



**Universidade Federal de Uberlândia  
Faculdade de Engenharia Elétrica  
Pós-graduação em Engenharia Elétrica**

DANIEL DE OLIVEIRA FERREIRA

**PROPOSTA DE UM SISTEMA DE IOT PARA REDES DE DISTRIBUIÇÃO:  
ENTIDADES VIRTUAIS, VISUALIZAÇÃO DE INFORMAÇÃO E  
INTELIGÊNCIA ARTIFICIAL APLICADAS A *SMART GRIDS***

Uberlândia  
2022

DANIEL DE OLIVEIRA FERREIRA

PROPOSTA DE UM SISTEMA DE IOT PARA REDES DE DISTRIBUIÇÃO:  
ENTIDADES VIRTUAIS, VISUALIZAÇÃO DE INFORMAÇÃO E  
INTELIGÊNCIA ARTIFICIAL APLICADAS A *SMART GRIDS*

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Uberlândia, como exigência parcial para a obtenção do título de mestre em Ciências.

Orientador: Prof. Dr. Alan Petrônio Pinheiro

Banca Examinadora:

\_\_\_\_\_  
Alan Petrônio Pinheiro (orientador), Dr. – UFU

\_\_\_\_\_  
Alexandre Cardoso, Dr. – UFU

\_\_\_\_\_  
Lorenço Santos Vasconcelos, Dr. – UFU

\_\_\_\_\_  
Ricardo Augusto Souza Fernandes, Dr. – UFSCar

Uberlândia, 04 de agosto de 2022

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU  
com dados informados pelo(a) próprio(a) autor(a).

F383  
2022 Ferreira, Daniel de Oliveira, 1995-  
Proposta de um Sistema de IoT para Redes de  
Distribuição: Entidades Virtuais, Visualização de  
Informação e Inteligência Artificial Aplicadas a Smart  
Grids [recurso eletrônico] / Daniel de Oliveira  
Ferreira. - 2022.

Orientador: Alan Petrônio Pinheiro.  
Dissertação (Mestrado) - Universidade Federal de  
Uberlândia, Pós-graduação em Engenharia Elétrica.  
Modo de acesso: Internet.  
Disponível em: <http://doi.org/10.14393/ufu.di.2022.447>  
Inclui bibliografia.  
Inclui ilustrações.

1. Engenharia elétrica. I. Pinheiro, Alan Petrônio,  
1982-, (Orient.). II. Universidade Federal de  
Uberlândia. Pós-graduação em Engenharia Elétrica. III.  
Título.

CDU: 621.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:  
Gizele Cristine Nunes do Couto - CRB6/2091  
Nelson Marcos Ferreira - CRB6/3074



**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**  
 Coordenação do Programa de Pós-Graduação em Engenharia Elétrica  
 Av. João Naves de Ávila, 2121, Bloco 3N - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902  
 Telefone: (34) 3239-4707 - www.posgrad.feelt.ufu.br - copel@ufu.br



### ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Elétrica				
Defesa de:	Dissertação de Mestrado, 777, PPGEELT				
Data:	quatro de agosto de dois mil e vinte e dois	Hora de início:	09:00	Hora de encerramento:	11:30
Matrícula do Discente:	11922EEL002				
Nome do Discente:	Daniel de Oliveira Ferreira				
Título do Trabalho:	Proposta de um Sistema de IoT para Redes de Distribuição: Entidades Virtuais, Visualização de Informação e Inteligência Artificial Aplicadas a Smart Grid				
Área de concentração:	Processamento da Informação				
Linha de pesquisa:	Redes de Computadores				
Projeto de Pesquisa de vinculação:	Coordenador do projeto: Alan Petrônio Pinheiro. Título do projeto: Estudo e desenvolvimento piloto de novos modelos de serviços e infraestrutura de TIC voltados ao uso de antenas de telecomunicações da rede de distribuição da CEB alinhados ao cenário de SG e IoT. Agência financiadora: ANEEL. Número do processo na agência financiadora: 00310-00009787/2018-36. Vigência do projeto: 12/2018 a 12/2021				

Reuniu-se por meio de videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta: Professores Doutores: Alexandre Cardoso - FEELT/UFU; Lorenço Santos Vasconcelos - FEELT/UFU; Ricardo Augusto Souza Fernandes - UFSCar; Alan Petrônio Pinheiro - FEELT/UFU, orientador do candidato.

Iniciando os trabalhos o presidente da mesa, Dr. Alan Petrônio Pinheiro, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

**Aprovado.**

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Alan Petronio Pinheiro, Professor(a) do Magistério Superior**, em 04/08/2022, às 12:22, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Alexandre Cardoso, Professor(a) do Magistério Superior**, em 04/08/2022, às 15:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ricardo Augusto Souza Fernandes, Usuário Externo**, em 04/08/2022, às 16:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Lorenzo Santos Vasconcelos, Professor(a) do Magistério Superior**, em 04/08/2022, às 17:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://www.sei.ufu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3788611** e o código CRC **C7BB1639**.

# Agradecimentos

Agradeço à Deus pela oportunidade de existir neste universo incrível, e pelas inúmeras graças e aprendizados que me proporcionou nesta jornada. Agradeço também a minha família, e em especial: ao meu Pai, que construiu um império de simplicidade e, com o ensino fundamental, mostrou-me desde cedo como um engenheiro deve ser criativo e proativo; à minha mãe, que com serenidade e fé perseverou orando e banhando por nós todos os dias; ao meu irmão Guilherme que me ensinou matemática, antes da hora, informática, nutrição, a ter foco, a estudar e superar as adversidades com união; ao meu irmão Matheus que me ensinou a tirar um tempo para realizar os sonhos, e que caridade deve estar sempre presente; a Laila, que mesmo de longe continuou torcendo e compartilhando as experiências ~~anormais~~ típicas da engenharia; a Marlene Tolentino por rezar e torcer para que tudo desse certo; ao meu nobre gato e companheiro de estudo Stefan, presente de Deus que aliviou momentos cansativos com seu ~~sono~~ espírito felino e fraterno; e a Nathan, que não me deixou ficar ~~quieto~~ sedentário durante a trajetória do mestrado.

Agradeço aos meus professores pelos conhecimentos e companhia, e em especial ao Dr. Alan Petrônio, que confiou em mim mesmo diante de batalhas improváveis e longas, tranquilizando e encorajando! Agradeço ainda à Universidade Federal de Uberlândia, e assim, a todos que dela fizeram ou fazem parte, pois se tornou segunda casa. Que honra poder fazer parte desta universidade que leva ensino de qualidade a tantos jovens.

Agradeço ainda aos colegas e amigos, especialmente a Willian, Leandro, Ubaldo, Caio, Guilherme e Gabriel, que fizeram parte do dia a dia do mestrado de perto. Agradeço também a Fernanda, que foi e é exemplo de pesquisadora e cientista. E tantos outros que de alguma forma contribuíram com sua energia para que eu aprendesse tanto e chegasse aqui!

Essa pesquisa foi financiada pelo projeto de P&D ANEEL nº 05160-1805/2018, uma parceria entre a Companhia Energética de Brasília (CEB) e Universidade Federal de Uberlândia (UFU), e com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo número 135168/2019-8. Desta forma, agradeço às agências de fomento, e em especial ao CNPq, que possibilitam diversos pesquisadores contribuir com a ciência.

---

*Pesquisa realizada na:*



Universidade Federal de Uberlândia  
Programa de pós-graduação em Engenharia Elétrica



CePEDRI - Centro de P&D em Redes Inteligentes  
[www.cepedri.eletrica.ufu.br](http://www.cepedri.eletrica.ufu.br)

*Financiamento e apoio:*



Agência Nacional de Energia Elétrica  
Programa de Pesquisa e Desenvolvimento



Conselho Nacional de Desenvolvimento  
Científico e Tecnológico

---

Esta pesquisa foi parcialmente financiada pela Companhia Energética de Brasília por meio do projeto de P&D ANEEL nº 05160-1805/2018.

# Resumo

A quantidade de dados produzidos e armazenados vem crescendo expressivamente. Impulsionada por fatores relacionados à internet das coisas (IoT), esta tendência deve continuar. No setor elétrico, a implantação das *smart grids* (SGs) almeja fornecer novos serviços e superar desafios, como diminuir perdas não técnicas, integrar fontes distribuídas e intermitentes de energia, bem como atender à crescente demanda. Para que as SGs possam alcançar estes objetivos, é necessário que os dados coletados sejam transformados em informações úteis, e apresentados de forma eficiente. Neste contexto, este trabalho apresenta uma proposta de sistema de monitoramento e análise para redes de distribuição e SGs, aplicando conceitos relacionados a entidades virtuais, técnicas de visualização de informação e algoritmos de inteligência artificial. São utilizadas técnicas de visualização em árvore, mapa e coordenadas paralelas. Com isso, cenários de monitoramento com mais de um milhão de entidades são apresentados, confirmando a eficiência da proposta. Além disso, redes neurais recorrentes são utilizadas para realizar a previsão de carga. Para a avaliação dos modelos de previsão são utilizadas diversas métricas de erro, em conjunto com a técnica de visualização de coordenadas paralelas. Assim, a aplicação implementada materializa diversos conceitos e funcionalidades importantes para sistemas de monitoramento e análise no contexto de *smart grids*.

**Palavras-chave:** Internet das coisas. Entidades virtuais. Redes de distribuição. Visualização de informação. Redes neurais artificiais. Previsão de carga.



# Abstract

The amount of generated and stored data has been growing significantly. Driven by factors related to the internet of things (IoT), this trend is expected to continue. In the electricity sector, the implementation of smart grids (SGs) aims to provide new services and overcome challenges, such as reducing non-technical losses, integrating distributed and intermittent energy sources, as well as meeting the growing demand. For SGs to achieve these goals, the collected data must be transformed into useful information and presented efficiently. In this context, this work presents a proposal for a monitoring and analysis system for distribution networks and SGs, applying concepts related to digital twins, information visualization techniques, and artificial intelligence algorithms. A tree visualization, a map, and the parallel coordinates technique are used. Then, monitoring scenarios with more than a million entities are presented, confirming the efficiency of the proposal. In addition, recurrent neural networks are used to perform load prediction. For the evaluation of the forecast models, several error metrics are used, along with the parallel coordinates visualization technique. Thus, the implemented application materializes several important concepts and functionalities for monitoring and analysis systems, in the context of smart grids.

**Keywords:** Internet of things. Digital twins. Distribution grid. Information Visualization. Artificial Neural Networks. Load forecast.

# Lista de figuras

Figura 1 – Paradigmas de IoT.....	25
Figura 2 – Arquitetura de 3 camadas para sistemas de IoT.....	26
Figura 3 – Arquitetura de 4 camadas baseada em serviços para sistemas de IoT. ....	27
Figura 4 – Arquitetura de 5 camadas com <i>middleware</i> para sistemas de IoT. ....	27
Figura 5 – Arquitetura de 5 camadas com <i>edge computing</i> para sistemas de IoT. ....	28
Figura 6 – Tendências no contexto de IoT. ....	29
Figura 7 – Diagrama simplificado de uma <i>smart grid</i> . ....	31
Figura 8 – Esquema de um sistema elétrico de potência. ....	34
Figura 9 – Visualização em coordenadas paralelas. ....	40
Figura 10 – Visualização em árvore explícita: diagrama nó-link, na aplicação VPlant.....	42
Figura 11 – Visualização hierárquica implícita. ....	42
Figura 12 – Esquema de rede neural em camadas. ....	51
Figura 13 – Principais elementos da aplicação e sistemas adjacentes.....	57
Figura 14 – Fotos do dispositivo (a) medindo dados de um transformador de distribuição, (b) em detalhe e (c) dentro da caixa de proteção. ....	59
Figura 15 – Elementos e organização da aplicação VPlant.....	60
Figura 16 – Classes de entidades virtuais da aplicação VPlant <i>back-end</i> . ....	62
Figura 17 – Elementos gráficos iniciais da aplicação VPlant. ....	64
Figura 18 – Painel com detalhes da entidade virtual. ....	65
Figura 19 – Árvore de relacionamentos. ....	66
Figura 20 – Visualização em coordenadas paralelas na aplicação. ....	67
Figura 21 – Seleção de linhas com dois <i>brushes</i> . ....	69
Figura 22 – <i>Layout</i> em coordenadas paralelas após utilização da função “Filtrar”. ....	69
Figura 23 – Seções de configuração do painel de análise de entidades a jusante. ....	70
Figura 24 – Visualização de dados coletados em coordenadas paralelas.....	71
Figura 25 – Painel para previsão de dados. ....	72
Figura 26 – Seção “Seleção de dados” do painel de previsão de dados.....	72
Figura 27 – Seção “Modelos” do painel de previsão de dados. ....	74
Figura 28 – Parte superior da subseção “Métricas de treino”, do painel de previsão de dados. ....	75
Figura 29 – Gráfico de linha com previsões para o conjunto de validação.....	76

Figura 30 – Erros e atributos em coordenadas paralelas. ....	76
Figura 31 – Evento em uma única entidade. ....	79
Figura 32 – Eventos em um transformador de distribuição e em 16 unidades consumidoras a jusante do transformador. ....	79
Figura 33 – Eventos em 112 transformadores de distribuição e cerca de 1500 unidades consumidoras, relacionados a um circuito de média tensão. ....	80
Figura 34 – Eventos em cerca de 10 mil entidades a jusante de uma subestação. ....	81
Figura 35 – Eventos em quase 30 mil transformadores de distribuição. ....	81
Figura 36 – Painel com coordenadas paralelas e dados da BDGD. ....	83
Figura 37 – Coordenadas paralelas com eixos selecionados. ....	84
Figura 38 – Coordenadas paralelas com dados de unidades consumidoras. ....	84
Figura 39 – Eixo de carga instalada com novo intervalo. ....	85
Figura 40 – Dados selecionados para análise com coordenadas paralelas. ....	85
Figura 41 – Dados de um transformador em coordenadas paralelas. ....	86
Figura 42 – Dados de 0 a 5 horas em coordenadas paralelas. ....	87
Figura 43 – Painel com o conjunto de dados e o modelo selecionados para o primeiro transformador. ....	88
Figura 44 – Métricas de erro e gráfico de previsões em função do tempo na aplicação VPlant, para o primeiro transformador. ....	89
Figura 45 – Coordenadas paralelas mostrando quartis do erro absoluto relativo para o primeiro transformador. ....	90
Figura 46 – Coordenadas paralelas mostrando três <i>outliers</i> do erro absoluto relativo e classe de cor por erro percentual para o primeiro transformador. ....	90
Figura 47 – Coordenadas paralelas mostrando amostras com mais de 15% de erro percentual absoluto e classes de cor por erro absoluto para o primeiro transformador. ....	91
Figura 48 – Coordenadas paralelas mostrando as amostras com mais de 10% de erro e classes de cor por potência ativa total para o primeiro transformador. ....	92
Figura 49 – Métricas de erro e gráfico de previsões em função do tempo para o modelo com entradas de 7 dias. ....	93
Figura 50 – Métricas de erro e gráfico de previsões em função do tempo para o modelo com entradas de 3 dias. ....	94

# Lista de abreviaturas e siglas

AdaGrad	<i>Adaptive gradient algorithm</i>
Adam	<i>Adaptive moment estimation</i>
AMI	<i>Advanced metering infrastructure</i> – infraestrutura de medição avançada
ANEEL	Agência Nacional de Energia Elétrica
ANN	<i>Artificial neural network</i> – rede neural artificial
API	<i>Application programming interface</i> – interface de programação de aplicativos
AR	<i>Augmented reality</i> – realidade aumentada
ARIMA	<i>Autoregressive integrated moving average</i> – autorregressivo integrado de médias móveis
BNDES	Banco Nacional do Desenvolvimento
CEB	Companhia Energética de Brasília
CNN	<i>Convolutional neural network</i> – rede neural convolucional,
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
DIC	Duração de interrupção individual por unidade consumidora
DTI	<i>Digital twin instance</i> – instância de gêmeo digital
DT	<i>Digital twin</i> – gêmeo digital
ET	Estação transformadora
FIC	Frequência de interrupção individual por unidade consumidora
GPU	<i>Graphics processing unit</i> – unidade de processamento gráfico
GRU	<i>Gated recurrent unit</i>
HTTP	<i>Hypertext transfer protocol</i>
IA	Inteligência artificial
IED	<i>Intelligent electronic device</i> – dispositivo eletrônico inteligente
InfoVis	<i>Information visualization</i> – visualização de informação
IoT	<i>Internet of things</i> – internet das coisas
JADE	<i>Java Agent Development</i>
LNS	<i>LoRaWAN Network Server</i>
LoRa	<i>Long Range</i>
LPWAN	<i>Low power wide area network</i>
LSTM	<i>Long short-term memory</i>

LTLF	<i>Long-term load forecasting</i> – previsão de carga de longo prazo
MAE	<i>Mean square error</i> – erro médio absoluto
MAPE	<i>Mean absolute percentage error</i> – erro médio percentual absoluto
MASE	<i>Mean absolute scaled error</i> – erro médio absoluto em escala
MCTIC	Ministério da Ciência, Tecnologia, Inovações e Comunicações
MdRAE	<i>Median relative absolute error</i> – mediana do erro absoluto relativo
ME	<i>Mean error</i> – erro médio
MLP	<i>Multilayer perceptron</i>
MQTT	<i>Message queuing telemetry transport</i>
MRAE	<i>Mean relative absolute error</i> – erro médio absoluto relativo
MSE	<i>Mean square error</i> – erro médio quadrático
MTLF	<i>Medium-term load forecasting</i> – previsão de carga de médio prazo
MVC	<i>Model view controller</i>
NA	Normalmente aberta
NF	Normalmente fechada
NP	<i>Network protectors</i>
P&D	Pesquisa e desenvolvimento
PCs	<i>Parallel coordinates</i> – coordenadas paralelas
PLC	<i>Power line communication</i>
PRODIST	Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional
RIAPS	<i>Resilient Information Architecture Platform for Smart Grid</i>
RMSE	<i>Root mean square error</i> – raiz do erro médio quadrático
RMSProp	<i>Root mean squared propagation</i>
RNA	Rede neural artificial – <i>artificial neural network</i>
RNN	<i>Recurrent neural network</i> – rede neural recorrente
SE	Subestações
SEMS	<i>Sustainable Energy Management System</i>
SEP	Sistema elétrico de potência
SGD	<i>Stochastic gradient descent</i> – gradiente descendente estocástico
SG	<i>Smart grid</i>
sMAPE	<i>Symmetric mean absolute percentage error</i> – erro médio percentual absoluto simétrico
SM	<i>Smart meter</i>

STLF	<i>Short-term load forecasting</i> – previsão de carga de curto prazo
SVR	<i>Support vector regression</i>
TCP	<i>Transmission control protocol</i> – protocolo de controle de transmissão
UCBT	Unidade consumidora de baixa tensão
UC	Unidade consumidora
UDP	<i>User datagram protocol</i> – protocolo de datagramas do usuário
UNTRD	Unidade transformadora de distribuição
VPlant	<i>Virtual Power Plant</i>
VSTLF	<i>Very short-term load forecasting</i> – previsão de carga de curtíssimo prazo

# Lista de símbolos

ENE_06	Consumo de energia no mês de junho
ENE_11	Consumo de energia no mês de novembro
ENE_12	Consumo de energia no mês de dezembro
$e_t$	Erro de previsão para um instante $t$
$e_t(\tau)$	Erro residual para um instante $t$ e horizonte de previsão $\tau$
$e_t^*$	Erro de um algoritmo de referência para um instante $t$
CAR_INST	Carga instalada
ENE_M_R	Razão entre a potência média consumida e a potência nominal
POT_NOM	Potência nominal do transformador
PP2	Poder público distrital ou estadual
$t$	instante de tempo
$y_t$	uma amostra da série $y$ num instante $t$
$\hat{y}_t$	estimação de uma amostra da série $y$ num instante $t$

# Sumário

<b>Lista de figuras .....</b>	<b>9</b>
<b>Lista de abreviaturas e siglas.....</b>	<b>11</b>
<b>Lista de símbolos.....</b>	<b>14</b>
<b>Sumário .....</b>	<b>15</b>
<b>1      <b>Introdução.....</b></b>	<b>17</b>
1.1    APRESENTAÇÃO .....	17
1.2    OBJETIVOS DA PESQUISA .....	18
1.3    LIMITAÇÕES, CONVENÇÕES E ESCOPO.....	19
1.4    JUSTIFICATIVA .....	21
1.5    CONTRIBUIÇÃO.....	22
1.6    ORGANIZAÇÃO DO TEXTO .....	22
<b>2      <b>Fundamentação teórica.....</b></b>	<b>24</b>
2.1    IOT E <i>SMART GRIDS</i> .....	24
2.1.1 <i>Internet of things</i> .....	24
2.1.1.1    Elementos e arquiteturas .....	25
2.1.1.2    Tendências para os sistemas de IoT .....	29
2.1.2 <i>Smart grids</i> .....	30
2.1.2.1    Elementos básicos.....	30
2.1.2.2    Benefícios da implantação de <i>smart grids</i> .....	32
2.1.3 <i>Virtualização: entidades virtuais</i> .....	32
2.2    REDES DE DISTRIBUIÇÃO .....	34
2.2.1 <i>Sistemas elétricos de potência</i> .....	34
2.2.2 <i>Elementos dos sistemas de distribuição</i> .....	35
2.2.2.1    Dispositivos de proteção e controle.....	36
2.2.2.2    Transformadores.....	36
2.2.2.3    Reguladores de tensão.....	36
2.2.3 <i>Tipos de redes de distribuição</i> .....	37
2.2.3.1    Radial.....	37
2.2.3.2    Radial com recurso.....	37
2.2.3.3    Primário seletivo .....	37
2.2.3.4    Primário operando em malha aberta .....	38
2.2.3.5 <i>Spot network</i> .....	38
2.3    FERRAMENTAS DE VISUALIZAÇÃO DE INFORMAÇÃO .....	38
2.3.1 <i>Coordenadas paralelas</i> .....	39
2.3.2 <i>Visualizando relações e dados hierárquicos</i> .....	40
2.4    PREVISÃO DA DEMANDA DE CARGA.....	43
2.4.1 <i>Horizontes, seleção de variáveis e agregação de cargas em previsões</i> .....	43
2.4.2 <i>Avaliação das previsões</i> .....	44



