



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
MESTRADO ACADÊMICO EM ENGENHARIA ELÉTRICA

KAROL DAMASCENO CYSNE

**APLICAÇÃO DE ALGORITMOS INTELIGENTES NA PREVISÃO DO MONTANTE
DE USO DO SISTEMA DE TRANSMISSÃO PARA OTIMIZAÇÃO DO CONTRATO DE
USO DO SISTEMA DE TRANSMISSÃO EM DISTRIBUIDORAS DE ENERGIA
ELÉTRICA**

FORTALEZA

2023

KAROL DAMASCENO CYSNE

APLICAÇÃO DE ALGORITMOS INTELIGENTES NA PREVISÃO DO MONTANTE DE
USO DO SISTEMA DE TRANSMISSÃO PARA OTIMIZAÇÃO DO CONTRATO DE USO
DO SISTEMA DE TRANSMISSÃO EM DISTRIBUIDORAS DE ENERGIA ELÉTRICA

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia Elétrica do Programa de Pós-Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de Concentração: Sistemas de Energia Elétrica.

Orientador: Prof. Dr. Giovanni Cordeiro Barroso.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- C998a Cysne, Karol Damasceno.
Aplicação de Algoritmos Inteligentes na Previsão do Montante de Uso do Sistema de Transmissão para Otimização do Contrato de Uso do Sistema de Transmissão em Distribuidoras de Energia Elétrica / Karol Damasceno Cysne. – 2024.
85 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Fortaleza, 2024.
Orientação: Prof. Dr. Giovanni Cordeiro Barroso.
1. MUST. 2. CUST. 3. Algoritmos Inteligentes. 4. Modelagem de Equações Estruturais. 5. Previsão de Energia Elétrica. I. Título.

CDD 621.3

KAROL DAMASCENO CYSNE

APLICAÇÃO DE ALGORITMOS INTELIGENTES NA PREVISÃO DO MONTANTE DE
USO DO SISTEMA DE TRANSMISSÃO PARA OTIMIZAÇÃO DO CONTRATO DE USO
DO SISTEMA DE TRANSMISSÃO EM DISTRIBUIDORAS DE ENERGIA ELÉTRICA

Dissertação apresentada ao Curso de Mestrado Acadêmico em Engenharia Elétrica do Programa de Pós-Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de Concentração: Sistemas de Energia Elétrica.

Aprovada em: 08/02/2024.

BANCA EXAMINADORA

Prof. Dr. Giovanni Cordeiro Barroso (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dra. Ruth Pastôra Saraiva Leão
Universidade Federal Ceará (UFC)

Prof. Dra. Raquel Cristina Filiagi Gregory
Universidade Federal Ceará (UFC)

Prof. Dr. Nivalde José de Castro
Universidade Federal do Rio de Janeiro (UFRJ)

À minha família. Minha mãe, por ser meu exemplo diário de força e fé. Meu pai, por ser minha segurança nos momentos de dificuldade. Meu marido, por despertar a melhor parte de mim. Meu filho, para que um dia se orgulhe de mim.

AGRADECIMENTOS

À Universidade Federal do Ceará, pela oportunidade dada para cursar o Curso de Mestrado Acadêmico em Engenharia Elétrica.

Ao Prof. Dr. Giovanni Cordeiro Barroso, pela excelente orientação, sugestões e críticas ao longo de todo processo de produção do trabalho.

Aos professores participantes da banca examinadora Prof. Dra. Ruth Pastôra Saraiva Leão, Prof. Dra. Raquel Cristina Filiagi Gregory e ao Prof. Dr. Nivalde José de Castro pelo tempo, pelas valiosas colaborações e sugestões.

Aos colegas, Letícia de Oliveira Santos e Francisco Alexandre Andrade de Souza pelo tempo concedido nas reuniões, pelas sugestões, críticas e por todo o apoio no desenvolvimento do trabalho.

À minha família, por sempre ter acreditado em mim, me dando todo apoio psicológico com muita paciência e amor.

Aos colegas do trabalho, pelas sugestões recebidas e compreensão pelas inúmeras horas dedicados ao curso de mestrado.

"Se, a princípio, a ideia não é absurda, então não há esperança para ela." (Albert Einstein)

RESUMO

Os agentes geradores, as distribuidoras e os consumidores de energia elétrica conectados ao Sistema Interligado Nacional (SIN) possuem um vínculo contratual, o Contrato de Uso do Sistema de Transmissão (CUST), com o Operador Nacional do Sistema (ONS). O CUST permite que o ONS gerencie e promova a otimização da operação do sistema eletroenergético e quantifique os custos incorridos pelo uso da rede básica para que sejam rateados entre os usuários do sistema. No CUST, os agentes geradores informam a sua máxima potência elétrica injetável no sistema enquanto as distribuidoras e os consumidores informam anualmente seus Montantes do Uso do Sistema de Transmissão (MUST), definidos de acordo com os valores de máxima demanda em cada ponto de conexão com a rede básica e horário de contratação para os quatro anos civis subsequentes. De forma a estabelecer o valor do MUST a ser contratado, a maioria das concessionárias de energia elétrica se baseia em estimativas derivadas de métodos simplificados e/ou estatísticos a partir da base histórica de dados de demanda de energia elétrica. Neste trabalho é proposta a aplicação de algoritmos inteligentes para previsão dos MUST, a fim de se obter valores mais próximos da realidade e que fundamentem uma otimização no CUST de forma a se obter o menor valor financeiro associado. Para isto, são levadas em consideração variáveis climáticas, econômicas, temporais e operacionais destacadas como fatores que influenciam a demanda de energia elétrica, através do uso da Modelagem de Equações Estruturais (MEE). Após ser feito um estudo de correlação entre variáveis, foi proposta a aplicação de algoritmos inteligentes para prever os dados de demanda de energia elétrica e, a partir disso, otimizar os valores de contratação do CUST. Os resultados indicaram que o algoritmo de melhor desempenho foi o *Random Forest Regressor* com 71,1% de precisão, resultando em uma ganho financeiro de 12% com a otimização em relação à estratégia de contratação tradicional.

Palavras-chave: MUST; CUST; algoritmos Inteligentes; modelagem de equações estruturais; previsão de energia elétrica

ABSTRACT

Generating agents, distributors and consumers of electricity connected to the National Interconnected System (SIN) have a contractual link, the CUST, with the National System Operator (ONS). CUST allows the ONS to manage and promote the optimization of the operation of the electroenergetic system and quantify the costs incurred when using the basic network so that they can be apportioned among the system's users. In CUST, generating agents report their maximum electrical power injectable into the system while distributors and consumers report annually their MUST, defined according to the maximum demand values at each connection point with the core network and hiring schedule for the subsequent four calendar years. In order to establish the MUST value to be contracted, most electricity concessionaires rely on estimates derived from simplified and/or statistical methods based on historical electricity demand data. This work proposes the application of intelligent algorithms to predict MUST, in order to obtain values closer to reality and that support an optimization in CUST in order to obtain the lowest associated financial value. To this end, climatic, economic, temporal and operational variables are taken into account, highlighted as variables that influence the demand for electrical energy, through the use of MEE. After carrying out a correlation study between variables, the application of intelligent algorithms was proposed to predict electrical energy demand data and, based on this, optimize the CUST contracting values. The results indicated that the best performing algorithm was *Random Forest Regressor* with 71.1% accuracy, resulting in a financial gain of 12% with optimization in relation to the traditional hiring strategy.

Keywords: MUST, CUST, intelligent algorithms; structural equation modeling; electricity forecast

LISTA DE FIGURAS

Figura 1 – Fases e atividades do processo de revisão sistemática	21
Figura 2 – Quantitativo Anual de publicações relacionadas aos temas da pesquisa. . .	24
Figura 3 – Quantidade de estudos excluídos.	25
Figura 4 – Visão geral dos contratos firmados entre os agentes.	32
Figura 5 – Demonstrativo dos encargos e ultrapassagens.	37
Figura 6 – Exemplo do cálculo de sobrecontratação.	37
Figura 7 – Exemplo de relação causal interpretada pelo pacote <i>Structural Equation Models Optimization</i> (SEMOPY).	40
Figura 8 – Principais modelos de <i>Machine Learning</i>	43
Figura 9 – Exemplo de problema de classificação supervisionado	44
Figura 10 – Exemplo de problema de regressão supervisionado	44
Figura 11 – Diagrama da metodologia empregada.	51
Figura 12 – Transformadas de <i>Fourier</i> dos dados de tempo.	55
Figura 13 – Modelo de relação causal entre variáveis e demanda de energia elétrica. . . .	56
Figura 14 – Mapa de Correlação de <i>Spearman</i>	57
Figura 15 – <i>Pipeline</i> da aplicação dos algoritmos regressores	59
Figura 16 – Linha do tempo de contratação	59
Figura 17 – Curva de tendência dos dados de demanda medida no período escolhido. . .	61
Figura 18 – Dados de máxima demanda e geração medidos e mensais no período de 2018 a 2022.	62
Figura 19 – Dados de máxima demanda medidos e valores de contrato no período de 2018 a 2022 para horário de fora ponta.	63
Figura 20 – Dados de máxima demanda medidos e valores de contrato no período de 2018 a 2022 para horário de ponta.	63
Figura 21 – Modelo de equação estrutural utilizado.	64
Figura 22 – Curvas de máxima demanda prevista pelo algoritmo e máxima demanda medida real na Fora Ponta	67
Figura 23 – Curvas de máxima demanda prevista pelo algoritmo e máxima demanda medida real na Ponta	67
Figura 24 – Curvas de máxima demanda mensal medida e contratada com os respectivos valores de encargos calculados para fora ponta no cenário 1	68

Figura 25 – Curvas de máxima demanda mensal medida e contratada com os respectivos valores de encargos calculados para ponta no cenário 1	69
Figura 26 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 1	69
Figura 27 – Curvas de máxima demanda mensal prevista e contratada otimizada com os respectivos valores de encargos calculados para fora ponta no cenário 2	70
Figura 28 – Curvas de máxima demanda mensal prevista e contratada otimizada com os respectivos valores de encargos calculados para ponta no cenário 2	70
Figura 29 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 2	71
Figura 30 – Curvas de máxima demanda mensal realizada em 2022 e contratada otimizada com os respectivos valores de encargos calculados para fora ponta no cenário 3	72
Figura 31 – Curvas de máxima demanda mensal realizada em 2022 e contratada otimizada com os respectivos valores de encargos calculados para ponta no cenário 3	72
Figura 32 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 3	73
Figura 33 – Curvas de máxima demanda mensal prevista para 2022 e contratada reduzido com os respectivos valores de encargos calculados para fora ponta no cenário 4	75
Figura 34 – Curvas de máxima demanda mensal prevista para 2022 e contratada reduzido com os respectivos valores de encargos calculados para ponta no cenário 4	75
Figura 35 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 4 com os dados de previsão	76
Figura 36 – Curvas de máxima demanda mensal realizada para 2022 e contratada reduzido com os respectivos valores de encargos calculados para fora ponta no cenário 4	77
Figura 37 – Curvas de máxima demanda mensal realizada para 2022 e contratada reduzido com os respectivos valores de encargos calculados para ponta no cenário 4	77
Figura 38 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 4 com os dados reais	78

LISTA DE TABELAS

Tabela 1 – Parâmetros utilizados nos algoritmos.	58
Tabela 2 – Índices de avaliação do modelo de equação estrutural escolhido.	65
Tabela 3 – Métricas de avaliação dos algoritmos regressores sem a técnica de <i>grid search</i> . 65	
Tabela 4 – Métricas de avaliação dos algoritmos regressores com a técnica de <i>grid search</i> . 66	
Tabela 5 – Valores contratados no CUST por cenário analisado	74
Tabela 6 – Comparativo financeiro dos cenários analisados	80

LISTA DE QUADROS

Quadro 1 – Mapeamento de palavras-chave para <i>string</i> de busca.	23
Quadro 2 – Trabalhos selecionados.	25
Quadro 3 – Modelos de algoritmos inteligentes utilizados nos trabalhos selecionados. . .	29
Quadro 4 – Atributos utilizados nos trabalhos selecionados.	29
Quadro 5 – Principais índices de avaliação para o ajuste da MEE.	41
Quadro 6 – Resumo da extração dos dados por fonte.	53
Quadro 7 – Correlações entre variáveis que foram utilizadas no modelo testado.	58
Quadro 8 – Variações do modelo de equações estruturais simulados no SEMOPY	64

LISTA DE ABREVIATURAS E SIGLAS

R^2	<i>R-Squared</i>
ACL	Ambiente de Contratação Livre
ACR	Ambiente de Contratação Regulado
AdaBoost	<i>Adaptive Boosting</i>
ADEUST	Adicional dos Encargos de Uso do Sistema de Transmissão
AIC	<i>Akaike Information Criteria</i>
ANEEL	Agência Nacional de Energia Elétrica
CCEE	Câmara de Comercialização de Energia Elétrica
CCT	Contrato de Conexão às Instalações de Transmissão
CNN	<i>Convolutional Neural Network</i>
CPST	Contrato de Prestação do Serviço de Transmissão
CUST	Contrato de Uso do Sistema de Transmissão
DIT	Demais Instalações de Transmissão
EUST	Encargos de Uso do Sistema de Transmissão
GLS	<i>General Least Squares</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IBRE	Instituto Brasileiro de Economia
IGP-M	Índice Geral de Preços – Mercado
INMET	Instituto Nacional de Meteorologia
IPCA	Índice Nacional de Preços ao Consumidor Amplo
KNN	<i>K-Nearest Neighbors</i>
L-BFGS-B	<i>Broyden-Fletcher-Goldfarb-Shanno</i>
LR	<i>Linear Regression</i>
LSTM	<i>Long Short Term Memory</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MEE	Modelagem de Equações Estruturais
MLP	<i>Multi-layered Perceptrons</i>
MLR	<i>Multiple Linear Regression</i>
MLW	<i>Wishart Maximum Likelihood</i>

MSE	<i>Mean Squared Error</i>
MUST	Montantes do Uso do Sistema de Transmissão
ONS	Operador Nacional do Sistema Elétrico
PIM	Plataforma Integrada de Medição
PIS	Parcela de Ineficiência por Sobrecontratação
PIU	Parcela de Ineficiência por Ultrapassagem
PRORET	Procedimentos de Regulação Tarifária
RAP	Receita Anual Permitida
REN	Resolução Normativa
RF	<i>Random Florest</i>
RFC	<i>Random Forest Classifier</i>
RFR	<i>Random Forest Regressor</i>
RMSE	<i>Root Mean Square Error</i>
RMSEA	<i>Root-Mean-Square Error of Aproximation</i>
RNA	Redes Neurais Artificiais
SEMOPY	<i>Structural Equation Models Optimization</i>
SEP	Sistema Elétrico de Potência
SLSQP	<i>Sequential Least-Squares Quadratic Programming</i>
SMF	Sistema de Medição de Faturamento
SVM	<i>Support Vector Machines</i>
SVR	<i>Support Vector Regressor</i>
TUST	Tarifa do Uso do Sistema de Transmissão
TUST-FR	Tarifa do Uso do Sistema de Transmissão de Fronteira
TUST-RB	Tarifa do Uso do Sistema de Transmissão de Rede Básica
ULS	<i>Unweighted Least Squares</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Motivação	18
1.2	Objetivos	19
1.3	Organização do Trabalho	19
2	REVISÃO BIBLIOGRÁFICA	21
2.1	Planejamento	21
2.1.1	<i>Objetivo da revisão sistemática</i>	21
2.1.2	<i>Protocolo</i>	22
2.1.3	<i>Avaliação do protocolo</i>	24
2.2	Condução	24
2.2.1	<i>Síntese dos dados após a aplicação dos critérios de seleção e exclusão</i>	25
2.3	Resultados da Revisão Sistemática	26
3	FUNDAMENTAÇÃO TEÓRICA	30
3.1	Condições Gerais de Acesso ao Sistema de Transmissão	30
3.2	Regras Gerais de Contratação de Uso do Sistema de Transmissão	32
3.2.1	<i>Encargos e Eficiência da Contratação em Caráter Permanente</i>	34
3.3	Fundamentos Básicos da Modelagem de Equações Estruturais	37
3.3.1	<i>Pacote SEMOPY</i>	39
3.4	Modelos de Machine Learning	41
3.4.1	<i>Aprendizado supervisionado</i>	42
3.4.2	<i>Aprendizado não supervisionado</i>	42
3.4.3	<i>Aprendizado por reforço</i>	43
3.4.4	Algoritmos	45
3.4.4.1	<i>Regressão Linear</i>	45
3.4.4.2	<i>Máquina de Vetor de Suporte</i>	45
3.4.4.3	<i>K-ésimo Vizinho mais Próximo</i>	46
3.4.4.4	<i>Floresta aleatória</i>	46
3.4.4.5	<i>Perceptron Multicamadas</i>	46
3.4.4.6	<i>Memória longa de curto prazo</i>	47
3.4.5	Métricas de avaliação de desempenho	47

3.4.5.1	<i>R-Quadrado</i>	48
3.4.5.2	<i>Erro Quadrático Médio</i>	48
3.4.5.3	<i>Raiz do erro quadrático médio</i>	49
3.4.5.4	<i>Erro Absoluto Médio</i>	49
3.4.5.5	<i>Erro Percentual Absoluto Médio</i>	50
4	METODOLOGIA	51
4.1	Bases de Dados	51
4.2	Pré-Processamento dos Dados	54
4.2.1	<i>Transformação dos dados de tempo</i>	54
4.2.2	<i>Normalização da base de dados</i>	55
4.3	Modelagem de Equações Estruturais	55
4.4	Algoritmos de Regressão	58
4.5	Otimização do CUST	60
5	RESULTADOS	61
5.1	Análise Exploratória dos Dados	61
5.2	Modelo de Equação Estrutural Definido	63
5.3	Algoritmos de Regressão	65
5.4	Análise da Contratação do Uso do Sistema de Transmissão	67
5.4.1	<i>Cenário 1</i>	68
5.4.2	<i>Cenário 2</i>	69
5.4.3	<i>Cenário 3</i>	71
5.4.4	<i>Cenário 4</i>	73
5.4.5	<i>Cenário 4 - Dados Previstos</i>	74
5.4.6	<i>Cenário 4 - Dados Reais</i>	76
6	CONCLUSÕES E TRABALHOS FUTUROS	79
6.1	Trabalhos Futuros	81
	REFERÊNCIAS	82

1 INTRODUÇÃO

O Sistema Elétrico de Potência (SEP) brasileiro possui um grande número de geradores de energia elétrica, milhares de quilômetros de linhas de transmissão e um grande número de subestações e centros de controle. Na década de 1990, o governo implementou medidas de reformas estruturais e institucionais no setor elétrico com o objetivo de atrair investimentos privados (Mendonça *et al.*, 2021).

Com isso, houve um processo de desestatização das empresas que compunham o sistema elétrico de potência e surgiram algumas reformas econômicas dispostas a aumentar a eficiência e a competitividade no mercado. A ideia é que o Estado passasse a ter o papel de apenas regulador da atividade econômica (Rudnik *et al.*, 1996).

A partir dessa década passou-se a ser exigido que todas as entidades atuantes em geração, distribuição e comercialização fossem associadas à Câmara de Comercialização de Energia Elétrica (CCEE), existindo duas opções para comercialização de energia: o Ambiente de Contratação Livre (ACL) e o Ambiente de Contratação Regulado (ACR).

Na estrutura atual do setor elétrico, a maioria absoluta dos geradores são livres para negociar seus preços, seja diretamente com consumidores livres, seja por meio de leilões regulados. Já as transmissoras e distribuidoras têm seus preços regulados pela Agência Nacional de Energia Elétrica (ANEEL) portanto, não são livres para praticar os preços que desejam, inserindo-se no contexto dos contratos de concessão, que usualmente contam com mecanismos de revisões e reajustes tarifários periódicos (Lima, 2009).

Sendo assim, a transmissão de energia elétrica representa um monopólio natural, onde a tarifação do uso da rede afeta diretamente a remuneração das empresas concessionárias e os custos dos agentes do mercado (Faria, 2013).

Desta forma, é necessário que haja um arcabouço regulatório que apoie as transações de mercado entre os agentes e represente bem os procedimentos operacionais físicos existentes.

De acordo com o Operador Nacional do Sistema Elétrico (ONS), o negócio de transmissão de energia elétrica é bastante consolidado do ponto de vista legal e contratual, sendo completamente ancorado nos Procedimentos de Rede e em contratos específicos. Primeiramente, a ANEEL, formaliza através dos leilões e dos contratos de concessões a delegação do Poder Concedente para que as empresas de transmissão de energia elétrica possam explorar, como prestadoras de serviço público, as respectivas instalações de transmissão (ONS, 2023).

Uma vez definidas as empresas prestadoras do serviço público e as instalações de

transmissão, o ONS firma o Contrato de Prestação do Serviço de Transmissão (CPST). Esse documento estabelece os termos, condições técnicas e comerciais para a prestação dos serviços de transmissão, em que concessionárias de transmissão assumem a responsabilidade pela operação e manutenção e recebem a Receita Anual Permitida (RAP) pela disponibilização das instalações (ONS, 2023).

Além disso, nesse instrumento os agentes de transmissão autorizam o ONS a praticar todos os atos necessários para representá-los perante os usuários da rede de transmissão no Contrato de Uso do Sistema de Transmissão (CUST) (ONS, 2023).

O CUST estabelece os termos e condições que irão regular o uso da Rede Básica, permitindo que o ONS gerencie e promova a otimização da operação do sistema eletroenergético e quantifique os custos incorridos pelo uso da Rede Básica para que sejam rateados entre os usuários do sistema. Esse contrato é celebrado por centrais geradoras, unidades consumidoras, importadores, exportadores de energia e distribuidoras que se conectam à Rede Básica ou pelas Demais Instalações de Transmissão (DIT) compartilhadas por mais de uma distribuidora (ONS, 2023).

Tendo em vista que a energia ainda não pode ser armazenada em quantidade considerável, faz-se necessário que haja um equilíbrio justo entre a geração e a demanda de energia elétrica. Sendo assim, uma previsão precisa da demanda de energia elétrica pode impactar significativamente a operação, manutenção e gerenciamento do sistema de energia (Alhussein *et al.*, 2020).

Sendo assim, por meio do CUST os agentes geradores devem informar a sua máxima potência elétrica injetável no sistema e as distribuidoras e os consumidores devem informar anualmente seus valores de máxima demanda em cada ponto de conexão com a rede básica e horário de contratação para o ano vigente e os quatro anos civis subsequentes (ANEEL, 2020).

O custo associado ao CUST, num determinado regime, pode ser dividido em duas partes: uma fixa e outra variável. O custo fixo é determinístico e por esta razão, existem incentivos econômicos para que os agentes determinem os valores dos contratados no CUST o mais próximo possível das reais necessidades futuras. Caso os valores de MUST contratados não sejam respeitados devido uma avaliação inadequada, a ANEEL aplica multas e penalidades para os agentes pela ocorrência da ineficiência operacional (Milhorange; Silva, 2023).

As penalidades são definidas segundo a Resolução Normativa (REN) nº 905/2020, que estabelece as condições gerais para contratação do acesso aos sistemas de transmissão de

energia elétrica, e ocorrem sempre que há uma subcontratação ou sobrecontratação em relação aos valores de demanda medidos em cada ponto de conexão com a rede básica e horário de contratação em cada ano vigente.

Sob a ótica das distribuidoras de energia, a possibilidade de ter que arcar com custos de penalidades pode fazer com que a empresa haja de uma maneira mais conservadora no momento de estabelecimento dos valores contratuais. Desta forma, a empresa pode, levando em consideração as regras de contratação explicadas no Capítulo 3 e com o objetivo de eliminar possíveis multas por ultrapassagem, decidir em contratar mais MUST do que o realmente necessário (Lima *et al.*, 2021).

