação será usada nos cálculos (ou seja, espera-se que os dados sejam centralizados).
copy_X	True	Se True, X será copiado; caso contrário, ele poderá ser substituído.
N_jobs	None	O número de jobs/trabalhos a serem usados para o cálculo. Isso só fornecerá aceleração no caso de problemas suficientemente grandes, ou seja, se primeiro $n_targets > 1$ e, em segundo lugar, X for disperso. Nenhum significa 1.
positive	False	Quando definido como True, força os coeficientes a serem positivos. Esta opção só é compatível com arrays densos.

Fonte: Autor

Tabela 11. Parâmetros de configuração utilizado no método de *Support Vector Regression*

Parâmetro	Valor	Descrição
kernel	rbf	Especifica o tipo de kernel a ser usado no algoritmo.
degree	3	Se $\gamma = 'scale'$ for passado então ele usa $1 / (n_features * X.var())$ como valor de gama, se 'auto', usa $1/n_features$ se for float, não deve ser negativo.
gamma	scale	Coefficiente de kernel para 'rbf', 'poly' e 'sigmoid'.
coef0	0.0	Termo independente na função do kernel. É significativo apenas em 'poly' e 'sigmoid'.
epsilon	0.1	Épsilon no modelo epsilon-SVR. Ele especifica o tubo epsilon dentro do qual nenhuma penalidade está associada na função de perda de treinamento com pontos previstos dentro de uma distância epsilon do valor real.
max_iter	-1	Limite rígido para iterações no solucionador ou -1 para nenhum limite.
shrinking	True	Se deve usar a heurística de redução.

Fonte: Autor

As variáveis disponíveis para os experimentos com os métodos de regressão, após o processo de normalização, são apresentadas na Figura 47. Estas, foram selecionadas com base em sua relevância para a modelagem e previsão do preço de energia, garantindo que contemplem os principais fatores que influenciam a dinâmica do mercado.

Figura 47. Variáveis disponíveis para os experimentos pelos métodos de regressão.

```
FEATURES_LPC = ['CE_SECO_MIN_MAX', 'CMO_SECO_MIN_MAX',
'EA_SECO_MIN_MAX', 'GE_SECO_N_MIN_MAX', 'DM_SECO_MIN_MAX',
'PLD_SECO_MIN_MAX', 'CI_SECO_MIN_MAX', 'PIB_MIN_MAX']
```

Fonte: Autor

Após a normalização dos dados, calculou-se a importância de cada atributo no processo de predição, analisando o fator de importância em relação à variável alvo, que é o

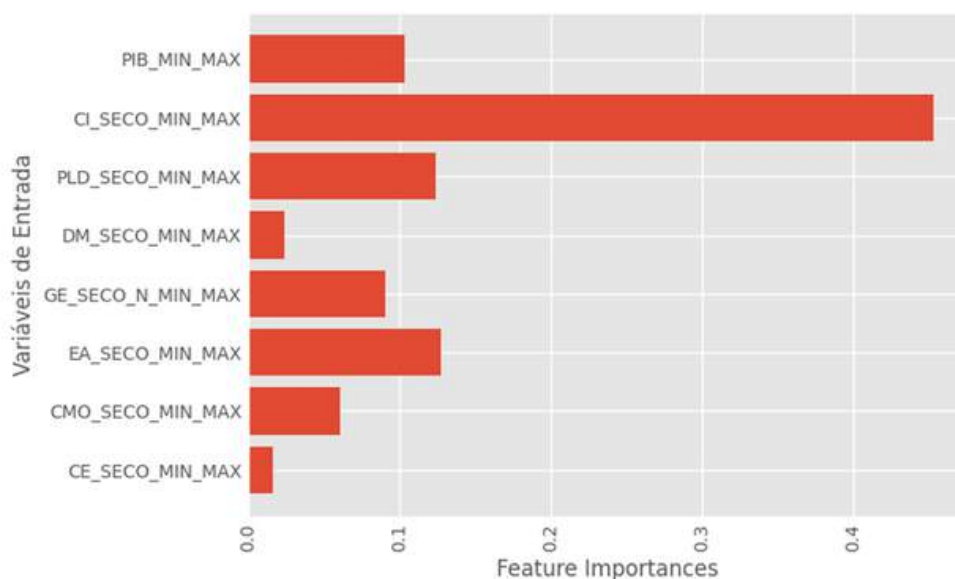
Tabela 12. Parâmetros de configuração utilizado no método de *Extreme Gradient Boosting*

Parâmetro	Valor	Descrição
booster	gbtree	Qual reforço usar. Pode ser ‘gbtree’, ‘gblinear’ ou ‘dart’; ‘gbtree’ e ‘dart’ usam modelos baseados em árvore, enquanto ‘gblinear’ usa funções lineares.
eta	0.3	Redução do tamanho do passo usada na atualização para evitar overfitting. Após cada etapa de boosting, podemos obter diretamente os pesos das novas features, e o ‘eta’ reduz os pesos das features para tornar o processo de boosting mais conservador.
max_depth	9	Profundidade máxima de uma árvore. Aumentar esse valor tornará o modelo mais complexo e com maior probabilidade de overfit.
min_child_weight	1	Soma mínima do peso da instância (hessian) necessária em um filho.
n_estimators	500	A função para medir a qualidade de uma divisão. Os critérios suportados são “squared_error” para o erro quadrático médio, que é igual à redução da variância como critério de seleção de recursos e minimiza a perda L2 usando a média de cada nó terminal, “friedman_mse”, que usa o erro quadrático médio com a pontuação de melhoria de Friedman para potencial splits, “absolute_error” para o erro médio absoluto, que minimiza a perda L1 usando a mediana de cada nó terminal, e “poisson” que usa a redução no desvio de Poisson para encontrar splits.
eval_metric	MAE	Métricas de avaliação para dados de validação, uma métrica padrão será atribuída de acordo com o objetivo.
subsample	0.8	Proporção de subamostras das instâncias de treinamento. Defini-lo como 0,5 significa que o XGBoost criaria aleatoriamente amostras com metade dos dados de treinamento antes de cultivar árvores. e isso evitará overfitting. A subamostragem ocorrerá uma vez em cada iteração de reforço.
learning_rate	0.1	A taxa de aprendizagem reduz a contribuição de cada árvore em “learning_rate”. Existe uma compensação entre taxa de aprendizagem e “n_estimadores”. Os valores devem estar no intervalo [0.0, inf].

Fonte: Autor

preço de energia de longo prazo fornecido pela empresa de consultoria DCIDE. O resultado dessa análise é ilustrado na Figura 48. Como pode ser observado, a variável que se destacou como a mais importante foi a Capacidade Instalada no submercado Sudeste/Centro-Oeste (CI_SECO_MIN_MAX).

Figura 48. Resultado da avaliação do fator de importância entre as variáveis de entrada com o preço no longo prazo fornecido pela Dcide



Fonte: Autor

Após a análise das variáveis disponíveis para os experimentos, foram selecionadas as cinco que apresentaram o maior fator de importância. A escolha dessas variáveis foi baseada em critérios de relevância para a modelagem preditiva, garantindo que apenas os atributos mais influentes fossem considerados. Na Figura 49 são listadas as variáveis em questão.

Figura 49. Variáveis selecionadas como entrada para os experimentos dos métodos de regressão.

```
FEATURES_LPC = ['EA_SECO_MIN_MAX', 'GE_SECO_N_MIN_MAX',
                'PLD_SECO_MIN_MAX', 'CI_SECO_MIN_MAX', 'PIB_MIN_MAX']
```

Fonte: Autor

A organização dos experimentos seguiu o roteiro descrito a seguir: foram definidos três cenários para avaliar o desempenho das Redes Bayesianas Dinâmicas em comparação com outras técnicas previamente mencionadas, variando o horizonte de projeção entre 3 e 6 anos. Esse arranjo permite analisar o impacto das inferências feitas pelas RBDs no momento da projeção, destacando a relevância do conhecimento tácito do especialista na análise.

A base de dados utilizada compreende uma série histórica de 2012 a 2023, sendo dividida em conjuntos de treinamento e teste. Para as previsões com horizonte de 6 anos, os dados de teste correspondem ao período de 2012 a 2017, enquanto os dados de treinamento abrangem de 2018 a 2023, conforme ilustrado na Figura 50.

Figura 50. Série histórica do banco de dados: Divisão entre conjunto de treinamento e teste para cenários de projeções de 6 anos.



Fonte: Autor

Nos cenários com previsões de 3 anos, o conjunto de teste abrange o período de 2012 a 2020, enquanto o conjunto de treinamento corresponde aos anos de 2021 a 2023, conforme mostrado na Figura 51.

Figura 51. Série histórica do banco de dados: Divisão entre conjunto de treinamento e teste para cenários de projeções de 3 anos.