2.4.2.1	Erro residual e erro de previsão .....	44
2.4.2.2	Métricas de erro dependentes de escala .....	45
2.4.2.3	Métricas de erro percentuais e relativas.....	46
2.4.2.4	Métricas de erro em contextos específicos.....	47
2.4.3	<i>Pré-processamento de dados: desafios e procedimentos</i> .....	48
2.4.3.1	Normalização por amplitude.....	48
2.4.3.2	Dados ausentes ou inconsistentes .....	49
2.4.3.3	Aspectos temporais.....	49
2.5	REDES NEURAI ARTIFICIAIS .....	50
<b>3</b>	<b>Estado da arte e pesquisas correlatas .....</b>	<b>52</b>
3.1	SISTEMAS E FERRAMENTAS COMPUTACIONAIS PARA <i>SMART GRIDS</i> .....	52
3.2	TÉCNICAS DE VISUALIZAÇÃO DE INFORMAÇÃO NO CONTEXTO DE <i>SMART GRIDS</i> .....	53
3.3	ALGORITMOS DE REDES NEURAI PARA PREVISÃO NO CONTEXTO DE <i>SMART GRIDS</i> .....	55
<b>4</b>	<b>Material e métodos .....</b>	<b>57</b>
4.1	MATERIAL .....	58
4.1.1	<i>Linguagens de programação, frameworks e bibliotecas</i> .....	58
4.1.2	<i>Dados: fontes e ferramentas</i> .....	58
4.1.3	<i>Infraestrutura e hardware</i> .....	59
4.2	DESENVOLVIMENTO DA APLICAÇÃO .....	60
4.2.1	<i>Back-end: gerenciamento de entidades virtuais e algoritmos</i> .....	61
4.2.2	<i>Elementos básicos para monitoramento</i> .....	63
4.2.3	<i>Representação da rede de distribuição com visualização em árvore</i> .....	65
4.2.4	<i>Layout para visualização em coordenadas paralelas</i> .....	67
4.2.5	<i>Previsão de carga</i> .....	71
<b>5</b>	<b>Resultados e discussões .....</b>	<b>77</b>
5.1	CASOS DE USO.....	77
5.1.1	<i>Monitoramento</i> .....	77
5.1.1.1	Escalabilidade.....	77
5.1.1.2	Visualização de estados.....	78
5.1.2	<i>Análises de dados em coordenadas paralelas</i> .....	82
5.1.2.1	Dados estáticos .....	82
5.1.2.2	Dados coletados.....	85
5.1.3	<i>Previsão de carga</i> .....	87
5.1.3.1	Caso 1: <i>outliers</i> de erro absoluto relativo e análise de erros percentuais.....	87
5.1.3.2	Caso 2: medidas escassas e intervalo de entrada dos modelos .....	92
5.2	DISCUSSÕES.....	95
<b>6</b>	<b>Conclusão e trabalhos futuros.....</b>	<b>98</b>
6.1	CONCLUSÃO .....	98
6.2	TRABALHOS FUTUROS .....	99
	<b>Referências .....</b>	<b>101</b>
	<b>Apêndice A: Lista de publicações .....</b>	<b>118</b>

# 1 Introdução

## 1.1 Apresentação

A quantidade de dados criados, coletados ou copiados deve chegar a impressionantes 175 ZB em 2025, o que equivale a cinco vezes mais que em 2018 [1]. Dentre os fatores que impulsionam esta tendência de crescimento, estão as tecnologias ligadas ao conceito de internet das coisas (*internet of things* – IoT). Seja com a implantação de grandes projetos, como cidades inteligentes, ou com a popularização de pequenos dispositivos, essas tecnologias devem viabilizar a coleta de grandes volumes de dados.

Com a difusão e integração de elementos como sensores e atuadores, por meio de redes de comunicação e sistemas de computação, as tecnologias de IoT viabilizam a prestação de melhores serviços e produtos, e geram novos modelos de negócio [2]. Além disso, a utilização destas tecnologias busca solucionar diversos desafios, atuais e futuros. Por isso, vêm sendo aplicadas em diversas áreas.

Nas cidades, podem melhorar a mobilidade, a segurança pública, serviços de saúde, e aumentar a eficiência energética por exemplo [3]. No setor agropecuário, podem auxiliar no uso eficiente dos recursos naturais e insumos, bem como na automação de tarefas [4]. Nas indústrias podem aumentar a integração de fornecedores, a rastreabilidade de produtos e a eficiência em diversos processos [4].

No setor elétrico, os conceitos de IoT se concretizam nas chamadas *smart grids* (SGs), as redes inteligentes. Estas redes representam uma expressiva evolução das redes elétricas atuais, com a introdução de funções como monitoramento, controle à distância e automação em larga escala. Para isso, redes de comunicação devem ser utilizadas para permitir o sensoriamento e a atuação remota. Com os dados coletados, visa-se garantir a qualidade do fornecimento de energia elétrica, viabilizar novos serviços e solucionar desafios do setor.

Neste contexto, o emprego de tecnologias de IoT pode permitir a implementação de *microgrids* [5], diminuir as perdas não técnicas [6], além de auxiliar na manutenção dos indicadores de qualidade do fornecimento de energia elétrica. Essas tecnologias também podem auxiliar na solução de novos desafios do setor: a utilização de fontes renováveis e intermitentes de energia, que aumenta a complexidade do planejamento e operação dos sistemas de energia elétrica; a popularização de sistemas de geração distribuída, que representa um desafio aos

sistemas concebidos considerando um paradigma de geração centralizada; e a adoção de carros elétricos, que impulsiona ainda mais a crescente demanda por energia elétrica.

Para que a implantação dos sistemas de IoT possa alcançar os objetivos almejados, solucionando desafios e viabilizando novos produtos e serviços, é necessário que os dados coletados sejam transformados em informações úteis. Para isso, é fundamental que sejam utilizados algoritmos para análise e sintetização de novas informações, além de formas eficientes de visualização dos dados coletados e sintetizados.

Neste cenário, algoritmos de inteligência artificial (IA), como redes neurais, que “aprendem” por meio de dados, podem representar uma abordagem interessante, uma vez que aproveitam o volume de dados gerados em sistemas de IoT [7, 8]. Considerando a quantidade e a complexidade das informações geradas e sintetizadas, as técnicas de visualização de informação também representam um importante recurso nestes sistemas [9, 10]. Tais técnicas auxiliam no monitoramento e na análise de dados. E além de apresentarem as informações coletadas e sintetizadas, podem auxiliar ainda na avaliação dos algoritmos utilizados.

Assim, considerando a importância do setor elétrico, o potencial das *smart grids* frente aos desafios do setor, e a necessidade de utilizar algoritmos e formas de representação eficientes para o sucesso na implantação das SGs, neste trabalho, propõe-se implementar um sistema com funcionalidades de monitoramento, análise e visualização, baseado em conceitos de virtualização de entidades (*digital twins*, entidades virtuais, *digital shadows*). Para isso, são utilizadas técnicas de visualização de informação e algoritmos de redes neurais.

Este trabalho faz parte de um projeto de pesquisa e desenvolvimento (P&D), desenvolvido em parceria com a concessionária Neoenergia Distribuição Brasília. O projeto busca desenvolver um sistema piloto alinhado aos conceitos de *smart grids* e IoT. Para isso, outros trabalhos foram realizados com diferentes propósitos, como o estudo de tecnologias de comunicação [11], dispositivos [12], análises e algoritmos com foco em transformadores de distribuição [13] etc. Estas pesquisas se integram, e contribuem para o desenvolvimento de uma plataforma de IoT específica para o setor elétrico

## 1.2 Objetivos da pesquisa

Considerando a importância das *smart grids* para superar os desafios apresentados, e a complexidade envolvida na implementação de tais redes, com foco em transformar os dados coletados em informações úteis e apresentá-las de maneira eficiente, esta pesquisa tem o seguinte objetivo principal: apresentar uma proposta de sistema de monitoramento e análise

para sistemas de distribuição, considerando entidades virtuais, utilizando técnicas de visualização de informação, bem como algoritmos de inteligência artificial, para apresentar informações, coletadas e sintetizadas, de forma eficaz.

Desta maneira, em decorrência de tal objetivo e de outros aspectos do contexto deste trabalho, a pesquisa tem ainda os seguintes objetivos secundários:

- verificar a capacidade do sistema implementado de gerenciar e monitorar várias entidades virtuais, considerando o tamanho e a heterogeneidade de sistemas de distribuição reais;
- exemplificar como algumas técnicas de visualização de informação podem ser aplicadas no contexto de redes de distribuição e *smart grids*;
- explorar técnicas de inteligência artificial, verificando o potencial de aplicação destas técnicas no contexto de uma ferramenta computacional para *smart grids*;
- utilizar algoritmos de previsão de carga com o intuito de demonstrar como a aplicação implementada pode ser utilizada para gerar e apresentar informações úteis;
- verificar o potencial de aplicação de técnicas de visualização de informação, para avaliar algoritmos de previsão.

## 1.3 Limitações, convenções e escopo

Para que os objetivos propostos nesta pesquisa pudessem ser alcançados, foi necessário delimitar seu escopo e estabelecer convenções. Dentre as principais limitações, pode-se citar:

- utilizou-se dados de apenas dois elementos reais monitorados. Os dados para os outros elementos da rede foram restritos a uma base de dados geográfica da distribuidora (BDGD)<sup>1</sup>. Para compensar tal limitação e possibilitar, por exemplo, os testes com um número de entidades semelhante ao de redes de distribuição reais, foram utilizadas pseudomedidas;
- as entidades virtuais relacionadas aos elementos do sistema de distribuição não possuem todas as funcionalidades relacionadas ao conceito de *digital twins* (DTs), como capacidades plenas de atuação e modelos que consideram todas as características físicas das entidades reais;

---

<sup>1</sup> A BDGD é uma base de dados padronizada pela Agência Nacional de Energia Elétrica (ANEEL). Todas as concessionárias de energia elétrica enviam uma versão desta base anualmente, para a ANEEL, com dados de seus ativos. Mais informações sobre a estrutura da BDGD podem ser consultadas nos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST), módulo 10 [146].

- o dinamismo das conexões da rede elétrica de distribuição não pôde ser reproduzido, porque não se teve acesso aos dados supervisórios em tempo real. Assim, a representação da rede mostrada por meio da visualização em árvore (ver Figura 19) possui uma estrutura fixa, com base na BDGD;
- a complexidade das ligações da rede elétrica de distribuição não pôde ser representada em sua totalidade, uma vez que os dados cadastrais obtidos não possuíam as informações necessárias para a reconstrução total da estrutura da rede. Assim, a representação da rede mostrada por meio da visualização em árvore, apresenta os principais elementos da rede, mas não todos os tipos ou conexões reais entre os elementos. Ramais de ligação e configurações não radiais, por exemplo, não estão presentes na árvore;
- a implementação de uma aplicação com funções de *middleware* e, simultaneamente, de análise e visualização de dados, ao invés de distribuir tais funções em várias aplicações, não representa uma arquitetura ideal para sistemas de IoT, mas foi adotada por ser mais adequada à prototipação no contexto deste trabalho;
- não é objetivo deste trabalho encontrar os melhores modelos de previsão de carga, mas sim verificar o potencial da ferramenta para a execução e avaliação dos modelos, especialmente com o uso de técnicas de visualização. Para isso, utilizou-se os algoritmos de redes neurais recorrentes, que vêm sendo empregados na previsão de carga de curto prazo [14, 15, 16, 17, 18].

Com relação às convenções adotadas, pode-se citar:

- as medidas de curva de carga obtidas, de poucos elementos monitorados, são adequadas para o propósito deste trabalho;
- as entidades virtuais implementadas são adequadas ao escopo deste trabalho, uma vez que permitem que o sistema execute funcionalidades de monitoramento e análise;
- a hierarquia da rede, apresentada por meio da visualização em árvore, é suficiente para a utilização da ferramenta e apresentação de seu potencial, uma vez que são retratados os principais elementos da rede;
- a arquitetura adotada, com uma única aplicação, com funções de *middleware*, análise e visualização de dados, é adequada para exemplificar o uso de entidades virtuais, algoritmos de inteligência artificial e ferramentas de visualização em sistemas de IoT.

Além disso, simplifica a implementação do trabalho, permitindo que os objetivos principais sejam atingidos;

- os algoritmos de previsão utilizados não visam esgotar as possibilidades de algoritmos a serem integrados à aplicação, mas apenas ilustrar o potencial no contexto de IoT, em conjunto com técnicas de visualização.

## 1.4 Justificativa

A eletricidade possui notória importância para a humanidade. Esta importância pode ser observada pelo considerável e crescente volume de energia elétrica utilizado. A geração de energia elétrica global ultrapassou 26000 TWh em 2019, representando um aumento de cerca de 400% em relação a 1973 [19]. Em 2018, o consumo de energia aumentou cerca de 4%, o que representa o maior crescimento anual desde 2010 [20]. Observam-se ainda, complexas relações entre crescimento econômico, consumo de eletricidade e eficiência energética [21].

Por outro lado, questões ambientais relacionadas aos combustíveis fósseis impulsionam o uso de fontes alternativas e intermitentes de energia, bem como a adoção de carros elétricos. Desta forma, o setor elétrico possui um papel importante na redução das emissões de gases do efeito estufa [20].

Assim, é fundamental que novas pesquisas sejam realizadas para que o setor se torne ainda mais eficiente e supere os desafios atuais e futuros. Considerando as tecnologias de IoT como importantes alternativas para solução dos desafios do setor, pesquisas relacionadas ao contexto de *smart grids* também são de grande relevância.

E considerando que a geração de informação útil é vital para o sucesso na implantação de sistemas de IoT, a utilização de ferramentas computacionais e algoritmos para manipulação e extração de dados é um aspecto de estudo importante neste contexto. Além disso, formas eficientes para a representação destes dados constituem outro pilar para o sucesso dos sistemas de IoT.

Desta forma, justifica-se o trabalho proposto, que visa implementar uma aplicação com funções de monitoramento, análise e visualização de dados no contexto de sistemas de distribuição e *smart grids*.

## 1.5 Contribuição

Esta pesquisa pretende contribuir com um exemplo de aplicação para sistemas de distribuição, considerando três áreas de pesquisa: internet das coisas e *smart grids*; visualização de informação; e inteligência artificial. A aplicação implementada representa uma base para futuras pesquisas e sistemas, materializando uma ampla gama de conceitos relacionados ao contexto de SGs.

No contexto de internet das coisas, destaca-se a utilização de conceitos relacionados aos *digital twins*. Além disso, a implementação de funções de análise e visualização de dados representam uma importante contribuição, uma vez que, dentre os principais objetivos de sistemas de IoT, está geração de informação útil a partir de dados coletados.

Considerando as *smart grids*, a aplicação de técnicas de visualização de informação e algoritmos de inteligência artificial representa uma contribuição deste trabalho. A necessidade de conhecimentos acerca de sistemas de distribuição, algoritmos de IA e técnicas de visualização de informação, inibe a realização de pesquisas abordando estes tópicos em conjunto. Por outro lado, a utilização destes algoritmos e técnicas neste contexto, pode tornar a implantação de SGs mais atrativa e difundida.

A aplicação de técnicas de visualização de informação para avaliação dos algoritmos de previsão utilizados, constitui outra contribuição. A utilização de coordenadas paralelas (PCs) para avaliação de modelos de previsão possui caráter inovador e relevante para múltiplos tópicos de estudo: regressão, inteligência artificial, previsão e visualização de informação. Neste sentido, as PCs permitem visualizar vários atributos, incluindo métricas de erro, em nível de amostras.

## 1.6 Organização do texto

Este trabalho é composto por seis capítulos. Após este capítulo introdutório, apresenta-se, no Capítulo 2, a fundamentação teórica sobre os principais tópicos abordados neste trabalho. São apresentados os conceitos de IoT, *smart grids* e virtualização de entidades, bem como aspectos básicos das redes de distribuição. Tópicos sobre ferramentas de visualização de informação, com foco nas técnicas implementadas neste trabalho, também são abordados. Por fim, apresentam-se contextos sobre algoritmos de previsão e redes neurais.

No Capítulo 3 apresenta-se uma breve relação dos trabalhos correlatos encontrados na literatura. Considerando a diversidade de tópicos abrangidos por este trabalho, as pesquisas

correlatas são agrupadas conforme os diferentes tópicos: ferramentas computacionais para *smart grids*; técnicas de visualização no contexto de SGs; e algoritmos de redes neurais para previsão de carga no contexto de SGs.

No Capítulo 4, o material e as metodologias utilizadas nesta pesquisa são apresentados. Entre os recursos utilizados, apresentam-se as linguagens, bibliotecas e ferramentas de programação utilizadas, as fontes de dados, bem como a infraestrutura fornecida pelo projeto que este trabalho faz parte. Os métodos empregados são apresentados conforme os elementos e funcionalidades da aplicação, como gerenciamento de entidades virtuais e algoritmos, a interface gráfica e as técnicas de visualização empregadas, e a funcionalidade de previsão de carga.

O Capítulo 5 inicia com a apresentação dos resultados alcançados neste trabalho. Em seguida, são realizadas discussões acerca destes resultados com relação aos encontrados na literatura e suas limitações.

Finalmente, o Capítulo 6 apresenta as conclusões obtidas, bem como as sugestões de trabalhos futuros, considerando as limitações e desafios encontrados neste trabalho.

No Apêndice A é mostrada a lista de publicações diretamente relacionadas a esta pesquisa.



## 2 Fundamentação teórica

Neste capítulo são apresentados os principais tópicos necessários para compreensão deste trabalho. Primeiramente, os conceitos relacionados a IoT, *smart grids* e virtualização de entidades são apresentados. Em seguida, tópicos sobre as redes de distribuição são apresentados.

Na subseção seguinte são apresentados alguns conceitos sobre ferramentas de visualização de informação, com destaque para as técnicas empregadas neste trabalho. Então, apresentam-se alguns aspectos a respeito da previsão de carga, como métricas utilizadas para qualificar os algoritmos de previsão. Por fim, apresenta-se uma breve introdução às redes neurais.

É importante destacar que uma apresentação profunda destes tópicos foge do escopo deste trabalho, podendo ser encontrada nas referências indicadas.

### 2.1 IoT e *smart grids*

#### 2.1.1 *Internet of things*

Conforme [22], a expressão internet das coisas (*internet of things* - IoT) tem sua origem relacionada a uma palestra realizada em 1999, sobre RFID e cadeias de suprimentos. Neste contexto, o autor reforça a importância das “coisas”, e destaca a limitação humana para coletar, fornecer e analisar dados.

Em duas décadas, a disponibilidade de sensores, atuadores, redes de comunicação, bem como a capacidade computacional de diversos dispositivos, aumentaram de tal forma, que o conceito de IoT se tornou amplo e contempla uma grande diversidade de tecnologias e áreas. No entanto, o objetivo original de otimizar diversos processos ao fazer com que objetos se comuniquem, permanece pertinente ao conceito [23].

Assim, o termo internet das coisas, definido de variadas formas na literatura, representa uma idealização em que tudo é conectado. Isto é, a “internet” antes restrita a poucos dispositivos, principalmente computadores, passa a estar disponível em vários objetos, o que viabiliza diversas funcionalidades, como o sensoriamento, o controle à distância, a automação de processos, e a integração de diversos dispositivos e pessoas [24].

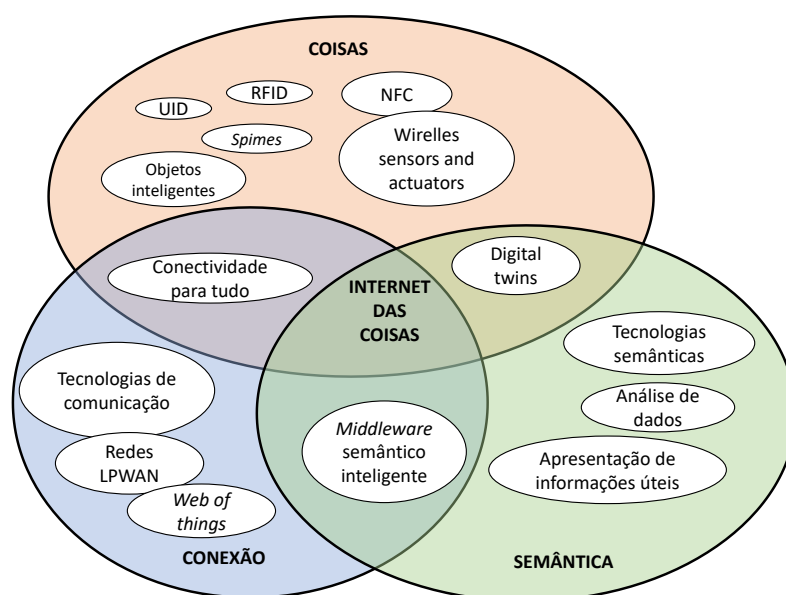
Neste sentido, IoT pode ser entendida como uma extensão da internet atual, em que diversos objetos passam a ter capacidade computacional e de comunicação. Tal expansão além de viabilizar diversas funcionalidades, também impõe desafios técnicos e sociais [25].

Além dos obstáculos que dificultam a consolidação e popularização dos sistemas de IoT, mais desafios surgem conforme os sistemas são implementados. A dificuldade para implantar sistemas de comunicação devido aos custos, à legislação ou à localização, e a segurança necessária para a proteção de dados e infraestruturas, são exemplos destes desafios [26, 25].

### 2.1.1.1 Elementos e arquiteturas

Em [27] os autores destacam, como aspecto básico de IoT, a presença massiva “coisas”, como sensores, atuadores e celulares, que são capazes de interagir e cooperar entre si para atingir objetivos comuns. Neste contexto, os autores definem três pilares de IoT: coisas, conexão e semântica. A conexão é necessária para que haja interação, enquanto a semântica representa os objetivos atingidos, isto é, as informações úteis geradas. Estes pilares são apresentados na Figura 1.