Entretanto, os erros de previsão têm uma influência considerável tanto na operacionalidade quanto no planejamento do sistema visto que é crucial projetar e desenvolver planos de programação e operação para uso eficiente da energia disponível (Alhussein *et al.*, 2020).

Portanto, o valor financeiro associado ao CUST desempenha um papel importante no planejamento da expansão do sistema de transmissão. A análise de investimentos, para atender a demanda total e garantir a confiabilidade do sistema, utiliza o CUST como dado de entrada (Milhorance; Silva, 2023).

Neste sentido, pode haver um aumento no custo da expansão da rede básica com o objetivo de atender os contratos que muitas vezes estão bem acima da previsão do mercado. Além disso, contratos conservadores geram encargos mensais com valores mais altos e consequentemente tarifas onerosas aos consumidores finais, visto que esses encargos são repassados para a tarifa.

1.1 Motivação

O uso do aprendizado de máquina como solução para o direcionamento de ações estratégicas dentro de organizações vem sendo muito discutido na literatura nos últimos anos. Pode-se afirmar que com o poder computacional atualmente disponível, a grande quantidade de dados gerados e armazenados e o surgimento de algoritmos mais eficientes, a aplicação de *Machine Learning* vem ganhando maior espaço na otimização dos processos, obtenção de vantagens competitivas de mercado, previsão de cenários futuros com base na análise de padrões de dados passados, entre outros.

Além disso, diferentes técnicas baseadas em *Machine Learning* estão sendo amplamente utilizadas por empresas de energia para prever a potência ou a energia necessária para

equilibrar entre geração e demanda (Jawad *et al.*, 2020).

Portanto, ao se realizar uma revisão sistemática da literatura sobre a associação dos temas que abordam os algoritmos inteligentes como solução para determinação do MUST ótimo a ser contratado pelas distribuidoras de energia, identifica-se uma lacuna com a obtenção de um número muito limitado de trabalhos.

Nesse contexto, entende-se que há uma oportunidade de pesquisa visto que, quanto mais precisa for a previsão dos valores de demanda futura, maior o benefício financeiro para os usuários do sistema elétrico.

1.2 Objetivos

- a) **Objetivo Geral:** Otimização da função de custo associado ao CUST das concessionárias de energia por meio da aplicação de algoritmos inteligentes na previsão dos MUST;
- b) **Objetivos Específicos:**
 - Realizar uma revisão literária sobre métodos utilizados para a previsão de demanda de energia elétrica e otimização do MUST com aplicação de algoritmos inteligentes;
 - A partir de dados históricos, construir um modelo de *Machine Learning* que tenha a capacidade de prever os valores de MUST mais compatíveis com o mercado de médio e longo prazo;
 - Baseando-se no modelo de *Machine Learning* de previsão, otimizar a função do custo total associado ao CUST;

1.3 Organização do Trabalho

Este trabalho é organizado em seis capítulos, incluindo o capítulo de introdução, conforme a seguir.

- **Capítulo 2:** Revisão literária dos trabalhos publicados que abordam os temas necessários para elaboração da dissertação;
- **Capítulo 3:** Fundamentação teórica referente aos assuntos principais requeridos para o desenvolvimento da pesquisa, que foram: condições gerais para contratação do acesso aos sistemas de transmissão de energia elétrica, aplicações de modelos de *Machine Learning*

para previsão de demanda de energia elétrica e Modelagem de Equações Estruturais (MEE).

- **Capítulo 4:** Apresentação da proposta de desenvolvimento desta dissertação;
- **Capítulo 5:** Apresentação da proposta de otimização do CUST, utilizando algoritmos de *Machine Learning* para previsão dos cenários futuros, fundamentados nas relações estabelecidas através do uso da MEE;
- **Capítulo 6:** Apresentação das conclusões do trabalho e trabalhos futuros.

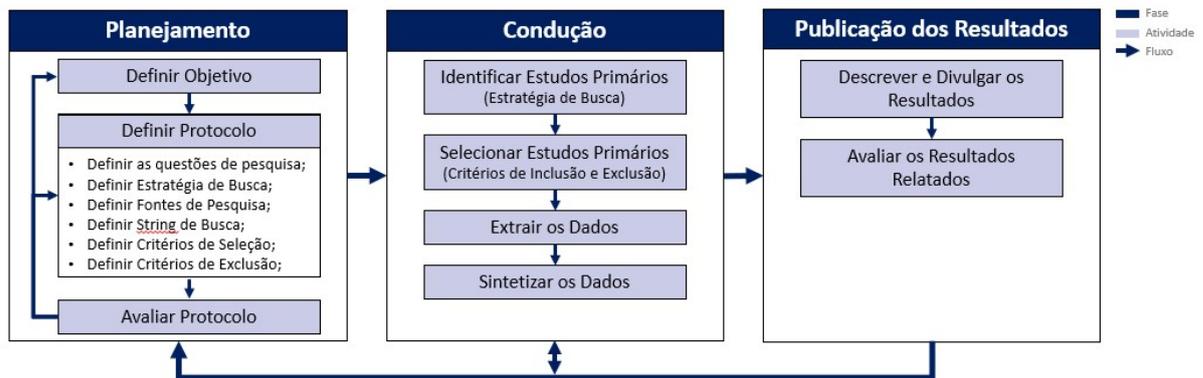
2 REVISÃO BIBLIOGRÁFICA

A revisão sistemática da literatura consiste em um processo de pesquisar, selecionar, avaliar, sintetizar e relatar as evidências científicas sobre uma determinada pergunta e/ou tópico. Nos dias atuais, a revisão sistemática é considerada uma maneira mais racional e menos tendenciosa de organizar, avaliar e integrar evidências científicas (Roever, 2020).

Antes de iniciar o planejamento da revisão, é fundamental identificar se já existem estudos secundários ou terciários em torno dos mesmos temas. Caso não haja um estudo secundário ou terciário sobre os temas e estes sejam de relevância para a comunidade científica das áreas relacionadas, justifica-se então a realização de uma revisão sistemática (Nakagawa *et al.*, 2017).

Nessa dissertação é utilizado o processo descrito na Figura 1 para elaboração da revisão sistemática da literatura.

Figura 1 – Fases e atividades do processo de revisão sistemática



Fonte: (Nakagawa *et al.*, 2017)

2.1 Planejamento

A fase de planejamento tem como objetivo identificar a real necessidade, ou seja, a motivação para a execução de uma revisão sistemática (Nakagawa *et al.*, 2017).

2.1.1 Objetivo da revisão sistemática

Como já foi enfatizado na Seção 1.1, o objetivo da revisão sistemática dessa dissertação é investigar a utilização de algoritmos inteligentes aplicados à previsibilidade dos MUST no Sistema de Transmissão para o processo de contratação do CUST, com um intervalo de tempo

de investigação entre 2017 a 2022.

2.1.2 *Protocolo*

Logo após identificar a necessidade da realização de uma revisão sistemática, é definido o protocolo da revisão que é o elemento essencial para sua execução. Vale destacar que a qualidade do protocolo impacta diretamente no resultado da revisão (Nakagawa *et al.*, 2017).

Desta forma, tem-se que:

- **Questões de Pesquisa:** O objetivo da seção Questões de pesquisa é descrever as questões referentes ao tópico de pesquisa abordado, que devem ser respondidas, neste caso:
 1. Quais os tipos de Aprendizado de Máquina utilizados e suas características para previsão dos Montantes de Uso do Sistema de Transmissão ou de demanda de energia elétrica?
 2. Quais são os parâmetros utilizados para previsão dos Montantes de Uso do Sistema de Transmissão ou de demanda de energia elétrica?
 3. Quais os modelos matemáticos que são utilizados para contratação dos Montantes de Uso do Sistema de Transmissão?
 4. Quais estratégias utilizadas na contratação dos Montantes de Uso do Sistema de Transmissão?
- **Estratégia de Busca:** Define qual o meio que será utilizado para realizar a revisão, neste caso, é por meio da aplicação de uma *string* de busca nas fontes de pesquisa escolhidas;
- **Fontes de Pesquisa:** Foram definidas as fontes: *Scopus, IEEEExplore, Engineering Village e Web of Science.*
- **CrITÉrios de Seleção:** Uma das características que difere a revisão sistemática de outros métodos de revisão é que esta segue um método rigoroso de seleção de estudos com base nos critérios de seleção, que abrangem tanto critérios de inclusão quanto de exclusão. A qualidade dos resultados da revisão sistemática depende dos estudos incluídos (Nakagawa *et al.*, 2017). Portanto, segue a definição dos critérios de seleção:
 1. Estudos revisados e publicados no intervalo de tempo de 2017 a 2022;
 2. Estudos escritos nas linguagens: Inglês e Português;
 3. Estudos que propõem modelos de previsão para demanda de energia elétrica ou MUST;
 4. Estudos que propõem estratégias de contratação para o CUST;

5. Estudos que descrevem técnicas de aprendizado de máquina ou de inteligência artificial para previsão de demanda de energia elétrica ou MUST.
- **Critérios de Exclusão:** Bem como os critérios de seleção, os critérios de exclusão são definidos para descartar os estudos. Sendo assim, tem-se os seguintes critérios de exclusão:
1. Estudos que não atendem os critérios de seleção;
 2. Estudos que não abordam MUST e CUST ou previsão de demanda de Energia Elétrica;
 3. Estudos que abordam a utilização de aprendizado de máquina ou inteligência artificial para previsão de assuntos não relacionados à demanda de energia elétrica; Livros, notas, dissertações, teses;
 4. Estudos apenas com resumo;
 5. Estudo do tipo *survey* e do tipo *review*.
- **Definição da *string* de busca:** É a combinação das palavras-chave e termos relacionados usando operadores lógicos de tal forma que a maior e melhor quantidade de estudos sejam encontrados.

As palavras-chave que compõem a *string* são apresentadas no Quadro 1.

Quadro 1 – Mapeamento de palavras-chave para *string* de busca.

Palavras - Chave	Sinônimos	Plural	Abreviações
Montante de Uso do Sistema de Transmissão	Amount of use of the transmission system	Montantes de Uso do Sistema de Transmissão	MUST
Contrato de Uso do Sistema de Transmissão	Contracted power of the transmission system	Contratos de Uso do Sistema de Transmissão	CUST
Algoritmo Inteligente	Inteligência Artificial	Algoritmos Inteligentes	IA
Aprendizado de Máquina	Machine learning		
Previsão de Energia Elétrica	Previsibilidade de Energia Elétrica / Forecast of Electric Energy	Previsões de Energia Elétrica	

Fonte: Elaborado pela própria autora.

Conforme apresentado no Quadro 1, segue a última versão da *string* utilizada: ((*Montante de Uso do Sistema de Transmissão*) OR (*Montantes de Uso do Sistema de Transmissão*) OR (*Amount of use of the transmission system*) OR “MUST” OR (*Contrato de Uso do Sistema de Transmissão*) OR (*Contratos de Uso do Sistema de Transmissão*) OR (*Contracted power of the transmission system*) OR “CUST”) AND ((*Algoritmo Inteligente*) OR (*Inteligência Artificial*) OR (*Algoritmos Inteligentes*) OR “IA” OR (*Aprendizado de Máquina*) OR (*Machine learning*)) AND ((*Previsão de Energia Elétrica*) OR (*Previsibilidade de Energia Elétrica*) OR (*Forecast of Electric Energy*) OR (*Previsões de Energia Elétrica*)).

2.1.3 Avaliação do protocolo

A avaliação do protocolo é essencial para melhorar a qualidade da revisão sistemática, entretanto, alguns autores consideram que as questões de pesquisa não devem ser alteradas após a avaliação do protocolo. Em decorrência disso, é sugerido que seja feito um teste piloto para refinamento do protocolo, em especial para que as questões de pesquisa caracterizem adequadamente o objetivo da revisão sistemática (Nakagawa *et al.*, 2017).

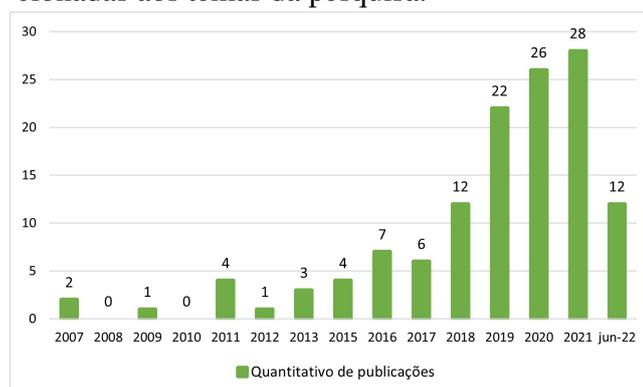
Sendo assim, algumas perguntas devem ser respondidas com base no protocolo construído. São essas:

1. Os objetivos da revisão estão claramente descritos? Sim.
2. As questões de pesquisa definidas são claras e objetivas? Sim.
3. As fontes a serem pesquisadas para identificar os estudos primários foram definidas? Sim.
4. A string de busca é devidamente derivada das questões de pesquisa e utiliza as palavras chave mapeadas? Sim.
5. Os critérios de inclusão estão descritos claramente? Sim.
6. Os critérios de exclusão estão descritos claramente? Sim.

2.2 Condução

Utilizando as definições descritas na subseção 2.1 pode-se ter uma visão quantitativa da produção mundial de estudos publicadas ao longo dos últimos anos. Após a realização da exclusão dos estudos duplicados, na Figura 2 é mostrado o quantitativo de publicações anuais, retratando um crescimento significativo nos últimos anos e justificando o período determinado dos critérios de seleção mostrados na subseção 2.1.

Figura 2 – Quantitativo Anual de publicações relacionadas aos temas da pesquisa.



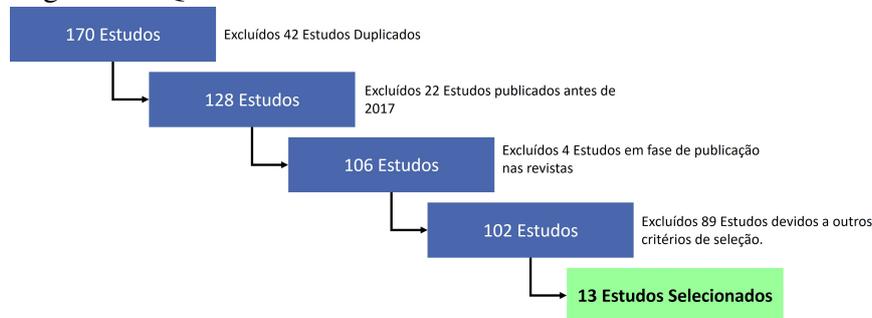
Fonte: Elaborado pela própria autora.

2.2.1 Síntese dos dados após a aplicação dos critérios de seleção e exclusão

Foram realizadas duas análises em todos os estudos encontrados, sendo a primeira com a leitura do título e resumo, e a segunda análise com a leitura superficial do texto geral (introdução, técnicas, resultados e conclusão), ambas com aplicação dos critérios de inclusão e exclusão a fim de selecionar os estudos que com maior aplicabilidade nesta dissertação.

Sendo assim, na Figura 3 é mostrada a quantidade de trabalhos excluídos em cada etapa realizada e no Quadro 2 é descrito um total de 13 trabalhos selecionados, dentre eles um estudo de 2008, que foi tratado como exceção pois trata especificamente de algoritmos inteligentes aplicados aos MUST.

Figura 3 – Quantidade de estudos excluídos.



Fonte: Elaborado pela própria autora.

Quadro 2 – Trabalhos selecionados.

ID	ANO	TÍTULO	AUTOR	PUBLICAÇÃO	PAÍS
1	2021	Rede Neural Convolucional Típica Suportada por Perfil de Carga para Previsão de Carga de Curto Prazo no Setor Industrial	Walsler T,Sauer A	Artigo de Jornal	Alemanha
2	2021	Modelos de Previsão de Carga Elétrica para Diferentes Modalidades de Geração: Uma Revisão	Azeem A,Ismail I,Jameel SM, Harindran VR	Artigo de Jornal	Irlanda
3	2021	Uma Previsão de Demanda de Carga de Curto Prazo Baseada no Método LSTM	Bodur I,Celik E,Ozturk N	Anais de Conferência	Turquia
4	2020	Uma Revisão Abrangente das Técnicas de Previsão de Carga Usando Modelos Preditivos Únicos e Híbridos	Al Mamun A,Sohel M,Mohammad N, Sunny MS, Dipta DR,Hossain E	Artigo de Jornal	EUA
5	2020	Modelagem de Machine Learning para Consumo de Energia dos Setores Residencial e Comercial	Nabavi SA,Aslani A,Zaidan MA, Zandi M,Mohammadi S,Motlagh NH	Artigo de Jornal	Irã
6	2020	Modelo de Previsão de Carga de Eletricidade Econômico Baseado em Aprendizado de Máquina Usando Parâmetros Meteorológicos Correlacionados	Jawad M,Nadeem MS,Shim SO, Khan IR,Shaheen A,Habib N,Hussain L,Aziz W	Artigo de Jornal	Arábia Saudita
7	2020	Modelo Híbrido CNN-LSTM para Previsão de Carga Doméstica Individual de Curto Prazo	Alhussein M,Aurangzeb K,Haider SI	Artigo de Jornal	Arábia Saudita
8	2019	Deep Learning CNN e AI-Tuned SVM para Previsão de Consumo de Eletricidade: Dados de Séries Temporais Multivariadas	Chan S,Oktavianti I,Puspita V	Anais de Conferência	EUA
9	2019	Um Modelo Híbrido de Deep Learning com Algoritmo Evolutivo para Previsão de Carga de Curto Prazo	Mamun AA,Hoq M,Hossain E,Bayindir R	Anais de Conferência	Romênia
10	2019	Aplicação de Deep Learnin Long Short-Term Memory na Previsão de Demanda de Energia	Al Khafaf N,Jalili M,Sokolowski P	Artigo de Jornal	Austrália
11	2018	Avaliando Técnicas de Previsão para Integração Prosumidores de Energia Doméstica em Redes Inteligentes	Petrican T,Vesa AV,Antal M,Pop C,Cioara T, Anghel I,Salomie I	Anais de Conferência	Romênia
12	2011	Determinação do MUST Ótimo para Empresas de Distribuição de Energia Elétrica	Luana M. Marangon Lima, Anderson R. de Queiroz, José W. Marangon Lima, Anselmo Ribeiro,Eduardo Elhage	Anais de Conferência	Brasil
13	2008	Previsão de Demandas para Controle dos Montantes de Uso do Sistema de Transmissão para Distribuidoras de Energia Elétrica	Vitor Hugo Ferreira Alexandre Rasi Aoki Fábio Alessandro Guerra	Anais de Conferência	Brasil

Fonte: Elaborado pela própria autora.

2.3 Resultados da Revisão Sistemática

Desde a reestruturação do setor elétrico brasileiro, quando a regulamentação passou a estabelecer que os custos incorridos pelo uso da Rede Básica em função dos MUST e da geração de cada agente fossem rateados proporcionalmente entre as partes integrantes do sistema, os agentes necessitam gerenciar suas contratações do sistema de transmissão de forma adequada.

A grande maioria dos trabalhos selecionados discutem modelos de aprendizado de máquina para previsão de demanda de energia elétrica a partir de dados históricos, além disso evidenciam as necessidades e as dificuldades encontradas no processo, bem como enfatizam a melhoria das previsões quando estão correlacionadas com outras variáveis, como temperatura ou clima.

No trabalho Ferreira *et al.* (2008), os autores propõem o desenvolvimento de uma metodologia baseada em Redes Neurais Artificiais (RNA) com a aplicação de inferência *bayesiana* com perceptrons de múltiplas camadas (*Multi-layered Perceptrons* (MLP)) para previsão dos MUST, assim como também realizam um comparativo com o modelo de máquina de vetor de suporte (*Support Vector Machines* (SVM)) a fim de avaliar o desempenho de ambos os modelos.

Em Mamun *et al.* (2020) os autores fazem um estudo abrangente e comparativo de modelos de previsão únicos e combinados, verificando métricas de desempenho com intuito de auxiliar pesquisadores na seleção de modelos adequados para previsão de demanda de energia elétrica.

Dentre os resultados do trabalho, os autores, Mamun *et al.* (2020), destacam que variáveis relacionadas à demanda utilizadas no modelo preditivo, como o tempo de previsão, o clima, a temperatura local e a economia, são fatores significativos na precisão das previsões de saída. Outros pontos destacados é que enquanto os modelos MLP apresentaram melhores valores de precisão, ou seja, baixa variação nos resultados a partir de diferentes entradas nos modelos de predição, os modelos SVM mostraram melhores valores de acurácia, ou seja, os valores previstos são mais próximos dos valores reais utilizados no algoritmo. Já os modelos combinados apresentaram bons resultados para precisão e acurácia devido o efeito cascata de modelo e, portanto, possuem os melhores resultados.

Como resultado, os autores confirmam o que foi mostrado no estudo Ferreira *et al.* (2008), que descreve que o modelo SVM apresentou resultados de previsão inferiores aos modelos de MLP combinados com a inferência *bayesiana*, apesar do modelo combinado ter necessitado de um maior esforço computacional.

Na publicação Jawad *et al.* (2020) são desenvolvidos modelos de previsão que correlacionam a carga elétrica com as variáveis meteorológicas de temperatura do ar, temperatura do ponto de orvalho, umidade relativa, cobertura de nuvens, velocidade do vento e precipitação, com o propósito de se obter a maior significância econômica, ou seja, o menor custo para atendimento à demanda elétrica prevista.

O algoritmo proposto por Jawad *et al.* (2020) tem como entrada séries temporais de dois anos em duas configurações distintas, a primeira somente com os dados históricos de carga elétrica e a segunda com os dados históricos de carga elétrica e parâmetros meteorológicos. São comparados cinco métodos de aprendizado de máquina: *Multiple Linear Regression* (MLR), *K-Nearest Neighbors* (KNN), *Support Vector Regressor* (SVR), *Random Florest* (RF) e *Adaptive Boosting* (AdaBoost).

Foram usadas três métricas para avaliação de desempenho dos modelos de previsão, conforme segue:

- **Mean Absolute Error (MAE):** Consiste na média das distâncias entre valores preditos e reais;
- **Root Mean Square Error (RMSE):** Métrica para problemas nos quais grandes erros não são tolerados, penaliza muito as previsões distantes da real;
- **Mean Absolute Percentage Error (MAPE):** Consiste em calcular o quanto o modelo avaliado erra sua previsão no percentual em média calculado do valor real, ou seja, é uma métrica que mostra a porcentagem de erro em relação aos valores reais.

O cálculo de significância econômica utiliza os valores das métricas de avaliação de cada configuração simulada como efeito comparativo entre os modelos. De acordo com a maior redução de erro na previsão no período específico, os modelos que apresentaram os melhores resultados foram KNN, SVR e MLR.

Contudo, dentre os resultados apresentados em Jawad *et al.* (2020), a maior contribuição é o comparativo entre as duas configurações, em que pode ser visto o benefício significativo de quando outras variáveis correlacionadas são combinadas com os dados de demanda elétrica nos modelos de previsão. Em todas as simulações os valores previstos apresentaram melhor precisão e acurácia.