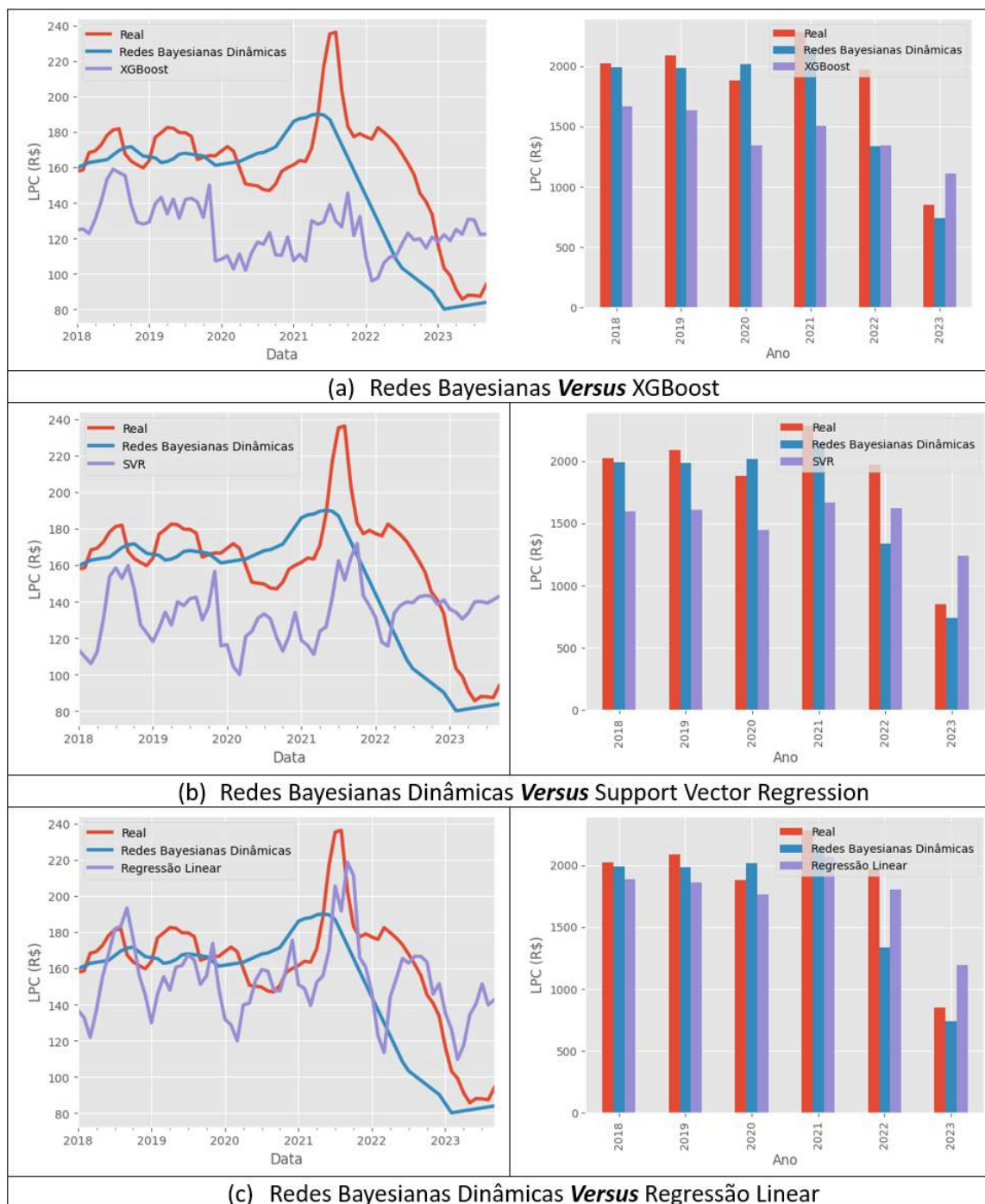


Fonte: Autor

Foram realizados três cenários de experimentos para a previsão da curva *forward* do preço da energia elétrica, considerando um horizonte de 6 anos. No cenário apresentado na Figura 52, não houve qualquer intervenção externa na Rede Bayesiana Dinâmica, ou seja,

nenhuma inferência foi realizada pelo usuário. Esse resultado foi gerado exclusivamente com base nos eventos históricos da série.

Figura 52. Resultados do Cenário 1: (a) RBD sem inferência de informações externa ao modelo comparando com os métodos: XGBoost; (b) *Support Vector Regression* e (c) Regressão Linear.



Fonte: Autor

Na Figura 52 (a), são apresentados os resultados da projeção da Rede Bayesiana

Dinâmica e do algoritmo XGBoost, juntamente com os dados reais fornecidos pela Dcide (Real). O gráfico de linha permite uma análise mais detalhada do comportamento dos preços em nível mensal, enquanto o gráfico de barras mostra os resultados agregados por ano.

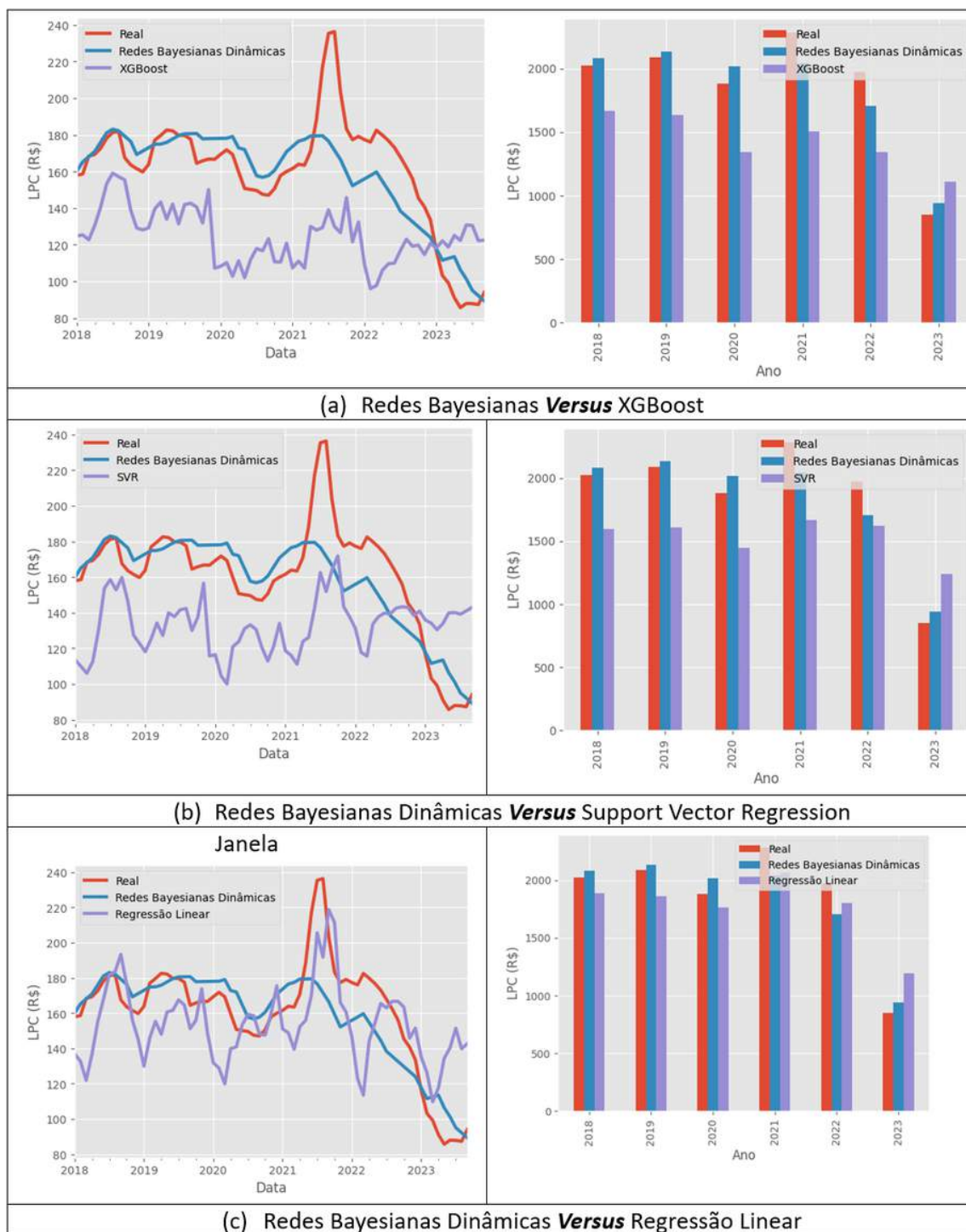
Em geral, a previsão da RBD reflete o comportamento das curvas do preço real, embora em alguns momentos esteja abaixo ou acima. Uma exceção notável ocorreu em 2021, quando houve um pico no preço da energia devido a crise hídrica; os reservatórios do Sudeste/Centro-Oeste, que são responsáveis por cerca de 70% do armazenamento do país, estavam em um nível mínimo histórico de 26% e não deixando de mencionar também o efeito da pandemia. Por outro lado, os resultados apresentados pelo XGBoost, ao longo de quase toda a série histórica, mostraram-se consistentemente abaixo dos preços reais, exibindo uma sazonalidade marcante. Os gráficos de barra anuais indicam que o ano com a maior discrepância entre o valor acumulado previsto para ambos os métodos e o real, foi 2022.

Na Figura 52 (b) e (c), que apresentam os resultados dos métodos Support Vector Regression (SVR) e Regressão Linear, respectivamente, observa-se que a Regressão Linear teve um desempenho superior ao das Redes Bayesianas Dinâmicas no ano de 2022, conforme indicado pelo gráfico de barras em (c).

Além disso, ao analisar os gráficos de linha, percebe-se que a Regressão Linear (c) apresenta um ajuste mais próximo da série real em determinados períodos, especialmente nos anos de alta volatilidade, como 2021 e 2022. Já o método SVR, mostrado na Figura (b), demonstra maior dificuldade em capturar variações acentuadas nos preços, resultando em previsões mais suavizadas em comparação às demais abordagens. Embora o cenário gerado pela RBD não tenha capturado o pico ocorrido em 2021, ao longo da série, o modelo consegue acompanhar o comportamento das curvas.

No estudo das variáveis, identificou-se que a Energia Armazenada no submercado Sudeste/Centro-Oeste apresenta uma correlação de 66% com o preço de energia elétrica no longo prazo. Com base nesse dado, foi conduzido o experimento do Cenário 2, conforme ilustrado na Figura 53. Durante determinados períodos no horizonte de projeção, foram inseridos ajustes no comportamento dos reservatórios. Isso significa que, de acordo com o número de categorias definidas pelo usuário, diferentes níveis de armazenamento foram representados. No processo de discretização foram utilizados 4 categorias que são: muito baixo, baixo, intermediário, alto e muito alto. O número de ajustes realizados variava conforme a escolha do usuário.

Figura 53. Resultados do Cenário 2: (a) RBD com inferência de informações externas na variável Energia Armazenada comparando com os métodos: XGBoost; (b) *Support Vector Regression* e (c) Regressão Linear.



Fonte: Autor

Ao analisar os resultados apresentados na Figura 53 a), observa-se que, nos anos iniciais, a Rede Bayesiana Dinâmica acompanhou a tendência dos valores reais. Em 2020

analisando a curva real, houve uma queda no primeiro semestre, seguida de uma retomada do crescimento no segundo semestre. É interessante notar que a RBD reproduziu esse mesmo comportamento, ainda que com valores ligeiramente superiores.

Nos dois últimos anos, a RBD apresentou um desempenho superior ao do Experimento do cenário 1, especialmente no primeiro semestre de 2022. Já os resultados obtidos pelos outros métodos permaneceram os mesmos do cenário 1. Isso demonstra que a metodologia proposta permite a geração de diferentes cenários de preços de energia elétrica toda vez que se realizar inferências em alguns atributos do modelo.

Comparando os resultados do Experimento do cenário 1 (Figura 52) com os do cenário 2 (Figura 53), percebe-se uma melhora significativa nos resultados da RBD, apenas com a inserção de informações em uma única variável do modelo.

Para demonstrar que, à medida que são inseridas informações externas e mais variáveis no modelo de projeção, diferentes cenários da curva de preço são gerados, evidenciando a flexibilidade e a capacidade de adaptação do método a essas alterações. Dessa forma, o modelo tende a retornar cenários de curvas com uma taxa de acerto aprimorada.