Figura 1 – Paradigmas de IoT.



Fonte: adaptado de [27].

De maneira complementar, em [28] os autores apresentam três categorias de elementos importantes em sistemas de IoT: *hardware*, que consiste em sensores, atuadores, dispositivos de comunicação embarcados etc.; *middleware*, representado por armazenamento e ferramentas computacionais para condicionamento e análise de dados; e apresentação, que consiste em

ferramentas de visualização e interpretação, preferencialmente acessíveis em várias plataformas e adequadas às aplicações.

No que se refere a arquiteturas, o modelo de três camadas mostrado na Figura 2 está presente em diversas fontes [29, 30, 31, 32, 12]. Tal modelo é genérico, mas também flexível para representar os sistemas de IoT, uma vez que está diretamente relacionado aos elementos básicos destes sistemas, destacados anteriormente.

Figura 2 – Arquitetura de 3 camadas para sistemas de IoT.



Fonte: o autor.

A camada física, ou de percepção, representa os sensores, atuadores e outros dispositivos. Ou seja, é a camada em que os dados são coletados, e interações com o mundo real são realizadas.

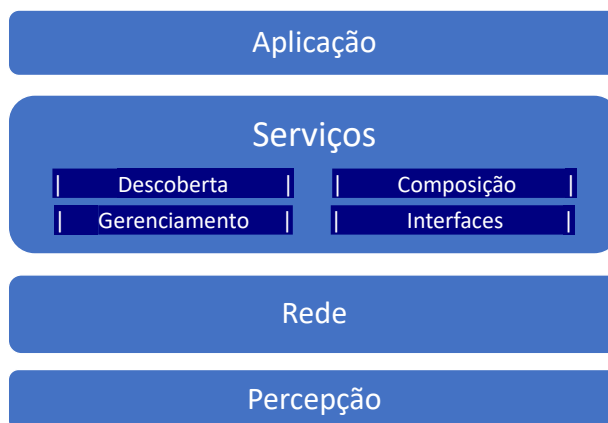
Já a camada de rede, consiste em infraestruturas, tecnologias e protocolos de comunicação, que podem apresentar elevado grau de diversidade [31]. Assim, conecta o mundo real com o virtual, além de viabilizar a interação entre os próprios dispositivos.

A camada de aplicação, por sua vez, representa os elementos utilizados para execução de diversas funções requisitadas pelos usuários, como consultas a banco de dados, *backup*, realização de análises, e execução de algoritmos relacionados aos diversos campos de aplicação de IoT, como saúde, *smart grids*, agricultura, indústria e automação residencial [29, 31, 32, 12].

Algumas camadas foram adicionadas a este modelo, com o intuito de representar outros elementos e funcionalidades úteis em sistemas de IoT. O modelo de quatro camadas baseado em serviços, apresentado na Figura 3, visa coordenar diferentes serviços e facilitar o reuso de *software* e *hardware* [32]. A camada de serviços é comumente dividida em duas subcamadas,

de composição e de gerenciamento [31, 30], ou ainda em quatro, considerando também as subcamadas de descoberta e de interfaces de serviços [31, 12, 32].

Figura 3 – Arquitetura de 4 camadas baseada em serviços para sistemas de IoT.



Fonte: o autor.

A arquitetura de cinco camadas da Figura 4 foi idealizada considerando a crescente necessidade de escalabilidade, interoperabilidade e confiabilidade em sistemas de IoT [32, 33]. Esta arquitetura possui duas camadas adicionais em relação ao modelo genérico de três camadas: *middleware*, também chamada de camada de processamento [29]; e a camada de *business*.

Figura 4 – Arquitetura de 5 camadas com *middleware* para sistemas de IoT.



Fonte: o autor.

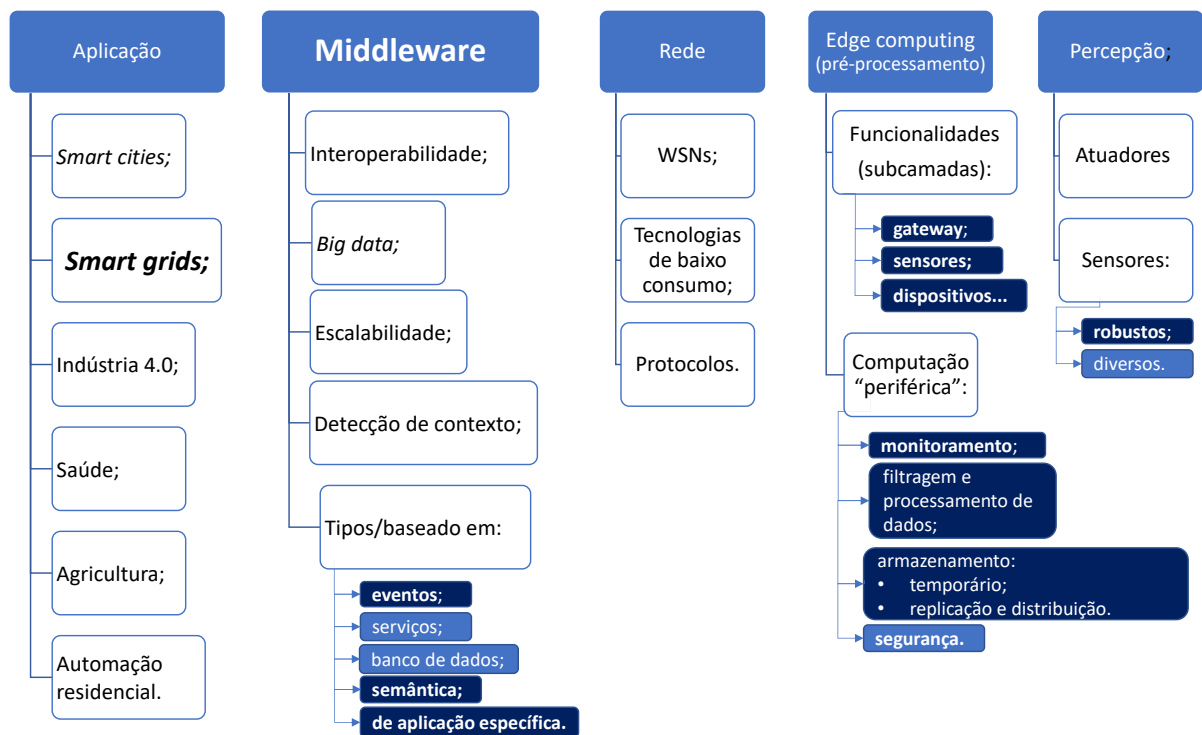
Em geral, o *middleware* é um serviço, ou um conjunto de serviços, que fornece abstração entre as tecnologias de IoT e as aplicações [32]. Isto é, fornece uma interface padronizada que tornam os diferentes tipos de comunicação, transparentes às aplicações. Este elemento normalmente processa grandes volumes de dados. Possui funções como agregar, filtrar, padronizar e armazenar dados enviados pelos dispositivos [29, 32]. A camada de serviços, da arquitetura anterior, é frequentemente associada à camada de *middleware*, uma vez que as

ambas as camadas visam contribuir para aumentar a interoperabilidade entre sistemas e o reuso de componentes [31, 12, 30, 29].

Outras arquiteturas em destaque nas pesquisas sobre IoT, são as que incorporam elementos de pré-processamento, representados pelos termos *edge computing* e *fog computing* [29]. Nelas, *gateways*, sensores e outros dispositivos possuem funções como filtrar, processar ou armazenar dados, descentralizando, em certa medida, o processamento realizado no *middleware*. Isto favorece a escalabilidade e pode aumentar a eficiência dos sistemas, uma vez que, por exemplo, dados desnecessários deixam de ser enviados aos servidores.

A Figura 5 mostra um exemplo de arquitetura com elementos de *edge computing*. Nesta figura estão presentes outras camadas e elementos de sistemas de IoT, e são destacados alguns aspectos presentes no contexto deste trabalho, com intuito exemplificativo. Estes aspectos serão mostrados na Seção 4.

Figura 5 – Arquitetura de 5 camadas com *edge computing* para sistemas de IoT.



Fonte: o autor.

Sobre os tipos de *middleware* mostrados na Figura 5, temos: baseado em eventos, em que componentes geram e recebem notificações, numa arquitetura *publish-subscribe*; baseado em serviços, em que módulos independentes disponibilizam serviços por meio de interfaces, frequentemente com repositório e descoberta de serviços; baseado em banco de dados, em que a rede de dispositivos é uma base de dados virtual que pode ser consultada; semântico, que

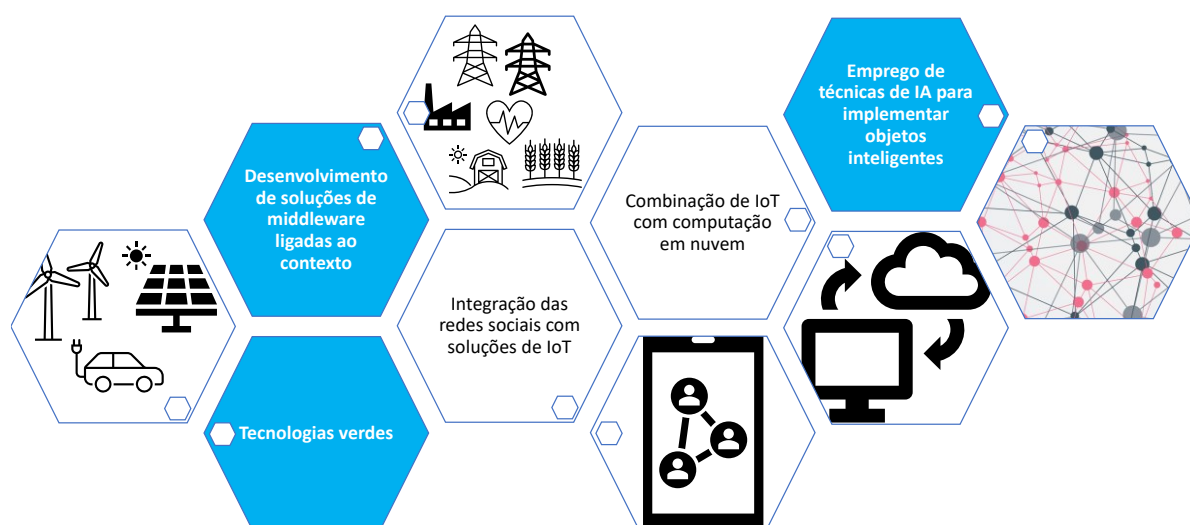
viabiliza a interoperabilidade entre diferentes dispositivos e protocolos por meio de adaptadores; e de aplicação específica, utilizado para que aja uma melhor integração entre o *middleware* e as aplicações [29]. É interessante observar que estas classificações não são rígidas nem mutuamente exclusivas.

### 2.1.1.2 Tendências para os sistemas de IoT

Em [34], os autores destacam cinco tendências para os sistemas de IoT, mostradas na Figura 6. Dentre tais tendências, o desenvolvimento de soluções de *middleware* e a aplicação de inteligência artificial estão diretamente relacionadas ao contexto deste trabalho.

Além disso, o desenvolvimento de tecnologias “verdes” está indiretamente relacionado, uma vez que a aplicação de tecnologias de IoT em sistemas de energia elétrica, viabiliza uma melhor integração e utilização de fontes de energia renováveis [35], como a solar e a eólica. Além disso, auxilia nos desafios relacionados à utilização de carros elétricos e à demanda de energia elétrica.

Figura 6 – Tendências no contexto de IoT.



Fonte: adaptado de [34, 24].

No Brasil, a fim de consolidar uma visão estratégica com relação às tecnologias de IoT, o Banco Nacional do Desenvolvimento (BNDES) e o Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC) viabilizaram a realização de um estudo intitulado “Internet das Coisas: um plano de ação para o Brasil” [36]. Nesse estudo, considerando critérios relacionados à demanda, à oferta e à capacidade de desenvolvimento, foram selecionadas quatro verticais prioritárias: cidades, saúde, rural e indústrias [26, 37]. O estudo aponta que, em tais verticais, as tecnologias de IoT terão um impacto econômico considerável em 2025: nas cidades

este impacto seria de US\$ 13 a 27 bilhões; na agricultura seria de US\$ 5 a 21 bilhões; na saúde de US\$ 5 a 39 bilhões; e na indústria de US\$ 11 a 45 bilhões [26].

Além disso, foram elencados os principais desafios para o desenvolvimento da IoT no Brasil: capital humano capacitado; inovação e inserção internacional do país no setor; infraestruturas de conectividade e interoperabilidade; e questões regulatórias relacionadas à segurança e privacidade [26].

Outras iniciativas governamentais para incentivar o desenvolvimento de tecnologias e sistemas de IoT incluem o Decreto nº 9.854 de 2019 [38] e a Lei nº 14.108 de 2020 [39].

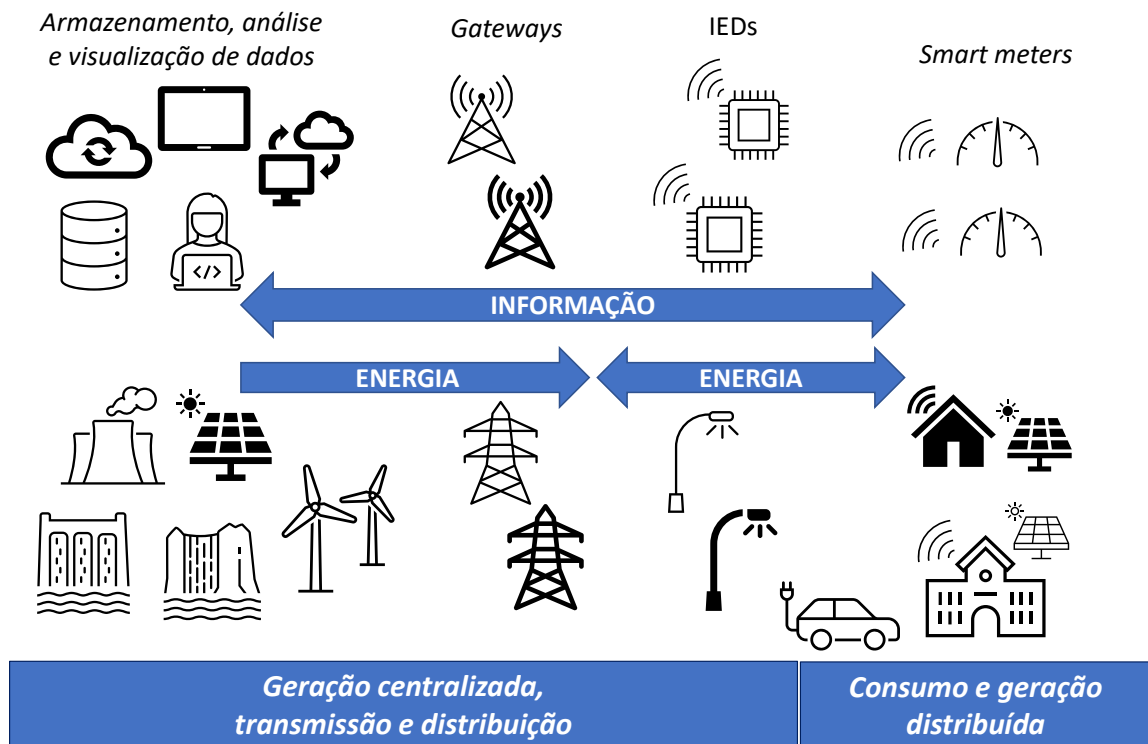
### **2.1.2 Smart grids**

*Smart grids* consistem em redes de comunicação integradas às redes de energia elétrica, para coletar e analisar dados de equipamentos, linhas de transmissão e distribuição, bem como de unidades consumidoras (UCs) [40]. Como outros campos de aplicação de IoT, a implementação de SGs visa gerar informações úteis e possibilitar melhores serviços por meio da disponibilização e análise de dados, automação etc. [3]. Assim, *smart grids* representam a aplicação de IoT no setor elétrico, e possuem, além dos elementos genéricos de sistemas de *internet das coisas*, aspectos e componentes únicos relacionados aos sistemas de energia elétrica.

#### **2.1.2.1 Elementos básicos**

A utilização de redes de comunicação e dispositivos inteligentes em *smart grids*, permite que haja um fluxo de dados bidirecional entre as concessionárias e os consumidores de energia elétrica. Assim, capilariza o monitoramento, estendendo-o da geração e transmissão, à distribuição e consumo.

Na Figura 7 é mostrado um diagrama com alguns dos elementos presentes em *smart grids*. Na parte superior da figura são mostrados os elementos que viabilizam a coleta, transporte, armazenamento, análise e visualização de dados. Ou seja, os elementos que estabelecem o fluxo bidirecional de dados e transformam a rede de energia elétrica em “inteligente”. A infraestrutura formada por estes elementos é chamada de *advanced metering infrastructure* (AMI) [41].

Figura 7 – Diagrama simplificado de uma *smart grid*.

Fonte: o autor.

Os *smart meters* (SMs) permitem as concessionárias coletarem dados instantâneos de demanda dos consumidores. Assim, facilitam a implantação de tarifas em tempo real e permitem aos consumidores participarem ativamente da rede, seja adaptando sua demanda em tempo e volume, ou por meio da geração distribuída [42]. Os SMs representam, desta forma, uma evolução em relação aos medidores de energia convencionais, e são um dos elementos base para as AMIs [43].

Além das unidades consumidoras, outros elementos da rede podem ser monitorados: transformadores de distribuição, subestações, religadores, reguladores de tensão etc. Para isso, são utilizados dispositivos com capacidades de comunicação, computação, sensoriamento e/ou atuação, como os chamados *intelligent electronic devices* (IEDs) [44].

Para que os dados das UCs e de outros elementos da rede de distribuição possam chegar às aplicações nos servidores, é necessário que se utilize algum meio de comunicação. Dentre as alternativas de comunicação para as *smart grids*, estão a fibra óptica, a tecnologia *power line communication* (PLC), e as redes *wireless*, com destaque para as *low power wide area networks* (LPWANs) [45, 23].

As LPWANs são um grupo de tecnologias e protocolos que permitem uma comunicação de longo alcance, isto é, na ordem de quilômetros, e um baixo consumo de energia [23]. O



longo alcance e o baixo custo em relação a outras alternativas de comunicação, tornam as LPWANs extremamente interessantes para a implantação de *smart grids*, uma vez que a rede de distribuição é extensa, e o custo se torna um fator limitante para monitorar elementos numerosos, como unidades consumidoras e transformadores de distribuição.

Algumas destas tecnologias operaram em topologia estrela, em que *gateways* recebem mensagens de dispositivos, e as redirecionam para os servidores. Estes elementos também são chamados de concentradores, e estão representados na Figura 7.

Na parte inferior da Figura 7, são mostrados os elementos que permitem o fluxo de energia. Os carros elétricos e as fontes de energia renováveis, intermitentes e/ou distribuídas, têm seu uso impulsionado por questões ambientais. Por outro lado, representam novos desafios para o setor elétrico.

A geração distribuída resulta em um fluxo de energia também bidirecional em algumas partes da rede elétrica, que antes estavam sujeitas a um fluxo unidirecional. Já os carros elétricos tendem a estimular ainda mais a crescente demanda por energia elétrica, além de tornar os padrões de consumo mais complexos, uma vez que, por exemplo, possuem mobilidade.

### **2.1.2.2 Benefícios da implantação de *smart grids***

O setor elétrico enfrenta diversos desafios, como integrar as fontes de energia renováveis, mitigar as perdas não técnicas, e manter – ou ainda melhorar – os índices de qualidade. Além disso, devido a fatores como o aumento do uso de fontes renováveis e distribuídas de energia, e a adoção de carros elétricos, o setor elétrico precisa superar desafios ainda mais complexos.

Diante destes desafios, a adoção de tecnologias relacionadas à IoT, como o telemonitoramento, permite uma melhor integração de fontes intermitentes de energia, além de auxiliar na geração de energia distribuída. Neste contexto, viabilizam inclusive a precificação dinâmica da energia elétrica considerando dados de demanda e disponibilidade instantâneos [3]. Além disso, possibilitam monitorar aspectos de qualidade de energia elétrica, detectar comportamento anormais, como fraudes, e realizar desligamentos e religamentos remotos e automáticos [3, 46, 47].

### **2.1.3 Virtualização: entidades virtuais**

Entidades virtuais, objetos virtuais ou gêmeos digitais (*digital twins* - DTs) representam um modo avançado e completo de se monitorar e gerenciar sistemas complexos [48]. DTs são,

portanto, um meio de realizar a fusão do mundo físico com o virtual, e fornecer uma interface eficiente entre os vários elementos dos sistemas de IoT [49, 50].

O conceito de *digital twins* foi introduzido em 2003, e tem sua origem relacionada à Michael Grieves e John Vickers [49, 51]. Com o advento de tecnologias relacionadas à IoT, este conceito evoluiu. A coleta de dados das entidades, antes realizadas manualmente ou com pouca frequência, passou a ser realizada por sensores com medidas frequentes e até instantâneas [51]. Além disso, os *digital twins*, antes apenas descritivos, passaram a ter capacidade de atuação.

Para implementar DTs em *smart grids*, o primeiro passo consiste em criar o modelo para as instâncias de gêmeos digitais (*digital twin instances* - DTIs) com base nas medidas feitas nos elementos da rede [52, 53]. Dentre as possíveis informações de interesse pode-se citar a potência ativa, a potência reativa, o fator de potência, o consumo de energia, a temperatura de operação no caso de equipamentos de potência, índices de distorção harmônica etc.

O segundo passo consiste na integração do modelo com os elementos reais da rede, utilizando sensores e atuadores. Os sensores são utilizados para coletar as informações, enquanto os atuadores permitem ações de controle e otimização. Neste contexto, é necessário fornecer uma infraestrutura para processar e agregar os dados de monitoramento recebidos dos sensores [52].

O terceiro passo visa fornecer o suporte tecnológico necessário para que sejam utilizados algoritmos com os dados coletados. Neste contexto, alguns desafios estão relacionados ao fato de *smart grids* gerarem uma expressiva quantidade de dados heterogêneos, com potencial presença de ruído [48, 52].