Walser; Sauer (2021) propuseram um modelo híbrido de aprendizado de máquina, combinando um perfil de carga típico baseado em árvore de decisão com uma rede neural convolucional que extrai recursos de entradas multidimensionais com medições das duas se-

manas anteriores para gerar previsões. Além disso, o estudo utiliza variáveis adicionais, como temperatura local, umidade e radiação. Os resultados foram demonstrados com um comparativo entre três perfis de carga diferentes em três indústrias. Apesar do modelo não prever a longo prazo, evidencia as vantagens da combinação de variáveis correlacionadas aos dados de demanda e as oportunidades que existem com algoritmos híbridos.

Outra perspectiva é mostrada através do trabalho Lima *et al.* (2021), que apresenta dois modelos matemáticos com aplicação da otimização estocástica e otimização inteira mista, algo mais relacionado ao problema de determinação do MUST pelas distribuidoras para a contratação.

A priori, os dados de demanda de energia das subestações são divididos através de um processo de agrupamento realizado por meio do modelo *K-means*, de modo a criar grupos com subestações de perfis de demanda fortemente correlacionadas. Uma vez formados os grupos, simula-se o fluxo de potência para obtenção dos fluxos nas interligações da distribuidora com a rede básica. Com isso, utilizam-se esses fluxos como parâmetros de entrada para os modelos de otimização.

Como resultado, tem-se que a utilização do método de otimização inteira mista é mais indicada para análises de cenários determinísticos, enquanto que o método da otimização estocástica é mais indicado quando considera-se a incerteza nas demandas das subestações. Entretanto, ambos os modelos apresentam soluções similares, o que valida a hipótese de independência entre as cargas que estão fisicamente próximas.

No Quadro 3 são mostrados graficamente os modelos debatidos em cada um dos trabalhos selecionados com base nos índices informados no Quadro 2, com adição do índice 14 que corresponde a esta dissertação (em verde). Pode-se notar que os métodos mais abordados dentre os trabalhos selecionados foram SVM, *Long Short Term Memory* (LSTM) e *Convolutional Neural Network* (CNN).

Quadro 3 – Modelos de algoritmos inteligentes utilizados nos trabalhos selecionados.

MODELOS																		
ID	LSTM	SVM	CNN	MLP	DL	ANN	RNN	SARIMA	MLR	LMLR	NARX	KNN	SVRL	SVRR	SVRP	ADABOOST	RFR	K-MEANS
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		

Fonte: Elaborado pela própria autora.

Sob o ponto de vista dos atributos utilizados nos modelos de previsão dos trabalhos selecionados, e com base nos mesmos índices da Quadro 2, a Quadro 4 mostra graficamente quais os atributos utilizados em cada um dos trabalhos selecionados, bem como dos quais àqueles que foram utilizados nesta dissertação em verde.

Quadro 4 – Atributos utilizados nos trabalhos selecionados.

ATRIBUTOS										
ID	TEMPERATURA	UMIDADE	VELOCIDADE DO VENTO	CHUVA	RADIAÇÃO GLOBAL	RADIAÇÃO DIFUSA	TARIFA	CLIMA	DADOS DEMOGRÁFICOS	DIA ÚTIL / NÃO ÚTIL
1										
4										
5										
3										
6										
7										
8										
9										
10										
9										
11										
13										
12										
14										

Fonte: Elaborado pela própria autora.

Diante dos principais tópicos debatidos acerca dos temas relacionados, nesta dissertação é proposta a aplicação de algoritmos inteligentes a fim de otimizar o custo associado ao CUST e para tanto foi feito um estudo de correlação entre os parâmetros e os dados históricos de demanda de energia elétrica.

A diferença desta dissertação para as demais referências, é que neste trabalho é utilizado a MEE para preestabelecer uma relação causal dos fatores que influenciam a demanda de energia elétrica e, com isso, utilizar algoritmos de *machine learning* para realizar uma previsão de MUST mais realista e compatível com o mercado de energia a fim de resultar na otimização do custo associado ao CUST.

3 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são abordados conceitos fundamentais sobre as condições gerais para contratação do acesso aos sistemas de transmissão de energia elétrica, os fundamentos básicos da MEE utilizada para estabelecer as relações entre as variáveis de influência e os valores de demanda a serem previstos e os princípios dos modelos de *Machine Learning* usados para construção do algoritmo proposto de previsão.

3.1 Condições Gerais de Acesso ao Sistema de Transmissão

Com o objetivo de estabelecer os requisitos da contratação do acesso, do uso e da conexão aos sistemas de transmissão de energia elétrica, a ANEEL promulgou a REN n°905 em 2020, a resolução normativa que estabelece as regras dos serviços de transmissão de energia elétrica para centrais geradoras, consumidores, distribuidoras, importadores e exportadores de energia (ANEEL, 2020). Na sequência, são sintetizadas as principais informações destacadas na Resolução Normativa REN n°905/2020.

Por definição, as instalações de transmissão são objetos de contrato de concessão para prestação do serviço público de transmissão de energia elétrica, além das autorizadas por resolução específica da ANEEL e das que tenham sido cedidas, doadas ou transferidas às transmissoras.

Para terem acesso às instalações de transmissão, os agentes devem firmar o CUST com o ONS, estabelecendo as condições técnicas e as obrigações relativas ao uso do sistema de transmissão, e o Contrato de Conexão às Instalações de Transmissão (CCT) com as transmissoras, estabelecendo as responsabilidades pela implantação, operação e manutenção das instalações de conexão e os respectivos encargos de conexão, quando couber, bem como as condições comerciais e financeiras, em cada ponto de conexão.

Este trabalho tem como foco as distribuidoras de energia elétrica, desta forma, entende-se por ponto de conexão o local da conexão que, para efeito do acesso, os valores de MUST devem ser contratados e verificados.

Tais valores são contabilizados para fim de faturamento através da leitura efetuada pelo responsável pela instalação do respectivo Sistema de Medição de Faturamento (SMF), e ocorre em intervalos de integralização de 15 minutos.

Dentre as condições técnicas e comerciais a serem observadas, o CUST deve dispor

dos MUST contratados nos horários de ponta e fora de ponta, bem como as condições e antecedência mínima para a solicitação de alteração dos valores de uso contratados. Além disso, deve dispor da administração, pelo ONS, da cobrança e liquidação dos Encargos de Uso do Sistema de Transmissão (EUST) e a execução do sistema de garantias por conta e ordem das transmissoras, assim como as penalidades por atraso no pagamento.

Interpreta-se como EUST, os montantes financeiros definidos e homologados pela ANEEL relativos ao uso dos pontos de conexão, devidos pela distribuidora à transmissora e considerados no cálculo da tarifa do consumidor final da concessionária a partir da respectiva prestação do serviço, sem efeitos retroativos.

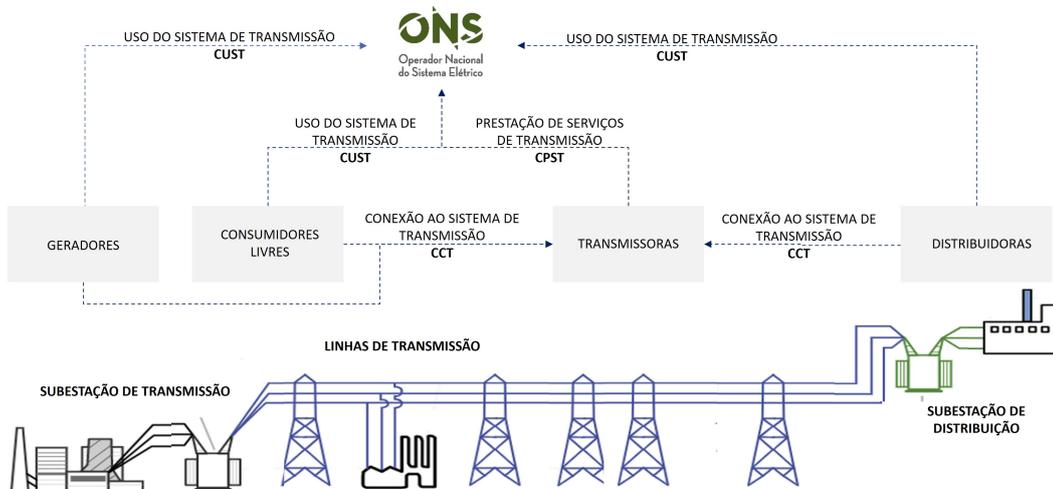
Os valores do EUST deverão ser suficientes para a prestação do serviço de transmissão, observando a RAP para as empresas de transmissão, a parcela do orçamento anual do ONS e a compensação de déficit ou superávit do exercício anterior, contabilizado anualmente pelo ONS e aprovado pela ANEEL.

Além do contrato firmado com o ONS, a distribuidora deve celebrar também o CCT com a transmissora, dispondo:

- Da descrição detalhada dos pontos de conexão, incluindo o conjunto de equipamentos necessários para a interligação elétrica das instalações;
- Dos valores dos EUST, estabelecidos pela ANEEL por meio de resolução homologatória;
- Das penalidades pelo atraso no pagamento dos EUST;
- Da data de entrada em operação das instalações do ACESSANTE;
- Da data de início da prestação dos serviços;
- Do prazo de vigência.

Na Figura 4 é mostrada a visão geral dos contratos firmados entre os agentes de acordo com as resoluções normativas da ANEEL.

Figura 4 – Visão geral dos contratos firmados entre os agentes.



Fonte: Elaborado pela própria autora baseado nos Procedimentos de Rede.

3.2 Regras Gerais de Contratação de Uso do Sistema de Transmissão

Os CUST celebrados em caráter permanente por distribuidoras deverão conter seus MUST para os quatro anos civis subsequentes, sendo livre a declaração do quarto ano. Tais valores devem ser informados ao ONS até o dia 31 de outubro de cada ano, passando a vigorar a partir de 1º de janeiro do ano subsequente. Caso não sejam informados até a data limite do ano corrente, o ONS considera os valores constantes de cada CUST vigente, o mesmo se aplica ao valor contratado para o terceiro ano como o de contratação para o quarto ano e uma cláusula de renovação automática (ANEEL, 2020).

Os MUST a serem contratados devem ser informados nos horários de contratação, em cada um dos pontos de conexão utilizados e que fazem fronteira com a rede básica (ANEEL, 2020).

Os horários de contratação são definidos como o período horário para operação dos agentes de distribuição divididos em:

- **Horário de Ponta:** Corresponde ao período composto pelas três horas diárias consecutivas em que ocorre a máxima demanda diária da distribuidora. No Ceará, corresponde ao período das 17h:30min às 20h:30min.
- **Horário Fora Ponta:** Período composto pelo conjunto das horas diárias consecutivas e complementares àquelas definidas no horário de ponta.

Os valores contratados podem ser alterados de três formas: reduzidos de forma não onerosa, reduzidos de forma onerosa e aumentados.

As solicitações de aumento estão limitadas em quatro vezes para o ano civil em curso e devem ser realizadas com antecedência mínima de noventa dias em relação à data de início do aumento pretendido e mediante parecer de acesso específico (ANEEL, 2020).

As reduções de forma não onerosa ocorrem desde que sejam realizadas até no máximo 10% do montante previamente contratado para o mesmo ano civil e para os 3 (três) anos civis subsequentes (ANEEL, 2020).

Com isso, as reduções em valores superiores a 10% se dão de forma onerosa e têm seus valores calculados multiplicando-se a Tarifa do Uso do Sistema de Transmissão (TUST) vigente no primeiro mês da redução onerosa e o MUST a ser reduzido que exceder 10%, por horário de contratação, até o final do terceiro ano civil subsequente, com a liquidação ocorrendo na primeira apuração mensal de serviços e encargos subsequentes, e valores não repassados ao usuário final, tornando uma opção pouco atraente para as distribuidoras (ANEEL, 2020).

A TUST é dividida entre a parcela referente à Tarifa do Uso do Sistema de Transmissão de Rede Básica (TUST-RB) e a parcela referente à Tarifa do Uso do Sistema de Transmissão de Fronteira (TUST-FR), conforme segue:

- A TUST-RB é calculada conforme descrito nas Regras de Transmissão e nos Procedimentos de Regulação Tarifária (PRORET) que leva em conta as parcelas da RAP associadas às instalações de Rede Básica;
- A TUST-FR será aplicável apenas à distribuidora que utilize transformadores de potência com tensão primária igual ou superior a 230 kV e tensões secundárias e terciárias inferiores a 230 kV, bem como às respectivas conexões e demais equipamentos ligados ao terciário, em caráter exclusivo ou compartilhado, ou que se conecte em Demais Instalações de Transmissão DIT sob responsabilidade da transmissora, em caráter compartilhado e levará em conta as parcelas da RAP associadas a estas instalações, sendo rateada pelos MUST contratados pela respectiva distribuidora nos horários de ponta e fora de ponta.

Existe ainda a possibilidade do CUST ser celebrado em caráter flexível, quando houver o uso da capacidade remanescente do sistema de transmissão por tempo determinado (ANEEL, 2020). Neste caso, o MUST contratado em caráter flexível deve ser único para cada CUST e ser contratado simultaneamente ao mesmo, por ponto de conexão e horário de contratação (ANEEL, 2020).

A TUST a ser aplicada à contratação em caráter flexível deve ser estabelecida a partir da TUST calculada para os contratos em caráter permanente conforme mostrado nas Equações

3.1 e 3.2 para horários de ponta e fora ponta, respectivamente.

$$TUST_{p-flexivel} = TUST_p \cdot \left(\frac{MUST_p + MUST_{p-flexivel}}{MUST_p} \right) \quad (3.1)$$

$TUST_{p-flexivel}$ = TUST Ponta para CUST celebrado em caráter flexível;

$TUST_p$ = TUST Ponta para CUST celebrado em caráter permanente;

$MUST_{p-flexivel}$ = MUST Ponta contratado em caráter flexível;

$MUST_p$ = MUST Ponta contratado em caráter permanente;

$$TUST_{fp-flexivel} = TUST_{fp} \cdot \left(\frac{MUST_{fp} + MUST_{fp-flexivel}}{MUST_{fp}} \right) \quad (3.2)$$

$TUST_{fp-flexivel}$ = TUST Fora Ponta para CUST celebrado em caráter flexível;

$TUST_{fp}$ = TUST Fora Ponta para CUST celebrado em caráter permanente;

$MUST_{fp-flexivel}$ = MUST Fora Ponta contratado em caráter flexível;

$MUST_{fp}$ = MUST Fora Ponta contratado em caráter permanente;

Os EUST referentes às contratações em caráter flexível, serão devidos apenas nos dias em que ocorrer o uso, por horário de contratação, e sobre o EUST total contratado em caráter flexível.

3.2.1 Encargos e Eficiência da Contratação em Caráter Permanente

O ONS é o responsável pela apuração, administração da cobrança, liquidação dos serviços e pelos valores de EUST mensais que são atribuídos às distribuidoras de forma proporcional às suas demandas máximas de potência em cada ponto de conexão e horário de contratação, a partir do produto entre a TUST e os valores plenos contratados dos MUST. A Equação 3.3 apresenta o cálculo anual dos valores de EUST (ANEEL, 2020).

$$EUST = \sum_{i=1}^{12} [TUST_{fp} \cdot MUST_{fp}(i) + TUST_p \cdot MUST_p(i)] \quad (3.3)$$

Em que:

i = Mês

$MUST_{fp}$ = MUST Contratado Mensal Fora Ponta em caráter permanente

$MUST_p$ = MUST Contratado Mensal Ponta em caráter permanente

$TUST_{fp}$ = TUST Fora Ponta para CUST celebrado em caráter permanente;

$TUST_p$ = TUST Ponta para CUST celebrado em caráter permanente;

Quando forem iguais a zero os MUST contratados, o ONS fica autorizado a administrar a cobrança diretamente dos encargos destinados a remunerar as parcelas de RAP associadas a estas instalações, sendo calculados com base em duodécimos destas receitas (ANEEL, 2020).

As diferenças entre os MUST contratados e os máximos verificados por meio da medição são apuradas e avaliadas para eficiência da contratação que são apuradas pelo ONS por horário de contratação e ponto de conexão, da seguinte forma:

- Mensalmente, quando houver ultrapassagem de demanda, caracterizada pela medição de demanda máxima em valor superior ao MUST contratado em caráter permanente; e
- Anualmente, quando houver sobrecontratação de demanda, caracterizada pela medição de demanda máxima anual em valor inferior a 90% do maior MUST contratado em caráter permanente no ano civil.

Nos meses em que houver a ultrapassagem do valor de MUST contratado, o ONS apurará o Adicional dos Encargos de Uso do Sistema de Transmissão (ADEUST) e a Parcela de Ineficiência por Ultrapassagem (PIU). Na Equação 3.4 e Equação 3.5 são mostrados como são calculados seus respectivos valores anuais.

$$ADEUST = \sum_{i=1}^{i=12} [TUST_{fp} \cdot (D_{maxfp}(i) - MUST_{fp}(i)) + TUST_p \cdot (D_{maxp}(i) - MUST_p(i))] \quad (3.4)$$

Em que:

i = Mês

$MUST_{fp}$ = MUST Contratado Mensal Fora Ponta

$MUST_p$ = MUST Contratado Mensal Ponta

D_{maxfp} = Demanda Máxima Mensal Medida Fora Ponta

D_{maxp} = Demanda Máxima Mensal Medida Ponta

$$PIU = \sum_{i=1}^{i=12} [3 \cdot TUST_{fp} \cdot (D_{maxfp}(i) - 1,10 \cdot MUST_{fp}(i)) + 3 \cdot TUST_p \cdot (D_{maxp}(i) - 1,10 \cdot MUST_p(i))] \quad (3.5)$$

Em que:

i = Mês

$MUST_{fp}$ = MUST Contratado Mensal Fora Ponta

$MUST_p$ = MUST Contratado Mensal Ponta

D_{maxfp} = Demanda Máxima Mensal Medida Fora Ponta

D_{maxp} = Demanda Máxima Mensal Medida Ponta

Nos anos em que as demandas máximas medidas não conseguem atingir 90% do valor de MUST contratado em nenhum dos meses, o ONS apurará a Parcela de Ineficiência por Sobrecontratação (PIS). Na Equação 3.6 é mostrado como deve ser calculado o valor anual do PIS.

$$PIS = [12 \cdot TUST_{fp} \cdot (0,90 \cdot MUST_{fp}(i) - D_{maxanofp}(i)) + 12 \cdot TUST_p \cdot (0,90 \cdot MUST_p(i) - D_{maxanop}(i))] \quad (3.6)$$

Em que:

i = Mês

$MUST_{fp}$ = MUST Contratado Mensal Fora Ponta

$MUST_p$ = MUST Contratado Mensal Ponta

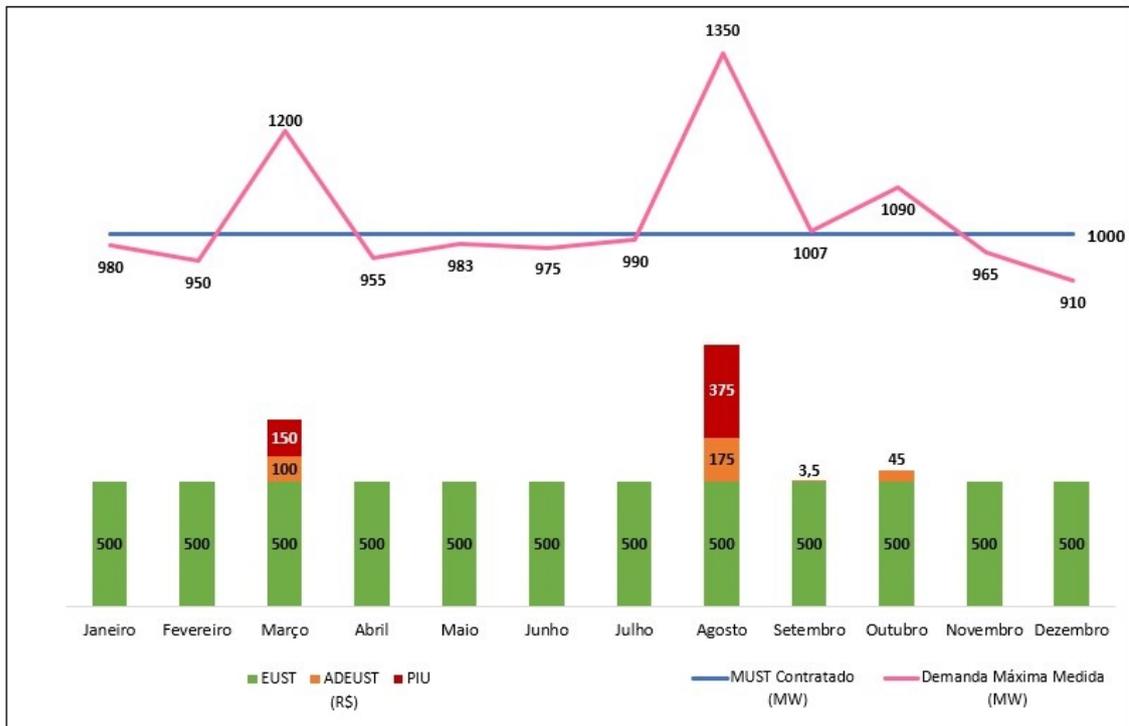
$D_{maxanofp}$ = Demanda Máxima Anual Medida Fora Ponta

$D_{maxanop}$ = Demanda Máxima Anual Medida Ponta

Com o intuito de facilitar o entendimento, nas figuras 5 e 6 são mostrados exemplos didáticos de aplicação das expressões para o cálculo dos encargos e a eficiência da contratação no CUST, levando em consideração uma $TUST = 0,5R\$/kW$

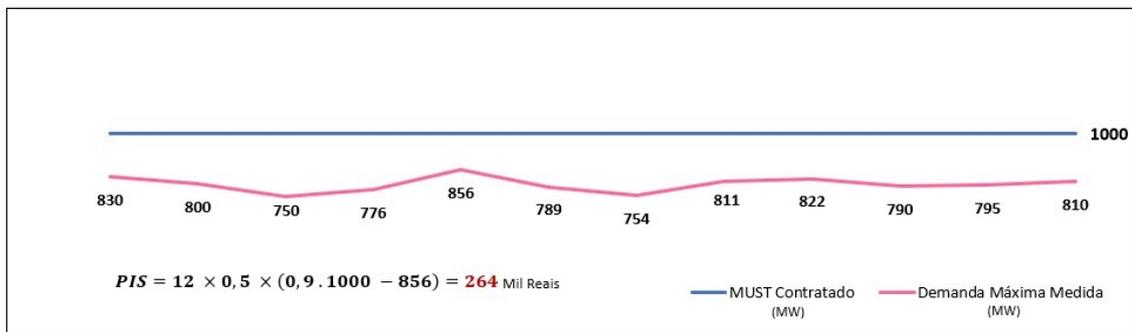
Observa-se que apesar dos valores referentes a parcela PIS serem calculados utilizando um fator multiplicativo igual a 12, não pode-se concluir que esta penalidade apresenta uma influência maior ou menor que a parcela PIU, que utiliza um valor multiplicativo igual a 3, nos encargos pagos pelas distribuidoras de energia elétrica. Vale ser destacado que as penalidades são calculadas em tempos diferentes, enquanto a parcela PIU é mensal, a parcela PIS é anual.

Figura 5 – Demonstrativo dos encargos e ultrapassagens.



Fonte: Elaborado pela própria autora.

Figura 6 – Exemplo do cálculo de sobrecontratação.



Fonte: Elaborado pela própria autora.

3.3 Fundamentos Básicos da Modelagem de Equações Estruturais

A Modelagem de Equações Estruturais (MEE) é uma técnica de modelagem estatística multivariada de caráter geral, que é amplamente utilizada nas Ciências Humanas e Sociais. As relações entre as construções teóricas são representadas por coeficientes de regressão ou coeficientes de trajetória entre variáveis observadas e/ou latentes. O modelo de equações estruturais implica uma estrutura para as covariâncias entre as variáveis observadas (Neves, 2018).