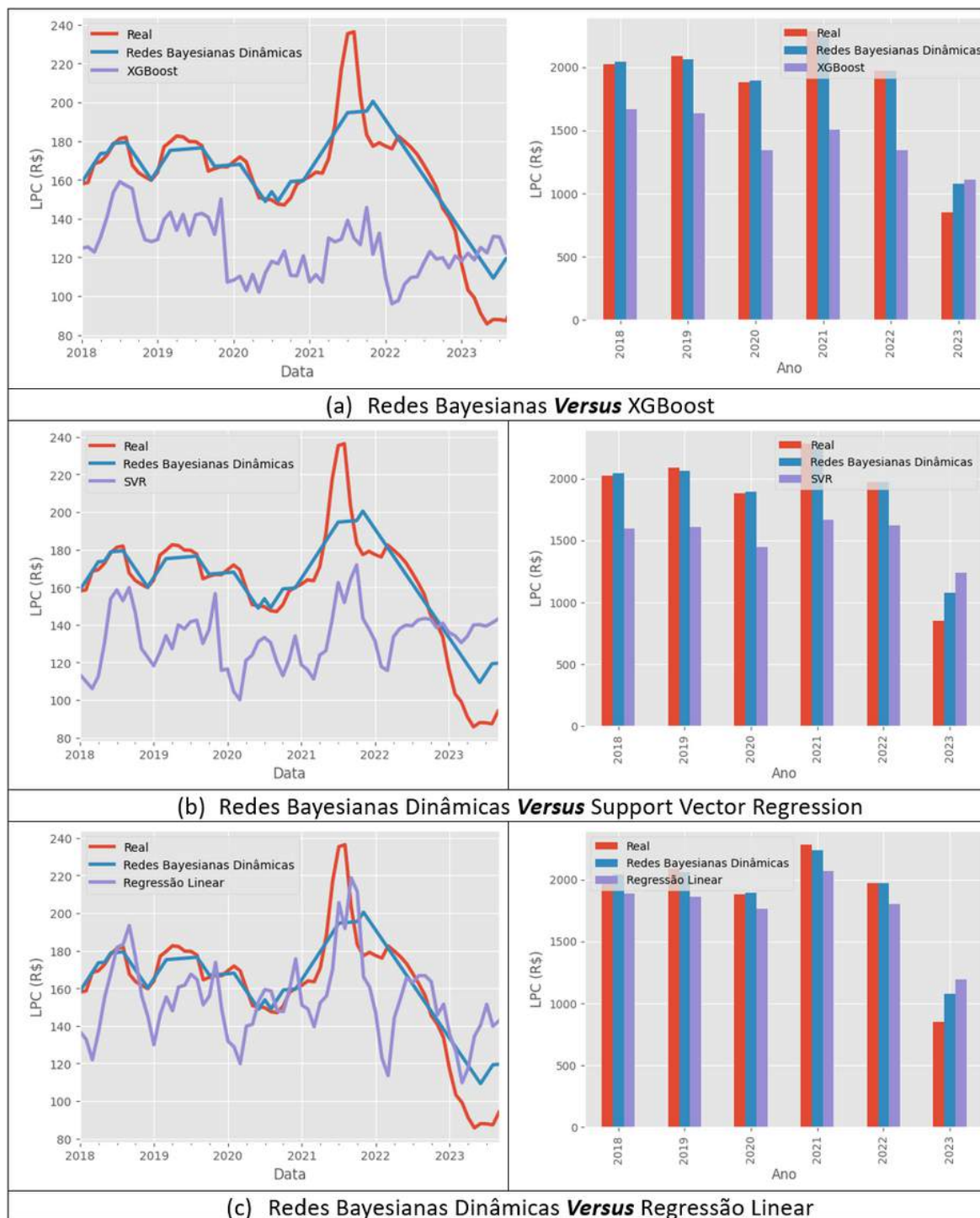
Os resultados obtidos ao utilizar essas inferências indicam que o usuário possui autonomia e responsabilidade sobre as projeções, uma vez que pode ajustar os parâmetros e incluir novas informações para refinar as estimativas. Comparando os diferentes experimentos, observa-se que a Rede Bayesiana Dinâmica melhora significativamente seus resultados quando há inserção de novas variáveis, demonstrando sua eficiência na captura de padrões complexos e na geração de previsões mais alinhadas ao comportamento real do mercado.

Por fim, foi realizada mais uma rodada de experimentos para o horizonte de 6 anos, conforme apresentado na Figura 54. Para esse experimento foram inseridos informações externas ao modelo nas variáveis Energia Armazenada, o Preço de Liquidação das Diferenças.

Conforme discutido anteriormente, as projeções de longo prazo do PLD não seguem o padrão histórico dos preços de energia fornecidos pela Dcide, como demonstrado na Figura 38 da Seção 5.1, que trata do Estudo das Variáveis. No entanto, quando combinado com outras variáveis, o PLD pode influenciar significativamente os resultados das previsões realizadas pela RBD.

O Experimento do cenário 3, apresentado na Figura 54 a), indica que, ao incorporar apontamentos significativos (ou seja, próximos aos eventos reais), é possível obter resultados mais precisos em cenários de projeção de preços. Neste experimento, foram inseridos dados em duas variáveis do modelo com uma frequência maior do que no experimento do cenário 2.

Figura 54. Resultados do Cenário 3: (a) RBD com inferência de informações externas nas variáveis Energia Armazenada e o Preço da Liquidação das Diferenças comparando com os métodos: XGBoost; (b) *Support Vector Regression* e (c) Regressão Linear.



Fonte: Autor

Os resultados gerados pela Rede Bayesiana Dinâmica demonstraram um desempenho positivo, especialmente nos primeiros quatro anos, capturando bem as tendências

observadas no histórico real de preços. Embora o modelo não tenha identificado o comportamento atípico do segundo semestre de 2021, ele conseguiu acompanhar a tendência de queda dos preços ao longo de 2022.

A Figura 54 também apresenta comparações entre a RBD e outros modelos, como XGBoost, Support Vector Regression e Regressão Linear. Neste âmbito, observa-se que a RBD acompanha de forma mais consistente o comportamento real dos preços em comparação com os outros métodos testados. O XGBoost, por exemplo, apresenta maior divergência em relação aos valores reais, especialmente em momentos de maior volatilidade. Já o SVR e a Regressão Linear demonstram um ajuste inferior à RBD, falhando em capturar adequadamente picos e quedas abruptas nos preços de energia.

Com o objetivo de aprofundar a análise e demonstrar a importância dos apontamentos na precisão das previsões do modelo, foram realizados três novos experimentos considerando um horizonte de 3 anos. Esses experimentos visam avaliar o impacto da inserção de diferentes níveis de informação no desempenho da Rede Bayesiana Dinâmica e comparar os resultados obtidos com os métodos XGBoost, *Support Vector Regression* e Regressão Linear.

A Figura 55 apresenta os resultados do cenário de previsão de preços de energia para o período de 2021 a 2023. O conjunto de dados de treinamento abrangeu o período de 2012 a 2020, conforme ilustrado na Figura 51.

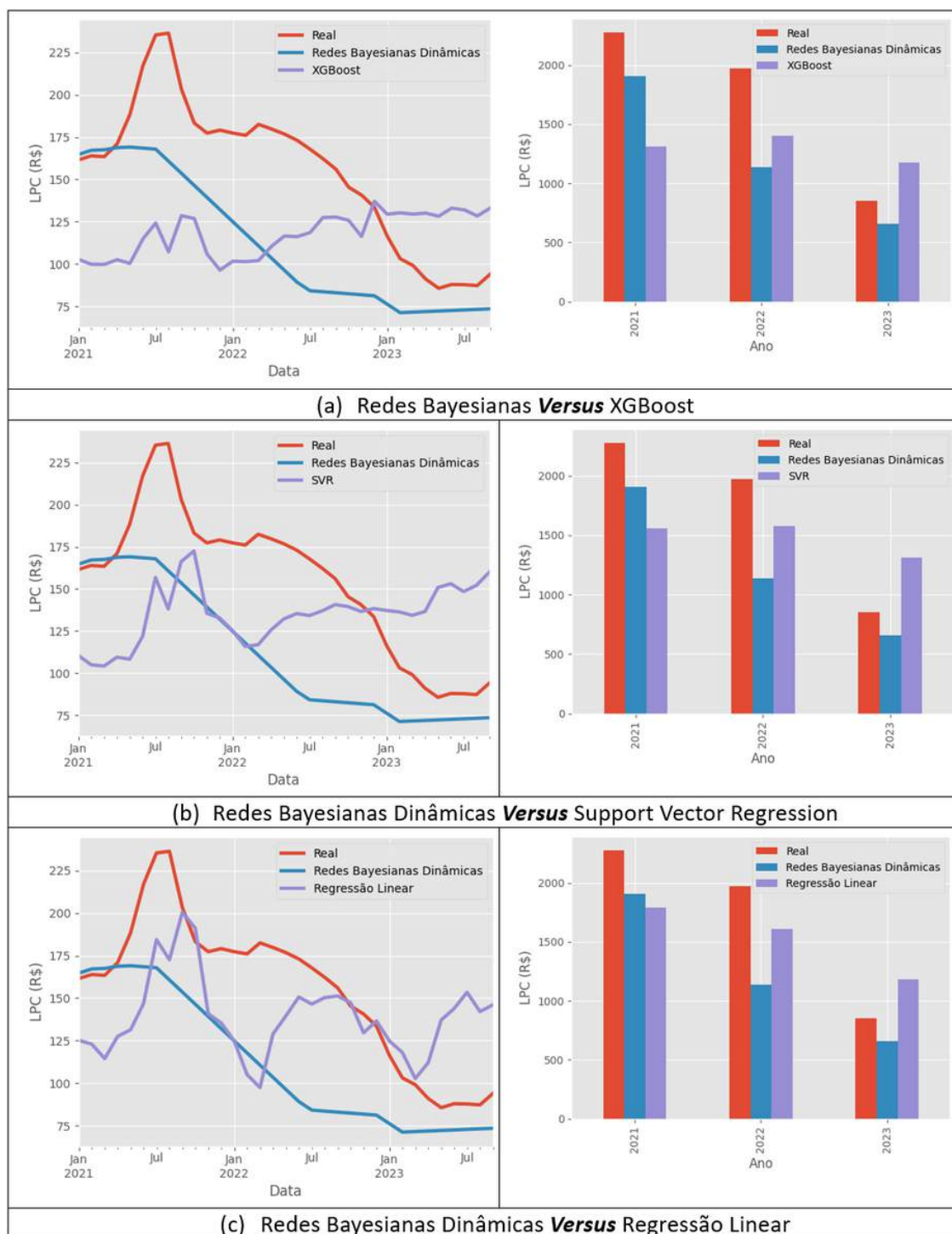
Como apresentado na Figura 55 a), no primeiro semestre de 2021, a Rede Bayesiana Dinâmica acompanhou inicialmente a tendência do preço real, mas posteriormente apresentou uma leve subestimação, com a curva mostrando uma queda. Apesar dos valores inferiores, a RBD manteve a coerência com a trajetória geral do preço real ao longo do período analisado.

O método XGBoost, por outro lado, exibiu uma sazonalidade crescente, mas não conseguiu capturar corretamente a tendência do preço real. Seu comportamento se mostrou mais instável, apresentando desvios significativos, especialmente em momentos de alta volatilidade.