Ao sincronizar com precisão as DTIs e seus objetos reais, informações e funcionalidade interessantes podem ser obtidas, como o monitoramento de elementos da rede, previsão de energia, funções relacionadas à segurança e proteção, detecção de erros etc. [52].

Neste contexto, tecnologias e algoritmos de *machine learning* podem viabilizar funções importantes em SGs, por meio dos dados coletados e disponibilizados por DTs, como previsão de carga eficiente, otimizações, suporte a decisões e detecção de padrões [52]. Ainda neste contexto, tecnologias de realidade aumentada (*augmented reality* - AR) e/ou visualização de informação (*information visualization* – InfoVis) podem ser utilizadas para melhorar o entendimento acerca dos elementos do sistema elétrico [54, 55, 52].

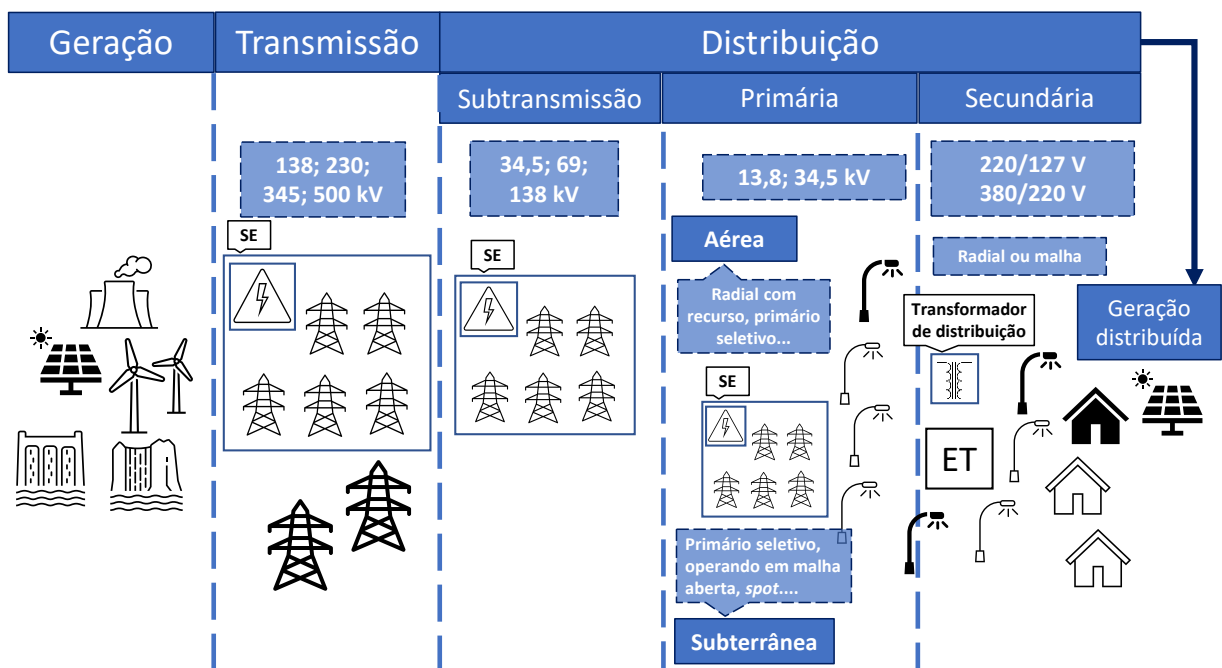
## 2.2 Redes de distribuição

### 2.2.1 Sistemas elétricos de potência

Os sistemas elétricos de potência (SEPs) consistem em um grande conjunto de infraestruturas, equipamentos e processos que possibilitam fornecer energia elétrica às unidades consumidoras, geralmente, de maneira interrupta e padronizada [56, 57, 58]. Os SEPs podem ser divididos em três subsistemas: geração, transmissão e distribuição [56, 57].

A Figura 8 apresenta um esquema de um SEP. Tal figura, não representa a totalidade e a complexidade destes sistemas. Busca, por outro lado, representar a rede elétrica de maneira simplificada, apresentando alguns aspectos importantes, principalmente sobre sistemas de distribuição. Os níveis de tensão padronizados para cada subsistema, por exemplo, são mostrados na figura, mas há níveis de tensão não padronizados utilizados nestes subsistemas [56].

Figura 8 – Esquema de um sistema elétrico de potência.



Fonte: o autor.

Na Figura 8, o termo geração se refere aos sistemas que realizam a conversão de diferentes formas energia, por exemplo, mecânica ou luminosa, para energia elétrica. Em usinas hidrelétricas, termelétricas ou nucleares, por exemplo, utiliza-se máquinas elétricas rotativas, geradores síncronos ou alternadores, que são movimentados por turbinas, hidráulicas ou a vapor, convertendo o conjugado mecânico em energia elétrica [56]. Na geração fotovoltaica,

painéis convertem a energia luminosa para elétrica diretamente, devido ao efeito fotovoltaico [59].

Para permitir que a geração ocorra de maneira centralizada e/ou distante dos centros de consumo, como em grandes usinas hidrelétricas, o subsistema de transmissão é utilizado para transportar a energia. Os níveis de tensão obtidos na geração de energia elétrica não são adequados para o transporte de uma potência elevada por longas distâncias. Por isso, são utilizadas subestações (SEs) elevadoras, que elevam a tensão para, por exemplo, 500kV, conforme a Figura 8 [56].

No Brasil, este sistema é extenso: somente a rede elétrica básica de transmissão tem cerca de 141756 km [60]. E além de viabilizar hidrelétricas distantes dos locais de consumo, permite que haja uma “cooperação” entre regiões em que a seca e o período chuvoso são diferentes [56]. Isto aumenta a confiabilidade e reduz os custos do sistema elétrico.

Por meio das linhas de transmissão, a energia chega às SEs de subtransmissão, que abaixam a tensão para níveis de 34,5 a 138 kV. A energia então segue para as SEs de distribuição, ou para os clientes atendidos em tensão de subtransmissão. As subestações e as linhas de subtransmissão fazem parte do chamado subsistema de subtransmissão [56].

Nas SEs de distribuição, a tensão é novamente ajustada, neste caso para níveis de 13,8 ou 34,5 kV. Então, a energia segue para os transformadores de distribuição ou unidades consumidoras, por meio dos alimentadores, que são as redes elétricas destinadas a transportar a energia em média tensão, e fazem parte da distribuição primária. Entre os consumidores normalmente atendidos nesses níveis, temos indústrias e *shopping centers* [56].

Por sua vez, os transformadores de distribuição fornecem os níveis de tensão normalmente utilizados em casas, 127 e 220 V, ou 220 e 380 V, por meio da rede distribuição secundária. Estes níveis são classificados como baixa tensão [56].

### **2.2.2 Elementos dos sistemas de distribuição**

Para que a energia seja fornecida nos diferentes níveis tensão e com qualidade e segurança, são necessários diferentes dispositivos. Entre os vários elementos que compõem os sistemas de distribuição, podemos citar os transformadores, disjuntores, chaves seccionadoras, chaves fusíveis e reguladores de tensão.

### **2.2.2.1 Dispositivos de proteção e controle**

Entre os dispositivos utilizados para a proteção, podemos citar os fusíveis, os disjuntores, os relés e os religadores. Geralmente, estes dispositivos têm a função de desconectar o trecho em que houve alguma falha, evitando ou minimizando danos em razão de sobrecorrentes.

A chave fusível tem baixo custo e interrompe a passagem de corrente permanentemente em caso de sobrecorrente. Os disjuntores, além de protegerem contra sobrecorrentes, são utilizados para controlar a passagem de energia, uma vez que podem ser ligados e desligados. Os relés podem proteger as cargas de outras anomalias, como sobrecarga, sobretensão etc. [61].

Os religadores automáticos permitem o reestabelecimento de energia de maneira automática, quando o curto-circuito deixa de atuar no trecho que o regulador está localizado. Para isso, os reguladores realizam tentativas de reconexão e, caso o curto-circuito permaneça após algumas tentativas, desconectam-se até que o problema seja resolvido. Neste caso, precisam ser reconectados manualmente. Os intervalos de tempo entre as tentativas normalmente podem ser ajustados [61].

### **2.2.2.2 Transformadores**

Os transformadores são elementos que possibilitam elevar e abaixar a tensão em sistemas de corrente alternada. Nos sistemas de distribuição, normalmente são abaixadores, e estão presentes em diferentes estágios: nas subestações, onde normalmente fornecem média tensão; ou em estações transformadoras (ETs) e em postes, nos quais, geralmente, transformadores de distribuição são alimentados em média tensão e fornecem baixa tensão. Existem transformadores de diferentes tamanhos, potências e aplicações [62].

### **2.2.2.3 Reguladores de tensão**

A tensão varia ao longo da rede elétrica. Fatores como perdas devido à extensão da rede, ou o acionamento de cargas de elevada potência, são exemplos de causas deste problema. Um dos equipamentos utilizados para que os níveis de tensão se mantenham adequados, conforme a legislação, são os reguladores de tensão. Basicamente, estes equipamentos são autotransformadores que podem elevar ou abaixar a tensão em diferentes níveis, conforme são alterados os seus *taps*. Podem ter ajuste de tensão automático [63].

### **2.2.3 Tipos de redes de distribuição**

Além da classificação inerente ao nível de tensão, as redes primárias e secundárias podem ser classificadas em aéreas e subterrâneas, e possuir diferentes configurações, com o intuito de alcançar a confiabilidade necessária para cada trecho da rede.

#### **2.2.3.1 Radial**

A configuração mais simples, chamada de radial, consiste num trecho com um único suprimento. Desta forma, esta configuração está mais sujeita a interrupções. Quando a agregação ou a densidade de cargas é maior, torna-se viável implementar suprimentos alternativos, isto é, caminhos que permitam o fornecimento de energia em caso de falhas no trecho de suprimento usual.

#### **2.2.3.2 Radial com recurso**

As redes primárias possuem um tronco principal e derivações. Estas derivações normalmente são protegidas por fusíveis e possuem dispositivos que possibilitam isolar trechos, como chaves de seccionamento, viabilizando manutenções e outros procedimentos. Além destas chaves normalmente fechadas (NFs), com o intuito de possibilitar transferências de carga, é usual utilizar chaves normalmente abertas (NAs) entre circuitos ou trechos diferentes [56].

Em caso de falha, chaves NF podem ser abertas para isolar os trechos com defeito e chaves NA podem ser fechadas para reestabelecer a energia para trechos a jusante. Esta configuração é chamada de primário radial com recurso, ou radial com socorro, e geralmente é utilizada apenas em redes aéreas [56].

É importante ressaltar que os circuitos que passam a alimentar outras cargas devem ser dimensionados para isso. Desse modo, o carregamento em operação normal dos circuitos deve ficar limitado, de tal forma que ao receber novas cargas, não ultrapasse o limite térmico do circuito. Para aumentar o carregamento em operação normal, podem ser utilizados mais de um circuito de “socorro”, dividindo a carga transferida [56].

#### **2.2.3.3 Primário seletivo**

Com o mesmo objetivo, isto é, de aumentar a confiabilidade, utilizam-se primários seletivos. Nesta configuração, utiliza-se um circuito duplo, e as derivações são ligadas a ambos por meio de chaves de transferência. Tais chaves normalmente são de transferência automática, isto é, em caso de tensão nula nos terminais de um circuito de entrada, verificam o

funcionamento normal do trecho a jusante, e o transferem automaticamente para o outro circuito de entrada. Esta configuração é utilizada em redes aéreas e subterrâneas [56].

#### **2.2.3.4 Primário operando em malha aberta**

Com um custo mais elevado, é utilizado em regiões com maior densidade de carga e geralmente em redes subterrâneas. Nesta configuração, as derivações são agrupadas em barramentos, e ligadas ao alimentador por meio de dois disjuntores. O alimentador é seccionado por uma chave NA em um ponto estratégico. Em caso de falhas, alguns disjuntores abrem e isolam o trecho com defeito. Então, a chave NA é fechada de tal forma que os barramentos desenergizados passam a ser alimentados pelo outro trecho de alimentador [56].

#### **2.2.3.5 Spot network**

Nesta configuração, cada transformador de distribuição tem dois ou três circuitos de alimentação. Estes circuitos podem derivar-se de SEs distintas. Os transformadores são conectados por meio de chaves denominadas *network protectors* (NP), que se abrem na presença de correntes reversas, isolando os trechos com defeitos. Com confiabilidade e custos elevados, é utilizada em áreas de elevada densidade de carga, como o Plano Piloto de Brasília [56].

## **2.3 Ferramentas de visualização de informação**

Ferramentas de visualização de informação são representações visuais interativas de dados abstratos, que auxiliam em tarefas como otimização, reconhecimento de padrões e monitoramento. Elas possibilitam uma assimilação mais rápida, reduzindo a busca por informações, e permitindo uma análise mais eficiente de dados [9].

Para que estas técnicas sejam aplicadas com sucesso, é necessário que as aplicações apresentem bons mecanismos de usabilidade e interação [64]. Por isso, além de implementar novas visualizações, pesquisadores buscam otimizar os modelos já propostos, desenvolvendo, por exemplo, novas formas de interação e animações. Avaliar o desempenho das visualizações em diversas tarefas e campos de aplicação, também é um passo importante, uma vez que permite selecionar as melhores técnicas para cada aplicação. Em [65], os autores avaliam diversas técnicas para tarefas de agrupamento, visualização de correlação, tendências e *outliers*.

Assim, as várias técnicas de visualização de informação são adequadas para diferentes tarefas e contextos de aplicação. Quando os dados possuem baixa dimensionalidade, isto é,

poucos atributos, técnicas como *scatter plots*, histogramas, *box plots*, gráficos de linha, pizza e barras, são comumente empregadas com sucesso, inclusive em *smart grids* [10]. Caso os dados possuam alta dimensionalidade, isto é, muitos atributos, técnicas como curvas de Andrews [66], matrizes de *scatter plots* [67] e coordenadas paralelas são algumas das técnicas utilizadas.

Caso os dados possuam características hierárquicas, *treemaps*, árvores e gráficos *sunburst* são exemplos de visualizações adequadas [68, 69]. Além disso, visualizações que não foram originalmente projetadas para dados hierárquicos, também têm sido adaptadas com novos recursos visando a utilização neste cenário, como as coordenadas paralelas [70, 71].

Em análises que possuem informações geográficas, os mapas são componentes essenciais. No entanto, podem ser combinados a vários recursos como legendas, *glyphs* e animações, além de poderem ser associados a outras visualizações [72].

### 2.3.1 Coordenadas paralelas

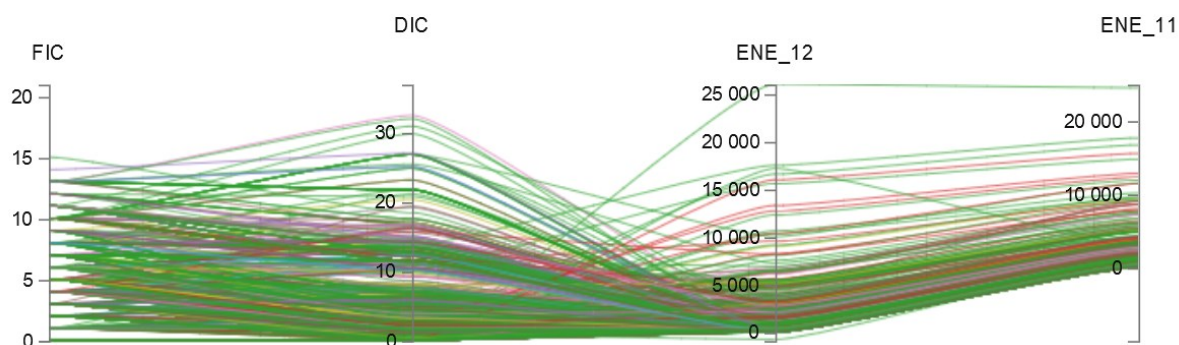
Popularizadas por Inselberg [73], inicialmente com foco em geometrias hiperdimensionais, e por Wegman [74], para análise de dados com elevada dimensão, as coordenadas paralelas (PCs) possuem sua origem relacionada a diversas fontes [75, 76, 77]. Wegman [74] relaciona ainda as PCs com curvas de Andrews [66], que de fato têm campos de aplicação em comum com as PCs [10].

Essa técnica de visualização consiste basicamente em representar as dimensões dos dados em eixos paralelos, ao contrário dos eixos ortogonais presentes em planos cartesianos. Desta forma, cada amostra é representada por uma *polyline*, cujos vértices são posicionados nos eixos paralelos, nas coordenadas que representam os valores dos atributos da amostra.

A Figura 9 mostra um exemplo de coordenadas paralelas, em que os eixos verticais representam os índices de frequência e duração de interrupção individual por unidade consumidora (FIC e DIC), e o consumo de energia nos meses de dezembro e novembro (ENE\_12 e ENE\_11). Neste gráfico, cada linha representa uma unidade consumidora.



Figura 9 – Visualização em coordenadas paralelas.



Fonte: o autor.

É possível observar correlações entre as variáveis por meio de PCs, assim como é feito com as matrizes de *scatter plots* [78, 79]. Para isto, analisa-se os padrões em inclinações e cruzamentos entre as linhas. No entanto, estudos sugerem que as matrizes de *scatter plots* apresentam melhores resultados para essa tarefa, o que pode ser compensado com uma preparação do usuário para utilizar as coordenadas paralelas [79]. Por outro lado, a dificuldade de relacionar uma amostra em vários *scatter plots*, pode inviabilizar a identificação de relações entre várias amostras, e de tendências em grupos, o que é diferente para as PCs, visto que uma amostra é representada por uma *polyline*, apresentando continuidade.

Na Figura 9 é possível observar que várias linhas se cruzam entre os eixos DIC e ENE\_12. Isso representa uma tendência de maiores índices DIC para consumidores com menor consumo, e vice-versa, considerando os dados visualizados na parte superior destes eixos. No entanto, é necessário cautela ao generalizar observações como esta. Neste caso, a densidade de linhas nas regiões inferiores dos eixos DIC e ENE\_12 é grande, e dificulta a visualização de muitas instâncias. Assim, é interessante que se realize uma análise com foco nestes intervalos.

Com as coordenadas paralelas, também é possível realizar a separação de *clusters* conforme mostrado em [74]. A identificação de *outliers* também pode ser feita. Na Figura 9, por exemplo, é possível visualizar uma amostra com comportamento incomum, que possui um consumo muito menor em novembro do que em dezembro, representada por um segmento de linha decrescente entre o último par de eixos. A maioria dos segmentos entre os eixos ENE\_12 e ENE\_11 tem inclinação crescente.

### 2.3.2 Visualizando relações e dados hierárquicos

Além de auxiliar na análise dos atributos das amostras dos conjuntos de dados, algumas técnicas de visualização visam mostrar relações entre tais amostras. Estas relações podem ter

semântica variada: relações hierárquicas, como parte e subparte; expressar conectividade, como cidades ligadas por estradas ou computadores em rede; indicar sequencialidade; ou ainda similaridade entre atributos, temporais ou espaciais por exemplo [80].

Além disso, as relações podem ser classificadas de diversas maneiras: simples ou complexas, considerando o número de entidades envolvidas, uni ou bidirecionais, ponderadas ou não etc. [80].

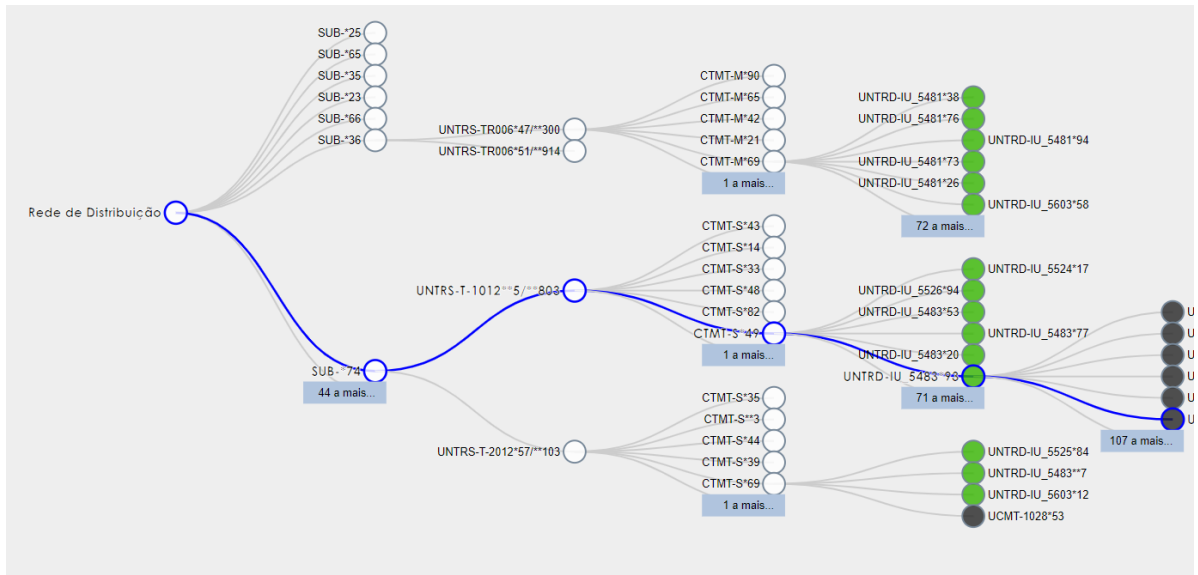
Um grafo é uma estrutura que representa relações de instâncias num conjunto de dados. De maneira formal, um grafo  $G = (V, E)$  é formado por dois conjuntos: o conjunto de vértices ou nós  $V$ ; e o conjunto de arestas ou conexões  $E$ . Cada aresta tem um ou dois vértices associados a ela [81, 82]. O pares de vértices associados às arestas podem ser considerados ordenados ou não, o que resulta, respectivamente, em grafos orientados ou não direcionados [83, 82]. Uma definição mais rigorosa de grafo pode ser encontrada em [84].