Pode ser vista como uma combinação de análise fatorial e regressão, conforme é explicado abaixo:

- **Análise fatorial:** é uma técnica estatística multivariada não determinística que permite a mensuração de variáveis latentes, construtos não observados de forma direta, a partir de um conjunto de variáveis manifestas, ou seja, observadas diretamente;
- **Análise de Trajetórias:** é uma extensão da análise de regressão linear de mínimos quadrados. Essa extensão permite a decomposição de efeitos estatísticos entre: efeito direto e efeitos indiretos.

Os modelos de equações estruturais são, na maioria das vezes, visualizados por um diagrama de trajetórias. O modelo estatístico geralmente pode ser representado por um conjunto de equações matriciais. No início da introdução dessa técnica nas pesquisas sociais e comportamentais, os softwares normalmente demandavam configurações que especificassem o modelo em termos dessas matrizes. Assim, os pesquisadores tinham que representar a matriz a partir do diagrama de trajetórias e fornecer aos softwares uma série de matrizes para os diferentes conjuntos de parâmetros, como cargas fatoriais e coeficientes de regressão (Neves, 2018).

Isso não se faz mais necessário hoje, pois os modelos podem ser representados a partir de um conjunto de comandos, neste trabalho é utilizado o pacote SEMOPY em *Python*.

De acordo com (Neves, 2018), a modelagem de equações estruturais pode ser utilizada de três diferentes formas:

- a) **Abordagem estritamente confirmatória:** na qual se testa um modelo teórico previamente especificado, concluindo-se por sua aceitação ou refutação;
- b) **Abordagem de modelos alternativos:** na qual se faz uma análise comparativa da qualidade de ajuste de dois ou mais modelos teóricos previamente especificados;
- c) **Abordagem de desenvolvimento de modelos:** na qual há um primeiro passo semelhante ao item a), porém, no caso da refutação do modelo especificado, se parte para a busca de um modelo com melhor qualidade de ajuste, em geral mais simples do que o modelo original.

Como dito acima, a MEE combina uma análise de regressão com uma análise fatorial. A análise de regressão é a técnica de análise multivariada mais utilizada por ser uma técnica de análise determinística, na qual se busca observar as covariâncias, ou “efeitos”, de uma ou mais variáveis independentes sobre uma variável dependente (Neves, 2018).

O modelo de análise de regressão pode ser expresso a partir de estruturas matriciais. Assim, na forma de matrizes, um sistema de equações de regressão múltipla pode ser apresentado pela Equação 3.7:

$$Y = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,k} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,k} \end{bmatrix} \cdot \begin{bmatrix} \beta_{1,1} \\ \beta_{2,1} \\ \vdots \\ \beta_{m,1} \end{bmatrix} + \begin{bmatrix} e_{1,1} \\ e_{2,1} \\ \vdots \\ e_{m,1} \end{bmatrix} \rightarrow Y = X\beta + e_i \quad (3.7)$$

Em que:

Y = Matriz com os valores de variáveis dependentes dos regressores

X = Matriz com os valores de variáveis independentes

β = Coeficientes de regressão

e = Erro

Para esta dissertação, a MEE foi aplicada na abordagem de desenvolvimento de modelos. Como pode ser visto no Capítulo 4, um modelo intuitivo foi testado inicialmente e após a sua análise, houve uma busca de um modelo com melhor qualidade de ajuste e que represente bem as influências que as variáveis exercem na demanda de energia elétrica e, portanto, auxiliando na construção de um algoritmo de previsão.

3.3.1 Pacote SEMOPY

O SEMOPY é um pacote desenvolvido na linguagem *Python* para ajudar pesquisadores que empregam técnicas de MEE. Em termos de desempenho, a ferramenta foi criada para ser uma versão melhorada do pacote LAVAAN, desenvolvido para linguagem *R*.

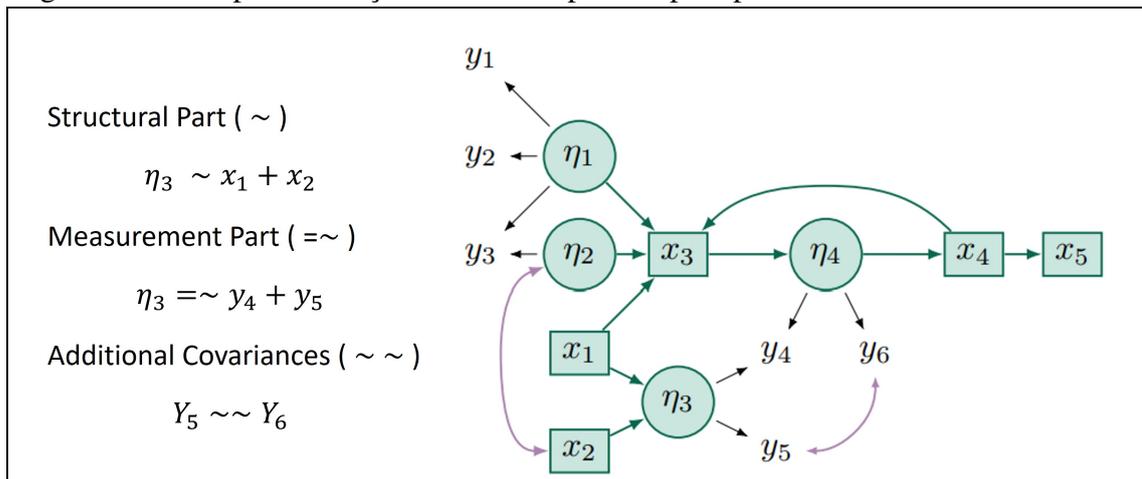
A ideia é estruturar uma relação causal no modelo da Equação 3.7. Desta forma, a sintaxe do SEMOPY divide o modelo em três partes (Georgy; Anna, 18/02/2020):

- **Parte estrutural:** Na qual descreve a relação das variáveis dependentes com as variáveis latentes ou independentes e seus coeficientes de regressão;
- **Parte de medição:** Na qual é definida a relação entre as variáveis latentes e as variáveis medidas;
- **Parte de covariâncias:** Na qual são especificadas as covariâncias existentes entre pares de variáveis.

Na Figura 7 são mostradas as relações construídas no SEMOPY a partir de uma relação causal. Como pode ser visto, a variável erro não precisa ser descrita de forma explícita, pois é uma medida que é dada no relatório final do SEMOPY.

Além disso, o SEMOPY permite estabelecer várias relações para uma mesma variável de forma individual, sem deixar de interpretá-las em conjunto. No exemplo, η_3 é uma variável latente dependente das variáveis de medição y_4 e y_5 , que possui uma covariância com y_6 , e também, com as variáveis independentes x_4 e x_5 .

Figura 7 – Exemplo de relação causal interpretada pelo pacote SEMOPY.



Fonte: (Georgy; Anna, 18/02/2020)

O pacote SEMOPY possui formas diferentes para abordagem e definição dos parâmetros que devem ser escolhidas de acordo com a situação problema de cada usuário (Georgy; Anna, 18/02/2020):

1. *Unweighted Least Squares* (ULS)
2. *General Least Squares* (GLS)
3. *Wishart Maximum Likelihood* (MLW)

Neste trabalho, foram utilizadas as abordagens GLS e MLW.

Além dos diferentes métodos de abordagem para definição dos coeficientes de regressão, o pacote possui também várias formas de otimização da solução para relação analisada (Georgy; Anna, 18/02/2020):

1. *Sequential Least-Squares Quadratic Programming* (SLSQP)
2. *Broyden-Fletcher-Goldfarb-Shanno* (L-BFGS-B)
3. *Stochastic optimisation method Adam*
4. *Stochastic Nesterov Accelerated Gradient*
5. *Stochastic Gradient Descent*

Para este trabalho, foram utilizados os métodos de otimização SLSQP e L-BFGS-B, sendo o primeiro escolhido por ter apresentado melhor performance quando comparado com o

segundo, de acordo com os índices de avaliação mostrados no Quadro 5.

A literatura recomenda que sejam utilizados vários índices de avaliação da abordagem e da otimização da MEE, sendo assim, este trabalho utiliza como referência o Quadro retirado de Amorim *et al.* (2012) e mostrado no Quadro 5 como métrica de comparação entre os modelos de abordagem e otimização na MEE.

Quadro 5 – Principais índices de avaliação para o ajuste da MEE.

Medida de Ajuste	Critério para Avaliação	Comentários
Qui-Quadrado	$\rho^2 > 0,05$	Com grandes amostras e menores desvios, o ajuste pode ser estatisticamente significativo.
Qui-Quadrado/gl	$\rho^2/gl < 2$: Excelente Ajuste $3 < \rho^2/gl < 5$: Ajuste Bom $\rho^2/gl > 5$: Ajuste Pobre	Compensa para amostras grandes.
Jorskog Sorbom Índice de Bondade	GFI = 1: Ajuste Perfeito GFI > 0,9: Objetivo	Refere-se à proporção da covariância observada que é explicada pela covariância do modelo.
Jorskog Sorbom Índice de Bondade Ajustado	AGFI = 1: Ajuste Perfeito AGFI > 0,9: Objetivo	Inclui ajuste para modelos complexos.
Índice de Comparação do Ajuste de Bentler	CFI = 1: Ajuste Perfeito CFI > 0,9: Objetivo	Corresponde melhor ajuste do modelo aos dados quando as variáveis são independentes.
Índice de Tucker-Lewis	TLI = 1: Ajuste Perfeito TLI > 0,9: Objetivo	Inclui ajuste para modelos complexos.
Raiz do Erro Quadrático Médio de Aproximação	RMSEA < 0,05	Reflete a diferença média entre a covariância observada e a do modelo.
Hoelter's	$n > 200$	Estima o tamanho da amostra suficiente para o ajuste adequado do modelo.

Fonte: (Amorim *et al.*, 2012)

Além dos valores de referência dos índices mostrados no Quadro 5, neste trabalho também foram considerados os valores descritos em Marôco (2021) para os índices *Root-Mean-Square Error of Aproximation* (RMSEA), *Akaike Information Criteria* (AIC) e *Log Likelihood*.

3.4 Modelos de *Machine Learning*

O *Machine Learning* surgiu da necessidade de processar e obter informação útil a partir de uma quantidade significativa de dados brutos. Desta forma, pode-se entender que o *Machine Learning* busca descobrir um modelo matemático, por meio de algoritmos, que expliquem o relacionamento entre os dados, além de estudar formas de automatização de tarefas manuais.

Em *Machine Learning*, o aprendizado é o objetivo principal e possui diversos ramos de estudo, cada um com suas metodologias e aplicações, o que vai diferenciar a metodologia usada são os tipos de dados disponíveis e quais são os objetivos a serem alcançados com o aprendizado.

Sendo assim, pode-se dividir os tipos de aprendizado em: Aprendizado supervisionado, Aprendizado não supervisionado e Aprendizado por reforço (Ignacio, 2021).

3.4.1 *Aprendizado supervisionado*

O aprendizado é realizado de maneira supervisionada quando se tem acesso a uma amostra do problema com informações sobre determinado objeto de estudo, os chamados dados de entrada, e o resultado obtido com essas informações, chamados dados de saída. Em outras palavras, há presença de uma variável resposta para guiar o aprendizado, ou seja, o conjunto de treino é da forma $S = (x_1, y_1), \dots, (x_m, y_m)$, em que para cada ponto de entrada se tem um valor x_i e um valor de saída y_i (Ignacio, 2021).

No processo de treinamento, o algoritmo recebe os valores de entrada, e caso ele não apresente a saída correta, continua-se o treinamento para minimização do erro. Assim, ao longo deste processo, espera-se que o algoritmo cometa cada vez menos erros sobre a amostra de treino disponível, e que isto se reflita em pontos aos quais ele nunca teve acesso (Ignacio, 2021).

3.4.2 *Aprendizado não supervisionado*

O aprendizado é dito não supervisionado quando não se tem acesso a um conjunto de amostras contendo a relação entre pontos de entrada e de saída, e o objetivo do aprendizado consiste em extrair padrões dos dados disponíveis. Nesse caso, existe uma estrutura do espaço de entrada tal que certos padrões ocorrem mais frequentemente do que outros, e quer-se entender esses padrões (Ignacio, 2021).

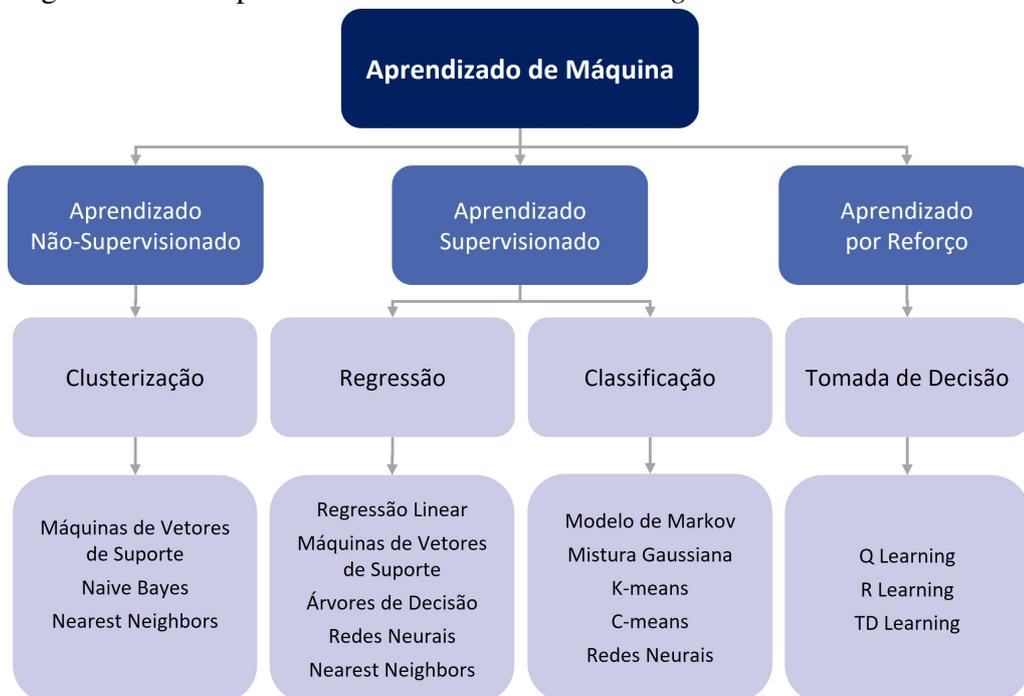
Em Estatística, isso é chamado de estimação da densidade, e um dos métodos de estimação de densidade mais conhecido é a análise de conglomerados, em inglês *clustering*, cujo objetivo é encontrar agrupamentos das entradas e em vista disso, busca-se observar algumas similaridades entre os objetos e incluí-los em grupos apropriados (Ignacio, 2021).

3.4.3 Aprendizado por reforço

O aprendizado por reforço é um paradigma de aprendizado baseado no processo de decisão de *Markov*. Esta técnica consiste em ensinar um agente a interagir com o meio, através de um conjunto finito de ações, a fim de que ele atinja um objetivo definido. Neste contexto, o processo de treinamento constitui-se em aplicar um valor de punição para cada ação realizada pelo agente, de modo que o agente busque encontrar as ações que minimizam a punição recebida (Ignacio, 2021).

Na Figura 8 são mostrados exemplos de modelos de *Machine Learning* para cada um dos tipos de aprendizado.

Figura 8 – Principais modelos de *Machine Learning*



Fonte: Elaborado pela própria autora.

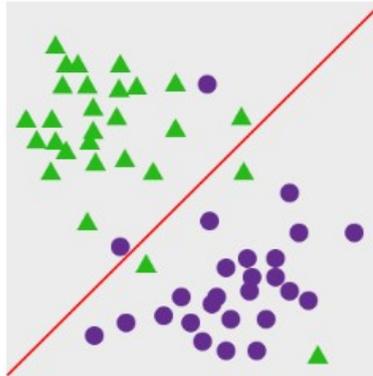
O modelo de aprendizado supervisionado é o mais amplamente utilizado, principalmente em dois tipos de problemas: Qualitativos e Quantitativos.

Variáveis quantitativas são aquelas que assumem valores numéricos, como idade e altura, enquanto que variáveis qualitativas (ou categóricas) assumem seu valor em uma classe dentre um conjunto de classes disponível, como por exemplo o sexo de uma pessoa e a marca de um produto (Ignacio, 2021).

Em geral, se diz que um problema é de classificação se a variável de resposta é qualitativa. Neste caso, a partir do processo de treinamento e dependendo de características

mapeadas previamente, consegue-se categorizar, por exemplo, se uma pessoa X_i possui alto risco de infarto (categoria Y_1) ou baixo risco de infarto (categoria Y_0). Na Figura 9 é exemplificado um conjunto de dados representativo de um problema de classificação, onde os triângulos e bolas representam as pessoas que possuem alto risco de infarto e baixo risco de infarto, respectivamente.

Figura 9 – Exemplo de problema de classificação supervisionado

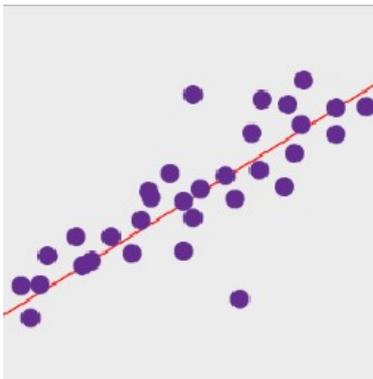


Fonte: (LeanSaude, 2023)

Por outro lado, se estamos interessados em prever um valor quantitativo com a tentativa de estimar uma função contínua desconhecida, tem-se um problema de regressão .

Por exemplo, a previsão de preços de um determinado produto no mercado de ações, para um conjunto de preços $S = x_1, \dots, x_m$, consegue-se obter uma função contínua $F(x)$ conhecida em que se é possível ter uma estimativa dos próximos preços de ações. Na Figura 10, a $F(x)$ é representado por uma reta que segue a mesma tendência do conjunto de preços do exemplo.

Figura 10 – Exemplo de problema de regressão supervisionado



Fonte: (LeanSaude, 2023)

3.4.4 Algoritmos

Para cada um dos problemas de *Machine Learning* existem diversos algoritmos que podem ser utilizados. Entretanto, vale a pena mencionar o conhecido teorema “não existe almoço grátis”, ou seja, não existe um algoritmo de aprendizado que seja superior a todos os demais quando considerados todos os problemas possíveis. A cada problema, os algoritmos disponíveis devem ser experimentados a fim de identificar aqueles que obtêm melhor desempenho através das métricas de avaliação de desempenho (Escovedo; Koshiyama, 2020).

3.4.4.1 Regressão Linear

Formalmente, a Regressão Linear (*Linear Regression* (LR)) modela a relação entre a variável de resposta Y e as variáveis preditoras X , e corresponde ao problema de estimar uma função a partir de pares entrada-saída, considerando que Y pode ser explicado por uma combinação linear de X . Assim, a solução de um problema de regressão consiste em encontrar valores para os coeficientes de regressão de forma que a reta (ou plano/hiperplano) se ajuste aos valores assumidos pelas variáveis no conjunto de dados (Escovedo; Koshiyama, 2020).

A saída do modelo é um valor numérico contínuo que deve ser o mais próximo possível do valor desejado, e a diferença entre esses valores fornece uma medida de erro do algoritmo. Se a equação de regressão aproxima suficientemente bem os dados de treinamento, então ela pode ser usada com novos dados (para os quais não conhecemos o valor de Y) para estimar Y a partir do valor das variáveis X , assumindo uma relação linear entre estas variáveis. Em suma, a Regressão Linear procura pelos coeficientes da reta que minimizam a distância dos objetos à reta (Escovedo; Koshiyama, 2020).

3.4.4.2 Máquina de Vetor de Suporte

O algoritmo Máquina de Vetor de Suporte (*Support Vector Machine* -SVM), é um dos algoritmos mais populares efetivos para problemas de clusterização, a classificação para aprendizados de máquina não supervisionado, além de também poder ser usado para problemas de regressão. Apesar de o treinamento do SVM geralmente ser lento, esses modelos exigem poucos ajustes e tendem a apresentar boa acurácia, conseguindo modelar fronteiras de decisão complexas e não lineares (Escovedo; Koshiyama, 2020).

Resumidamente, o SVM realiza um mapeamento não linear (utilizando funções

kernel) para transformar os dados de treino originais em uma dimensão maior, buscando nesta nova dimensão um hiperplano que separe os dados linearmente de forma ótima. Com um mapeamento apropriado para uma dimensão suficientemente alta, dados de duas classes poderão ser sempre separados por um hiperplano. O SVM encontra este hiperplano usando vetores de suporte (exemplos essenciais para o treinamento) e margens, definidas pelos vetores de suporte (Escovedo; Koshiyama, 2020).

3.4.4.3 *K-ésimo Vizinho mais Próximo*

O algoritmo K-ésimo Vizinho mais Próximo (K-Nearest Neighbors - KNN) é um algoritmo de simples entendimento e que funciona muito bem na prática, podendo ser utilizado tanto para problemas de Classificação quanto para problemas de Regressão. Sua ideia principal é considerar que os exemplos vizinhos são similares ao exemplo cuja informação se deseja inferir. O KNN considera que os registros do conjunto de dados correspondem a pontos no espaço R_n , em que cada atributo corresponde a uma dimensão deste espaço (Escovedo; Koshiyama, 2020).

3.4.4.4 *Floresta aleatória*

A Floresta aleatória (Random Florest-RF) é inspirado na forma de como os humanos tomam decisão e, por este motivo, é um dos modelos mais simples de se entender. Uma das principais vantagens deste algoritmo é a apresentação visual da informação, facilitando o entendimento pelo ser humano. As árvores podem ser usadas para problemas de Classificação (Árvores de Classificação - *Random Forest Classifier* (RFC)) ou Regressão (Árvores de Regressão - *Random Forest Regressor* (RFR)). De forma resumida, uma árvore de decisão usa amostras das características dos dados para criar regras de decisão no formato de árvore, mapeando os dados em um conjunto de regras que podem ser usadas para uma decisão (Escovedo; Koshiyama, 2020).

O RF costuma ter bons resultados e boa interpretabilidade, e pode realizar automaticamente a seleção de variáveis para compor suas estruturas (Escovedo; Koshiyama, 2020).

3.4.4.5 *Perceptron Multicamadas*

O perceptron multicamadas (Multi-layer Perceptrons-MLP) é uma RNA que foi desenvolvida com o objetivo de solucionar os problemas mais complexos, como não-linearidade,

até então não resolvidos pelos outros modelos de RNA (Gusman, 2011).

As RNA são modelos computacionais fundamentados na estrutura neural biológica do cérebro humano, onde os neurônios estão conectados uns aos outros através de sinapses, formando assim uma grande rede de processadores chamada rede neural (Gusman, 2011).

Desta forma, as RNA são constituídas por conjuntos de unidades de processamento conectadas entre si, denominadas neurônios artificiais. Cada neurônio artificial calcula a soma das entradas ponderadas pelos pesos sinápticos, além de um valor constante. O valor encontrado é fornecido a uma função não-linear denominada função de ativação, que por sua vez gera o valor de saída Y do neurônio, que será fornecido a outros neurônios da rede neural artificial (Gusman, 2011).

O diferencial da rede MLP é ser formada por mais de uma camada. Normalmente a arquitetura mais encontrada é de uma camada de entrada, uma ou mais camadas intermediárias e uma camada de saída (Gusman, 2011).