Avaliando os outros dois métodos, a Regressão Linear demonstrou um desempenho relativamente melhor, conseguindo acompanhar a tendência do preço real em determinados períodos. Entretanto, seu ajuste ainda é limitado, falhando em capturar variações mais abruptas. Já o Support Vector Regression (SVR) apresentou um desvio considerável em relação à curva real, não conseguindo acompanhar o comportamento de flutuação dos preços.

Nos experimentos 5 e 6, foram realizadas inferências utilizando diferentes combinações de variáveis. No Experimento 5, apenas na variável Energia Armazenada foi inserido inferências, enquanto no Experimento 6 foram incorporadas tanto Energia Armazenada

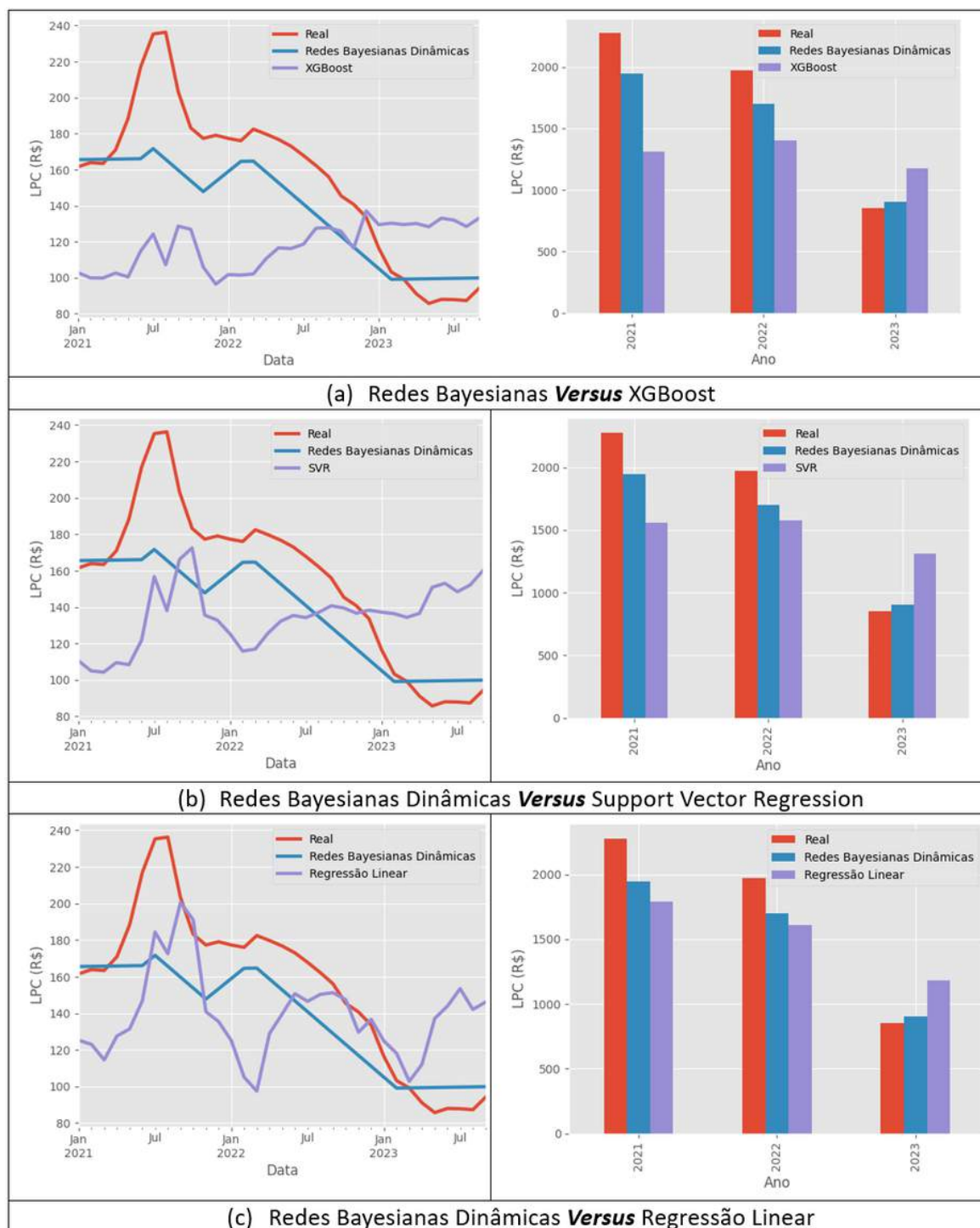
Figura 55. Resultados do Cenário 4: (a) RBD sem inferência de informações externas ao modelo comparando com os métodos: XGBoost; (b) *Support Vector Regression* e (c) Regressão Linear no horizonte de 3 anos.



Fonte: Autor

quanto o Preço de Liquidação das Diferenças. Os resultados desses experimentos estão ilustrados nas Figuras 56 e 57, respectivamente.

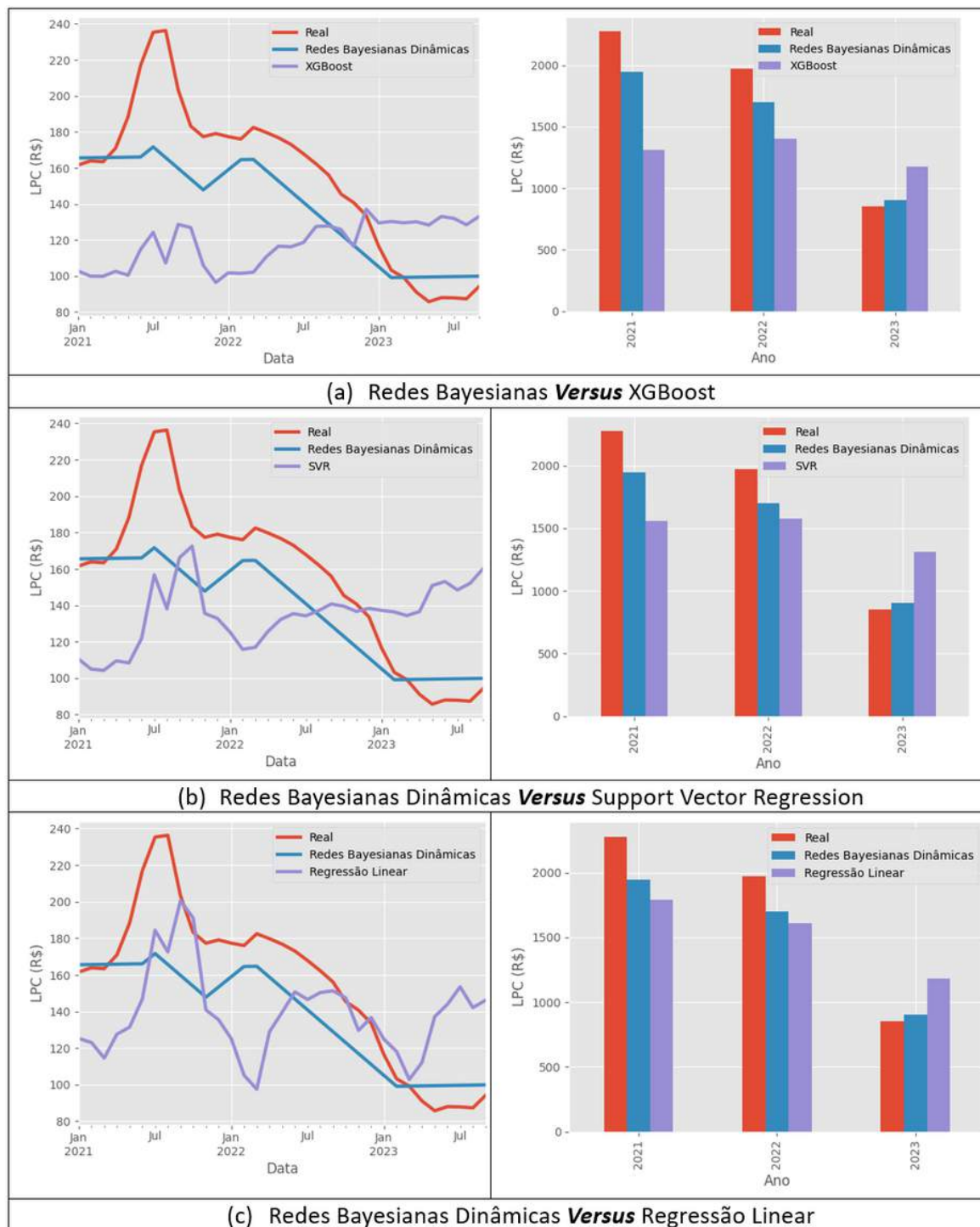
Figura 56. Resultados do Cenário 5: (a) RBD com inferência de informações externas na variável Energia Armazenada comparando com os métodos: XGBoost; (b) *Support Vector Regression* e (c) Regressão Linear no horizonte de 3 anos.



Fonte: Autor

Como discutido anteriormente, a inclusão de um maior volume de informações ao modelo possibilita a projeção de uma variedade de cenários. Em relação à confiabilidade

Figura 57. Resultados do Cenário 6: (a) RBD com inferência de informações externas nas variáveis Energia Armazenada e o Preço de Liquidação das Diferenças comparando com os método: XGBoost; (b) *Support Vector Regression* e (c) Regressão Linear no horizonte de 3 anos.



Fonte: Autor

dessas informações, é importante notar que quanto mais precisas e próximas da realidade forem, mais acurados serão os cenários projetados

Para avaliar estatisticamente o desempenho da Rede Bayesiana Dinâmica em relação aos métodos XGBoost, Support Vector Regression e Regressão Linear nos horizontes temporais de 3 e 6 anos, utilizou-se a Média Percentual Absoluta do Erro (MAPE) como métrica de comparação. Esse indicador permite quantificar a precisão das previsões, fornecendo uma medida relativa do erro em relação aos valores reais.

Média Percentual Absoluta do Erro do inglês (Mean Absolute Percentage Error - MAPE) é uma métrica estatística amplamente utilizada para avaliar a precisão de modelos preditivos, especialmente em séries temporais e previsão de preços, é definida pela Equação 5.1. Seu principal objetivo é quantificar o erro relativo entre os valores previstos e os valores reais, expressando esse erro em termos percentuais.

$$MAPE = \frac{1}{n} \left| \sum_{i=1}^n \frac{y_i - x_i}{y_i} \right| 100 \quad (5.1)$$

onde: y_i representa o valor real no instante i ; x_i representa o valor previsto no instante i e; n é o número total de observações.

Os resultados comparativos da taxa de erro entre os diferentes métodos são apresentados na Tabela 13. A análise dos valores obtidos com a Rede Bayesiana Dinâmica para os horizontes de 3 e 6 anos reforça as discussões anteriores desta seção sobre a importância de inserir *insights* no horizonte de projeção.