Ao aplicar técnicas de visualização de informação em dados hierárquicos, um tipo de grafo ocupa lugar de destaque: as árvores, grafos conexos e acíclicos. Um grafo é dito conexo se existe ao menos um caminho entre cada par de nós, passando ou não por vários vértices [82]. E é acíclico caso não haja mais que um caminho entre dois vértices [83]. Ou seja, uma árvore, formalmente, tem um e apenas um caminho entre dois nós.

Diversas técnicas de visualização foram desenvolvidas para representar estruturas de dados em árvores [80]. É interessante destacar, que estas técnicas também têm sido utilizadas para a visualização de grafos mais complexos, com ciclos [85] ou com múltipla herança [86] por exemplo. Para isso são utilizados artifícios como a introdução de nós duplicados, conexões extras, ou ainda recursos de animação e foco.

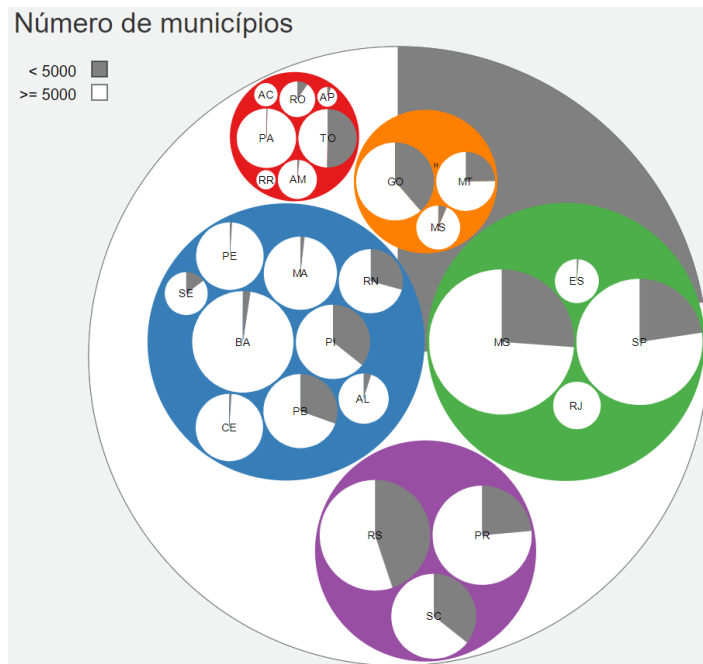
Estas técnicas são geralmente classificadas quanto à dimensionalidade, em 2D ou 3D, quanto ao modo de representação das conexões, em explícito ou implícito, ou quanto ao alinhamento dos nós, em radial, paralelo ou livre [87, 80]. Quanto ao modo de representação das conexões, a técnica é explícita se as conexões são representadas por linhas, como mostrado na Figura 10. Caso a relação hierárquica seja representada por meio do posicionamento dos nós, como mostrado na Figura 11, são técnicas de visualização em árvore implícitas [87]. Há técnicas que agregam ambas as formas de representação [88].

Figura 10 – Visualização em árvore explícita: diagrama nó-link, na aplicação VPlant.



Fonte: o autor.

Figura 11 – Visualização hierárquica implícita.



Fonte: o autor.

Entre as principais técnicas de visualização em árvore, podemos destacar, no grupo das técnicas com *links* implícitos, *treemaps* [89], *sunbursts* [90] e *icicle plots* [91]. Entre as que apresentam as conexões explícitas, temos o diagrama nó-link, mostrado na Figura 10. Uma variação deste diagrama, chamada de *cone tree* [92], utiliza três dimensões. Com funcionalidades de rotação, translação e zoom, pode-se representar uma expressiva quantidade

de dados em árvores 3D. No entanto, este tipo de visualização apresenta algumas desvantagens, como a possibilidade de sobreposição e oclusão de instâncias.

## 2.4 Previsão da demanda de carga

Nesta subseção, apresenta-se uma breve revisão, sobre aspectos relacionados aos métodos de previsão utilizados para a estimação de carga neste trabalho. Neste sentido, são abordadas algumas características dos tipos de previsão de carga, métricas para avaliar os modelos, e abordagens para o pré-processamento de dados.

### 2.4.1 Horizontes, seleção de variáveis e agregação de cargas em previsões

Considerando os horizontes de previsão, podemos classificar as previsões como de curtíssimo (*very short-term load forecasting* - VSTLF), curto (*short-term load forecasting* - STLF), médio (*medium-term load forecasting* - MTLF) e longo prazo (*long-term load forecasting* - LTLF) [93, 15]. Os autores definem horizontes de previsão ligeiramente diferentes para cada uma destas classificações. No entanto, a importância e a diferença de cada tipo de previsão são comumente destacadas.

O horizonte de previsão para a VSTLF é definido de um minuto a meia hora em [15] e se estende até um dia segundo [93]. Para este tipo de previsão, o histórico de valores da demanda pode ser suficiente para prever os valores futuros. As variáveis meteorológicas, como a temperatura, geralmente são úteis, mas opcionais ao estabelecer os atributos de entrada do modelo [93]. É interessante destacar, que a importância desse tipo de previsão deve aumentar com o aumento do uso de fontes de energia intermitentes.

Para a STLF, o horizonte de previsão é normalmente definido de horas a semanas [14, 93, 15]. As variáveis meteorológicas têm uma importância maior neste tipo de previsão, que é útil em atividades de compra e venda de energia [93].

Em MTLF, consideram-se horizontes de um mês a alguns anos e, além de variáveis meteorológicas, as econômicas passam a ter relevância [93]. Este tipo de previsão é importante para a alocação de recursos na geração de energia elétrica. Também é utilizado em outros contextos, como planejamento financeiro [93].

Por fim, a LTLF é caracterizada por previsões de anos a décadas, e é importante para o planejamento de expansão de infraestruturas de geração, transmissão e distribuição [93, 15]. É interessante destacar que, devido às diversas incertezas e variáveis envolvidas em horizontes

de previsão muito longos, as previsões deste tipo consistem geralmente em simulações de diversos cenários [93].

Desta maneira, o horizonte de previsão desejado deve ser considerado para a correta seleção das variáveis do modelo. Além disso, diferentes aplicações exigem diferentes horizontes de previsão, requisitos de precisão e taxas de atualização.

Além dos horizontes de previsão, o nível de agregação de cargas afeta diretamente os requisitos e o desempenho de algoritmos de previsão. O consumo agregado para elementos “centrais”, como subestações, é geralmente mais previsível que o consumo referente a elementos mais periféricos, como transformadores de distribuição ou unidades consumidoras [94].

Assim, com o aumento do uso de fontes intermitentes e distribuídas, como a solar, a tarefa de previsão de carga se torna ainda mais importante e desafiadora, seja pela aleatoriedade, ou pela desagregação da geração. Por outro lado, a crescente disponibilidade de dados, impulsionada por tecnologias relacionadas às *smart grids*, constitui oportunidades para pesquisas relacionadas à previsão de carga.

## 2.4.2 Avaliação das previsões

Com o intuito de comparar os métodos e verificar a qualidade das previsões realizadas, é necessário utilizar algumas métricas. Destacam-se neste sentido, o erro médio absoluto (*mean square error* – MAE), a raiz do erro médio quadrático (*root mean square error* – RMSE) e o erro médio percentual absoluto (*mean absolute percentage error* – MAPE), como alguns dos erros largamente utilizados na literatura [95, 96, 97]. Cada métrica possui alguns aspectos específicos, que podem ser considerados vantagens ou desvantagens. Por isso, uma análise conjunta de tais métricas pode proporcionar melhores avaliações [98].

### 2.4.2.1 Erro residual e erro de previsão

No contexto dos algoritmos de previsão, é interessante distinguir entre o erro residual e o erro de previsão [97]. O erro residual, para um instante  $t$ , consiste na diferença do valor  $y_t$  de uma série, e o valor  $\hat{y}_t$  obtido ao utilizar um modelo que foi ajustado para a série  $y$ , num intervalo de tempo que contém o instante  $t$ . Já o erro de previsão, é a diferença entre  $y_t$  e a previsão  $\hat{y}_t(t - \tau)$ , realizada com base em valores anteriores ao instante  $t - \tau$ . Neste contexto, o parâmetro  $\tau$  está diretamente relacionado ao horizonte de previsão. Os erros residual e de previsão são respectivamente definidos pelas equações 2.1 e 2.2 [97].

$$e_t = y_t - \hat{y}_t . \quad 2.1$$

$$e_t(\tau) = y_t - \hat{y}_t(t - \tau) . \quad 2.2$$

Isto é, o erro chamado de residual neste trabalho, não está relacionado a uma previsão para um instante futuro. Por outro lado, reflete a capacidade do modelo em aderir a um conjunto de dados. Poderia também ser chamado de erro de “treino”.

Já o erro de previsão, é calculado visando verificar a capacidade do modelo em prever amostras, com base em informações passadas. É interessante destacar que os erros residuais costumam ser menores que os erros de previsão.

Neste trabalho, os erros residual e de previsão serão diferenciados quando necessário. A simbologia  $e_t$  será utilizada de forma genérica quando o contexto for aplicável a ambos.

#### 2.4.2.2 Métricas de erro dependentes de escala

Com o intuito de sumarizar os erros para vários instantes de tempo, é necessário utilizar artifícios como a média, a mediana etc. Assim, por meio da média dos valores absolutos dos erros instantâneos, o MAE, uma das métricas de erro mais utilizadas, é definido pela expressão abaixo [95, 99].

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t| \quad 2.3$$

É interessante destacar que o módulo dos erros é utilizado para que não haja anulação entre erros positivos e negativos, o que ocorreria com o erro médio (*mean error* - ME)

$$ME = \frac{1}{n} \sum_{t=1}^n e_t . \quad 2.4$$

Apesar de não representar a magnitude dos erros devido a esta anulação, o ME indica se um modelo está com “viés”. Ao utilizar um modelo para previsão, caso o ME se afaste de zero, a série provavelmente mudou sua tendência, de tal forma que o modelo não foi capaz de reproduzir [97].

O erro médio quadrático (*mean square error* – MSE) e o RMSE são definidos pelas expressões abaixo.

$$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2 . \quad 2.5$$

$$RMSE = \sqrt{MSE} . \quad 2.6$$

O MSE é uma estimativa da variância dos erros de previsão [97]. Fatores como a facilidade de interpretação e de cálculo, são algumas das vantagens do MAE em relação ao

MSE e ao RMSE. O MAE é um limite superior para RMSE quando  $|e_t| > 0$  e inferior quando  $|e_t| < 0$ . Assim, o erro médio quadrático e sua raiz, ponderam com um peso maior os erros de maior magnitude [99].

### 2.4.2.3 Métricas de erro percentuais e relativas

Todas as métricas de erro definidas na Subseção 2.4.2.2 são dependentes da escala dos dados. Isso pode inviabilizar comparações entre modelos com dados em escalas diferentes. Se a grandeza analisada é uma medida de comprimento, por exemplo, errar um centímetro pode ser adequado para uma medida de 50 cm, mas inadmissível para uma medida de dois centímetros.

O erro médio percentual absoluto (*mean absolute percentage error* – MAPE) indica o percentual de erro a cada amostra, em relação ao respectivo valor instantâneo, sendo definido pela expressão [97, 100]:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{y_t} \right| . \quad 2.7$$

Assim, o MAPE independe da magnitude dos valores instantâneos numa série. Desta forma, em contextos nos quais o erro deva ser ponderado em relação à sua respectiva amostra, o MAPE pode ser uma medida efetiva. Além disso, é uma métrica que, assim como o MAE, possui um significado diretamente relacionado ao erro.

É interessante notar que a métrica não é válida para  $y_t = 0$ , e produz valores excessivamente grandes caso  $y_t$  se aproxime de 0 [101]. No entanto, pode-se adaptar tal formulação, de tal forma que valores nulos ou muito pequenos numa série, sejam substituídos por um valor proporcional ao erro, e compatível com erros de outras amostras.

Por outro lado, também é interessante destacar que, para um mesmo valor de erro, o fato desta métrica considerar mais, o que for referente a uma amostra de menor magnitude, pode ser uma grande desvantagem. Isto pode levar a conclusões equivocadas, com uma valorização de erros pequenos, em detrimento de erros maiores. Alguns autores enunciam este aspecto informando que o MAPE pondera mais erros positivos do que negativos [102], o que é questionado por outros autores [98, 101].

As desvantagens do MAPE levaram à criação de métricas como o erro médio percentual absoluto simétrico (*symmetric mean absolute percentage error* – sMAPE) [103]. No entanto, esta métrica normalmente não resolve o problema relacionado aos valores de pequena magnitude [98, 104] e, ao contrário do MAPE, não é “simétrica”. Isto é, considerando  $y_t$  e  $|e_t|$  constantes, resulta em valores de erro menor para valores positivos de erro ( $y_t > \hat{y}_t$ ) [101, 98].

Outra estratégia que visa possibilitar a comparação da previsão com conjuntos de dados em escalas diferentes, é utilizar a razão do erro produzido por um determinado modelo, e do erro produzido por um algoritmo de referência. Normalmente, este algoritmo de referência consiste simplesmente em utilizar o valor atual de uma série, como estimativa para o próximo valor [101]. Utilizando esta estratégia em conjunto com outras técnicas de agregação de erros instantâneos, como média, mediana ou média geométrica, obtém-se várias métricas de erro. Considerando a média dos erros absolutos, por exemplo, define-se o erro médio absoluto relativo (*mean relative absolute error* – MRAE) [104, 102]

$$MRAE = \frac{1}{n} \sum_{t=1}^n \left| \frac{e_t}{e_t^*} \right|, \quad 2.8$$

em que  $e_t^*$  é o erro obtido ao se utilizar o algoritmo de referência. Desta forma, valores de erro menores que um, representam um desempenho melhor que o algoritmo de referência. De forma análoga, utilizando-se a mediana, pode-se definir a mediana do erro absoluto relativo (*median relative absolute error* – MdRAE) [104]. A métrica MdRAE sofre menor influência de erros *outliers*, isto é, valores com elevada ou baixa magnitude.

Outra alternativa proposta para comparação entre previsões com diferentes grupos de dados, é o erro médio absoluto em escala (*mean absolute scaled error* – MASE) [104]

$$MASE = \frac{1}{n} \sum_{t=1}^n \frac{|e_t|}{\frac{1}{n-1} \sum_{i=2}^n |y_i - y_{i-1}|}. \quad 2.9$$

O MASE pode ser utilizado para séries com valores nulos e não é distorcido por amostras de pequena magnitude.

#### 2.4.2.4 Métricas de erro em contextos específicos

As medidas de erro usualmente empregadas podem não ser adequadas ou suficientes para todos os contextos. Em sistemas de energia elétrica, as magnitudes dos erros de previsão, principalmente em momentos de máxima demanda, podem levar a sérios problemas. Caso a demanda supere a capacidade de fornecimento de energia, apagões e falhas em equipamentos podem levar a prejuízos expressivos.

Desta forma, é interessante considerar métricas como o erro máximo, ou o erro no ponto de máximo consumo, na avaliação de métodos de previsão. Em [105], os autores utilizam a média do erro máximo para verificar se os erros das previsões eram compatíveis com as reservas operativas do sistema.



### 2.4.3 Pré-processamento de dados: desafios e procedimentos

Para aplicar algoritmos de previsão, é comum que alguns aspectos como a escala dos dados, informações faltosas, escala temporal não homogênea, estacionariedade, seleção de atributos, presença de *outliers* e formato dos dados de entrada do modelo, devam ser observados. Nesta subseção, são apresentados alguns destes aspectos, bem como procedimentos utilizados para analisar ou adequar os dados considerando tais aspectos.

#### 2.4.3.1 Normalização por amplitude

Dentre as operações mais comuns na preparação de dados para algoritmos de previsão, está a normalização por amplitude, que geralmente é útil quando os atributos possuem diferentes escalas [106]. Este procedimento pode evitar que um atributo predomine sobre outro, mas nem sempre é necessário ou recomendado utilizá-lo [107]. Além disso, mesmo em modelos com um único atributo, a normalização pode facilitar a comparação entre previsões para diferentes grupos de dados, bem como a utilização de modelos pré-treinados. Para algoritmos de redes neurais, também pode melhorar o desempenho de modelos, acelerando o treinamento, ou compatibilizando o intervalo da função de ativação da última camada e as variáveis *target*.

Dentre as formas de normalização, podemos destacar: a normalização por escala, comumente chamada de “*min-max*”, que visa redefinir os valores mínimo e máximo de atributos; e por padronização, que visa redefinir medidas de posição e espalhamento, como a média e o desvio padrão respectivamente [106].

O primeiro tipo, por escala, consiste em mapear cada valor  $x$  de um atributo num conjunto de dados, por meio da expressão [106, 107]

$$x_n = (max_n - min_n) \frac{x - min_o}{max_o - min_o} + min_n, \quad 2.10$$

em que  $x_n$  é o valor mapeado de  $x$  na nova escala de amplitudes do atributo,  $max_n$  e  $min_n$  são os valores máximo e mínimo desejados para o atributo na nova escala, e  $max_o$  e  $min_o$  são os valores de máximo e mínimo do atributo na escala original.

Já a normalização por padronização pode ser definida por [106, 107]

$$x_n = \frac{x - \mu}{\sigma}, \quad 2.11$$

em que  $\mu$  e  $\sigma$  são, respectivamente, a média e a variância do atributo a ser normalizado.

### 2.4.3.2 Dados ausentes ou inconsistentes

Algumas amostras de um conjunto de dados podem não apresentar todos os atributos. Numa série temporal, pode haver intervalos com dados incompletos ou totalmente ausentes. Este cenário é comum e desafiador em sistemas de IoT com condições severas ou redes restritas, como o LoRa [12].

Dentre as alternativas para contornar este problema, estão o destarte de amostras com informações ausentes, utilizar algoritmos resilientes a este problema, ou preencher as informações faltosas. Para preencher os dados ausentes, pode-se escolher um valor que represente a ausência, mas o sucesso de tal abordagem dependerá de como o modelo vai interpretar este valor. Pode-se ainda, estimar os valores ausentes por meio de medidas como média e mediana, ou por meio dos valores de outros atributos e amostras semelhantes [106].

Outro problema comum em bases de dados, é a presença de dados inconsistentes, isto é, amostras com atributos que não estão de acordo com alguma regra [106]. No contexto de *smart grids*, valores negativos de carga, exceto em casos de geração, são um exemplo deste problema. A integração entre diferentes bases e versões de aplicações, o que também é comum em sistemas de IoT, favorece a ocorrência deste tipo de problema.

Quando algumas amostras apresentam dados que divergem da distribuição do atributo, também podem ser consideradas inconsistentes ou *outliers*. Retirar ou retificar estas amostras, pode melhorar o desempenho do modelo, mas também pode fazer com que informações importantes não sejam consideradas [106].

Neste sentido, em [14] os autores utilizam a média e a mediana como métodos de pré-processamento na aplicação de redes LSTM para previsão de carga. Para isto, sintetizam os valores de carga horária das últimas semanas em apenas uma semana de dados. Para este procedimento, os autores observaram que a utilização da mediana resultou em melhores previsões do que a média, sendo menos sensível a *outliers*.

### 2.4.3.3 Aspectos temporais

No caso de algoritmos de previsão, definir o tamanho do intervalo de tempo considerado para as entradas em cada previsão, tem papel semelhante à seleção de variáveis. Neste contexto, cada valor de uma série pode ser considerado um atributo para a previsão dos próximos valores. Assim, definir este intervalo de tempo, é estabelecer qual o tamanho do histórico necessário e adequado para a realização de cada previsão. Além disso, o intervalo de tempo das amostras utilizadas para o treinamento do modelo, normalmente deve considerar a estacionariedade da série para garantir um bom desempenho na previsão.

Além de definir o tamanho do histórico utilizado em cada previsão, a janela de dados utilizada no treinamento, e a resolução ou granularidade temporal necessária para atingir um objetivo, também é comum, principalmente em sistemas de sensoriamento, como os sistemas de IoT e *smart grids*, que sejam necessárias estratégias para solucionar outra questão relacionada ao tempo: a escala temporal não homogênea, isto é, o período para coletar cada medida pode mudar dinamicamente.

Como apresentado na Subseção 4.2.5, uma metodologia simples foi adotada para diminuir os efeitos de *outliers*, dados faltosos e escala temporal não homogênea neste trabalho.

## 2.5 Redes neurais artificiais

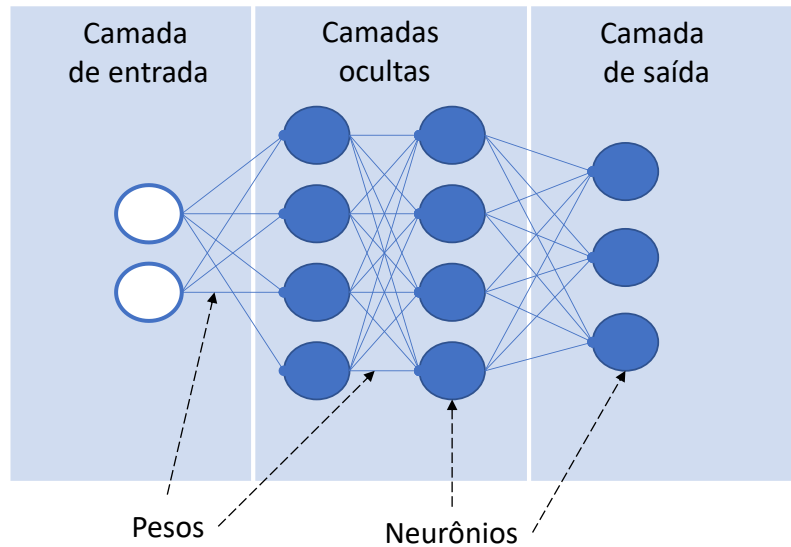
Redes neurais artificiais (RNAs), *artificial neural networks* (ANNs), são estruturas compostas por pequenas unidade de processamento, chamadas de neurônios, conectadas por *sinapses*, e representadas matematicamente por matrizes de pesos, que visam emular o paralelismo, a conectividade e a capacidade de aprendizado por exemplos, de sistemas biológicos [108].