3.4.4.6 Memória longa de curto prazo

As redes de memória longa de curto prazo (Long Short-Term Memory -LSTM) são um tipo especial de rede neural, pois são capazes de aprender conexões de longo prazo de forma recorrente. Dessa maneira, elas têm um incrível poder de predição e funcionam muito bem em uma variada gama de problemas, sendo amplamente usadas atualmente.

A diferença desse modelo para os outros modelos de RNA, é que enquanto nos modelos tradicionais a única informação passada de um instante de tempo para o próximo é uma saída h , correspondente a uma memória de curto prazo, o modelo LSTM, necessita da adição de uma memória de longo prazo. Em síntese, a arquitetura de LSTM consiste em um conjunto de células de memória conectadas de forma recorrente.

Este modelo é bem adequado para classificar, processar e prever séries temporais com intervalos de tempo de duração desconhecida.

3.4.5 Métricas de avaliação de desempenho

Após a criação de um modelo, é necessário analisá-lo em termos de seu desempenho. No caso da regressão, as métricas utilizadas são, entre outras, *R-Squared* (R^2), MAE, MAPE, *Mean Squared Error* (MSE), RMSE, sendo que uma métrica não exclui a outra, mas são complementares, e o uso de cada uma depende da natureza do problema.

Os erros médios quadráticos e absolutos captam melhor as tendências, enquanto o erro mediano absoluto é resistente a anomalias (Pascual, 2018).

O erro médio quadrático tende a ter erros mais uniformes e é a estratégia mais utilizada para avaliar regressões (Pascual, 2018).

3.4.5.1 R-Quadrado

Amplamente usado em modelos de regressões o R-Quadrado (R^2), ou Coeficiente de Determinação, é uma métrica que visa expressar a quantidade da variância dos dados que é explicada pelo modelo construído. Em outras palavras, essa medida calcula qual a porcentagem da variância que pode ser prevista pelo modelo de regressão e, portanto, nos diz o quão “próximo” as medidas reais estão do nosso modelo (Pascual, 2018).

O valor do R^2 varia de 0 a 1, quanto mais ajustado o modelo, maior é o valor do R^2 , geralmente é representado em porcentagem e é calculado a partir da Equação 3.8.

$$R^2 = 1 - \frac{\text{Variância Residual}}{\text{Variância Total}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.8)$$

Em que:

\hat{y} = Valor predito

y_i = Valor real

\bar{y} = Valor médio das amostras

n = Número de amostras

3.4.5.2 Erro Quadrático Médio

É uma das métricas mais utilizadas, o erro quadrático médio (*Mean Squared Error-MSE*) consiste na média do erro das previsões ao quadrado. Em outras palavras, pega-se a diferença entre o valor predito pelo modelo e o valor real, eleva-se o resultado ao quadrado, faz-se a mesma coisa com todos os outros pontos, soma-os, e divide-se pelo número de elementos preditos. Um ponto negativo do uso dessa métrica é sua falta de interpretabilidade direta, uma vez que, para a predição de valores de unidade u , a unidade do MSE seria u^2 (Pascual, 2018).

Essa métrica apresenta valor mínimo 0 e sem valor máximo, quanto maior esse número, pior o modelo e pode ser descrito pela Equação 3.9.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2 \quad (3.9)$$

Em que:

\hat{y} = Valor predito

y_i = Valor real

n = Número de amostras

3.4.5.3 Raiz do erro quadrático médio

Tendo em vista a diferença de unidades da métrica no MSE, a raiz do erro quadrático médio (*Root Mean Square Error*-RMSE) entra como uma forma de melhorar a interpretabilidade da métrica, acertando a unidade. Entretanto, essa medida, assim como o MSE, penaliza predições muito distantes dos valores reais (Pascual, 2018).

Da mesma forma que a métrica MSE, essa métrica apresenta valor mínimo 0 e sem valor máximo, quanto maior esse número, pior o modelo e pode ser descrito pela Equação 3.10.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y} - y_i)^2} \quad (3.10)$$

Em que:

\hat{y} = Valor predito

y_i = Valor real

n = Número de amostras

3.4.5.4 Erro Absoluto Médio

O erro absoluto médio (*Mean Absolute Error*-MAE) consiste na média das distâncias entre valores preditos e reais. Diferentemente do MSE, sua interpretação é mais intuitiva, com a mesma unidade dos valores trabalhados. É uma métrica sólida para modelos que devem prever muitos dados ou dados sazonais.

Essa métrica apresenta valor mínimo 0 e não apresenta valor máximo, quanto maior esse número, pior o modelo e pode ser encontrada por meio da Equação 3.11.

$$MAE = \frac{1}{n} \sum_{i=1}^n |(\hat{y} - y_i)| \quad (3.11)$$

Em que:

\hat{y} = Valor predito

y_i = Valor real

n = Número de amostras

3.4.5.5 Erro Percentual Absoluto Médio

Em contraste com as métricas anteriores, essa medida exprime uma porcentagem, obtida através da divisão da diferença entre predito (\hat{y}) e real pelo valor real (y). Por se tratar de uma porcentagem, essa métrica torna-se extremamente intuitiva, tanto para a interpretação do programador, quanto para a comunicação de resultados com pessoas sem conhecimento técnico (Pascual, 2018).

O erro percentual absoluto médio (Mean Absolute Percentage Error-MAPE), assim como outras métricas, indica que quanto menor o valor, mais preciso é o modelo de regressão.

Essa métrica também apresenta valor mínimo 0 e não apresenta valor máximo e pode ser encontrada por meio da Equação 3.12.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}}{y_i} \right| \quad (3.12)$$

Em que:

\hat{y} = Valor predito

y_i = Valor real

n = Número de amostras

4 METODOLOGIA

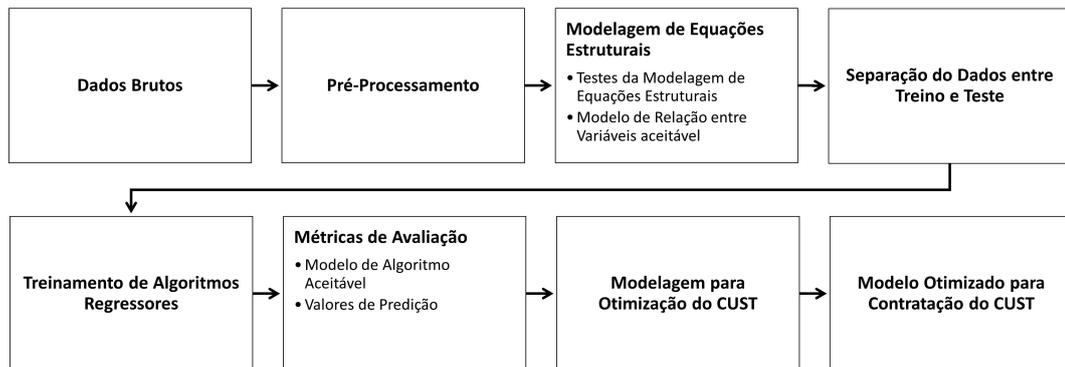
A metodologia proposta neste trabalho almeja minimizar o custo associado ao CUST por meio da previsão dos valores de MUST utilizando algoritmos regressores combinados com o estudo prévio dos fatores que influenciam a demanda de energia elétrica utilizando MEE.

Conforme é mostrado no diagrama da Figura 11, inicialmente, foram selecionados dados brutos de demanda de energia elétrica e variáveis supostamente correlacionadas. Tais dados, depois de pré-processados foram modelados e avaliados em cenários diferentes para estabelecer um modelo de equação estrutural satisfatório.

Com base na relação de dados pré-estabelecida por meio da MEE, os dados foram divididos em treino e teste e aplicados em modelos de algoritmos regressores de acordo com as técnicas pré-selecionadas. As métricas foram calculadas em cada técnica e comparadas para definição do modelo regressor mais aceitável.

A partir dos valores de MUST previstos, foi implementada uma função de otimização para o CUST, onde se espera alcançar o menor custo associado para contratação de acordo com valores futuros de MUST encontrados e com isso, uma tomada de decisão estratégica pela área responsável de contratação da empresa de distribuição de energia.

Figura 11 – Diagrama da metodologia empregada.



Fonte: Elaborado pela própria autora.

4.1 Bases de Dados

A base de dados utilizada neste trabalho foi construída a partir de seis fontes de dados distintas:

1. **Plataforma Integrada de Medição (PIM):** Tem acesso ao histórico dos valores de MUST medidos e de geração, ambos em megawatt, em cada ponto de conexão da concessionária

local de energia. Os dados são integralizados a cada 15min;

2. **Instituto Nacional de Meteorologia (INMET):** Como foi verificado através da revisão sistemática, existe um número significativo de pesquisas que evidenciam uma forte relação entre a demanda de energia elétrica e variáveis meteorológicas, como precipitação, pressão atmosférica e outros. Sendo assim, foi utilizado o histórico de dados com fatores meteorológicos do INMET integralizados a cada hora;
3. **Instituto Brasileiro de Geografia e Estatística (IBGE):** Tendo em vista que este trabalho visa a otimização do custo associado a um contrato com alto teor financeiro, foram utilizados os dados históricos mensais do IBGE para construção da base de dados com os valores do Índice Nacional de Preços ao Consumidor Amplo (IPCA), o qual é um dos índices de inflação mais tradicionais e importantes do Brasil, e que mede a variação dos preços de um conjunto de produtos e serviços vendidos no varejo e consumidos pelas famílias brasileiras;
4. **Instituto Brasileiro de Economia (IBRE):** Sob a mesma justificativa do uso do IPCA e como outra opção de índice econômico, foi extraído o histórico dos valores mensais do Índice Geral de Preços – Mercado (IGP-M) da fonte de dados do IBRE e adicionado também à base de dados deste trabalho. Apesar de ambos os índices econômicos possuírem o mesmo objetivo de observar tendências de inflação, diferentemente do IPCA, que considera, principalmente, a variação de preços para o consumidor final, o IGP-M considera também a variação de preços nos estágios de produção;
5. **CUST de anos anteriores:** Retirado do histórico de valores contratados por ponto de conexão em cada semestre;
6. **Resoluções Homologatórias ANEEL:** Retirado do histórico de valores de TUST por ponto de conexão em cada semestre.

Para simplificar a construção do modelo foram adotadas algumas premissas.

Normalmente, uma distribuidora de energia elétrica possui diversos pontos de conexão com a rede básica, entretanto os pontos de conexão são muitas vezes separados geograficamente de forma significativa e por isso apresentam diferentes características, por exemplo: existe locais em que a rede pode receber mais transferências de carga do que em outras regiões onde a demanda de energia elétrica possui um crescimento mais consolidado.

Neste trabalho é apresentado um modelo piloto que leva em consideração dados de uma rede de subtransmissão real da concessionária local e que representa uma região cuja série

temporal dos dados de geração e demanda é bem definida e sem influência de transferências ou inserção de novas fontes de energia elétrica no período avaliado. A previsão de valores de MUST com transferências de carga entre as regiões com pontos de conexão distintos é objetivo de trabalhos futuros e, portanto, está fora do escopo deste trabalho.

A região escolhida para a construção do projeto piloto atende uma área extremamente industrial, são mais de doze clientes significativos, além disso possui mais de uma geração centralizada conectada no sistema de subtransmissão da distribuidora, a região abrange uma área litorânea e com diversos pontos turísticos. O perfil de demanda de energia elétrica apresenta seus picos nos meses de março e dezembro, refletindo os maiores feriados do ano brasileiro, carnaval e ano novo.

Os dados foram extraídos para o período dos anos de 2018 a 2022, e com intervalos conforme a disponibilidade das fontes dos dados. Foi a partir de 2018 que a região passou a ser atendida de forma concreta pelo ponto de conexão escolhido e assim os dados de demanda já levam em consideração as gerações locais e o seu crescimento é apenas da ordem natural, aumento de demanda de clientes, novos clientes, expansão de fábricas e outros.

Durante a pandemia do COVID-19, a rede escolhida apresentou o crescimento de demanda constante. Tal observação pode ser feita, pois neste trabalho utilizam-se dados da rede em alta tensão, onde as cargas apresentam uma demanda contratada com perfis contínuos apesar da pandemia e a rede é mais estável às mudanças no perfil típico de cargas conectadas em baixa tensão.

No Quadro 6 é mostrado um resumo dos parâmetros utilizados para construção da base de dados e sua respectiva codificação usada na implementação do algoritmo, bem como o intervalo de tempo e a fonte da extração.

Quadro 6 – Resumo da extração dos dados por fonte.

Codificação	Fontes de Dados	Dados	Intervalo
D	PIM	Demanda Medida	15min
G		Geração Medida	
Hora		Hora	Exato
Ano		Ano	
Mes		Mês	
Dia		Dia	
FM1		INMET	Precipitação
FM2	Pressão atmosférica		
FM3	Radiação		
FM4	Temperatura de ar		
FM5	Temperatura do ponto de orvalho		
FM6	Umidade		
FM7	Velocidade do vento		
I1	IBGE	IPCA	Mensal
I2	IBRE	IGP-M	
C	CUST anteriores	Contratado	Semestral
Tx	Resoluções Homologatórias ANEEL	TUST	

Fonte: Elaborado pela própria autora.

4.2 Pré-Processamento dos Dados

A fase de pré-processamento pode ser dividida em limpeza dos dados, transformação dos dados e redução ou ampliação dos dados.

Inicialmente, na etapa de limpeza dos dados, foram removidos todos os registros contendo valores inválidos ou zerados.

Com todos os valores de registros válidos, optou-se pela repetição dos dados nas séries a fim de nivelar os intervalos de dados entre as variáveis. Neste trabalho, não foram utilizadas técnicas de preenchimento de dados no nivelamento de intervalos de tempo.

Desta forma, as séries com intervalos de extração de dados a cada 15 minutos tiveram seus dados mantidos, enquanto que as demais séries tiveram seus dados repetidos, de forma que a base de dados ficou com os intervalos entre os dados de 15 minutos cada.

Este período foi escolhido devido ser o mesmo da apuração mensal do CUST, ou seja, os valores de MUST avaliados para o cálculo financeiro ocorre também a cada 15 minutos como pode ser visto na Seção 3.1.

A última etapa de transformação dos dados foi dividida em duas partes: Transformação dos dados de tempo e normalização da base de dados.

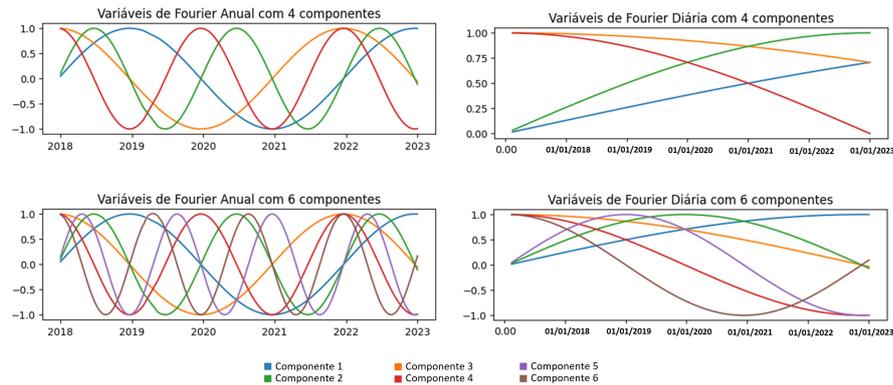
4.2.1 Transformação dos dados de tempo

Para a utilização dos dados temporais de dia e mês no modelo de previsão optou-se por transformar os dados discretos de tempo em dados contínuos por meio da Transformada de *Fourier* diárias ou anuais. Com isso, evita-se que o modelo de previsão interprete o tempo como um dado discretizado quando na realidade é algo contínuo.

Desta forma, foi aplicada a Transformada de *Fourier* nos dados de tempo e calculadas as variáveis anuais e as variáveis diárias, ambas com 4 e 6 componentes, conforme é mostrado na Figura 12.

Nas simulações foram obtidos os melhores resultados quando as variáveis de *Fourier* diárias com 6 componentes foram utilizadas no modelo.

Figura 12 – Transformadas de *Fourier* dos dados de tempo.



Fonte: Elaborado pela própria autora.

4.2.2 Normalização da base de dados

A normalização dos dados é necessária quando se tem parâmetros com intervalos e escalas muito diferentes, como é o caso deste trabalho.

Sendo assim, com o objetivo de alterar os valores das colunas numéricas no conjunto de dados para uma escala comum, sem distorcer as diferenças nos intervalos de valores, utiliza-se a técnica chamada de *Min-max Scaler* da biblioteca do *Scikit-learn* em Python.

A técnica baseia-se na Equação 4.1.

$$X_{MinMax} = \frac{X - X_{min}}{X_{max} - X_{min}}, \quad (4.1)$$

Em que:

X_{MinMax} = Vetor com escala entre 0 (zero) e 1 (um);

X = Vetor original;

X_{min} = O menor dos elementos do vetor X ;

X_{max} = O maior dos elementos do vetor X .

A normalização garante a integridade dos dados, evita que informações sem sentido sejam inseridas, permite organizar e dividir as tabelas da forma mais eficiente possível, diminuindo a redundância e permitindo a evolução do banco de dados com o mínimo de efeito colateral.

4.3 Modelagem de Equações Estruturais

Como foi dito na Seção 3.3, com a Modelagem de Equações Estruturais - MEE, é possível testar um modelo de relações entre variáveis que inicialmente foi determinado de forma

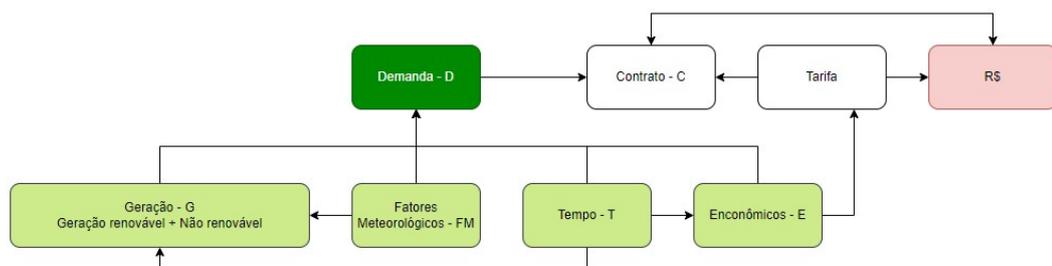
intuitiva.

Desta forma, os dados foram classificados em:

- **Fatores Meteorológicos:** Dados de precipitação, pressão atmosférica, radiação, temperatura do ar, temperatura do ponto de orvalho, umidade e velocidade do vento;
- **Geração:** Dados medidos das gerações conectadas ao sistema de subtransmissão;
- **Demanda:** Dados medidos de demanda;
- **Tempo:** Dados do ano, mês e dia;
- **Índices Econômicos:** Dados de IGP-M e IPCA.

Com a classificação dos dados foi possível ser construído um modelo intuitivo de relação causal conforme é mostrado na Figura 13.

Figura 13 – Modelo de relação causal entre variáveis e demanda de energia elétrica.



Fonte: Elaborado pela própria autora.

O modelo de relação causal, mostrado na Figura 13, serve como a versão inicial do modelo a ser testado no pacote SEMOPY. Em verde, pode ser visto a demanda e as variáveis que a influenciam, bem como a relação que os valores de demanda tem com o contrato e seu valor financeiro, nas cores branca e vermelha.

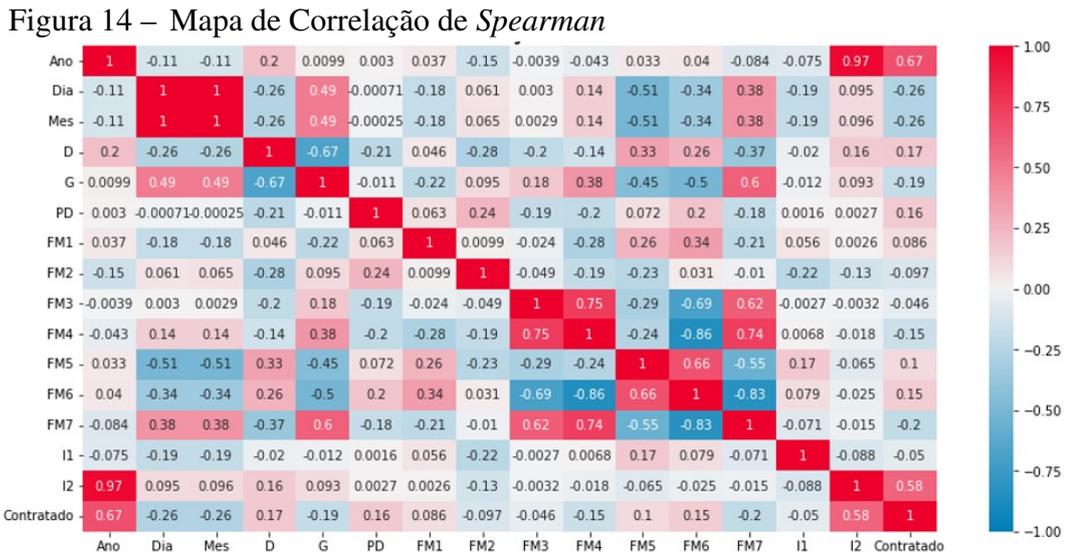
O que diferencia esta dissertação dos demais estudos apresentados no Capítulo 2, é que nos estudos apresentados são utilizados os dados de demanda de energia elétrica com as variáveis de forma independente e sem a apresentação de um estudo realizado previamente comprovando a existência de influência entre os dados selecionados.

Por exemplo, em Jawad *et al.* (2020) são desenvolvidos modelos de previsão que correlacionam os dados de demanda de energia elétrica apenas com os fatores meteorológicos, enquanto que neste trabalho é utilizado um conjunto de fatores meteorológicos, como também dados de geração, tempo e índices econômicos cuja correlações estão embasada pelo estudo do uso da MEE.

Para definição de um modelo inicial no pacote SEMOPY, além de definir uma

relação causal intuitiva, também é necessário informar as possíveis correlações existentes entre as variáveis que influenciam a demanda. Por exemplo, as medições de geração estão intimamente relacionadas aos dados de precipitação e velocidade do vento.

As correlações dessas variáveis foram baseadas no Mapa de Correlação de *Spearman* mostrado na Figura 14 e usando os códigos do Quadro 6.



Fonte: Elaborado pela própria autora.

Portanto observa-se que:

- **Coefficientes iguais ou próximos a zero:** Não há associação entre as duas variáveis analisadas;
- **Coefficientes com valores maior que 0:** Há uma associação positiva entre as variáveis analisadas, isto é, à medida que o valor de uma variável aumenta, o mesmo acontece com o valor da outra variável;
- **Coefficientes com valores menor que 0:** Há uma associação negativa entre as variáveis analisadas, isto é, à medida que o valor de uma variável aumenta, o inverso acontece com o valor da outra variável.

A partir da interpretação do mapa de correlação de *Spearman*, foram definidas as correlações entre as variáveis a serem utilizadas no modelo testado inicialmente no pacote SEMOPY. A partir dos resultados da MEE, saída da simulação desse modelo inicial, foram realizadas modificações no modelo, testando e simulando alternativas da versão inicial de acordo com os índices de avaliação do modelo mostrados no Quadro 5.

No Quadro 7 são mostradas as correlações entre variáveis que foram utilizadas no modelo testado.

Quadro 7 – Correlações entre variáveis que foram utilizadas no modelo testado.