Tabela 13. Tabela com a taxa MAPE dos experimentos entre a Rede Bayesiana Dinâmica e os métodos: XGBoost; *Support Vector Regression* e Regressão Linear nos horizontes de 3 e 6 anos

Técnica		3 anos	6 anos
Redes Bayesianas Dinâmicas	Padrão	26.54%	12.80%
	Inferência EA	12.38%	9.72%
	Inferência PLD/DA	7.39%	6.47%
XGBoost		24.98%	19.03%
Regressão Linear		17.07%	10.63%
Support Vector Regression		22.39%	17.53%

Fonte: Autor

Os resultados obtidos na configuração de inferência “Padrão” foram baseados exclusivamente na série histórica processada pela RBD. Em um segundo experimento, ajustes foram introduzidos na variável Energia Armazenada, resultando em uma redução significativa do erro. No horizonte de 3 anos, a taxa de erro caiu de 26,54% para 12,38%, enquanto no horizonte de 6 anos, a redução foi de 12,80% para 9,72%.

Na terceira etapa do experimento, a inclusão de informações externas adicionais combinando as variáveis Preço de Liquidação das Diferenças e Energia Armazenada,

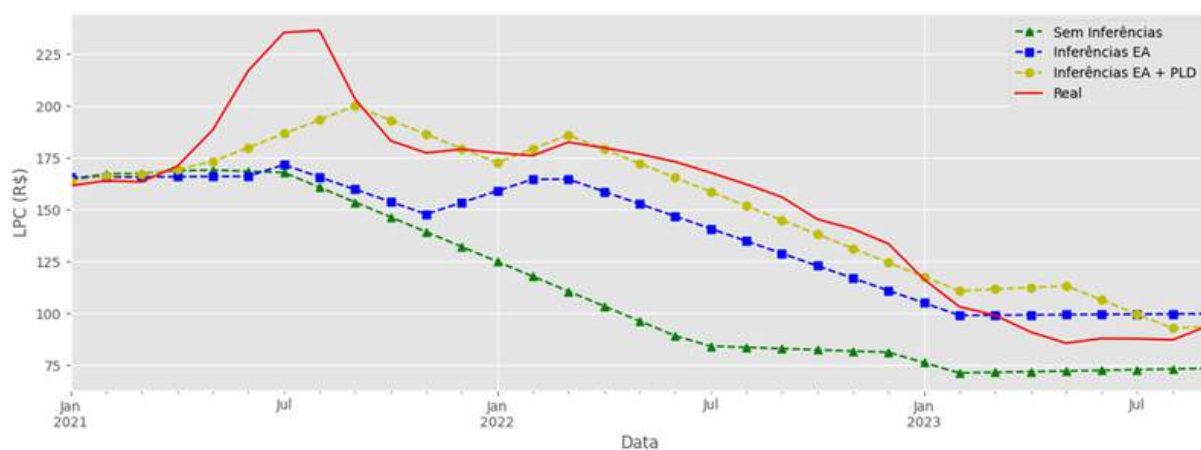
promoveu uma melhoria ainda mais expressiva na precisão do modelo. Nesse cenário, o erro foi reduzido para 7,39% no horizonte de 3 anos e 6,47% no horizonte de 6 anos, demonstrando a eficácia da incorporação de múltiplas informações nas variáveis para aprimorar as previsões de cenários de preços de energia.

Esses resultados evidenciam que a Rede Bayesiana Dinâmica se beneficia da inclusão de variáveis exógenas bem calibradas, tornando suas projeções mais robustas e alinhadas com a realidade do mercado de energia. Além disso, reforçam que, quanto maior a qualidade e a relevância das informações fornecidas ao modelo, mais precisos serão os cenários de curva *forward* gerados de preço de energia elétrica.

O objetivo de apresentar os resultados obtidos com os métodos XGBoost, *Support Vector Regression* e Regressão Linear, amplamente utilizados na análise de séries temporais, foi para justificar a escolha das Redes Bayesianas Dinâmicas. A comparação entre esses métodos permite evidenciar as vantagens da RBD na modelagem de cenários de previsão dos preços de energia elétrica, destacando sua capacidade de capturar dependências dinâmicas e lidar melhor com a incerteza inerente ao problema.

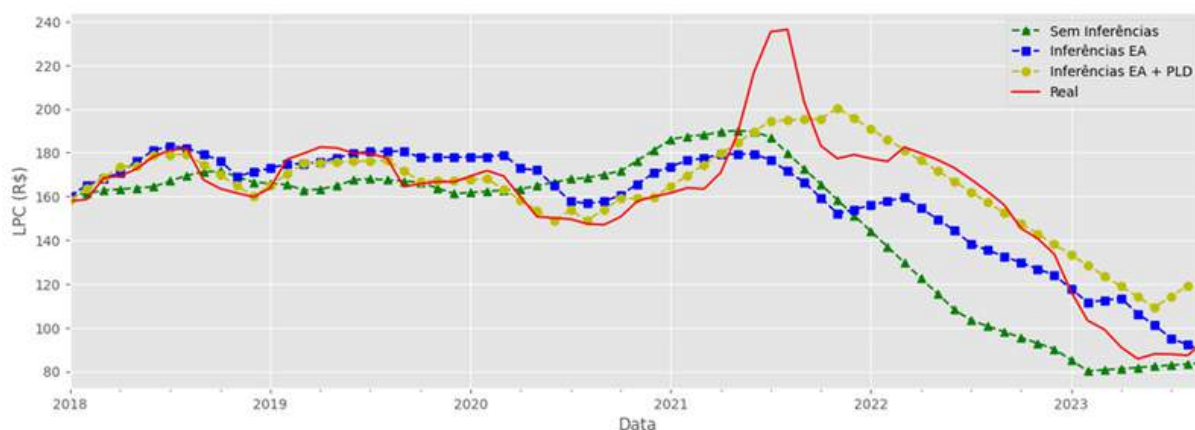
No decorrer dos experimentos, os resultados são apresentados individualmente para cada cenário analisado. Além disso, as Figuras 58 e 59 ilustram os cenários projetados para os horizontes de 3 e 6 anos, respectivamente.

Figura 58. Resultados dos Cenários de Previsão do Preço da Curva *Forward* no horizonte de 3 anos.



Fonte: Autor

Figura 59. Resultados dos Cenários de Previsão do Preço da Curva *Forward* no horizonte de 6 anos.



Fonte: Autor

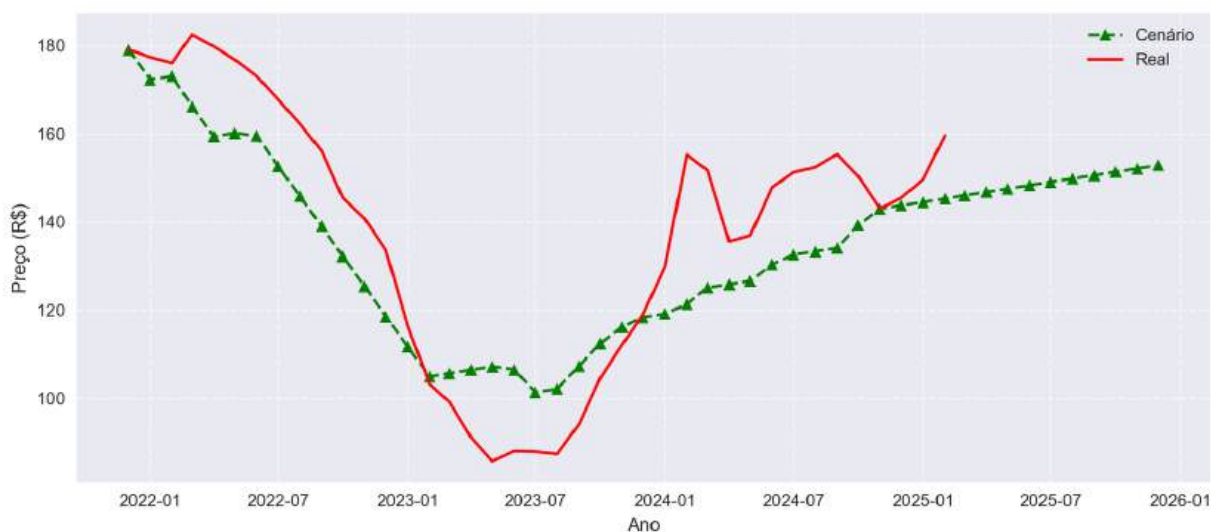
Os cenários modelados incluem:

- Cenário base (sem inferência do usuário): a projeção é realizada sem ajustes externos, utilizando apenas os dados históricos e a estrutura do modelo.
- Cenário com inferência da variável “Energia Armazenada”: incorpora informações sobre a energia armazenada para refinar as previsões.
- Cenário com múltiplas inferências: considera tanto a variável “Energia Armazenada” quanto o “Preço de Liquidação das Diferenças”, permitindo uma modelagem mais realista e próxima das dinâmicas do mercado.

Os experimentos realizados anteriormente foram projetados até setembro de 2023, pois, até então, os dados históricos de preços de energia de longo prazo fornecidos pela DCide estavam disponíveis para validação da metodologia proposta. Posteriormente, a série histórica foi atualizada até fevereiro de 2025, resultando em um banco de dados que agora abrange os anos de 2012 a 2025.

Com essa atualização, foram conduzidos dois novos cenários experimentais para horizontes de 4 anos. Para validação, utilizou-se o período de janeiro de 2022 a fevereiro de 2025, pois é até esse ponto que os dados reais estão disponíveis, conforme apresentado na Figura 60. O cenário gerado foi produzido automaticamente pelo modelo, sem necessidade de intervenção do usuário.

Figura 60. Resultados dos Cenários de Previsão do Preço da Curva *Forward* no horizonte de 4 anos.