As ANNs são frequentemente organizadas em camadas, de tal forma que os neurônios de uma camada são conectados aos neurônios da camada seguinte, conforme Figura 12. As conexões dos neurônios de uma rede neural estão relacionadas a valores numéricos, os pesos, que ponderam as entradas dos neurônios. Os valores dos pesos armazenam informações que são utilizadas pela rede neural para resolver algum problema.

Com base nos valores de entrada, após realizarem operações internas, os neurônios determinam suas saídas. Estas operações internas geralmente estão relacionadas às chamadas funções de ativação [109]. As saídas dos neurônios são então utilizadas como entradas para outros neurônios, ou como saídas da rede neural. A saída da rede é, desta forma, um valor numérico.

Em problemas de aprendizado supervisionado, esta saída é comparada à saída correta, considerando as entradas aplicadas à rede. Para esta categoria, isto é, aprendizado supervisionado, o processo de treinamento de uma rede neural consiste em ajustar seus pesos, para que possa produzir as saídas desejadas. Para isso, podem ser utilizados diversos algoritmos, como gradiente descendente, gradiente descendente estocástico (*stochastic gradient descent* – SGD), *adaptive gradient algorithm* (AdaGrad), *root mean squared propagation* (RMSProp) e *adaptive moment estimation* (Adam) [110, 111].

Figura 12 – Esquema de rede neural em camadas.



Fonte: o autor.

Além da capacidade de aprendizado dinâmico por meio de dados, as ANNs possuem, em geral, habilidades de generalização e de reproduzir relações não lineares. Devido a características como estas, vêm sendo aplicadas com sucesso em diversos campos e tarefas, como reconhecimento de padrões, regressão e previsão [112, 7]. No contexto de sistemas de IoT e *smart grids*, a quantidade considerável de dados coletados é um fator favorável à aplicação de redes neurais.

Para diferentes aplicações foram desenvolvidos vários tipos de ANNs. As redes convolucionais (*convolutional neural network* – CNNs), por exemplo, obtiveram grande sucesso em reconhecimento de padrões em imagens [113]. Este tipo de rede neural tem sua origem relacionada a autores como Kunihiko Fukushima [114], Alex Waibel [115], e Yann LeCun [116].

Já as redes neurais recorrentes (*recurrent neural networks* – RNNs), são especializadas em analisar dados de séries, como em problemas de previsão ou de processamento de linguagem natural [7]. A saída dos neurônios recorrentes, além de ser função de suas entradas, considera também suas saídas passadas. Entre as RNNs empregadas com sucesso, podemos citar os tipos *long short-term memory* (LSTM) [117] e *gated recurrent unit* (GRU) [118].

## 3 Estado da arte e pesquisas correlatas

Neste capítulo são apresentadas algumas pesquisas relacionadas aos objetivos propostos neste trabalho. Isto é, pesquisas relacionadas às ferramentas computacionais voltadas para sistemas de distribuição e *smart grids*, técnicas de visualização de informação no contexto de SGs, e redes neurais aplicadas para previsão neste contexto.

### 3.1 Sistemas e ferramentas computacionais para *smart grids*

Ferramentas computacionais no contexto de *smart grids* são implementadas com diversos propósitos, como simulação de SGs teóricas [119, 120], gerenciamento, análise ou visualização de dados [121, 122, 123, 124], e simular ou implementar funcionalidades de segurança [125]. Nesta subseção são apresentados alguns sistemas com foco similar aos objetivos deste trabalho: gerenciamento, análise e visualização de dados, além de geração de informação útil por meio de algoritmos, com aplicabilidade em sistemas de distribuição reais.

Em [121], o autor implementa um sistema multiagente de gerenciamento de *microgrid*. Neste sistema, elementos como consumidores e sistemas de geração distribuída são representados por agentes virtuais que atuam de maneira independente na manutenção do equilíbrio entre geração e consumo. Assim, compõem um sistema de controle descentralizado. O autor destaca como diferencial do trabalho, o fato de que a simulação realizada para validar o *framework* considera os sistemas de comunicação. O sistema foi implementado por meio *framework Java Agent Development (JADE)*.

Em [122] o autor implementa um sistema de gerenciamento e análise para sistemas de distribuição, com funções de monitoramento e localização de perdas não técnicas. Para isso, utiliza um algoritmo de estimação de estados e um mapa, com indicação dos elementos e conexões de um trecho de uma rede de distribuição. O trecho utilizado corresponde a um alimentador com 781 cargas de baixa tensão, 53 cargas de alta tensão e 47 transformadores de distribuição.

Em [123], os autores apresentam uma ferramenta, nomeada *Sustainable Energy Management System (SEMS)*, para gerenciamento de sistemas de distribuição, considerando

um distrito. São implementadas funções de otimização para controle e alocação de recursos, previsão e coordenação de serviços de diferentes ativos da rede. Com isso, os autores esperam que a ferramenta seja útil para auxiliar em processos de tomada de decisão.

O sistema apresentado em [123] é implementada em Python, e os autores indicam o uso de bibliotecas de código aberto, visando facilitar o reuso e a extensão do sistema. No entanto, não são fornecidos detalhes de implementação ou repositórios de código. São mostrados exemplos de previsão para geração de energia fotovoltaica e realocação de cargas, com gráficos de linha. Os modelos de previsão utilizados, ANN, *gradient boosting* e *random forest*, foram comparados por meio do MAE.

Em [124], um *framework* com funções *middleware*, isto é, controle do fluxo de dados, chamado de *Resilient Information Architecture Platform for Smart Grid (RIAPS)*, é apresentado. Os autores destacam vários aspectos da plataforma: as formas de conexão entre os atores disponíveis, como *publish-subscribe*, ou requisição-resposta de forma síncrona ou assíncrona; a gestão remota de aplicações, o gerenciamento de dados distribuído; e o compartilhamento de recursos entre os atores do sistema. Além disso, a resiliência a falhas de comunicação entre os atores também é destacada.

Em [126], os autores apresentam um sistema com funcionalidades de gerenciamento e análise de dados, com base numa arquitetura lambda [127]. Esta arquitetura se baseia em três camadas: *batch*, que é responsável por armazenar os dados coletados de maneira distribuída e extrair informações dos dados armazenados; *speed*, para processamento de dados em tempo real; e *erving*, que é responsável por mesclar os dados processados pelas camadas *batch* e *speed*. Assim, o sistema apresentado consiste em um *middleware*, que utiliza as aplicações como *Flume*, *Hadoop* e *Spark*.

Como exemplo de implementação, utilizam o algoritmo *K-means* para agrupamento de curvas de carga. Apresentam ainda visualizações das curvas de carga por meio do *Tableau* [128], mas não detalham as formas de interação ou se esta visualização poderia ser integrada ao sistema para utilização direta do usuário.

## **3.2 Técnicas de visualização de informação no contexto de *smart grids***

Em [55], os autores apresentam uma revisão sobre a visualização de dados em *smart grids*. Para isso, classificam as técnicas em três grupos: tradicionais, como histogramas,

gráficos de linha, pizza e barras; visualização de dados multivariados, que incluem coordenadas paralelas, técnicas radiais e *scatter plots*; e técnicas geográficas.

Ao comparar as técnicas utilizadas em outras áreas, os autores indicam que a aplicação de técnicas de visualização em *smart grids* ainda é incipiente. Além disso, os autores apontam duas tendências para as visualizações neste contexto. Primeiramente, indicam que a disponibilidade de recursos computacionais, e a necessidade de monitoramento e controle em um cenário mais complexo, incentivam a utilização de visualizações com mais recursos, incluindo técnicas com uma longa trajetória em outros campos de estudo. Em segundo lugar, indicam que outra tendência é a integração de várias técnicas em uma única visualização.

Em [10], os autores utilizam várias técnicas para mostrar dados de sistemas de distribuição: gráfico de barras; gráfico de superfície 3D; *scatter plots*; e coordenadas paralelas. No entanto, as técnicas são utilizadas por meio de programas de análise genéricos, com poucas possibilidades de interação, não sendo agregadas a uma ferramenta específica para *smart grids*. As coordenadas paralelas mostradas, por exemplo, apresentam um único eixo graduado para todos os atributos dos dados. Neste trabalho, os autores destacam ainda que a utilização de técnicas de visualização está mais madura para sistemas de transmissão, mas limitada para sistemas de distribuição.

A referência [129] apresenta métodos de visualização aplicados em sistemas de transmissão de energia elétrica, destacando artifícios que permitem a visualização de grandes conjuntos de dados e/ou atributos. As técnicas incluem diagramas de linha, mapas geográficos de calor e com gráficos de pizza, com versões 2D e 3D.

A direção do fluxo de potência é mostrada por meio de setas. Os autores indicam que este aspecto poderia ser melhor representado com o auxílio de animações. Para mostrar grandes conjuntos de dados, os autores limitam a quantidade de elementos mostrados simultaneamente ou recorrem à agregação de elementos próximos.

Em [130], também são apresentadas visualizações para sistemas de transmissão de energia elétrica. As técnicas mostradas incluem diagramas de linha, mapas de calor geográficos e *sparklines* – conjunto de gráficos de linha simplificados, utilizados de maneira semelhante a símbolos.

### 3.3 Algoritmos de redes neurais para previsão no contexto de *smart grids*

A manipulação dos dados coletados nas SGs ocorre com propósitos variados, e utiliza diversas variáveis. Entre as aplicações mais comuns e importantes está a previsão de carga. De maneira similar, trabalhos são desenvolvidos com o intuito de prever a geração de fontes intermitentes de energia [123, 131, 132] e os preços no mercado de energia [133]. A geração de curvas de carga para a realização de simulação [134] ou caracterização de perfis [135] também são tarefas pesquisadas. Além disso, procedimentos como completar dados faltosos, agrupar curvas de carga e corrigir dados, são úteis como métodos de pré-processamento para previsão de carga [136]. Em [137], os autores buscam melhorar a previsão de carga com a utilização de algoritmos de desagregação.

Assim, observa-se uma infinidade de algoritmos que buscam transformar em informações úteis, os dados coletados em SGs. Por isso, os algoritmos citados anteriormente nesta subseção, visam apenas ilustrar a diversidade deste contexto na literatura, uma vez que foram base para a escolha do algoritmo utilizado neste trabalho, e podem ser utilizados como referências em trabalhos futuros.

A seguir, são apresentados alguns trabalhos com contexto similar ao do algoritmo utilizado para a previsão de carga neste trabalho: previsão de curtíssimo ou curto prazo, utilizando modelos de redes neurais. Considerando a grande quantidade de dados gerados por sistemas de *smart grids*, estas redes são uma alternativa a métodos de previsão de carga de curto prazo clássicos, como o modelo autorregressivo integrado de médias móveis (*autoregressive integrated moving average* – ARIMA) [15]. Uma revisão sobre métodos de previsão de carga para baixa tensão, incluindo horizontes médios e longos, pode ser encontrada em [138].

As pesquisas utilizando ANNs para previsão de carga de curto prazo iniciaram-se há anos. Em [139], encontra-se um trabalho de 1995, em que os autores utilizaram um modelo *neural gas*, para agrupar os dados, e redes *multilayer perceptron* (MLP) para realizar a tarefa de regressão ou previsão. Com isto, os autores implementaram um modelo que poderia ser utilizado para mais de uma entidade ou localidade. As diferentes configurações testadas são comparadas por meio do MAPE.

Em [14], os autores utilizam redes LSTM para realizar previsões de curto prazo e verificam os efeitos de alguns métodos de pré-processamento. Ao invés de utilizar os dados históricos de carga como entradas diretas do modelo, utilizam agregações temporais, e entre unidades consumidoras com perfil similar. Os autores concluíram que a agregação por mediana



é melhor que a média, e que a agregação de perfis similares não melhorou a previsão. Já a agregação temporal, isto é, considerando a mediana dos perfis semanais de consumo, resultou em melhoria na previsão. Entre os principais resultados, está a diminuição de 11,89% para o RMSE, utilizando a mediana de seis semanas como entrada para o modelo.

Em [15], o autor utiliza modelos MLP para previsão de carga diária. Para cada dia da semana um modelo é otimizado. Os modelos são avaliados por meio do MAPE. O menor erro médio percentual absoluto foi de 0,14% para um modelo de quintas-feiras, e o maior foi de 2,34% para um modelo de segundas-feiras.

Em [16], redes LSTM são utilizadas para previsão de carga em resolução horária, para um horizonte de um dia. Os resultados obtidos são comparados com previsões realizadas por *support vector regression* (SVR). O modelo LSTM supera o SVR, com RMSE cerca de 10% menor.

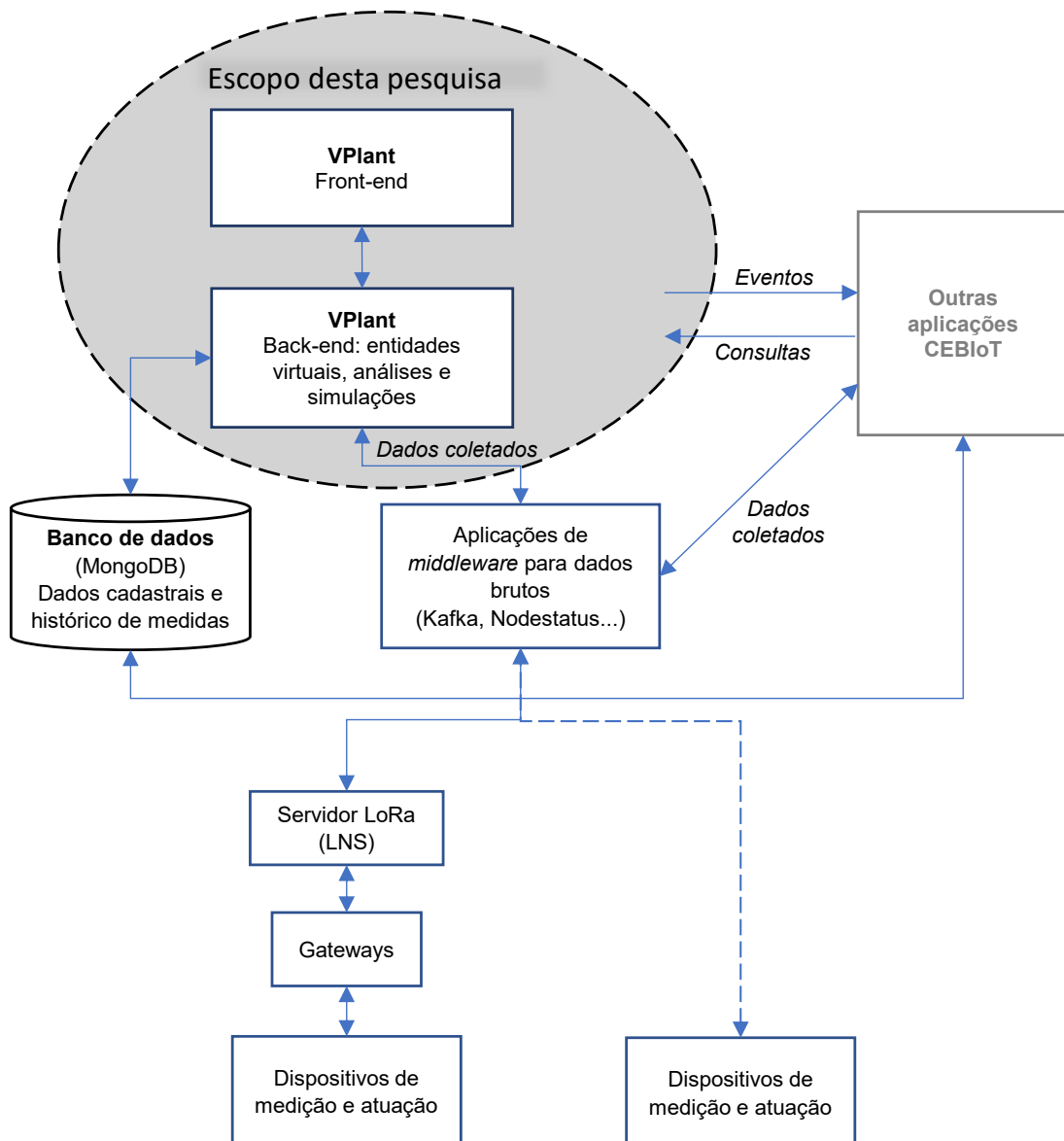
Em [17], os autores também empregaram redes LSTM para realizar a previsão de carga. Neste trabalho, destacam que as previsões dos modelos melhoraram com retirada das sazonalidades e da tendência de crescimento de demanda.

Em [18], os autores utilizam uma combinação entre modelos CNN, LSTM e ARIMA para produzirem um algoritmo de previsão. Então comparam, por meio das métricas MAPE E RMSE, o modelo proposto com modelos ARIMA e CNN-LSTM. O modelo apresentado superou os modelos de referência.

## 4 Material e métodos

Neste capítulo são apresentados os recursos e as metodologias utilizadas para o desenvolvimento da aplicação proposta. Na Figura 13 é apresentado um esquema com os principais elementos da aplicação, além de elementos adjacentes, isto é, presentes no contexto do projeto que esta aplicação faz parte. Assim, apenas os programas de *front-end* e *back-end* da aplicação **Virtual Power Plant (VPlant)** são produtos diretos deste trabalho. Tais elementos, bem como tecnologias e metodologias utilizadas para implementá-los, são descritos nas subseções seguintes.

Figura 13 – Principais elementos da aplicação e sistemas adjacentes.



Fonte: o autor.

## 4.1 Material

### 4.1.1 Linguagens de programação, *frameworks* e bibliotecas

Os programas desenvolvidos utilizaram majoritariamente a linguagem de programação *TypeScript*, que consiste em um superconjunto de *JavaScript*, adicionando tipagem estática opcional. O *back-end* foi implementado utilizando Node.js. Para os algoritmos de redes neurais, utilizou-se a biblioteca TensorFlow.js [140], que executa códigos em C++ para melhorar o desempenho na execução de algoritmos.

A comunicação entre *back-end* e *front-end* é realizada por meio do *hypertext transfer protocol* (HTTP) ou da tecnologia de *WebSockets*. Esta tecnologia permite uma comunicação *full-duplex* entre os programas, o que foi fundamental para implementar o sistema de notificação de eventos das entidades virtuais. Tal tecnologia também foi utilizada para permitir que outras aplicações pudessem consultar o *back-end* da aplicação VPlant.

O *front-end* foi implementado por meio do *framework* Angular. Para construir as visualizações em árvore e em coordenadas paralelas, utilizou-se a biblioteca D3.js [141]. A visualização em árvore foi inspirada em [142]. Já a visualização em coordenadas paralelas foi inspirada em [143].

Para que o mapa pudesse mostrar a quantidade de entidades presente em sistemas de distribuição reais, que podem chegar a milhões, utilizou-se uma biblioteca baseada em *WebGL*, chamada de *Leaflet.glify* [144]. A interface de programação de aplicativos (*application programming interface* – API) *WebGL* permite que navegadores de internet possam utilizar unidades de processamento gráfico (*graphics processing unit* – GPUs), o que acelera a execução de programas.

### 4.1.2 Dados: fontes e ferramentas

Os dados das entidades monitoradas são armazenados em um banco não relacional MongoDB [145]. A aplicação utiliza dados estáticos e coletados em tempo real. Os principais dados estáticos são provenientes de uma base de dados geográfica da distribuidora (BDGD), que é definida pelo módulo 10 dos Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) [146]. As BDGDs possuem dados dos ativos e demais entidades que compõem o sistema de distribuição das concessionárias brasileiras.

Além disso, no caso de uso com dados coletados, os dados são provenientes de dispositivos que se comunicam por meio de tecnologia *Long Range* (LoRa). As mensagens

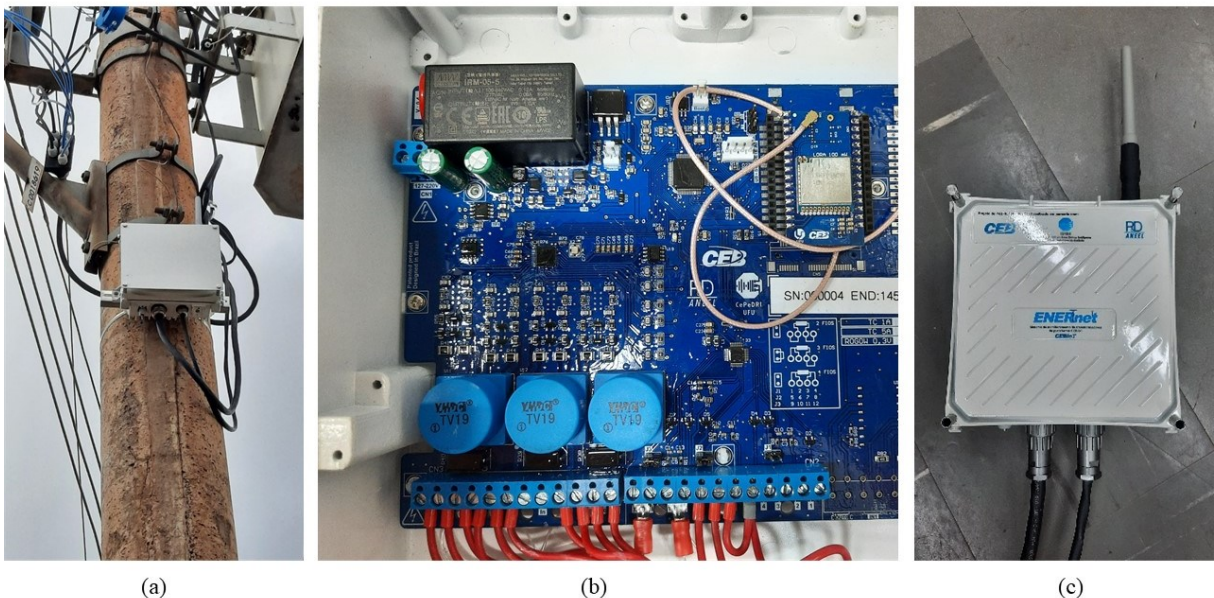
chegam aos *gateways* e são repassadas para o *LoRaWAN Network Server* (LNS). O LNS transmite as mensagens por meio do protocolo *message queuing telemetry transport* (MQTT), para aplicações que as deserializam e, por sua vez, encaminham tais mensagens a diversas aplicações de análise e monitoramento, como o VPlant.