Geração - G	Precipitação - FM1	Radiação - FM3	Temperatura do ar - FM4	Temperatura do Ponto de Orvalho - FM5	Umidade - FM6
Temperatura do ar - FM4	Umidade - FM6	Temperatura do ar - FM4	Temperatura do Ponto de Orvalho - FM5	Umidade - FM6	Velocidade do Vento - FM7
Temperatura do Ponto de Orvalho - FM5	Velocidade do Vento - FM7	Temperatura do Ponto de Orvalho - FM5	Umidade - FM6	Velocidade do Vento - FM7	
Umidade - FM6		Umidade - FM6	Velocidade do Vento - FM7		
Velocidade do Vento - FM7		Velocidade do Vento - FM7			

Fonte: Elaborado pela própria autora.

Após dezesseis testes realizados foi definido o modelo de equação estrutural, de acordo com os índices de avaliação do Quadro 5, que representa bem as influências que as variáveis exercem na demanda de energia elétrica. O modelo final definido está descrito e apresentado no Capítulo 5.

4.4 Algoritmos de Regressão

Após a análise da correlação das variáveis por meio da MEE, é feita uma avaliação de desempenho dos algoritmos de regressão, cujos resultados serão apresentados no Capítulo 5. Neste trabalho, foram utilizados os algoritmos LSTM, MLP, MLR, KNN e o RFR.

Além desses algoritmos, também foi testado o SVR, contudo o número de amostras utilizado neste trabalho é consideravelmente grande e o algoritmo apresenta uma característica de crescer cubicamente em relação ao número de amostras e por isso não converge para este caso específico.

Os parâmetros utilizados nos algoritmos que apresentaram melhores métricas são mostrados na Tabela 1. Para que fossem definidos os parâmetros de cada algoritmo adotou-se a técnica de *Grid Search*.

Tabela 1 – Parâmetros utilizados nos algoritmos.

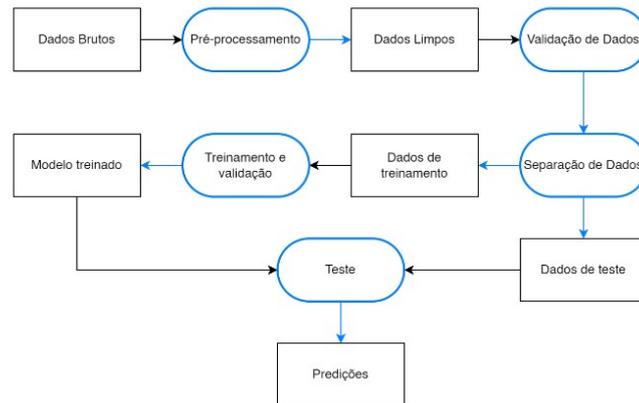
ALGORITMO	PARÂMETROS	VALORES UTILIZADOS
Random Forest Regressor	n_estimators	[50]
	max_features	[sqrt]
	min_samples_leaf	[2]
	min_samples_split	[8]
	bootstrap	[true]
Multilayer Perceptron	hidden_layer_sizes	[10,5,5]
	max_iter	[400]
	alpha	[0.01]
	tol	[0.000001]
	shuffle	[true]
	power_t	[1]
	learning_rate_init	[0.1]
	solver	[adam]
	activation	[tanh]
learning_rate	[constant]	
LSTM	units	[26]
	batch_size	[8]
	epochs	[5]
KNN Regression	n_neighbors	[100]

Fonte: Elaborado pela própria autora.

Em todos os algoritmos de regressão foram utilizados os dados de janeiro de 2018 a setembro de 2021 para treinamento e os dados de outubro de 2021 a dezembro de 2022 para teste, ou seja, o modelo deve prever um ano e três meses de dados de demanda de energia elétrica.

A Figura 15 mostra o *pipeline* com a sequência de etapas executadas na aplicação dos algoritmos regressores utilizadas nesta dissertação.

Figura 15 – *Pipeline* da aplicação dos algoritmos regressores

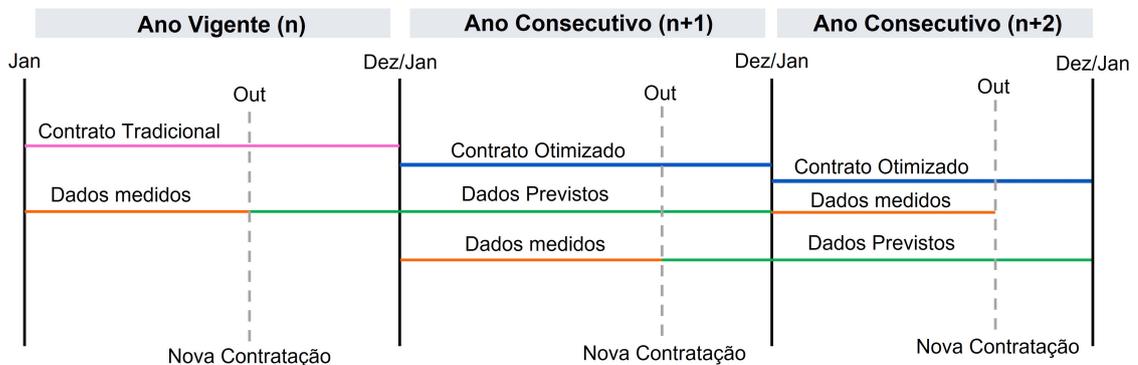


Fonte: Pipeline de Machine Learning - Site Dataside 2023

Os intervalos para treino e teste foram escolhidos conforme o período de contratação do CUST. Como a contratação é realizada em outubro do ano vigente para iniciar em janeiro do ano seguinte, conforme pode ser visto na Seção 3.1, essa previsão de um ano e três meses permite o usuário escolher estrategicamente o valor a ser contratado em seguida.

Na Figura 16, é ilustrada uma linha do tempo para exemplificar o período de contratação e a utilização dos dados de previsão e medição.

Figura 16 – Linha do tempo de contratação



Fonte: Elaborado pela própria autora.

Em rosa pode ser visto o contrato tradicional vigente do ano *n*. Neste contrato, as

distribuidoras de energia elétrica se baseiam no histórico de dados medidos, mostrados na cor laranja, para a contratação que acontece em outubro e se inicia no ano consecutivo $n + 1$.

A principal diferença para a proposta deste trabalho, é que ao invés de se basear nos dados medidos até outubro do ano vigente n para contratação, são utilizados os dados de um ano e três meses previstos através de algoritmos inteligentes, cor verde, para definição de um contrato otimizado iniciado no ano consecutivo $n + 1$, mostrado pela cor azul.

As métricas de avaliação foram aquelas apresentadas na Seção 3.4.5, o ambiente de desenvolvimento foi o *Google Colab* e a linguagem de programação utilizada foi o Python.

4.5 Otimização do CUST

Depois de definido o modelo com a melhor avaliação de performance, cria-se a função do custo associado ao CUST, mostrada na Equação 4.2, com o cálculo de todas as parcelas previstas e descritas na Seção 3.2.

$$Custo_{ano} = \min \left[\sum_{i=1}^{i=12} (EUST_{ponta} + ADEUST_{ponta} + PIU_{ponta}) + PIS_{ponta} + \sum_{i=1}^{i=12} (EUST_{foraponta} + ADEUST_{foraponta} + PIU_{foraponta}) + PIS_{foraponta} \right] \quad (4.2)$$

Com a função determinada, é feito o cálculo do custo associado ao CUST nos seguintes cenários para efeito de análise e comparação:

1. Valores de demanda medidos e valores contratados no CUST vigente;
2. Valores de demanda previstos pelo algoritmo de melhor performance e valores que devem ser contratados no CUST para minimizar o custo total, refletindo a otimização do CUST;
3. Valores de demanda medidos e valores que devem ser contratados no CUST para minimizar o custo total, refletindo a otimização do CUST diante das medições reais.

Para otimização dos valores de contratação foi utilizada a função *minimize* do pacote em Python chamado *optimize*. Como entrada da função foram colocados os valores de demanda previstos e como saída tem os valores ótimos calculados nos horários de ponta e fora ponta para que se obtenha o mínimo custo financeiro associado ao CUST e descrito pela Equação 4.2.

Desta forma, consegue-se fazer um comparativo entre o valor do custo total associado à forma tradicional de como é feita a contratação com o custo utilizando os valores de demanda preditos e otimização de contratação proposta, que é o objetivo final deste trabalho.

5 RESULTADOS

Neste Capítulo são apresentados e discutidos os resultados obtidos pela aplicação da metodologia descrita no Capítulo 4.

Inicialmente, foi feita uma análise exploratória e realizado o pré-processamento da base de dados. Em seguida, foi definida uma relação causal inicial intuitiva para teste e abordagem de desenvolvimento de modelos utilizando a MEE a fim de se comprovar as influências que as variáveis exercem na demanda de energia elétrica e auxiliar na qualidade dos algoritmos de previsão.

Com um modelo de equação estrutural aceitável, foi avaliado o desempenho de algoritmos de regressão para previsão dos dados de demanda de energia elétrica a serem utilizados para contratação do CUST.

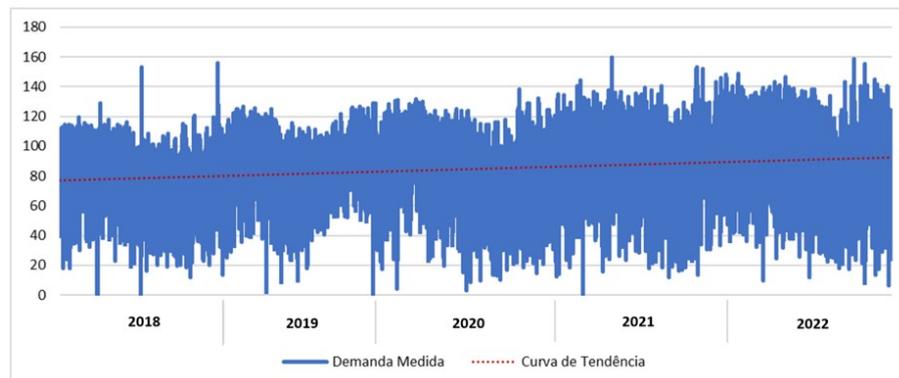
Após realizada a comparação entre modelos e alcançado uma previsão de dados aceitável, foi possível comparar a forma atual de contratação do CUST e desenvolver uma estratégia de otimização financeira do CUST.

5.1 Análise Exploratória dos Dados

Visando um melhor entendimento do comportamento dos dados, bem como da interpretação dos resultados, foi feita uma análise exploratória da base de dados.

Como pode ser visto na Figura 17, os dados de demanda da região escolhida apresentam uma curva de tendência ascendente que segue aumentando para o período de tempo escolhido.

Figura 17 – Curva de tendência dos dados de demanda medida no período escolhido.



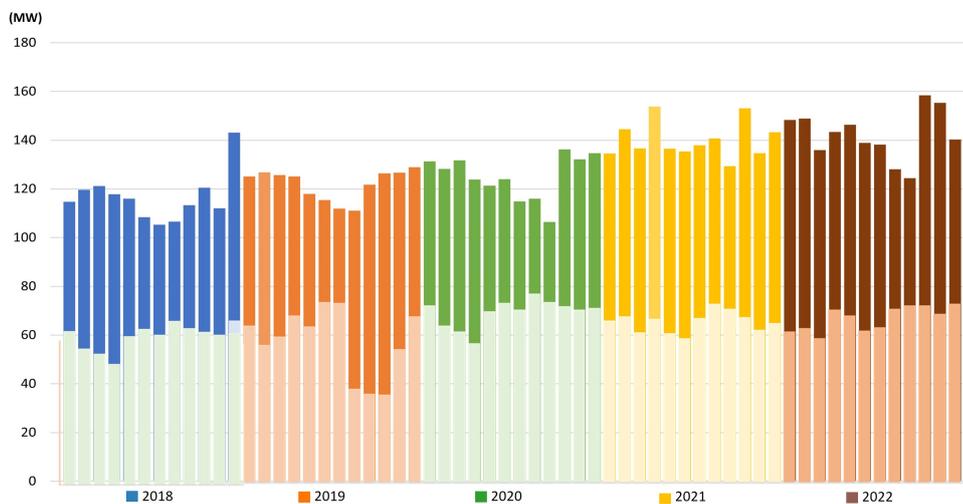
Fonte: Elaborado pela própria autora.

Utilizando como base a demanda máxima medida mensal a cada ano, foi possível discretizar a curva mostrada na Figura 17 e assim observar a sazonalidade dos dados de demanda. Portanto, na Figura 18 é evidenciado que, normalmente, nos meses de fevereiro e março do primeiro semestre e nos meses de outubro a dezembro do segundo semestre, há picos de demanda que podem ser justificados pelas zonas turísticas bastante frequentadas na região escolhida nos feriados anuais. Além disso, observa-se que nos meses de agosto e setembro tem-se uma redução significativa dessa demanda medida o que coincide com a baixa temporada.

Na Figura 18 é apresentada a demanda medida quando são levados em consideração os dados de geração máxima medidos mensalmente a cada ano e injetada diretamente na região, mostrados pela cor mais clara. Observa-se que a quantidade gerada permanece uniforme ao longo dos anos e, portanto, deve ser levada em consideração no momento da contratação.

Ressalta-se que caso a geração da própria região não seja considerada no momento de contratação do MUST, os contratos passariam a apresentar valores muito superiores à curva de demanda real e consequentemente com seu custo associado também muito maior.

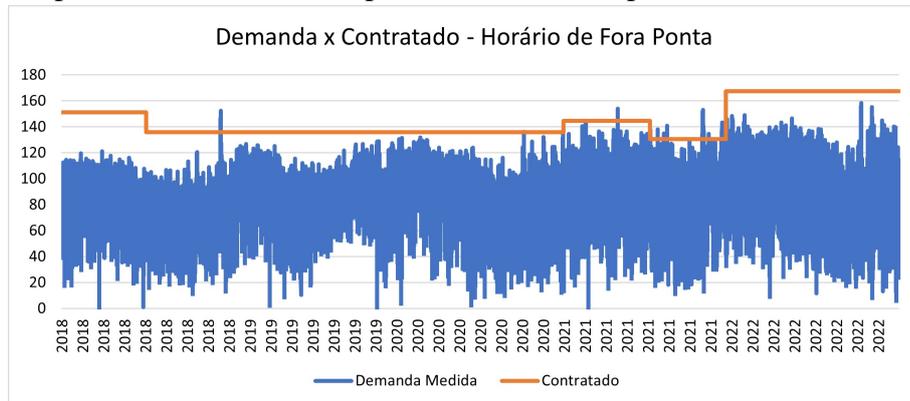
Figura 18 – Dados de máxima demanda e geração medidos e mensais no período de 2018 a 2022.



Fonte: Elaborado pela própria autora.

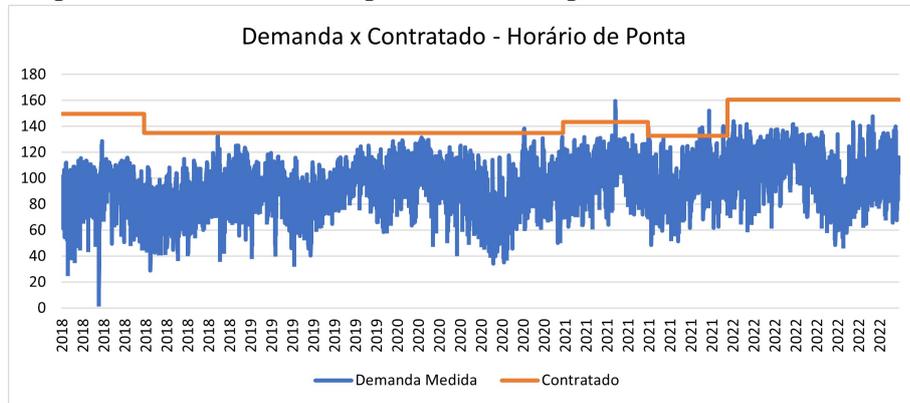
Através das Figuras 19 e 20, é mostrado o ponto de vista de contratação tradicional para os horários de Fora Ponta e Ponta respectivamente. Ou seja, pode ser observado nas figuras que os valores contratados são conservadores e levam em consideração a máxima demanda do ano subsequente visando que não haja multas e penalidades por subcontratação e sobrecontratação, resultando em um custo associado ao CUST recorrentemente maior que o desejado.

Figura 19 – Dados de máxima demanda medidos e valores de contrato no período de 2018 a 2022 para horário de fora ponta.



Fonte: Elaborado pela própria autora.

Figura 20 – Dados de máxima demanda medidos e valores de contrato no período de 2018 a 2022 para horário de ponta.



Fonte: Elaborado pela própria autora.

5.2 Modelo de Equação Estrutural Definido

Como foi dito no Capítulo 4, após ser testado um modelo inicial de uma relação causal intuitiva, foram realizados dezesseis testes com alternativas da versão inicial de acordo com os índices de avaliação do modelo mostrados no Quadro 5 até ser definido o modelo de equação estrutural que melhor representa as influências que as variáveis exercem na demanda de energia elétrica.

No Quadro 8 são apresentadas as variações do modelo simulado no SEMOPY. As células em azul mostram as variáveis utilizadas no modelo testado e as células cinzas as que foram descartadas.

O objetivo dos testes é verificar quais das variáveis representa mais influência na demanda de energia elétrica e a partir daí utiliza-las para determinar um modelo de equação estrutural que represente bem o problema em questão, baseando-se nas métricas do Quadro 5.

Quadro 8 – Variações do modelo de equações estruturais simulados no SEMOPY

ID	Período da Massa de Dados		Período de Previsão		Série de Fourier		Componentes de Fourier		Features											
	4 Anos	5 Anos	15 meses	6 meses	Annual	Diário	4 Comp.	6 Comp.	G	FM1	FM2	FM3	FM4	FM5	FM6	FM7	I1	I2	C	
Teste 1																				
Teste 2																				
Teste 3																				
Teste 4																				
Teste 5																				
Teste 6																				
Teste 7																				
Teste 8																				
Teste 9																				
Teste 10																				
Teste 11																				
Teste 12																				
Teste 13																				
Teste 14																				
Teste 15																				
Teste 16																				

Fonte: Elaborado pela própria autora.

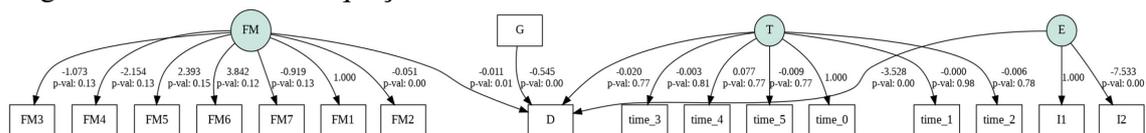
Como resultado, na Figura 21 é mostrado o modelo de equação estrutural que melhor representou as influências que as variáveis exercem na demanda de energia elétrica, sendo, portanto o modelo utilizado nos algoritmos de previsão.

Na Figura 21, pode ser observada a mesma estrutura matricial descrita pela Equação 3.7 apresentada no Capítulo 3. Sendo que:

- Y : A variável dependente de Demanda (D) de energia elétrica;
- x_i : As variáveis independentes de Tempo (T), de Fatores Meteorológicos (FM), Econômicos (E) e de Geração (G);
- β : Os valores são representados pelos números que se apresentam acima de cada seta;
- e_i : Erro representado pelo p -val, a probabilidade de se encontrar os resultados em um mundo no qual a hipótese nula seja verdadeira. Ou seja, é probabilidade de que não haja diferença para demanda de energia elétrica ao se utilizar a variável avaliada em cada seta.

A abordagem utilizada foi a Wishart Maximum Likelihood (MLW) e o método de otimização foi a Programação Quadrática de Mínimos Quadrados Sequenciais (SLSQP). Foram utilizados os dados de 2018 a 2022, a série de Fourier diária com seis componentes em substituição às variáveis relacionadas ao tempo, às variáveis relacionadas aos fatores meteorológicos e às relacionadas aos fatores econômicos.

Figura 21 – Modelo de equação estrutural utilizado.



Fonte: SEMOPY.

Na Tabela 2 são mostrados os índices de avaliação encontrados no modelo definido. Como dito no Capítulo 3, os valores de referência utilizados foram retirados de Amorim *et al.*

(2012) e para os índices RMSEA, AIC e *Log Likelihood* foram considerados os valores descritos em Marôco (2021).

Observa-se que na grande maioria dos índices de avaliação o modelo apresenta bons resultados, exceto para o AGFI e TLI, onde o valores ficaram cerca de 1,22% menor, o que não inviabiliza o modelo. Em relação aos índices AIC e *Log Likelihood*, os mesmos foram avaliados em comparação com os outros modelos.

Tabela 2 – Índices de avaliação do modelo de equação estrutural escolhido.

Índice	Valores do Modelo	Referência
Qui – Quadrado	109.831	>0,05
GFI	0.9198	>0,9
AGFI	0.8890	>0,9
CFI	0.9198	>0,9
TLI	0.8891	>0,9
RMSEA	0.0834	<0,10
AIC	106.6195	Quanto Menor em relação aos outros modelos, melhor é o índice
Log Likelihood	0.6902	Quanto Maior em relação aos outros modelos, melhor é o índice

Fonte: Elaborado pela própria autora.

5.3 Algoritmos de Regressão

Como descrito no Capítulo 4, foram utilizados os algoritmos de regressão LSTM, MLP, MLR, KNN e o RFR neste trabalho.

Em todos os algoritmos de regressão foram utilizados os dados dos anos de 2018 a 2022, sendo o período de treino com os dados de janeiro de 2018 a setembro de 2021, o período de teste com os dados de outubro de 2021 a dezembro de 2022 e com 1 ano e 3 meses de previsão.

A priori não foi utilizada técnica de *grid search* e os resultados das métricas de avaliação nos algoritmos regressores são mostrados na Tabela 3.

Tabela 3 – Métricas de avaliação dos algoritmos regressores sem a técnica de *grid search*.

Métrica	Multiple Linear Regression	Random Forest Regressor	Multilayer Perceptron	LSTM	KNN Regression
R^2	13.50%	61.60%	57.40%	60.90%	59.40%
MAE	0.105	0.508	0.077	0.073	0.524
MAPE	0.252	4.717	0.161	0.146	4.443
MSE	0.018	0.430	0.009	0.008	0.453
RMSE	0.135	0.656	0.095	0.091	0.673

Fonte: Elaborado pela própria autora.

Observa-se que o algoritmo que obteve os piores resultados da métrica R^2 foi o

Multiple Linear Regression e por isso não foi utilizada a técnica de *Grid Search* nesse algoritmo. Para os demais algoritmos, a técnica foi aplicada e os resultados das métricas de avaliação são mostrados na Tabela 4.

Tabela 4 – Métricas de avaliação dos algoritmos regressores com a técnica de *grid search*.

Métrica	Random Forest Regressor	Multilayer Perceptron	LSTM	KNN Regression
R^2	71.1%	66.2%	69.7%	62.0%
MAE	0.062	0.474	0.066	0.071
MAPE	0.127	7.592	0.145	0.152
MSE	0.006	0.378	0.006	0.008
RMSE	0.078	0.615	0.079	0.089

Fonte: Elaborado pela própria autora.

Com os resultados mostrados na Tabela 4 pode ser verificado o ganho da utilização da técnica de *grid search* para especificação dos hiperparâmetros dos algoritmos. Observa-se que em todos os casos os valores da métrica R^2 aumentaram, o que significa que as previsões ficaram mais próximas dos dados reais, sendo os algoritmos com os melhores resultados os RFR e LSTM.

Apesar de também haver uma melhoria na performance das demais métricas de avaliação de todos os algoritmos regressores, com exceção do MLP, o algoritmo que apresentou melhor resultado, foi o RFR. Portanto, o RFR foi escolhido para realizar as previsões de demanda necessárias na otimização do CUST desta dissertação.