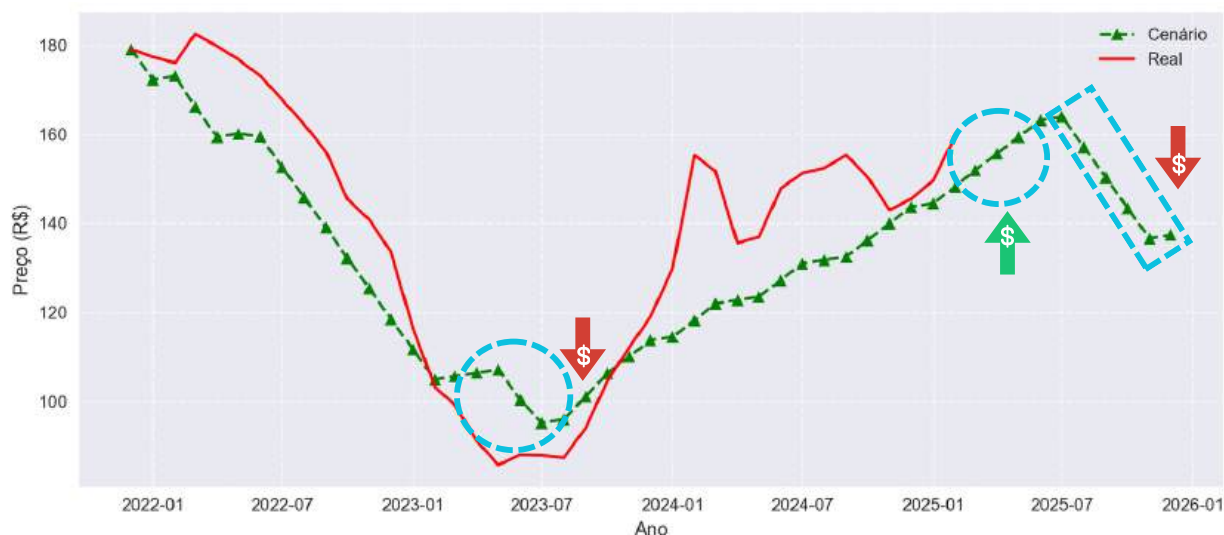


Fonte: Autor

A tendência geral do modelo projetado acompanha a direção dos preços reais, mas com algumas diferenças notáveis em períodos específicos. De 2022 a 2023, o modelo segue um padrão de queda semelhante ao observado nos preços reais, embora de forma mais suave e sem capturar totalmente as oscilações abruptas. A partir de mês 7/2023, a projeção mostra uma recuperação gradual e estável, enquanto os preços reais apresentam maior volatilidade com variações bruscas. Mas a projeção acompanha a tendência de crescimento.

Em seguida, foi gerado um novo cenário para o mesmo período anterior, porém com inferências na variável Energia Armazenada em três momentos distintos da projeção, conforme ilustrado na Figura 61. Isso confirma que, à medida que novas informações são inseridas no modelo, elas são incorporadas à projeção, resultando em ajustes que se propagam para os períodos futuros.

Figura 61. Resultados dos Cenários de Previsão do Preço da Curva *Forward* no horizonte de 4 anos, com inferências de informações externas na variável Energia Armazenada.



Fonte: Autor

Os dois momentos de tendência de queda (destacados em círculos azuis) ocorrem após a inclusão de informações que indicam alto nível de energia armazenada. Essa relação faz sentido, pois um maior volume de energia armazenada tende a indicar menos risco de escassez futura, o que afeta diretamente o preço de energia.

Já no terceiro ponto, que apresenta tendência de alta, a inferência de um baixo nível de energia armazenada, logo acontece o oposto. Pois com o nível baixo nos reservatórios, tende-se a utilizar mais energia elétrica das usinas térmicas, onde o custo operacional é maior, com isso o preço da energia tende a crescer.

O comportamento do modelo reforça a importância das variáveis exógenas no processo de previsão, demonstrando como diferentes cenários de armazenamento de energia podem afetar significativamente as projeções de preço no médio e longo prazo.

A Tabela 14 apresenta os valores da taxa erro para dois cenários de previsão de preço da energia elétrica no horizonte de quatro anos: um sem inferências e outro com inferências na variável Energia Armazenada.

Os resultados indicam que o cenário com inferência apresentou um erro MAPE de 9,19%, enquanto o cenário sem inferência teve um erro de 9,50%. Essa diferença, ainda que pequena, confirma a tendência observada em outros experimentos: à medida que informações mais precisas e relevantes são incorporadas ao modelo, a qualidade da previsão melhora, reduzindo o erro.

Tabela 14. Tabela com a taxa MAPE dos Experimentos de Cenários sem inferência e com inferências, na variável Energia Armazenada no horizonte de 4 anos.

Previsão utilizando RBD	4 anos
Cenário sem inferência	9.50%
Cenário com inferência	9.19%

Fonte: Autor

Os resultados obtidos demonstram-se promissores, evidenciando a relevância da participação dos usuários na geração de cenários de preços da curva *forward* de energia elétrica, por meio de possíveis ajustes e inferências. No entanto, mesmo sem essas intervenções, o modelo é capaz de gerar cenários que representam adequadamente os preços de mercado.

Além disso, a metodologia revela-se flexível, pois, diante de qualquer mudança no mercado de energia, bastaria adicionar novos atributos característicos para refletir tais alterações. Dessa forma, novas topologias poderiam ser geradas para representar o problema com maior precisão.

6 Considerações Finais

Conforme demonstrado na revisão da literatura, a comercialização de energia elétrica no Brasil enfrenta o desafio de prever valores no horizonte de longo prazo. Um exemplo disso é a assinatura de um contrato hoje, para entrega de energia, daqui a seis anos. Nesse período, diversos fatores podem impactar os preços, tornando essencial a utilização de ferramentas que auxiliem os agentes comercializadores a realizar previsões mais assertivas.

No Brasil, o preço de referência para contratos de energia é o PLD, fornecido pela Câmara de Comercialização de Energia Elétrica. No entanto, como demonstrado, sua aplicação em contratos de longo prazo é impraticável devido à volatilidade e a imprevisibilidade do mercado.

Neste contexto, foi proposta uma metodologia para a previsão de cenários de preço da curva *forward* de energia elétrica no horizonte de médio e longo prazo. As etapas dessa metodologia começam com o estudo das variáveis, a fim de identificar quais influenciam a formação do preço. A seleção dos atributos de maior impacto é realizada por meio de análise de correlação entre essas variáveis e a série histórica de preços de longo prazo fornecida pela Dcide, utilizando o coeficiente de correlação de Pearson.

Inicialmente, os dados analisados foram focados no submercado Sudeste/Centro-Oeste. Após a definição das variáveis é utilizado Algoritmo Genético para encontrar as melhores topologias que representam o problema em questão, onde essas topologias são entradas para o método de projeção. Por fim, emprega-se a técnica de Rede Bayesiana Dinâmica que leva em consideração o fator temporal, para as projeções de diversos cenários da curva *forward* de preços de energia elétrica, proporcionando uma modelagem mais robusta e adaptável às variações do mercado.

As técnicas de análises utilizadas permite ao analista inferir *insights* nas diferentes variáveis utilizadas, em qualquer período do horizonte de projeção no momento da geração dos cenários, proporcionando maior autonomia no uso da ferramenta.

Assim, foram realizados experimentos para os horizontes de 3, 4 e 6 anos em cenários de preço de energia elétrica, utilizando Redes Bayesianas Dinâmicas e os métodos de regressão XGBoost, *Support Vector Regression* e Regressão Linear. A partir dos resultados obtidos, é possível observar que: 1), com a RBD, para o mesmo horizonte de tempo, é possível gerar múltiplos cenários de preços, diferentemente dos demais métodos. 2), esses cenários se modificam à medida que o usuário fornece novos *insights* ao modelo, algo que os demais métodos não permitem. Dessa forma, os preços gerados nos cenários não são exclusivamente responsabilidade do método, mas também da interação do usuário. 3),

os resultados preliminares obtidos foram superiores aos dos demais métodos. Além disso, segundo analistas do mercado de projeções de energia elétrica no Brasil, as taxas de erro obtidas pelo modelo são promissoras.

É possível observar que o método proposto em comparação com os modelos computacional que são utilizados pela CCEE para suas projeções, flexibiliza a inferência de informações nas variáveis em qualquer tempo do horizonte de projeção, bem como o número de variáveis utilizadas são bem menores, com isso torna-se o modelo menos complexo e com bons resultados conforme já demonstrados.

6.1 Publicações

Nesta seção, apresenta-se as publicações desenvolvidas ao longo dos quatro anos de doutorado, divididas em duas categorias: i) publicações diretamente relacionadas aos resultados da pesquisa do doutorado, e, ii), publicações que contribuíram para outras linhas de pesquisa.

As publicações associadas à pesquisa do doutorado são:

- **Medium and Long Term Energy Forecasting Methods: A Literature Review. IEEE Access, 2025.** Josivan R. dos Reis, Jonathan M. Tabora, Matheus C. de Lima, Flávia P. Monteiro, Suzane C. A. Monteiro, Ubirtan H. Bezerra and Maria E. L. Tostes.
- **A Hybrid Methodology Using Machine Learning Techniques and Feature Engineering Applied to Time Series for Medium- and Long-Term Energy Market Price Forecasting. Energies, 2025.** Flávia P. Monteiro, Suzane Monteiro, Carlos Rodrigues, Josivan Reis, Ubiratan Bezerra, Maria E. Tostes and Frederico A. F. Almeida.
- **Demand-Side Management Optimization Using Genetic Algorithms: A Case Study. Energies, 2024.** <https://doi.org/10.3390/en17061463>. Lauro Correa dos Santos Junior, Jonathan Muñoz Tabora, Josivan Reis, Vinicius Andrade, Carminda Carvalho, Allan Manito, Maria Tostes, Edson Matos and Ubiratan Bezerra.

A seguir, são apresentadas as publicações realizadas em colaboração com outros pesquisadores, cujos trabalhos não estão diretamente alinhados à linha de pesquisa do doutorado. Essas contribuições refletem o envolvimento em diferentes áreas do conhecimento.

- **Prêmio pelo Artigo intitulado “Aplicativo Braille: uma ferramenta de acessibilidade pedagógica”, da categoria Tecnologias Educacionais Inovadores, com a 1ª colocação no InovaEduBr Summit, 2024.**

- **Assessing Accuracy in Capacitors Bank Estimations: A Case Study. IEEE Colombian Caribbean Conference, 2023.** Ayton do Nascimento, D. P. R. Carrera, J. Tabora, Bruno Santana De Albuquerque, Marcos Vinicius Pereira Braga, J. R. Dos Reis, Edson Ortiz de Matos, A. Manito, C. de Moura Carvalho and M. E. de Lima Tostes.
- **Augmented Reality Applications in Mathematics and Science: Exploring Pedagogical Viability in the Amazon Region. In: 2023 32nd Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEIE).** NASCIMENTO, ROBERTO ; REIS, JOSIVAN ; PRINTES, MARCOS ; BARONE, DANTE ; TABORA, JONATHAN ; TOSTES, MARIA.
- **Reconhecimento Automático de Caracteres Braille Utilizando Rede Neural Convolucional YOLO. RENOTE. REVISTA NOVAS TECNOLOGIAS NA EDUCAÇÃO, 2023.** Nascimento, R. P. ; REIS, J. R. ; ARAGAO, L. A. F. ; HEIDRICH, R. O. ; BARONE, D. A. C.
- **Avaliação de uma proposta de aprendizagem do sistema Braille para videntes”. In: Congresso sobre Tecnologias na Educação, 2023, Brasil. Anais do VIII Congresso sobre Tecnologias na Educação (Ctrl+e 2023).** NASCIMENTO, ROBERTO P. ; BARONE, DANTE A. C. ; REIS, JOSIVAN R. ; ARAÚJO, ARTHUR S. ; HEIDRICH, REGINA.
- **Tecnologia Assistiva: Um Relato de Experiência com Ênfase na Utilização de Ferramentas Colaborativas. XXVI Conferência Internacional sobre Informática na Educação. 1ed.Santiago de Chile: Nuevas Ideas en Informática Educativa, 2023, v. , p. 535-540.** NASCIMENTO, R. P. ; REIS, JOSIVAN R. ; SILVA, A. A. ; HENRIQUES, R. V. B. ; BARONE, D. A. C.
- **Aplicação do método YOLO para o reconhecimento de caracteres Braille.” XXVI Conferência Internacional sobre Informática na Educação. 1 ed.Santiago de Chile: Nuevas Ideas en Informática Educativa, 2023, v. 17, p. 564-569.** NASCIMENTO, R. P. ; ARAGAO, L. A. F. ; REIS, JOSIVAN R. ; HEIDRICH, R. O. ; BARONE, D. A. C.
- **Analisando os Dados do Percorso Acadêmico do Curso de Sistemas de Informação para Auxiliar na Revisão do PPC, no Simpósio Brasileiro de Educação em Computação, EDUCOMP, 2022.** Roberto P. Nascimento, Gabriela D. S. de Sousa, Mariana W. B. Nascimento, Josivan R. Reis Roberto P. Do Nascimento, Luiz H. C. Nunes, Mariana W. Buarque Do Nascimento, Carla M. C. Paxiúba, Josivan R. Reis.