Na atual implementação, estas aplicações podem ser um *middleware Apache Kafka* [147], ou um programa executado em Node.js, chamado *NodeStatus*, que foi implementado pelos pesquisadores do projeto. Enquanto o *Apache Kafka* transmite as mensagens em tópicos, com um protocolo binário próprio, sobre o protocolo de controle de transmissão (*transmission control protocol – TCP*), o *NodeStatus* envia as mensagens por meio do protocolo de datagramas do usuário (*user datagram protocol – UDP*) ou de *WebSockets*.

#### 4.1.3 Infraestrutura e hardware

Os dispositivos que coletaram os dados utilizados no caso de uso apresentado, foram implementados por outras pesquisas relacionadas ao contexto deste trabalho, e estão representados na Figura 14. É válido destacar que os dispositivos conectados à aplicação não precisam utilizar esta tecnologia de comunicação, uma vez que o VPlant não se comunica diretamente com eles. Isto é, tal aspecto é transparente para as aplicações que utilizam as aplicações de *middleware* de dados brutos.

Figura 14 – Fotos do dispositivo (a) medindo dados de um transformador de distribuição, (b) em detalhe e (c) dentro da caixa de proteção.



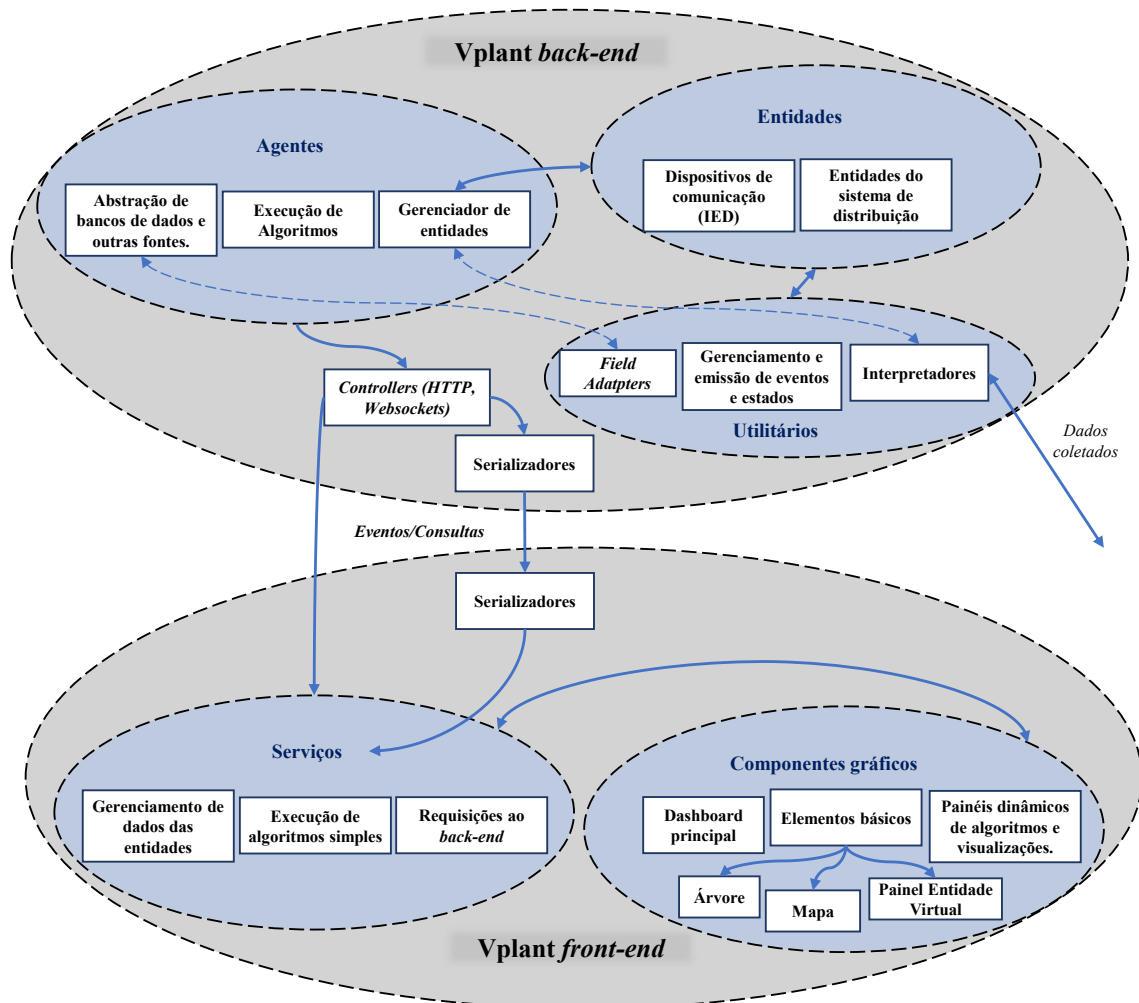
Fonte: .

Além disso, para execução de vários algoritmos e simulações, foram utilizados servidores disponibilizados por meio do projeto de pesquisa. Nestes servidores, as várias aplicações que compõem o sistema são virtualizadas em contêineres por meio da *Docker Engine* [148].

## 4.2 Desenvolvimento da aplicação

Nesta subseção são mostradas as metodologias empregadas para implementar os elementos e funcionalidades que compõem a aplicação VPlant. Na Figura 15, um diagrama com os principais elementos é apresentado, em que estes são agrupados conforme sua função na aplicação. Algumas interfaces gráficas criadas são mostradas nesta subseção para facilitar a compreensão acerca do processo de desenvolvimento da aplicação.

Figura 15 – Elementos e organização da aplicação VPlant.



Fonte: o autor.

Na Subseção 4.2.1 são apresentados os detalhes da implementação do programa *back-end*. Na Subseção 4.2.2 são apresentados os elementos básicos da interface gráfica. Nas subseções 4.2.3 e 4.2.4, são apresentadas as metodologias para a implementação das técnicas de visualização em árvore e em coordenadas paralelas respectivamente. Na Subseção 4.2.5 são apresentados os aspectos relacionados à função de previsão de carga implementada.

#### 4.2.1 *Back-end*: gerenciamento de entidades virtuais e algoritmos

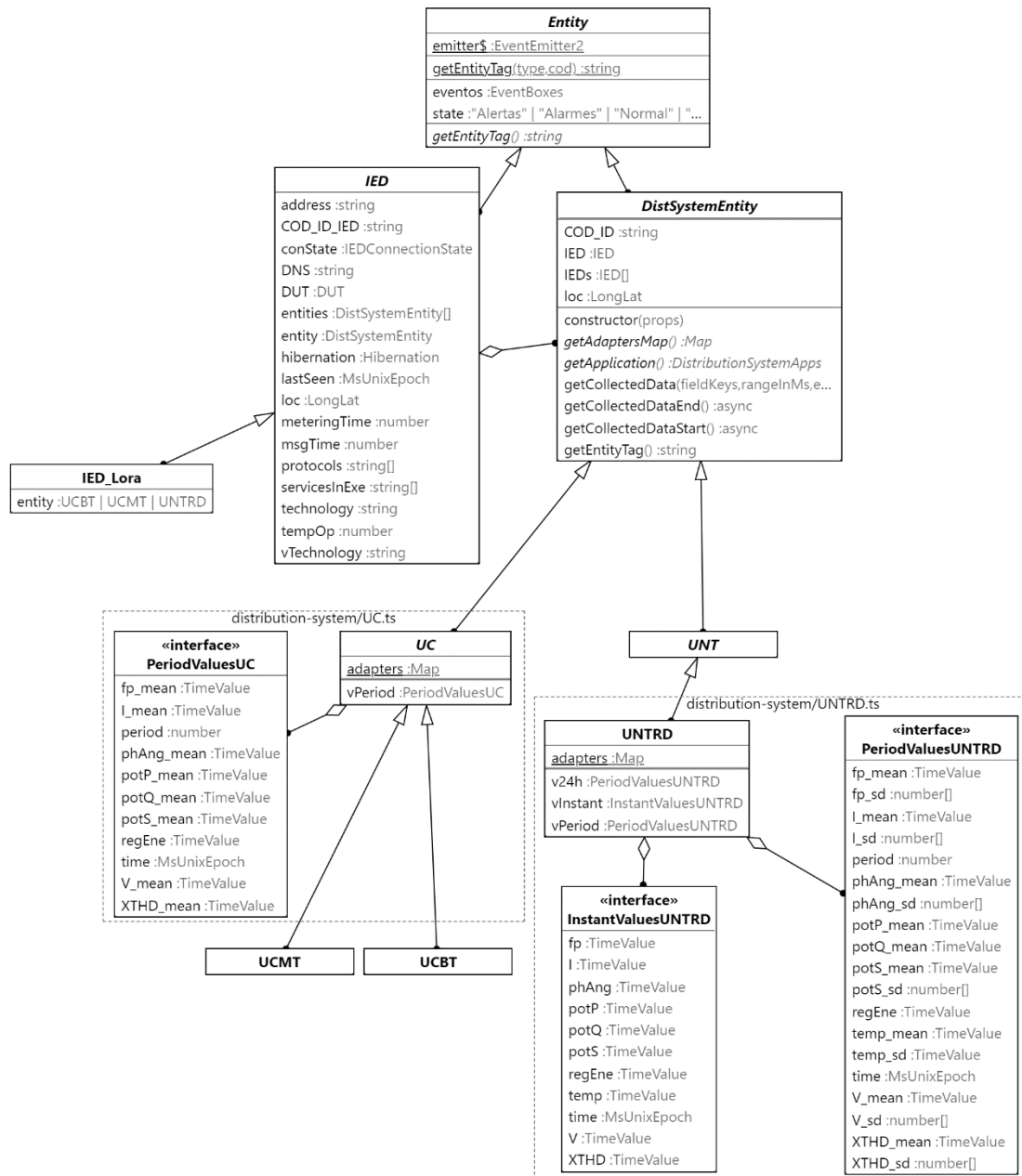
O programa VPlant *back-end* foi organizado de maneira semelhante ao padrão *model view controller* (MVC). Informações sobre este padrão podem ser encontrados em [149]. Na Figura 15, os agentes representam diversos serviços que executam algoritmos, acessam o banco dados e gerenciam as entidades virtuais. Os *controllers* representam elementos que permitem outros programas, como o VPlant *front-end*, realizarem requisições.

Para fornecer funcionalidades de monitoramento, análise e parametrização, o VPlant *back-end* gerencia objetos, equivalentes a entidades virtuais, que estão relacionados a elementos do sistema de distribuição, e aos dispositivos que se comunicam com o sistema, conforme Figura 15. Nesta figura, os dispositivos de comunicação são identificados pelo termo IED, em alusão aos *intelligent electronic devices* [150]. Por exemplo, no caso do transformador mostrado na Figura 14 (a), tanto o transformador quanto o dispositivo de comunicação, mostrado em destaque nas Figura 14 (b) e Figura 14 (c), teriam entidades virtuais correspondentes.

As relações entre as classes de entidades virtuais presentes no VPlant *back-end* são mostradas na Figura 16. Nesta figura, foram suprimidos métodos e atributos irrelevantes para a compreensão básica a respeito destas classes.

Para gerenciar estas entidades virtuais, criou-se um agente, que também é mostrado na Figura 15. O gerenciador de entidades permite que os pacotes com dados coletados possam atualizar o estado e os atributos das entidades virtuais, por meio dos interpretadores.

Os interpretadores são necessários para que a aplicação possa receber mensagens em diferentes formatos e adequar a estrutura de dados aos atributos das entidades de maneira padronizada. Assim, é possível receber dados para diferentes tipos de entidades, como transformadores de distribuição ou unidades consumidoras, ou ainda de diferentes fontes para dados, que possuam estruturas diferentes.

Figura 16 – Classes de entidades virtuais da aplicação VPlant *back-end*.

Fonte: o autor.

Além de atualizar os atributos das entidades, a aplicação emite eventos conforme as regras de cada entidade. Estes eventos geralmente estão relacionados à mudança de estado ou de um tipo de atributo da entidade. Isto permite que outras aplicações possam se inscrever nestes eventos, e receberem atualizações dos recursos de forma seletiva. Para isso, os eventos são representados por meio de *tags*, com a seguinte estrutura:

<tipo de entidade>.<código de entidade>.<tipo de recurso>.

Para que as aplicações requisitantes possam fazer inscrições genéricas, isto é, por exemplo, receber atualizações dos estados de todas as entidades ou de todos os transformadores, é possível utilizar um caractere asterisco como *wildcard*, em qualquer uma das partes da *tag*.

Visando permitir atualizações parciais dos atributos de uma entidade, utilizou-se uma estrutura que armazena *timestamps* sob demanda, isto é, somente em casos de atualização parcial de um atributo. Isto evita uma deterioração de desempenho desnecessária, uma vez que a maior parte dos dados chegam à aplicação agrupados. Para melhorar o desempenho, foram implementados serializadores para algumas mensagens, que convertem as informações para um formato binário e retiram redundâncias.

Para a execução dos algoritmos, a aplicação precisa buscar dados históricos em fontes que não estão padronizadas conforme as suas estruturas. Os dados coletados dos casos de uso apresentados, por exemplo, foram idealizados para outras aplicações, específicas para transformadores e unidades consumidoras. Assim, visando manter a compatibilidade entre as entidades virtuais, os dados externos são adequados com a utilização dos *field adapters*.

#### **4.2.2 Elementos básicos para monitoramento**

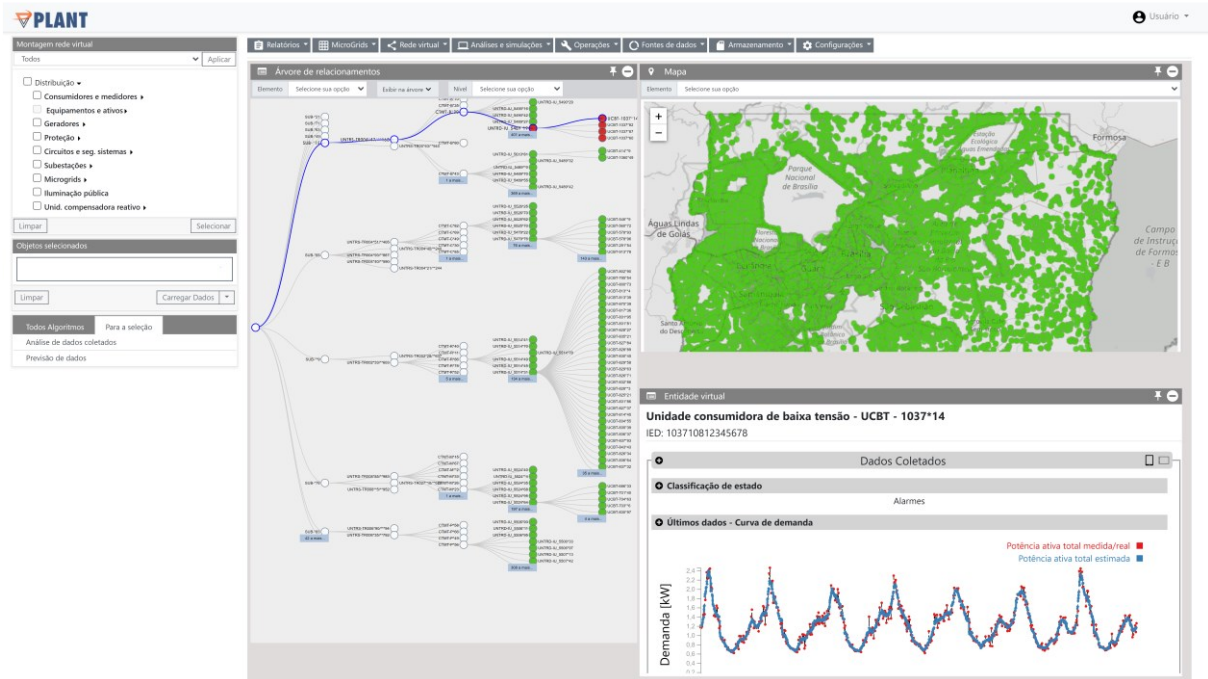
A Figura 17 mostra a interface gráfica da aplicação implementada. Alguns elementos mostrados nesta figura, como menus laterais e superiores, visam permitir que o usuário da ferramenta possa navegar entre as funcionalidades da aplicação. Os elementos centrais – árvore de relacionamentos, mapa e o painel da entidade virtual – permitem que o usuário possa monitorar as entidades da rede elétrica de distribuição em diferentes níveis e formas.

Cada ponto no mapa, ou nó na árvore, representa uma entidade virtual. As cores destes elementos indicam o estado de suas respectivas entidades. Por sua vez, os estados são definidos de acordo com a categoria dos eventos registrados para as entidades, considerando o de maior gravidade. É válido destacar, que todos os códigos de entidades mostrados nas interfaces foram mascarados.

Qualquer um dos painéis mostrados pode ser redimensionado, o que permite uma melhor visualização. O mapa e a árvore de relacionamentos têm funções de zoom, e podem ser utilizados para a seleção de entidades. O mapa foi inserido na aplicação para que o usuário pudesse assimilar aspectos geográficos das entidades com facilidade. Na verdade, constitui elemento indispensável no monitoramento de entidades da rede distribuição.



Figura 17 – Elementos gráficos iniciais da aplicação VPlant.

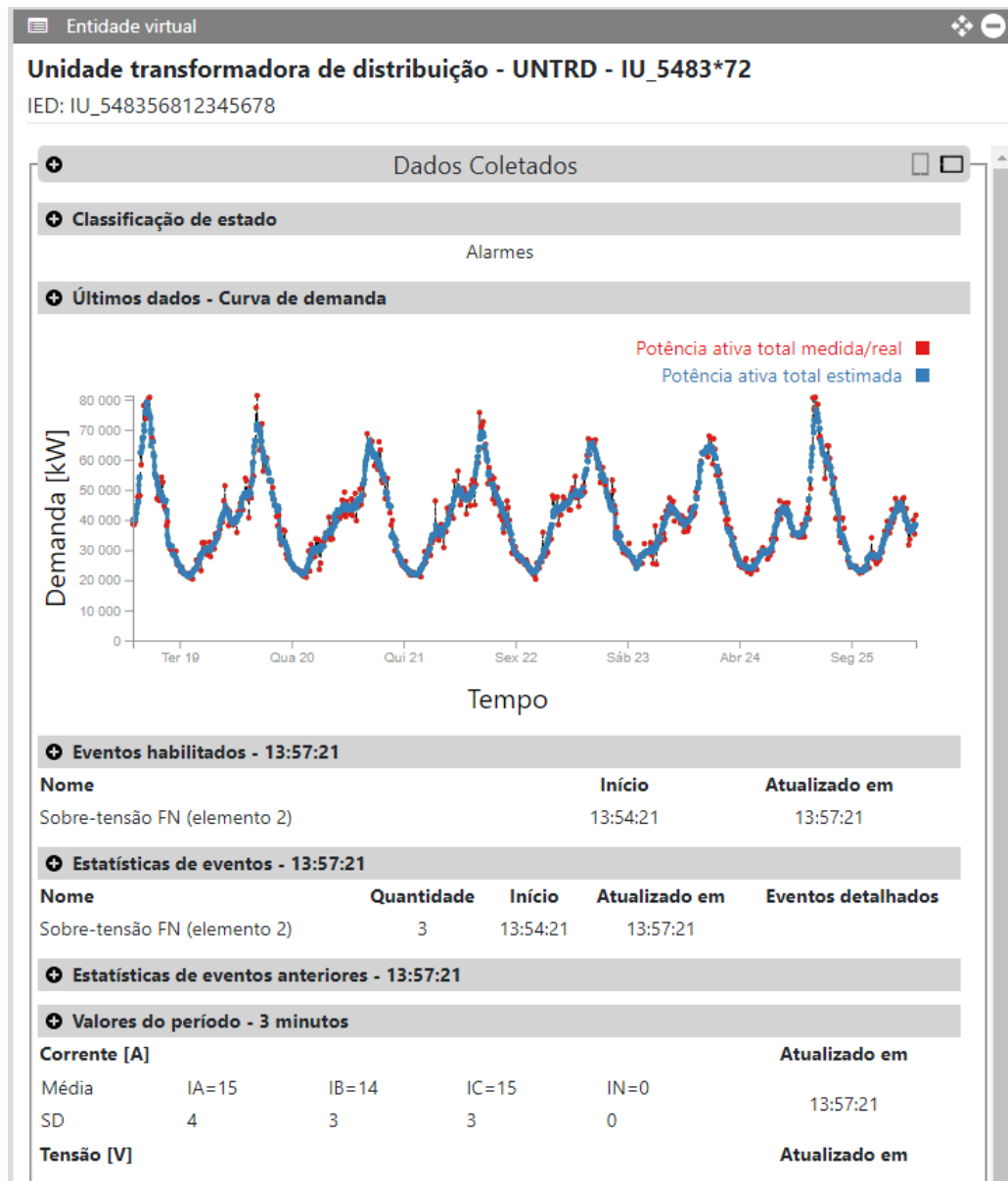


Fonte: o autor.

O painel de “Entidade Virtual”, mostrado no canto inferior direito da Figura 1, e em detalhes na Figura 18, mostra diversos detalhes da entidade virtual selecionada: o estado, um gráfico de linha com os últimos dados de potência ativa coletados para a entidade, e os demais atributos. As seções podem ser encolhidas, o que permite ao usuário visualizar as propriedades necessárias.

O gráfico de linha exhibe as medidas coletadas, mostradas em vermelho na Figura 18, e uma curva em azul, que representa a média semanal para cada horário e dia. Assim, além de completar informações em casos de dados faltosos, a linha azul representa uma curva de carga típica para a entidade, considerando as últimas semanas. O número de semanas e o intervalo de agregação considerados, dependem dos parâmetros enviados ao *back-end*.