Na Figura 22 são mostradas as curvas de máxima demanda prevista pelo algoritmo e máxima demanda medida para o horário Fora Ponta, nas cores amarelo e vermelho respectivamente. As colunas em verde representam a diferença percentual do módulo da diferença entre o valor medido e o valor previsto mostrando o quão distante a máxima mensal prevista de 2022 está da realidade. As mesmas curvas são mostradas para o horário de Ponta na Figura 23.

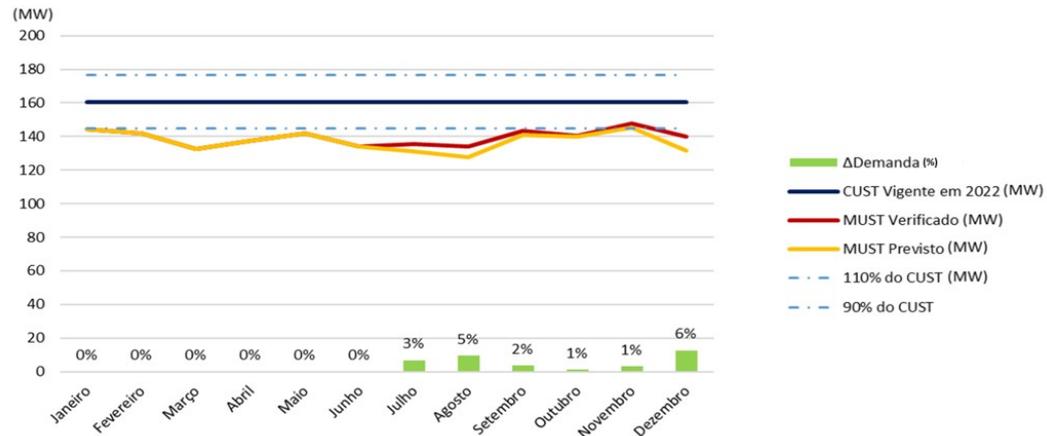
Observa-se que as previsões no segundo semestre do ano ficaram mais distantes do que no primeiro semestre, como consequência desse distanciamento tem-se que os custos financeiros associados na otimização do contrato possuirão um percentual de erro que representará um risco associado a estratégia de contratação proposta. O cálculo desse risco e sua significância financeira são estimados na Seção 5.4.

Figura 22 – Curvas de máxima demanda prevista pelo algoritmo e máxima demanda medida real na Fora Ponta



Fonte: Elaborado pela própria autora.

Figura 23 – Curvas de máxima demanda prevista pelo algoritmo e máxima demanda medida real na Ponta



Fonte: Elaborado pela própria autora.

5.4 Análise da Contratação do Uso do Sistema de Transmissão

Para a realização da análise do valor financeiro associado ao CUST foram simulados três cenários conforme descrito no Capítulo 4 e mostrados a seguir.

- **Cenário 1 ou Cenário Base:** Calculado o valor financeiro associado quando se utiliza os valores de demanda medidos e valores contratados no CUST vigente;
- **Cenário 2:** Calculado o valor financeiro associado quando se utiliza os valores de demanda previstos pelo algoritmo de melhor performance e valores que devem ser contratados no CUST para minimizar o custo total, refletindo uma otimização do CUST;
- **Cenário 3:** Calculado o valor financeiro associado quando se utiliza dos valores de demanda medidos e valores que devem ser contratados no CUST para minimizar o custo total, refletindo uma otimização do CUST diante das medições reais.

- **Cenário 4:** Com o resultado dos três cenários foi possível realizar um quarto cenário realista, em que se utilizou os valores de demanda medidos com a proposta de contratação otimizada, resultado do Cenário 3, e as regras previstas na legislação e descritas no Capítulo 3. Tal cenário possibilita de fato identificar o ganho real na estratégia de contratação otimizada e o risco associado comentado na Seção 5.3.

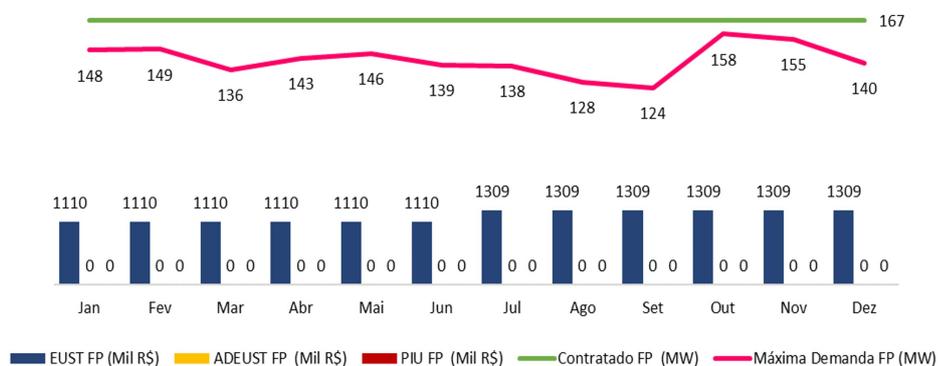
Em todos os cenários foram utilizados os valores de contratação e de demanda máxima mensal Megawatts (os valores financeiros calculados de EUST, ADEUST e PIU referentes ao período são apresentados na unidade de Mil reais).

5.4.1 Cenário 1

No cálculo financeiro para o cenário 1 foram utilizados os valores contratados no CUST vigente do ano de 2022, os valores de demanda máxima mensal medidos em 2022 e valores de tarifa de Rede Básica e de fronteira nos horários de contratação de Ponta e Fora Ponta do mesmo ano.

Nas Figura 24 e 25 são apresentadas as curvas de contratação e demanda, e o cálculo financeiro para o cenário 1 nos horários de fora ponta e ponta respectivamente. As curvas de contratação são mostradas em verde, enquanto as curvas de demanda máxima mensal medida são mostradas na cor rosa. Nas mesmas figuras também é possível verificar os valores de EUST, ADEUST e PIU calculados mensalmente nas cores azul, amarelo e vermelho respectivamente.

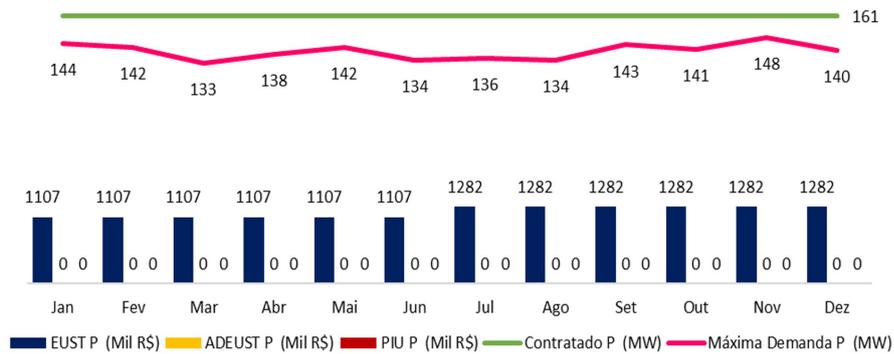
Figura 24 – Curvas de máxima demanda mensal medida e contratada com os respectivos valores de encargos calculados para fora ponta no cenário 1



Fonte: Elaborado pela própria autora.

Como dito nos capítulos anteriores, observa-se que a consequência da atual estratégia de contratação é não ter valores de ADEUST e PIU mensais para fora ponta e ponta.

Figura 25 – Curvas de máxima demanda mensal medida e contratada com os respectivos valores de encargos calculados para ponta no cenário 1

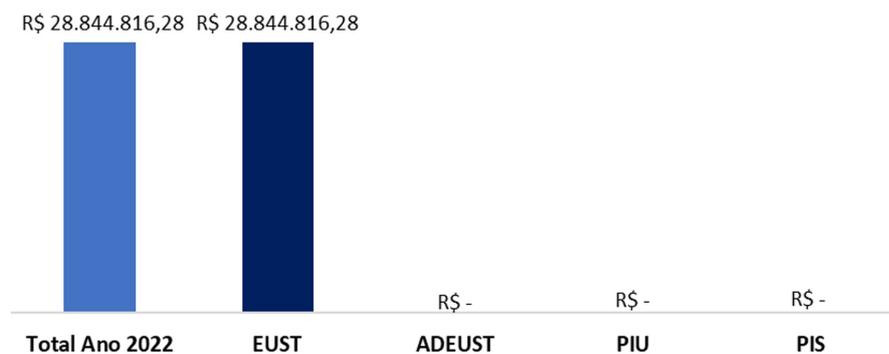


Fonte: Elaborado pela própria autora.

Portanto, pode-se verificar que a curva de contratação em verde é sempre superior à curva de demanda máxima mensal medida, em rosa, durante todo o ano, resultando apenas em valores de EUST mensais.

Para efeito de comparação com os outros cenários, na Figura 26 é mostrado o valor financeiro anual de 2022, em reais, associado ao contrato analisado no cenário 1. Observa-se que o valor total do contrato é igual ao valor total pago de EUST no ano.

Figura 26 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 1



Fonte: Elaborado pela própria autora.

5.4.2 Cenário 2

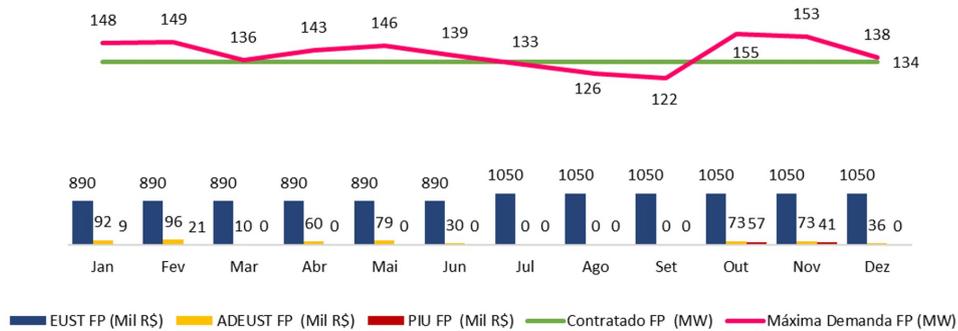
No cálculo financeiro para o cenário 2 foram utilizados os valores de demanda máxima mensal previstos pelo algoritmo para 2022 e os valores de tarifa de rede básica e de fronteira nos horários de contratação de Ponta e Fora Ponta do mesmo ano.

Como dito no Capítulo 4, os valores de contrato otimizado ponta e fora ponta foram

estabelecidos através da função *minimize* do Python que estabelece o valor de contrato ótimo para um menor custo associado ao CUST.

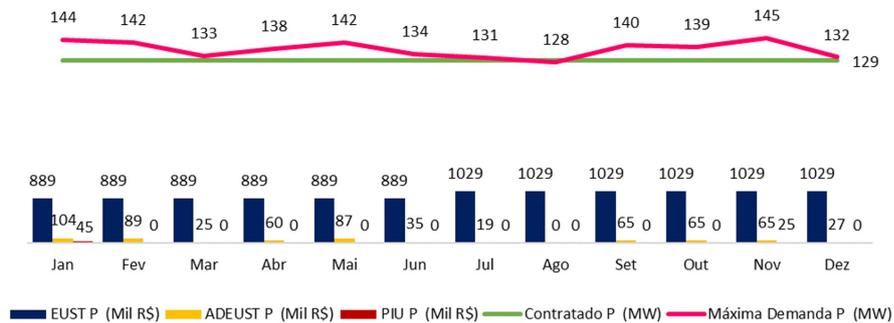
Nas Figura 27 e 28 são mostradas as curvas utilizadas e o cálculo financeiro para o cenário 2 nos horários de fora ponta e ponta respectivamente. As curvas de contratação otimizada são mostradas em verde, enquanto as curvas de demanda máxima mensal prevista são mostradas na cor rosa. Nas mesmas figuras também é possível verificar os valores de EUST, ADEUST e PIU calculados mensalmente nas cores azul, amarelo e vermelho respectivamente.

Figura 27 – Curvas de máxima demanda mensal prevista e contratada otimizada com os respectivos valores de encargos calculados para fora ponta no cenário 2



Fonte: Elaborado pela própria autora.

Figura 28 – Curvas de máxima demanda mensal prevista e contratada otimizada com os respectivos valores de encargos calculados para ponta no cenário 2



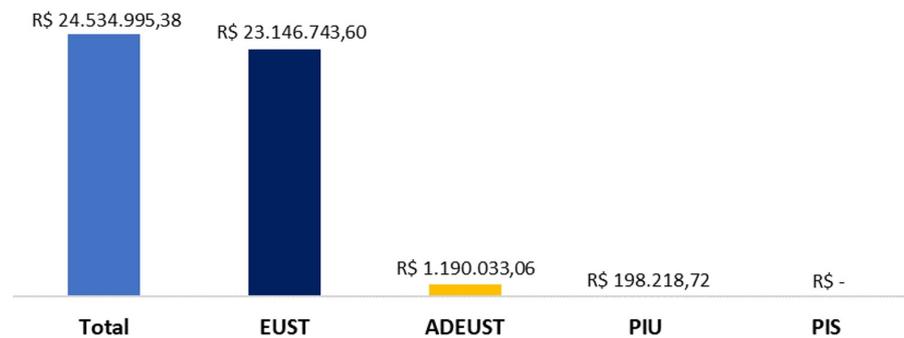
Fonte: Elaborado pela própria autora.

Observa-se que o valor ótimo contratado, resultado do algoritmo, foi de 134,3 MW para fora ponta e 129,0 MW para ponta. Vale destacar que o valor de EUST baixou, o que reflete o valor de contrato mais baixo quando comparado com o contrato tradicional.

Também foram calculados valores de ADEUST devidas de algumas ultrapassagens previstas na curva de demanda máxima. Além disso, o algoritmo previu ultrapassagens de contrato superior a 110% nos meses de janeiro, fevereiro, outubro e novembro para fora ponta e nos meses de janeiro e novembro para ponta, o que significa que para esse valor ótimo contratado, o algoritmo prevê que haja valores de PIU somente nesses meses do ano de 2022.

Assim como no cenário 1, na Figura 29 é mostrado o valor financeiro anual de 2022, em reais, associado ao contrato analisado no cenário 2.

Figura 29 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 2



Fonte: Elaborado pela própria autora.

Observa-se que o valor total do contrato é menor quando comparado com o valor da Figura 26, apesar de haver valores adicionais de ultrapassagens previstos. Portanto, pode-se dizer que se o valor contratado for reduzido para o valor ótimo, resultado do algoritmo, e a curva de demanda prevista se cumprir, é esperado um saldo positivo de 15% referente aos R\$ 4.309.820,90 reais de economia no ano de 2022.

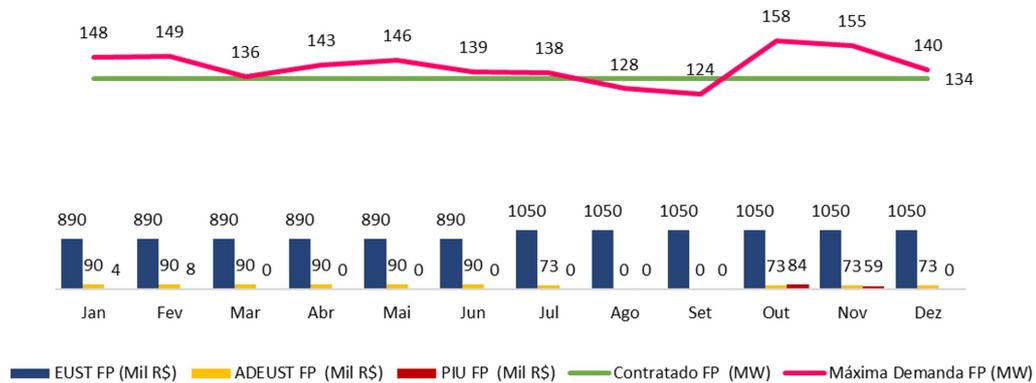
5.4.3 Cenário 3

Como foi dito na Seção 5.3, há uma diferença entre a curva de demanda prevista e a curva de demanda real, sendo assim, este cenário tem o propósito de verificar o cálculo do valor financeiro do CUST quando os valores contratados são reduzidos aos valores otimizados, mesmos do cenário 2, e a curva de demanda é a realizada em 2022.

Desta forma, no cálculo financeiro para o cenário 3 foram utilizados os valores de demanda máxima mensal realizada em 2022, os valores de tarifa de rede básica e de fronteira nos horários de contratação de Ponta e Fora Ponta do mesmo ano e os valores de contratação ótimos encontrados pelo algoritmo.

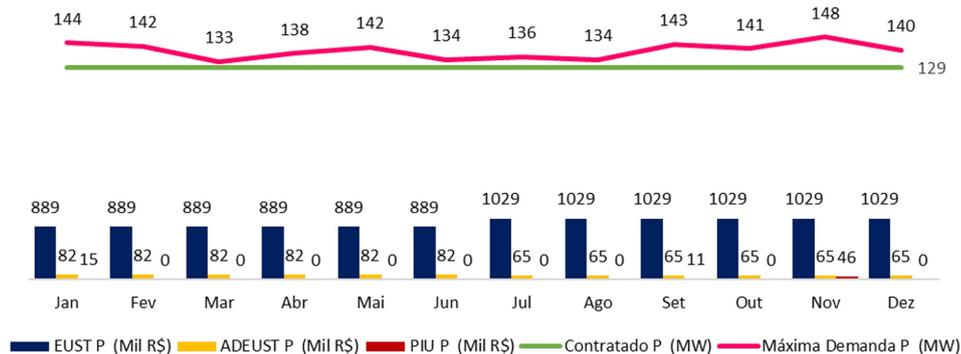
Nas Figura 30 e 31 são mostradas as curvas utilizadas e o cálculo financeiro para o cenário 3 nos horários de fora ponta e ponta respectivamente. As curvas de contratação otimizada são mostradas em verde, enquanto as curvas de demanda máxima mensal realizada são mostradas na cor rosa. Nas mesmas figuras também é possível verificar os valores de EUST, ADEUST e PIU calculados mensalmente nas cores azul, amarelo e vermelho respectivamente.

Figura 30 – Curvas de máxima demanda mensal realizada em 2022 e contratada otimizada com os respectivos valores de encargos calculados para fora ponta no cenário 3



Fonte: Elaborado pela própria autora.

Figura 31 – Curvas de máxima demanda mensal realizada em 2022 e contratada otimizada com os respectivos valores de encargos calculados para ponta no cenário 3

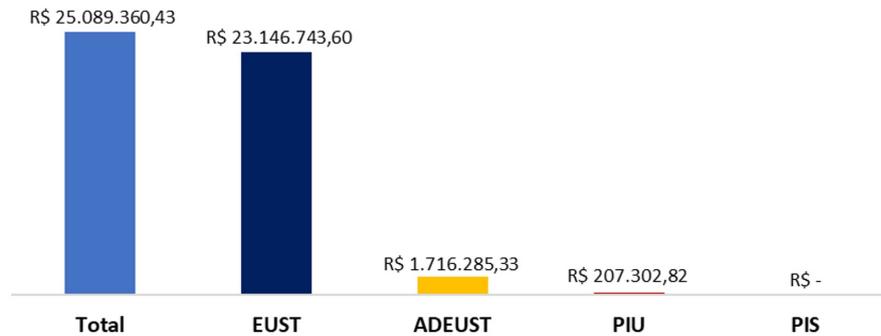


Fonte: Elaborado pela própria autora.

Observa-se que devido a curva de demanda realizada ser maior do que a curva de demanda prevista em alguns meses do ano, há um aumento no valor pago de ADEUST e PIU em ambos os horários de contratação. Em relação ao valor de EUST, ele permanece o mesmo valor quando comparado com o cenário 2 e mais baixo quando comparado com o contrato tradicional do cenário 1.

Assim como nos cenários 1 e 2, na Figura 32 é mostrado o valor financeiro anual de 2022, em reais, associado ao contrato analisado no cenário 3.

Figura 32 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 3



Fonte: Elaborado pela própria autora.

Pode ser verificado através da comparação entre os cenários 2 e 3 que o valor de ADEUST aumentou cerca de 44% e o valor de PIU aumentou cerca de 5% em relação ao valor previsto na Seção 5.4.2. Vale destacar que em relação ao valor total do contrato, o valor do cenário 3 é menor quando comparado com o valor da Figura 26 para o cenário 1, apesar de ser maior do que o valor previsto da Figura 29 para o cenário 2.

Portanto, pode-se dizer que se em 2022 o valor contratado tivesse sido reduzido para o valor ótimo, resultado do algoritmo, haveria uma economia de 13% referente aos R\$ 3.755.455,85, em reais, que poderiam ter sido economizados naquele ano.

5.4.4 Cenário 4

O propósito do cenário 4 é verificar, com base nas regras previstas na legislação e descritas no Capítulo 3, o comportamento do valor financeiro associado à estratégia de contrato real para a curva de demanda máxima prevista e para a curva de demanda máxima realizada.

Além disso, é realizado o cálculo do ganho financeiro real resultado da diferença entre o valor financeiro associado ao contrato tradicional e o valor financeiro associado ao contrato otimizado, bem como o cálculo do risco financeiro, resultado da diferença entre o valor total do custo previsto e o valor total do custo real para 2022.

Conforme especificado na resolução nº 905/2020 da ANEEL e comentado no Capítulo 3, para haver a redução não onerosa, premissa deste trabalho, a redução do valor contratado deve ser realizada em até no máximo 10% do montante previamente informado no contrato para

ambos os semestres de 2022.

O valor ótimo a ser contratado, resultado do algoritmo, é de 134,3 MW para fora ponta e 129,0 MW para ponta, entretanto, o valor do contrato previamente informado é de 167,5 MW para fora ponta e 161,6 MW para ponta, o que significa uma redução de aproximadamente 25% em ambos os horários de contratação.

Sendo assim, para não haver oneração por redução de valor de contrato, seria necessário realizar uma redução de 10% em outubro de 2021 em relação ao valor contratado para o primeiro semestre de 2022 e uma redução de mais 10% para o segundo semestre de 2022, totalizando uma redução total de no máximo 20% para 2022.

Na Tabela 5 é apresentado o resumo dos valores a serem contratados de acordo com a n° 905/2020 da ANEEL e a otimização prevista pelo algoritmo.

Tabela 5 – Valores contratados no CUST por cenário analisado

Semestre/Ano	Contrato Tradicional (MW) Cenário 1		Contrato Otimizado (MW) Cenário 2 e 3		Contrato Otimizado Legal (MW) Cenário 4	
	Ponta	Fora Ponta	Ponta	Fora Ponta	Ponta	Fora Ponta
	1° Semestre/2022	161,6	167,5	129,0	134,3	144,6
2° Semestre/2022	161,6	167,5	129,0	134,3	130,1	135,6

Fonte: Elaborado pela própria autora.

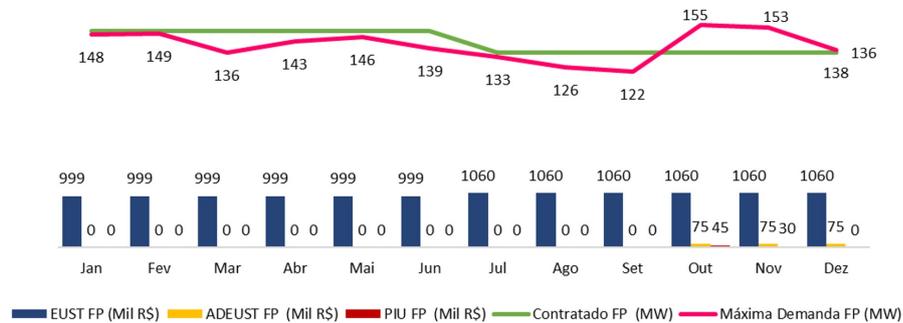
5.4.5 Cenário 4 - Dados Previstos

No cálculo financeiro para o cenário 4 com dados de previsão foram utilizados os valores de demanda máxima mensal prevista em 2022, os valores de tarifa de rede básica e de fronteira nos horários de contratação de Ponta e Fora Ponta do mesmo ano e os valores contratados da Tabela 5.