- **Muiraquitã Braille: Uma Proposta de Aplicativo para Aprendizagem e Tradução do Sistema Braille no Congresso Brasileiro de Informática na Educação. Apps.Edu, 2022.** Roberto P. Nascimento, Josivan R. Reis, Samuel O. Amorim, Marcos Printes, Regina Heidrich, Dante Augusto Couto Barone.
- **Analyzing Data from Computer Science Graduates to Assist in the Revision of the Course Pedagogical Project no 2021 XVI Latin American Conference on Learning Technologies (LACLO).** DO NASCIMENTO, ROBERTO P. ; NUNES, LUIZ H. C. ; DO NASCIMENTO, MARIANA W. BUARQUE ; PAXIUBA, CARLA M. C. ; REIS, JOSIVAN R.
- **Um método automatizado para realizar detecção e contagem do Pirarucu durante o manejo/An automated method to perform detection and counting of Pirarucu during management” no Brazilian Journal of Development. Brazilian Journal of Development, 2021.** DOI: <https://doi.org/10.34117/bjdv7n10-051>. Wanderlany F. de Abreu, Deam J. A. da Silva, Roberto P. do Nascimento, Rafael de L. Rocha, Carlos G. R. dos Santos, Márcio J. M. da Ponte, Josivan R. dos Reis.
- **Proposta de Framework para a Concepção de Ferramentas Voltadas a Domótica Assistiva de Baixo Custo no Anais Estendidos do XXI Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS Estendido 2021).** Gabriela D. Sousa, Danilo Ferreira, Cael Santos, Josivan R. Reis, Roberto P. Nascimento.
- **Desenvolvimento de Artefatos de Baixo Custo para Promover um Ambiente de Domótica Assistiva, no Anais Estendidos do XXI Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS Estendido 2021.** Danilo da S. Ferreira, Gabriela D. S. de Sousa, Josivan R. Reis, Roberto P. Nascimento.
- **Uma proposta de framework para simulação de casas inteligentes usando Unity no nais Estendidos do XVII Simpósio Brasileiro de Sistemas de Informação (Anais Estendidos do SBSI 2021).** Cael G. B. Santos, Gabriela D. S. Sousa, Josivan R. Reis, Roberto P. Nascimento.

Referências

- AMADEU, J. R. *Simulação estocástica de preços no mercado brasileiro de energia elétrica*. Dissertação (Dissertação de Mestrado) — Faculdade de Engenharia Civil, Arquitetura e Urbanismo, Universidade Estadual de Campinas, 2020. Citado na página 28.
- ANEEL, A. N. de E. E. *Atribuições da Agência Nacional de Energia Elétrica*. 2022. <<https://www.gov.br/aneel/pt-br/aceso-a-informacao/institucional/a-aneel>>. Acessado em: 15 de Fevereiro de 2024. Citado na página 40.
- ARA, A.; MAIA, R. O. e M. *Modelos de Vetores de Suporte*. 2023. <http://leg.ufpr.br/~ara/teach/svm/05_outrostipos.html>. Acessado em: 13 de abril de 2024. Citado na página 105.
- ARAUJO, F. V. de S. *Estimando a Curva Forward de Energia Elétrica no Brasil com um Modelo de Dois Agentes Utilizando Contratos por Diferença e Função ECP_G*. Tese (Doutorado) — Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brasil, 2020. Citado na página 29.
- BALAN, M. H.; ROSA, P. S.; FENILI, M.; CAMARGO, L. A. S.; BIASE, R. G.; RAMOS, D. S.; DIAS, M. M. Sistema integrado de apoio à decisão para definição da estratégia ótima de comercialização de energia elétrica de um agente gerador. *XXVI Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2022. Citado na página 29.
- BASTOS, J. P.; CUNHA, G. R. A.; BARROSO, L. A. Uma metodologia para a separação da comercialização de energia e lastro no brasil através da captura do valor econômico da escassez no mercado de eletricidade. *XXIV Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2017. Citado 2 vezes nas páginas 30 e 48.
- BESSEMBINDER, H.; BESSEMBINDER, H.; LEMMON, M. L.; LEMMON, M. L. Equilibrium pricing and optimal hedging in electricity forward markets. *Journal of Finance*, 2002. Citado na página 27.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDEMIR, T. B. *Redes Neurais Artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000. Citado 3 vezes nas páginas 59, 62 e 63.
- BRAGA, L. de A. *Um estudo sobre o mercado de energia elétrica no Brasil*. 34 f. Monografia (Especialização) — Escola de Minas, Universidade Federal de Ouro Preto, Minas Gerais, 2018. Citado na página 38.
- BRANDÃO, R. *Aperfeiçoamentos do Mercado Atacadista de Energia*. [S.l.: s.n.], 2019. Grupo de Estudos dos Setor Elétrico - GESEL p. Citado na página 42.
- BRUCE, A.; BRUCE, P. *Estatística prática para cientistas de dados*. [S.l.]: Alta Books, 2019. ISBN 978-8550806037. Citado na página 74.
- CABRAL, R. S.; SACCHI, R.; HANSEN, P.; LEME, H.; ROSA, L. F. S. de C.; MACIEL, D.; BARROSO, L.; LOBATO, M. V. de C. Gestão de risco na comercialização de energias: Situação atual e proposta de melhores práticas. *XVIII Seminário de Planejamento*

Econômico - Financeiro do Setor Elétrico (SEPEF), 2015. Citado 2 vezes nas páginas 16 e 28.

CAMARGO, L. A. S.; GUARNIER, E.; RAMOS, D. S. Estratégia ótima de contratação para consumidores livres, como “trade contratação imediata e postergação de decisão, ponderando incertezas nos preços de curto prazo e na precificação de contratos bilaterais”. *XXIII Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2015. Citado na página 30.

CARVALHO, G. P. de; VIANA, A. G.; MELLO, J. C. Os contratos derivativos e o setor elétrico – uma solução legítima para a segurança do mercado nacional. *XXVI Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2022. Citado na página 31.

CASTRO, N. de; BRANDÃO, R.; DANTAS, N. H. G. *Expansão do Mercado Livre e as Distribuidoras de Energia Elétrica*. [S.l.: s.n.], 2014. v. 62. Citado 2 vezes nas páginas 43 e 47.

CASTRO, N. de; BRANDÃO, R.; MACHADO, A. F.; GOMES, V. *Contribuições para o aperfeiçoamento do Mercado Atacadista de Energia Brasileiro*. [S.l.: s.n.], 2017. 1–80 p. ISBN 978-85-93305-31-3. Citado na página 42.

CASTRO, N. de; BRANDÃO, R.; MACHADO, A. F.; GOMES, V. *Reflexões sobre o mercado brasileiro de energia elétrica no atacado e a crise financeira recente*. [S.l.: s.n.], 2017. v. 74. 1–80 p. ISBN 978-85-93305-31-3. Citado 2 vezes nas páginas 42 e 47.

CASTRO, N. J. de; BRANDÃO, R. *Mercado Elétrico e Risco Financeiro*. [S.l.]: PUBL!T Soluções Editoriais, 2021. v. 1. 1–229 p. ISBN 9786586614428. Citado 2 vezes nas páginas 42 e 45.

CCEE, C. de Comercialização de E. E. 2023. <<https://www.ccee.org.br/>>. Acessado em: 11 de Agosto de 2023. Citado 3 vezes nas páginas 41, 44 e 45.

CHARNIAK, E. *Bayesians Netwoeks without Tears*. [S.l.]: IA Magazine, 1991. Citado na página 53.

CHEN, T.; GUESTIN, C. Xgboost: A scalable tree boosting system. p. 785–794, 2016. Citado na página 105.

CIVIL, P. da R. C. *Dos Princípios e Objetivos da Política Energética Nacional*. 1997. <https://www.planalto.gov.br/ccivil_03/leis/19478.htm>. Acessado em: 15 de Fevereiro de 2024. Citado na página 40.

CMSE, C. de Monitoramento do S. E. *Funções do Comitê de Monitoramento do Setor Elétrico*. 2024. <<https://www.gov.br/mme/pt-br/assuntos/conselhos-e-comites/cmse>>. Acessado em: 15 de Fevereiro de 2024. Citado na página 40.

COOPER, G. F.; HERSKOVITS, E. A bayesian method for constructing bayesian belief networks from databases. *Proceedings of the Seventh Annual Conference on Uncertainty in Artificial Intelligence*, 1991. Citado na página 56.

DCIDE. *Previsão de preço de energia*. 2024. <<https://www.dcide.com.br/>>. Acessado em: 19 de março de 2024. Citado 2 vezes nas páginas 68 e 74.

DRUCKER, H.; BURGESS, C. J. C.; KAUFMAN, L.; SMOLA, A.; VAPNIK, V. Support vector regression machines. MIT Press, v. 9, 1996. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/1996/file/d38901788c533e8286cb6400b40b386d-Paper.pdf>. Citado na página 105.

DU, X.; YU, C. L.; HAYES, D. J. Speculation and volatility spillover in the crude oil and agricultural commodity markets: A bayesian analysis. *Energy Economics*, v. 33, n. 3, p. 497–503, 2011. ISSN 0140-9883. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S014098831100017X>>. Citado na página 28.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Patter Classification*. [S.l.]: Wiley-Interscience, 2000. v. 2nd ed. edição. Citado 2 vezes nas páginas 53 e 79.

ELSHATER, A.; ABUSAADA, H. Notes on developing research review in urban planning and urban design based on prisma statement. v. 11, p. 1–8, 08 2022. Citado 2 vezes nas páginas 18 e 19.

EPE, E. de P. E. *Anuário Estatístico de Energia Elétrica 2023*. 2023. <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-160/topico-168/Fact>Acessado em: 16 de agosto de 2023. Citado na página 42.

EPE, E. de P. E. *Balanço Energético Nacional*. 2024. <<https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-ben/>>. Acessado em: 27 de fevereiro de 2024. Citado na página 86.

EPE, E. de P. E. *O papel da Empresa de Pesquisa Energética*. 2024. <<https://www.epe.gov.br/pt/a-epe/quem-somos>>. Acessado em: 15 de Fevereiro de 2024. Citado na página 41.

FELIZATTI, H. L.; HANSEN, P. M. *Diretrizes para gestão de risco de mercado no setor de energia elétrica*. [S.l.: s.n.], 2019. Citado na página 49.

FELIZATTI, H. L.; HANSEN, P. M.; HOTTA, L. K.; HERENCIA, M. E. Z. Curva forward no mercado de energia elétrica brasileiro: Construção, modelagem, previsão e simulação. 2019. Citado 2 vezes nas páginas 43 e 75.

FLETEN, S.-E.; LEMMING, J. Constructing forward price curves in electricity markets. *Energy Economics*, v. 25, n. 5, p. 409–424, 2003. ISSN 0140-9883. Citado na página 27.

FURNAS, E. *chamada pública para Projetos PD*. 2020. <<https://www.furnas.com.br/noticia/103/noticias/1553>>. Acessado em: 2 de abril de 2024. Citado na página 17.

GOMES, L. C. S.; PEREIRA, R. B.; EKEL, P.; CONCEIÇÃO, G. L.; SANTOS, A. A. M. dos; ANDRADE, A. S.; FILHO, L. B. G.; LIMA, M. O.; FONSECA, R. M.; MENDONÇA, M. D. O.; SILVA, G. R. L.; VIEIRA, D. A. G. Avaliação de risco de exposição ao mercado de curto prazo na aquisição de novos ativos de energia. *XXVI Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2022. Citado na página 30.

GOMES, P. C. T. *Modelos de Vetores de Suporte*. 2019. <<https://www.datageeks.com.br/xgboost/>>. Acessado em: 13 de abril de 2024. Citado na página 105.

HE, J.; LI, S.; QI, Z. *2024 3rd International Conference on Energy, Power and Electrical Technology (ICEPET)*, p. 1174–1177, 2024. Disponível em: <<https://api.semanticscholar.org/CorpusID:271867873>>. Citado na página 31.

IBGE, I. B. de Geografia e E. *Produto Interno Bruto - PIB*. 2024. <<https://www.ibge.gov.br/>>. Acessado em: 29 de fevereiro de 2024. Citado na página 87.

IBRAHIM, I. A. In: *Improved Grid Search Algorithm for Optimal LSTM Performance: A Case Study on Australian Electricity Price Forecasting*. [S.l.: s.n.], 2024. p. 1–5. Citado na página 31.

IEA, I. E. A. *Re-powering markets: Market design and regulation during the transition to low-carbon power systems*. [S.l.], 2016. Disponível em: <<https://ndcpartnership.org/toolbox/re-powering-markets-market-design-and-regulation-during-transition-low-carbon-power-systems>>. Citado na página 49.

JENSEN, F. *An Introduction to Bayesian Networks*. [S.l.]: UCL Press, 1996. Citado na página 56.

LAMPS, L. of A. M. P. S. *Ferramentas analíticas para previsão de preços de contratos futuros de energia (curva forward) em mercados hidrotérmicos com preços spot horários*. 2021. <<http://www.lamps.ind.puc-rio.br/projeto/pd-aneel-eneva-curvas-forward-para-mercados-de-eletricidade/>>. Acessado em: 18 de Outubro de 2023. Citado na página 49.

LEME, H.; HANSEN, P.; HOTTA, L. K.; ZEVALLOS, M. A forward price curve in the brazilian electricity market: Capture and analyses. 07 2015. Citado 2 vezes nas páginas 28 e 50.

MARQUES, N. L.; BRANDÃO, L. E. T.; GOMES, L. L.; ROSA, V. C. V. da. Modelo de apoio à tomada de decisão sobre operações de hedge na comercialização de energia elétrica. *XXVI Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2022. Citado na página 31.

MENDONÇA, M. D. O.; PINTO, P. H. A.; SANTOS, F. F. G.; PIRES, D. S. D. C.; VIEIRA, D. A. G.; LOBATO, M. V. D. C.; SILVA, G. R. L.; SALDANHA, R. R.; RESENDE, G. D.; SANTIAGO, F. P.; HORTA, G. D. L. Análise comparativa entre modelos de inteligência computacional para previsão do preço futuro no mercado de energia brasileiro. *XXV Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2019. Citado na página 29.

MIHAJLOVIĆ, V.; PETKOVIC, M. Dynamic bayesian networks: A state of the art. *Europhysics Letters (epl)*, 01 2001. Citado na página 58.

MME, M. de Minas e E. *instituições do setor elétrico brasileiro e as competências*. 2021. <<https://www.gov.br/mme/pt-br/assuntos/noticias/conheca-as-instituicoes-do-setor-eletrico-brasileiro-e-as-competencias-de-cada-uma>>. Acessado em: 10 de Fevereiro de 2024. Citado na página 39.

MOHER, D.; LIBERATI, A.; TETZLAFF, J.; ALTMAN, D. G. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *BMJ*, BMJ Publishing

Group Ltd, v. 339, 2009. Disponível em: <<https://www.bmj.com/content/339/bmj.b2535>>. Citado 2 vezes nas páginas 18 e 19.

MONTEIRO, M. D.; AGUIAR, A. S. de; CHÁVARRY, I. S. S. M.; FREIRE, A. C. B.; FERNANDES, C. A. C.; VALLADÃO, D. M.; MORAES, G. M. B. de. Ferramenta computacional para modelagem e previsão probabilística de curvas forward de eletricidade. *XXVI Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2022. Citado na página 29.

MORETTIN, P. A.; SINGER, J. da M. *Estatística e Ciência de Dados*. [S.l.]: LTC, 2022. v. 1ª edição. Citado na página 66.

NAMETALA, C. A. L. *Redes neurais atencionais aplicadas a modelagem e previsão de preços no mercado de eletricidade brasileiro*. Dissertação (Tese de Doutorado) — Escola de Engenharia de São Carlos, Universidade de São Paulo, 2023. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/18/18154/tde-16032023-161345/>>. Citado 2 vezes nas páginas 50 e 51.

ONS, O. N. do S. E. *Atuação do Operador Nacional do Sistema Elétrico*. 2024. <<https://www.ons.org.br/paginas/sobre-o-ons/atuacao>>. Acessado em: 02 de abril de 2024. Citado na página 41.

ONS, O. N. do S. E. *EVOLUÇÃO DA CAPACIDADE INSTALADA NO SIN*. 2024. <[ons.org.br](https://www.ons.org.br)>. Acessado em: 02 de abril de 2024. Citado na página 99.

PAPPAS, F.; SANTOS, M. L. L. Poder de mercado na formação de preços via oferta: análise de fatores de influência e métricas. *XXV Seminário Nacional de Produção e Transmissão de Energia Elétrica (SNPTEE)*, 2019. Citado na página 30.

PILIPOVIC, D. *Energy Risk: Valuing and Managing Energy Derivatives*. [S.l.]: McGraw Hill, 2007. v. 2nd ed. edição. ISBN 0071594477. Citado na página 48.

PORTO, T.; SACCHI, R. *Relatório do Grupo Temático – Mecanismos de Formação de Preço. Grupo de Trabalho Modernização do Setor Elétrico. Ministério de Minas e Energia*. [S.l.: s.n.], 2019. <<https://www.gov.br/mme/pt-br/assuntos/secretarias/secretaria-executiva/modernizacao-do-setor-eletrico/arquivos/pasta-geral-publicada/formacao-de-precos.pdf>>. Citado na página 48.

PREX, P. S. L. *Mercado Livre de Energia já está aberto para todo o Grupo*. 2024. <<https://www.mercadolivredeenergia.com.br/>>. Acessado em: 11 de março de 2024. Citado na página 46.

REIS, J. R. D.; TABORA, J. M.; LIMA, M. Carvalho de; MONTEIRO, F. P.; MONTEIRO, S. Cruz de A.; BEZERRA, U. H.; TOSTES, M. Emília de L. Medium and long term energy forecasting methods: A literature review. *IEEE Access*, v. 13, p. 29305–29326, 2025. Citado na página 32.

REPÚBLICA, P. da. *Medida Provisória nº 579, de 2012*. 2012. Disponível em: <<https://www.congressonacional.leg.br/materias/medidas-provisorias/-/mpv/107292>>. Acesso em: 27 nov 2013. Citado na página 39.

REPÚBLICA, P. da. *LEI Nº 12.783, DE 11 DE JANEIRO DE 2013*. 2013. Disponível em: <https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2013/lei/l12783.htm>. Acesso em: 27 nov 2013. Citado na página 39.

RUSSELL, S.; NORVIG, P. *Inteligência Artificial*. Rio de Janeiro: 3a edição, Elsevier, 2013. Citado na página 61.

SANTOS, R. R. dos; FIGER, V. *Transição energética: Brasil dribla a crise hídrica, mas permanece em alerta em relação a conta de luz*. 2022. [://portal.fgv.br/artigos/transicao-energica-brasil-dribla-crise-hidrica-mas-permanece-alerta-relacao-conta-luz](http://portal.fgv.br/artigos/transicao-energica-brasil-dribla-crise-hidrica-mas-permanece-alerta-relacao-conta-luz). Acessado em: 16 de fevereiro de 2024. Citado na página 47.

SCHMULLER, J.; BATISTA, S. *Análise estatística com R*. [S.l.]: Alta Books, 2019. Tradução da 2ª edição. Citado na página 66.

VIANA, A. G. *Leilões como mecanismo alocativo para um novo desenho de mercado no Brasil*. Dissertação (Tese de Doutorado) — Escola Politécnica, Universidade de São Paulo, 2018. Citado 6 vezes nas páginas 34, 35, 36, 37, 48 e 50.

WILEY, J.; SONS. *Applied Linear Regression*. [S.l.: s.n.], 2013. v. 62. ISBN 9781118386088. Citado na página 105.

WOOLDRIDGE, J. M. *Introdução à Econometria: uma Abordagem Moderna*. [S.l.]: 3ª edição, Cengage Learning, 2017. Citado na página 72.

YOUSEFI, A.; SIANAKI, O. A.; SHARAFI, D. Long-term electricity price forecast using machine learning techniques. In: *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*. [S.l.: s.n.], 2019. p. 2909–2913. Citado na página 29.