Com isso, nas Figura 33 e 34 são mostradas as curvas de demanda prevista, curva de demanda contratada reduzida, bem como o cálculo financeiro para o cenário 4 nos horários de fora ponta e ponta respectivamente. As curvas de contratação reduzida são mostradas em verde, enquanto as curvas de demanda máxima mensal prevista são mostradas na cor rosa. Nas mesmas figuras também é possível verificar os valores de EUST, ADEUST e PIU calculados mensalmente nas cores azul, amarelo e vermelho respectivamente.

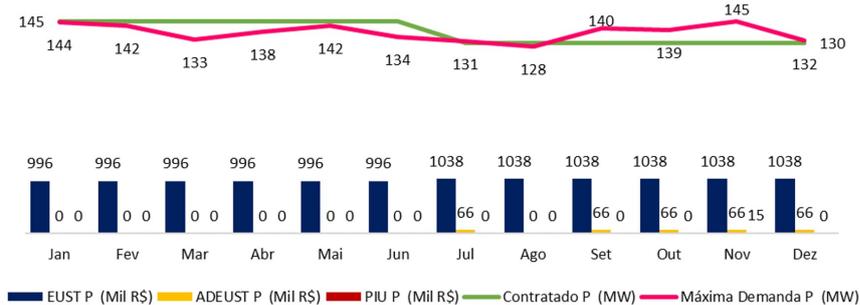
Observa-se que o valor contratado já começa reduzido no primeiro semestre de 2022 e no segundo semestre tem-se a segunda redução caracterizando assim a estratégia descrita anteriormente.

Figura 33 – Curvas de máxima demanda mensal prevista para 2022 e contratada reduzido com os respectivos valores de encargos calculados para fora ponta no cenário 4



Fonte: Elaborado pela própria autora.

Figura 34 – Curvas de máxima demanda mensal prevista para 2022 e contratada reduzido com os respectivos valores de encargos calculados para ponta no cenário 4



Fonte: Elaborado pela própria autora.

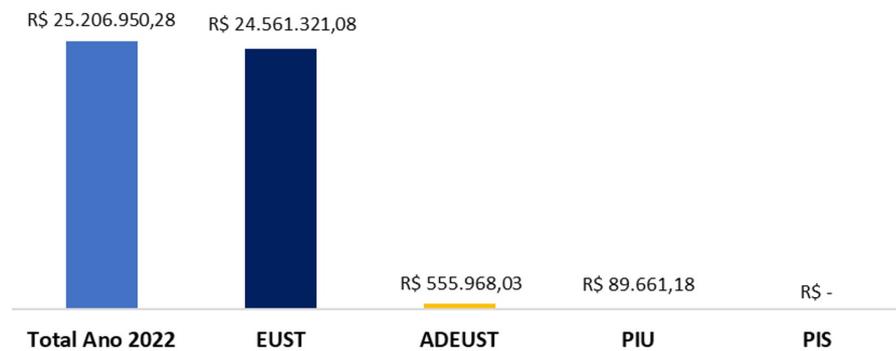
Sendo assim, vale salientar que os valores previstos de adicionais no primeiro semestre não foram mais mostrados devido o valor contratado estar maior que no cenário 2. Já no segundo semestre, os adicionais voltam a aparecer refletindo a redução mais próxima da esperada para otimização, contudo, seus valores são menores quando comparados com o cenário 2, visto que a redução dos 25% esperados não pode ser realizada afim de não haver oneração por redução prevista nos requisitos legais.

Assim como nos cenários anteriores, na Figura 35 é mostrado o valor financeiro anual de 2022, em reais, associado ao contrato analisado no cenário 4 para os dados de previsão.

Quando comparado com o cenário 2 hipotético, o cenário 4 utilizando dados de previsão tem o valor financeiro total associado ao contrato superior ao calculado no cenário 2, isto ocorre pois a legislação não permite que haja a redução total esperada pela otimização e nem que esta seja realizada de uma única vez.

Sob a mesma justificativa, o mesmo acontece com os adicionais do contrato que são

Figura 35 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 4 com os dados de previsão



Fonte: Elaborado pela própria autora.

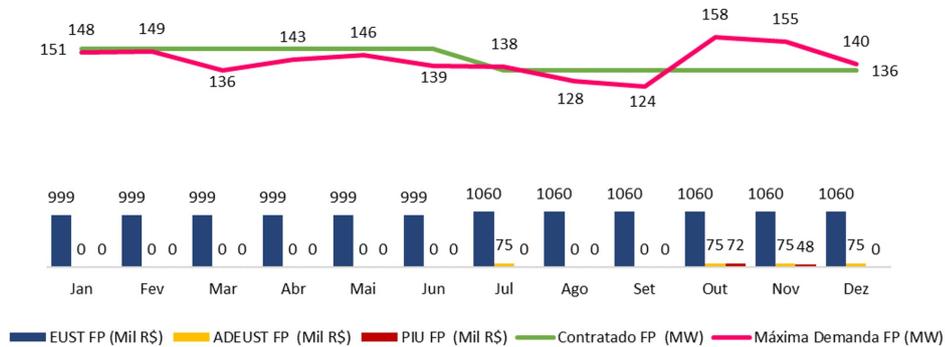
inferiores no cenário 4 em relação aos previstos no cenário 2 e o oposto acontece com os valores de EUST que aumentam no cenário 4, visto que esse encargo está relacionado com o valor do contratado mensal e neste cenário a redução do valor contratado é menor do que no cenário 2.

5.4.6 Cenário 4 - Dados Reais

No cálculo financeiro para o cenário 4 com dados reais foram utilizados os valores de demanda máxima mensal realizada em 2022, os valores de tarifa de rede básica e de fronteira nos horários de contratação de Ponta e Fora Ponta do mesmo ano e os valores contratados da Tabela 5.

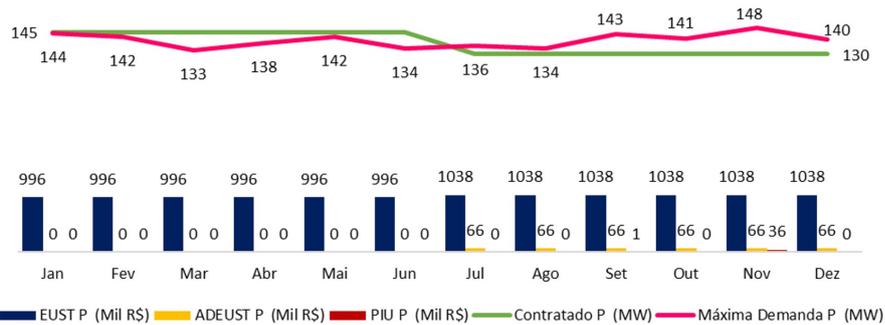
Com isso, nas Figura 36 e 37 são mostradas as curvas de demanda realizada, curva de demanda contratada reduzida, bem como o cálculo financeiro para o cenário 4 nos horários de fora ponta e ponta respectivamente. As curvas de contratação reduzida são mostradas em verde, enquanto as curvas de demanda máxima mensal realizada são mostradas na cor rosa. Nas mesmas figuras também é possível verificar os valores de EUST, ADEUST e PIU calculados mensalmente nas cores azul, amarelo e vermelho respectivamente.

Figura 36 – Curvas de máxima demanda mensal realizada para 2022 e contratada reduzido com os respectivos valores de encargos calculados para fora ponta no cenário 4



Fonte: Elaborado pela própria autora.

Figura 37 – Curvas de máxima demanda mensal realizada para 2022 e contratada reduzido com os respectivos valores de encargos calculados para ponta no cenário 4



Fonte: Elaborado pela própria autora.

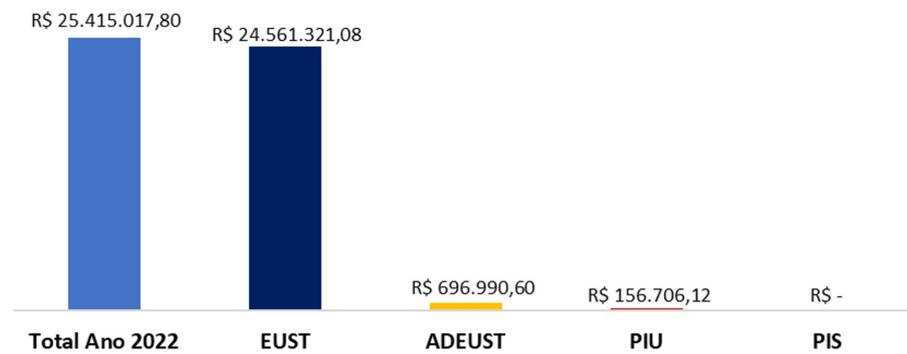
Observa-se que foi utilizada a mesma estratégia de contratação descrita anteriormente, o valor contratado já começa reduzido no primeiro semestre de 2022 e no segundo semestre tem-se a segunda redução.

Quando comparado com as curvas dos dados de previsão do cenário 4 percebe-se que os valores dos adicionais do primeiro semestre continuam não existindo e que há um aumento dos valores de ADEUST e PIU para o segundo semestre. Além disso, nos meses de julho e agosto é apresentado uma diferença entre os valores adicionais que não estavam previstos.

Já quando comparado com o cenário 3, tem-se que de forma geral os valores adicionais são menores, visto que a redução dos 25% esperados não pode ser realizada afim de não haver oneração por redução prevista nos requisitos legais.

Assim como nos cenários anteriores, na Figura 38 é mostrado o valor financeiro anual de 2022, em reais, associado ao contrato analisado no cenário 4 para os dados reais.

Figura 38 – Valor financeiro anual associado ao contrato em 2022 para o Cenário 4 com os dados reais



Fonte: Elaborado pela própria autora.

O ganho financeiro real resultado da diferença entre o valor financeiro associado ao contrato tradicional e o valor financeiro associado ao contrato otimizado, ambos utilizando os dados medidos de demanda de 2022, é cerca de 12%, referente a uma economia de R\$ 3.429.798,48, o que significa uma melhoria significativa do custo financeiro associado ao CUST com a aplicação do algoritmo proposto.

Devido a diferença entre os valores previstos e os valores realizados, o risco financeiro associado a aplicação do algoritmo proposto para otimização do CUST, resultado da diferença entre o valor total do contrato previsto e o valor total do contrato real, é de 0,82% referente ao valor de de R\$ 208.067,52.

6 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho foi porposta a aplicação de algoritmos inteligentes na previsão do MUST a fim de haver uma otimização da função de custo associado ao CUST das concessionárias de distribuição de energia elétrica.

A priori foi realizada uma revisão sistemática sobre os temas relacionados à aplicação de algoritmos inteligentes na previsão do MUST, sendo, desta forma, obtido um número limitado de trabalhos, indicando uma lacuna sobre o assunto e, portanto, uma oportunidade para o desenvolvimento deste trabalho.

O algoritmo desenvolvido foi dividido nas etapas de pré-processamento dos dados, análise de correlação dos dados, algoritmos regressores e otimização financeira.

Para realizar a análise de correlação dos dados, foi feito um estudo utilizando a Modelagem de Equações Estruturais MEE, para preestabelecer uma relação entre os dados históricos de demanda de energia elétrica com dados de fatores climáticos, econômicos, temporais e de geração. Como resultado, foi visto que há uma forte correlação entre os dados, auxiliando na seleção das melhores *features* para o modelo de *Machine Learning* utilizado na previsão de demanda.

Sendo assim, foram avaliados os desempenhos dos algoritmos de regressão LSTM, MLP, MLR, KNN e o RFR utilizando as métricas R^2 , MAE, MAPE, MSE e RMSE, que indicam que o algoritmo com melhor resultado de previsão foi o RFR com 71,1% de precisão.

Com os dados de previsão de demanda de energia elétrica, foi construída uma função do custo associado ao CUST anual de 2022 em quatro cenários diferentes, sendo calculados os valores da maneira tradicional no primeiro cenário, o segundo cenário utilizando os valores da previsão do algoritmo de regressão, o terceiro cenário utilizando os valores de contrato otimizado com base nos dados medidos e o quarto cenário utilizando os valores de contrato otimizado com base nos requisitos legais.

Na forma tradicional de contratação, a distribuidora de energia elétrica tende a contratar o máximo valor de demanda possível para que não sejam cobradas as penalidades e multas devido ultrapassagens. Contudo, o grande desafio é conseguir prever valores de demanda tais que se aproximem ao máximo das medições reais futuras e que embasem, estrategicamente, um valor de contrato que minimize os custos financeiros associados.

Desta forma, com a previsão dada pelos algoritmos regressores foi possível realizar uma minimização da função custo, cuja variável de entrada foi a demanda futura prevista e a

variável de saída foi o valor ótimo a ser contratado pelas distribuidoras de energia elétrica.

Com isso, foi possível calcular os valores financeiros associados ao CUST em quatro cenários diferentes conforme é mostrado na Tabela 6.

Tabela 6 – Comparativo financeiro dos cenários analisados

Cenário	Total CUST Ano 2022	EUST	ADEUST	PIU	PIS	Economia (R\$)	Economia (%)
Cenário 1 ou Cenário Base	R\$ 28.844.816,28	R\$ 28.844.816,28	R\$ -	R\$ -	R\$ -	R\$ -	0,0%
Cenário 2	R\$ 24.534.995,38	R\$ 23.146.743,60	R\$ 1.190.033,06	R\$ 198.218,72	R\$ -	R\$ 4.309.820,90	-14,9%
Cenário 3	R\$ 25.089.360,43	R\$ 23.146.743,60	R\$ 1.716.285,33	R\$ 207.302,82	R\$ -	R\$ 3.755.455,85	-13,0%
Cenário 4 (Dados Previsão)	R\$ 25.206.950,28	R\$ 24.561.321,08	R\$ 555.968,03	R\$ 89.661,18	R\$ -	R\$ 3.637.866,00	-12,6%
Cenário 4 (Dados Reais)	R\$ 25.415.017,80	R\$ 24.561.321,08	R\$ 696.990,60	R\$ 156.706,12	R\$ -	R\$ 3.429.798,48	-11,9%

Fonte: Elaborado pela própria autora.

A partir dos valores contratados da forma tradicional, foi calculado o valor financeiro associado ao CUST para o Cenário 1 ou Cenário Base. A partir deste cenário, foi possível analisar a economia financeira prevista para uma contratação otimizada.

Observa-se que no Cenário 2 calcula-se o valor financeiro idealizado, ou seja, com base nos dados de previsão de demanda decorrentes dos algoritmos de regressão, calcula-se qual o valor ótimo de contrato para que haja uma economia em comparação com o cenário 1.

Caso esse cenário fosse possível e a demanda medida fosse o mais próxima possível da prevista, o valor economizado com a otimização do contrato seria de R\$ 4.309.820,90, o equivalente a aproximadamente 15%.

Entretanto, há uma diferença entre os valores de demanda previstos pelo algoritmo e os valores de demanda realizados, resultando em uma economia menor do que a prevista pelo Cenário 2. O impacto financeiro dessa diferença pode ser visualizado através da análise do Cenário 3, onde há a utilização dos dados reais para o mesmo valor de contratação ótima do Cenário anterior.

Como consequência, tem-se uma economia de R\$ 3.755.455,85, o equivalente a aproximadamente 13% em comparação com o cenário 1 e com uma diferença próxima de -2% quando comparado com o valor esperado do Cenário 2.

Contudo, de acordo com a REN n°905/2020 não é possível que o contrato seja otimizado ao valor ideal calculado pelo cenário 2 sem oneração. Com isso, o Cenário 4 (Dados Previsão) calcula qual o valor economizado previsto, com base nos dados de previsão de demanda decorrentes dos algoritmos e no valor de contrato otimizado de acordo com as regras da redução não onerosa da resolução normativa.

Como resultado, tem-se um valor esperado de R\$ 3.637.866,00, o equivalente a aproximadamente 13% em comparação com o Cenário 1 e com uma diferença próxima de -2%

quando comparado com o valor esperado do Cenário 2.

Para a análise de um Cenário 4 mais realista, foi calculado qual o valor financeiro associado ao CUST que leva em consideração às regras da redução não onerosa da resolução normativa REN n°905/2020 e a diferença existente entre os dados de previsão e os dados de medição de demanda de energia elétrica.

Desta forma, viu-se que foi possível se obter um ganho financeiro de aproximadamente 12% referente à R\$ 3.429.798,48 de economia a partir do resultado da otimização em comparação com o contrato atual do Cenário 1.

Além disso, devido a diferença entre os valores reais e os valores previstos pelo algoritmo, foi calculado um risco financeiro de 0,82% em relação ao custo total associado ao CUST. Tal risco pode ser considerado pequeno quando comparado com o ganho financeiro previsto pela otimização.

6.1 Trabalhos Futuros

Algumas premissas foram adotadas nesta dissertação para elaboração desta proposta. Sendo assim, para trabalhos futuros recomenda-se que os próximos modelos de algoritmos possam prever transferências entre as regiões eletricamente conectadas, sem alterar o perfil de curva de demanda típico de determinada região analisada para sua contratação.

Sob o ponto de vista da base de dados utilizada, podem ser avaliadas outras técnicas de melhoria para aumento da precisão da previsão dos dados de demanda e obtenção de dados mais discretizados no tempo.

Além disso, esta dissertação não exaure todas as possibilidades de formas de otimização de contratação, com isso, há possibilidade de um aumento no ganho financeiro através da análise de uma contratação de caráter permanente associada a uma contratação de caráter flexível, não considerada neste trabalho.

REFERÊNCIAS

- ALHUSSEIN, M.; AURANGZEB, K.; HAIDER, S. I. **Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting**, IEEE Access, v. 8, p. 180544 – 180557, 2020.
- AMORIM, L. D. A. F.; FIACCONE, R. L.; SANTOS, C. A. de S. T.; MORAES, L. T. L. P. de; OLIVEIRA, N. F. de; OLIVEIRA, S. B.; SANTOS, T. N. L. dos. **Modelagem com Equações Estruturais: Princípios Básicos e Aplicações**, Departamento de Estatística da Universidade Federal da Bahia, 2012.
- ANEEL. **Resolução Normativa nº 905/2020, de 08 de dezembro de 2020**. Agência Nacional de Energia Elétrica - Módulo 5 - Dispõe sobre as regras dos serviços de transmissão de energia elétrica e dá outras providências. <https://www2.aneel.gov.br/cedoc/ren2020905.html>, 2020. Acesso em: 15/02/2024.
- ESCOVEDO, T.; KOSHIYAMA, A. **Introdução a Data Science — Algoritmos de Machine Learning e métodos de análise**, Ed. Casa do Código, 2020.
- FARIA, J. M. de. **Metodologia da Área de Influência para Cálculo da Tarifa de Uso do Sistema de Transmissão**, Universidade Federal de Itajubá, 2013.
- FERREIRA, V. H.; AOKI, A. R.; GUERRA, F. A.; ROCCO, S. M. de. **Previsão de Demandas para Controle dos Montantes de Uso do Sistema de Transmissão para Distribuidoras de Energia Elétrica**, XVIII Seminário Nacional de Distribuição de Energia Elétrica - Olinda - Pernambuco - Brasil, 2008.
- GEORGY, M.; ANNA, I. **semopy: A Python package for Structural Equation Modeling**, Structural Equation Modeling: A Multidisciplinary Journal, v. 27, p. 952–963, 18/02/2020.
- GUSMAN, C. S. de A. **Incerteza de Medição em Redes Neurais Artificiais Aplicadas à Manutenção Preditiva de Transformadores**, PUC-RIO, 2011.
- IGNACIO, L. F. F. **Aprendizado de máquina: da teoria à aplicação**, Departamento de Matemática - Universidade Federal Fluminense, 2021.
- JAWAD, M.; SAJJAD, M.; SHIM, S. o; RASOOL, I.; SHAHEEN, A.; HABIB, N.; HUSSAIN lal; AZIZ, W. **Machine Learning Based Cost Effective Electricity Load Forecasting Model Using Correlated Meteorological Parameters**, IEEE Access, v. 8, p. 146847 – 146864, 2020.
- LEANSAUDE. **Aprendizado Supervisionado vs. Aprendizado Não-Supervisionado em machine learning**. 2023. Disponível em: <https://www.leansaude.com.br/aprendizado-supervisionado-vs-aprendizado-nao-supervisionado-em-machine-learning/>. Acesso em: 23 set. 2023.
- LIMA, L. H. L. **Tarifação do Uso de Sistemas de Transmissão em Mercados Interligados de Energia Elétrica**, Universidade Federal de Itajubá, 2009.
- LIMA, L. M. M.; QUEIROZ, A. R. de; LIMA, J. W. M.; RIBEIRO, A.; ELHAGE, E. **Determinação do MUST Ótimo para Empresas de Distribuição de Energia Elétrica**, Congresso de Inovação Tecnológica em Energia Elétrica - CITENEL, VI, 2021.

MAMUN, A. A.; SOHEL, M.; MOHAMMAD, N.; SUNNY, M. S. H.; DIPTA, D. R.; HOSSAIN, E. A. **A Comprehensive Review of the Load Forecasting Techniques Using Single and Hybrid Predictive Models**, IEEE Access, v. 8, p. 134911 – 134939, 2020.

MARÔCO, J. **Livro Análise de Equações Estruturais Fundamentos teóricos, software aplicações**, ReportNumber, 2021.

MENDONÇA, A. K. de S.; SILVA, S. A. da; PEREIRA, L. Z.; BORNIA, A. C.; ANDRADE, D. F. de. **Analysis of the strategies for the commercialization of electric energy of generating agents in the short term environment**, Revista de Gestão, Educação e Tecnologia Ambiental, v. 25, 2021.

MILHORANCE, A.; SILVA, A. M. L. da. **Optimal Contracting of Transmission System Usage Combined With Subtransmission Planning via Cross-Entropy Method**, IEEE TRANSACTIONS ON POWER SYSTEMS, v. 38, n. 5, p. 4881–4889, 2023.

NAKAGAWA, E. Y.; FELIZARDO, K. R.; FERRAZ, S. C. P.; FERRARI, F. C. **Revisão Sistemática da Literatura em Engenharia de Software**, Elsevier, 2017.

NEVES, J. A. B. **Modelo de equações estruturais: uma introdução aplicada**, ENAP, 2018.

ONS. **A Contratação do Sistema de Transmissão**. Brasil: Sintegre, 2023. Disponível em: <https://www.ons.org.br/paginas/energia-no-futuro/transmissao>. Acesso em: 28 jul. 2023.

PASCUAL, C. **Tutorial: Understanding Regression Error Metrics in Python**. Dataquest, 2018. Disponível em: <https://www.dataquest.io/blog/understanding-regression-error-metrics/>. Acesso em: 23 ago. 2023.

ROEVER, L. **Guia Prático de Revisão Sistemática e Metanálise**, Thieme Revinter - Rio de Janeiro, 2020.

RUDNIK, H.; PALMA, R.; CURA, E.; SILVA, C. **Economically adapted transmission systems in open access schemes-application of genetic algorithms**, IEEE Transactions on Power Systems, v. 11, n. 3, p. 1427–1440, 1996.

WALSER, T.; SAUER, A. **Typical load profile-supported convolutional neural network for short-term load forecasting in the industrial sector**, Energy and AI - Elsevier, v. 5, p. 100104, 2021.