Figura 18 – Painel com detalhes da entidade virtual.

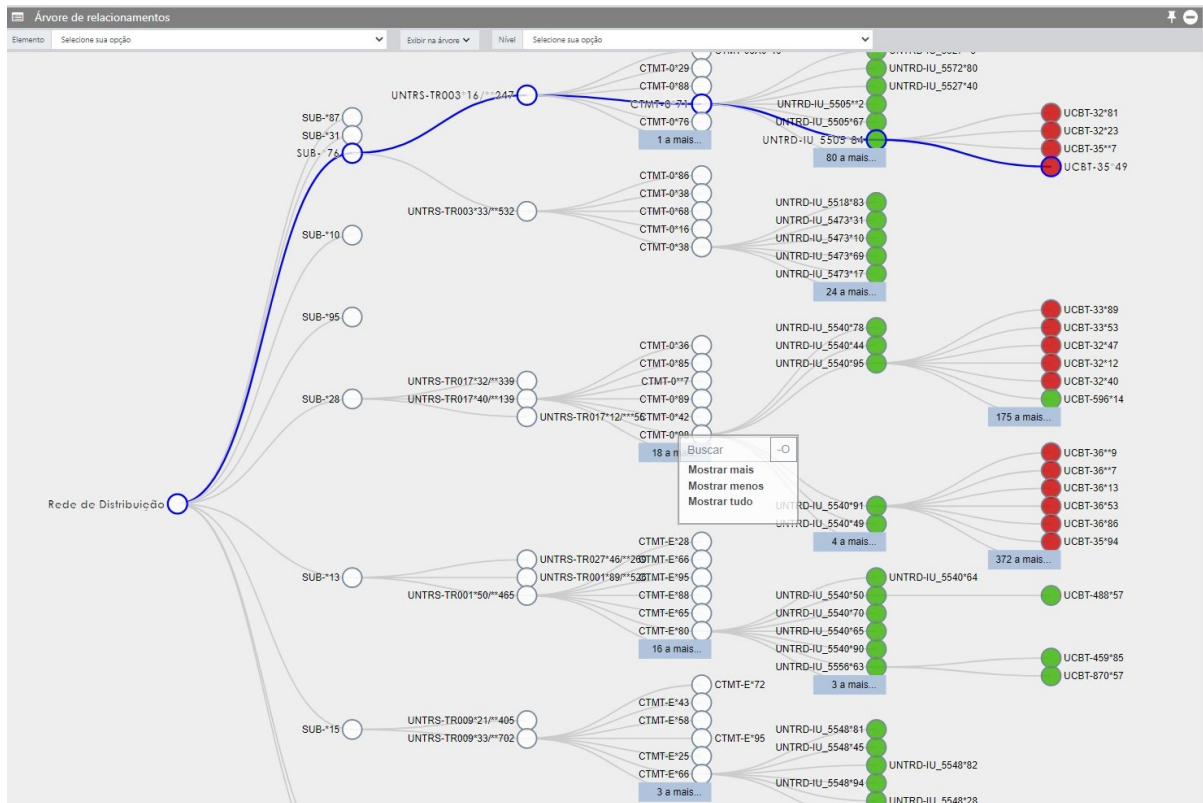


Fonte: o autor.

### 4.2.3 Representação da rede de distribuição com visualização em árvore

A árvore de relacionamentos, em destaque na Figura 19, contém nós representando os principais elementos da rede de distribuição elétrica: subestações, unidades transformadoras de subestação e distribuição, circuitos de média tensão e unidades consumidoras. Este *layout* foi inspirado em [142], e foi empregado na aplicação para permitir que o usuário pudesse conhecer e explorar a hierarquia da rede de distribuição com facilidade.

Figura 19 – Árvore de relacionamentos.



Fonte: o autor.

Cada nó é acompanhado do tipo e código da entidade que representa. O tipo das entidades corresponde às siglas definidas no módulo 10 do PRODIST [146], e são, desta forma, padronizadas. A hierarquia dos elementos mostrados na árvore foi definida pelos dados obtidos numa BDGD. Para isso, um algoritmo recursivo foi implementado, que agrupa os elementos inferiores, com base em atributos que identificam o elemento superior em comum.

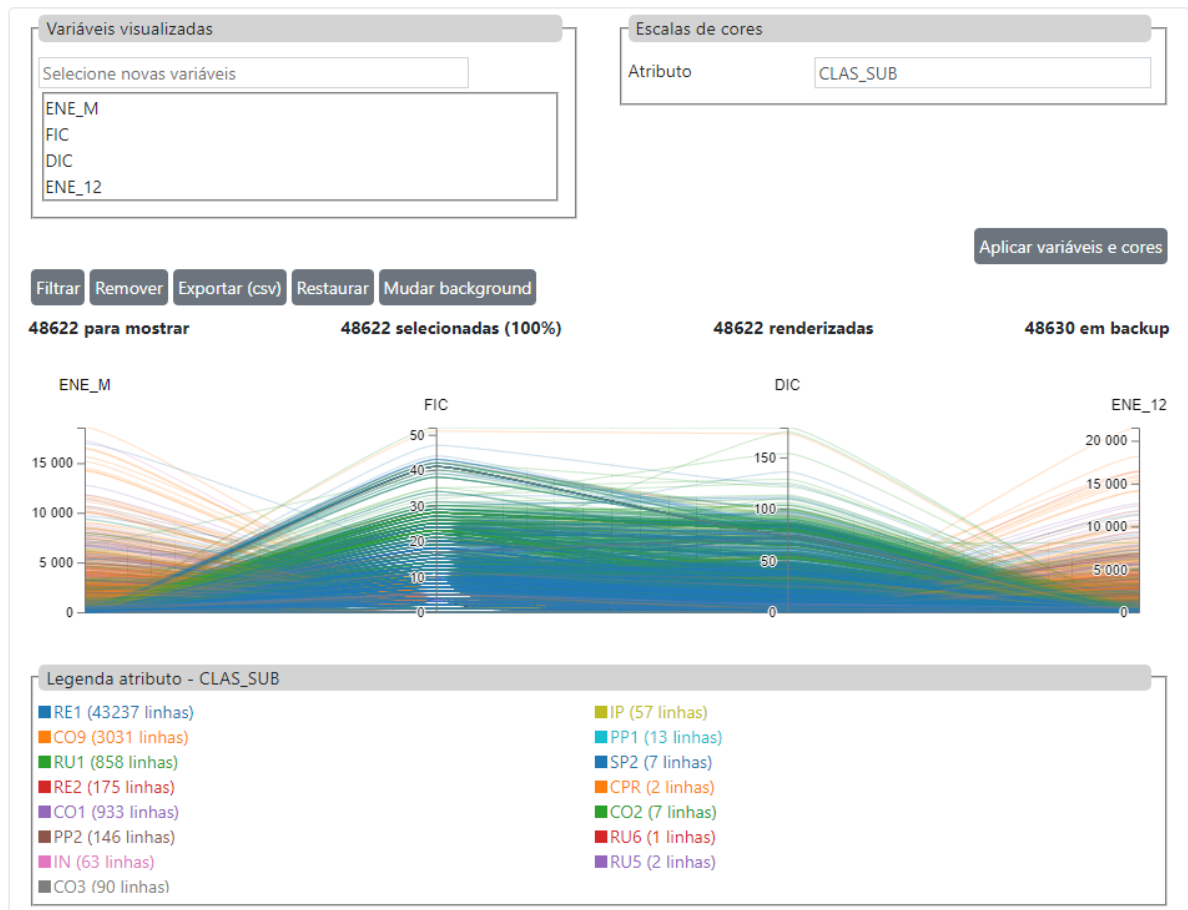
Ao clicar em uma entidade, na árvore ou no mapa, os elementos a jusante desta entidade são exibidos. Se o elemento selecionado no mapa não estiver visível na árvore, os nós a montante da entidade selecionada são expandidos em todos os níveis, de tal forma que o elemento passa a ser exibido.

Considerando o elevado número de entidades presentes em grandes sistemas de distribuição, foi necessário limitar o número de nós mostrados ao desaninhar os elementos da árvore. Por isso, caso o número de entidades filhas seja superior ao limite definido, um elemento com o número de entidades filhas excedentes é exibido. Ao clicar neste elemento, são apresentadas opções para que o usuário possa exibir as entidades desejadas, o que é mostrado no centro da Figura 19.

#### 4.2.4 *Layout* para visualização em coordenadas paralelas

Considerando a complexidade e volume de dados no contexto das *smart grids*, e até mesmo em sistemas de distribuição convencionais, implementou-se um *layout* em coordenadas paralelas. Neste sentido, tal visualização visa permitir que dados com elevada dimensão possam ser analisados. Este *layout* é mostrado na Figura 20, e foi inspirado em [143].

Figura 20 – Visualização em coordenadas paralelas na aplicação.



Fonte: o autor.

O *layout* implementado possui diversos recursos de interação e é capaz de exibir dezenas de milhares de instâncias. Na Figura 20, cada linha está relacionada a um elemento da rede elétrica. Como observado na imagem, visualizar um grande volume de dados pode resultar em um elevado grau de sobreposição. Por isso a interface renderiza as linhas em lotes, com uma animação, toda vez que uma nova seleção é feita. Isto permite o usuário ver algumas linhas, que ficariam ocultas devido à sobreposição. Além disso, os dados podem ser dinamicamente filtrados pelo usuário. Por outro lado, a densidade de linhas sobrepostas pode indicar padrões.

Na parte superior da Figura 20, há duas seções que mostram configurações para a visualização. Na seção “Variáveis visualizadas”, o usuário pode inserir e remover atributos

visualizados. Na seção “Escala de cores”, mostrada no canto superior direito da Figura 20, o usuário pode escolher o atributo que será utilizado para atribuição de cor às linhas visualizadas. Além disso, escolhe entre cinco tipos de escala de cor, sendo uma somente para atributos categóricos, e outras quatro para atributos numéricos.

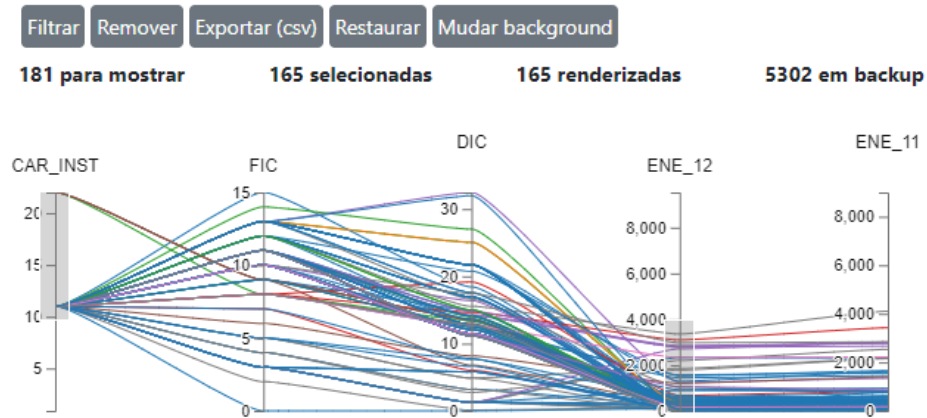
A escala de atributos categóricos atribui até dez cores diferentes para cada categoria do atributo, e repete as cores caso haja mais que dez categorias, conforme pode ser observado na legenda, mostrada na parte inferior da Figura 20. No caso dos atributos numéricos, o usuário pode escolher entre “Contínua”, “Discreta – quantis”, “Discreta – igualmente espaçada” e “Discreta – intervalos personalizados”.

Caso o usuário selecione a escala contínua, serão atribuídas cores que variam progressivamente em todo o domínio dos valores do atributo de cor. Caso a segunda opção seja selecionada, o usuário pode inserir o número de quantis em que os dados serão divididos. Escolhendo quatro, ter-se-ão quatro categorias de cor representando os quartis da distribuição de valores para o atributo selecionado.

Se o usuário escolher a opção “igualmente espaçada”, pode inserir o número de categorias desejadas. Desta forma, serão criadas classes de cores de mesma extensão, no intervalo de valores para o atributo de cor. Por fim, o usuário pode ainda definir os limites de cada classe de cor, com intervalos personalizados, escolhendo a quarta opção para atributos numéricos, em que os limites são definidos numa caixa de texto separados por “;”.

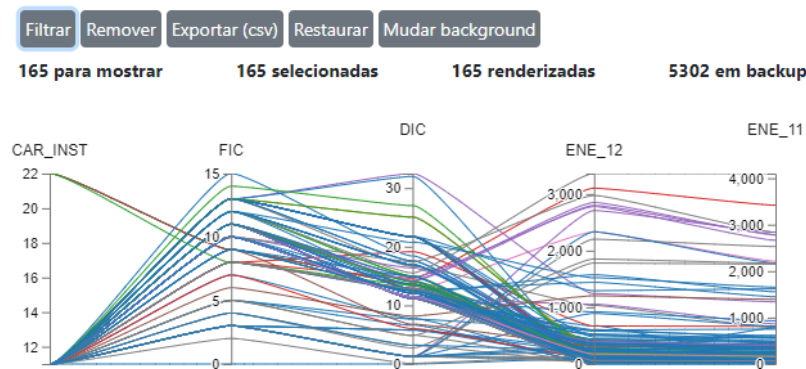
A interface contém ainda uma funcionalidade de seleção utilizando *brushes* em cada um dos eixos dos atributos, que são utilizados para restringir as linhas mostradas a intervalos escolhidos pelo usuário. Na Figura 21, dois *brushes* são utilizados simultaneamente para selecionar as unidades com carga instalada (CAR\_INST) maior que 10 kW e consumo em dezembro (ENE\_12) menor que 4000 kWh. A nomenclatura destas variáveis é a mesma empregada na BDGD, preservando a padronização indicada pela ANEEL [146].

Assim, como mostrado na Figura 21, os eixos mostram valores reais das grandezas, como consumo em kWh ou a carga instalada em kW, possibilitando ao usuário uma rápida associação da dimensão dos valores visualizados, além de facilitar a seleção e realização de filtros.

Figura 21 – Seleção de linhas com dois *brushes*.

Fonte: o autor.

Ainda na Figura 21, é possível observar vários botões acima da visualização: “Filtrar”, “Remover”, “Exportar (csv)”, “Restaurar” e “Mudar background”. O primeiro deles possibilita ao usuário observar apenas as linhas selecionadas, de tal forma que os eixos são alterados para os intervalos da seleção. A Figura 22 mostra o que foi obtido após filtrar os dados com a seleção apresentada na Figura 21. Observa-se que os valores de carga instalada estão acima de 10 kW, e o consumo em dezembro abaixo de 4000 kWh.

Figura 22 – *Layout* em coordenadas paralelas após utilização da função “Filtrar”.

Fonte: o autor.

O botão “Remover” por sua vez, faz o contrário do anterior, removendo o conteúdo selecionado e adequando os eixos aos novos intervalos. Todas essas mudanças na ferramenta acontecem com animações. Além disso é possível exportar a seleção no formato csv com o terceiro botão.

Em todos estes cenários, é possível acompanhar o número de linhas “a mostrar”, isto é, dentro dos limites dos eixos, bem como o número de amostras selecionadas, renderizadas e em backup, isto é, que podem retornar à visualização por meio do botão “Restaurar”. O último

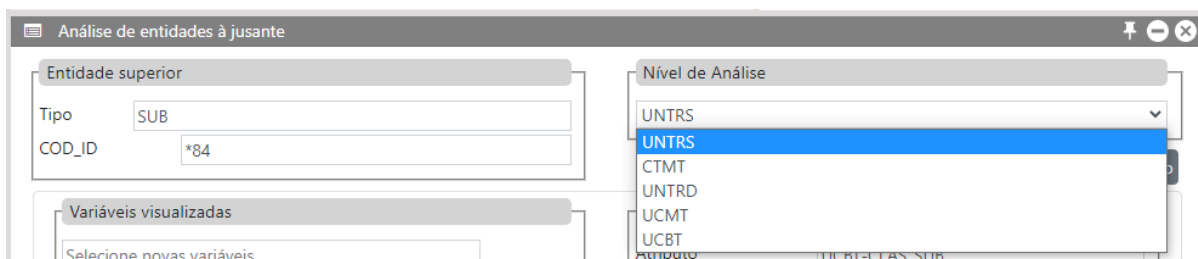
botão permite o usuário mudar a cor de fundo para preto, visando melhorar a visualização de certas cores.

As coordenadas paralelas foram utilizadas em diferentes contextos na aplicação: para visualizar dados estáticos provenientes da BDGD, o que pode visualizado nas figuras anteriores desta subseção; coletados por dispositivos, como abordado na Seção 5.1.2.2; e para a avaliação dos modelos de previsão de carga, como será detalhado na Subseção 4.2.5.

Em todos estes casos, para acessar as visualizações, seleciona-se uma entidade e em seguida, a opção no menu de algoritmos para a seleção, mostrado no canto inferior esquerdo da Figura 17. Desta forma, painéis dinâmicos são abertos.

No caso dos dados da BDGD, buscou-se explorar a hierarquia da rede, representada pela árvore de relacionamentos. Neste sentido, os dados visualizados correspondem às entidades a jusante da entidade selecionada. No painel dinâmico para esta funcionalidade, há duas seções de configuração adicionais, mostradas na Figura 23. Na seção “Nível de análise”, é possível escolher entre todos os tipos de entidades a jusante da entidade selecionada.

Figura 23 – Seções de configuração do painel de análise de entidades a jusante.



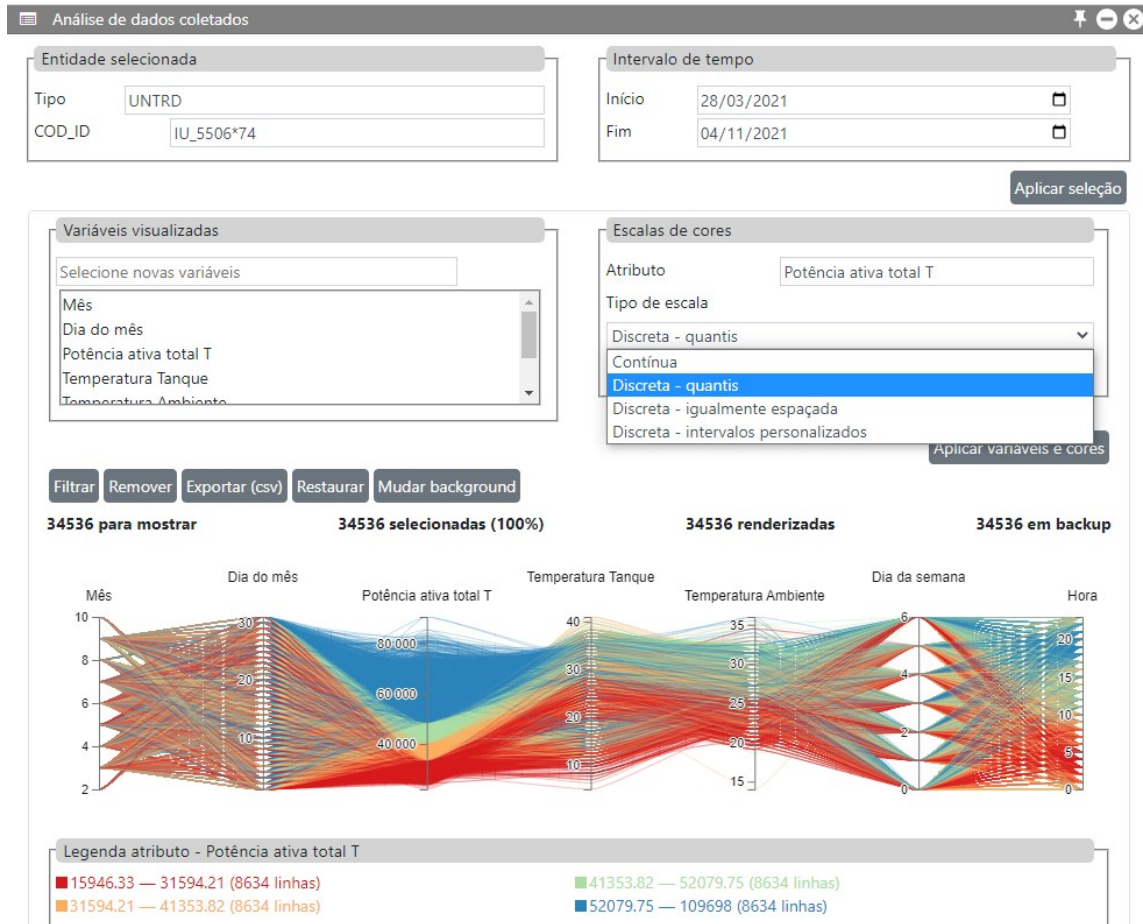
Fonte: o autor.

A Figura 24 mostra o painel dinâmico para a visualização de dados coletados com coordenadas paralelas. Neste painel, é possível selecionar o intervalo de tempo para as amostras a serem exibidas na visualização.

Para que os aspectos temporais destes dados pudessem ser explorados, criou-se eixos para representar atributos como hora, dia da semana, mês etc. Desta forma, pode-se por exemplo, selecionar facilmente medidas realizadas no sábado entre 10 e 15 horas.

Nesta figura, também é possível observar a escala de cores representando quartis da distribuição do atributo “Potência ativa total”. Desta forma, as três cores, concentradas na parte inferior das coordenadas, representam 75% dos dados, indicando que a maioria das amostras estão abaixo de 50 MW.

Figura 24 – Visualização de dados coletados em coordenadas paralelas.



Fonte: o autor.

#### 4.2.5 Previsão de carga

Com o intuito de exemplificar a utilização de algoritmos na aplicação, foram utilizados algoritmos de previsão de carga. A escolha de algoritmos de previsão como exemplo de implementação na aplicação, está relacionada à importância e à popularidade destes algoritmos para o setor elétrico. Tal importância pode crescer ainda mais com o crescente uso de fontes de energia renováveis intermitentes, como a solar e a eólica, e a popularização de carros elétricos.

O painel dinâmico implementado para as funções relacionadas à previsão de carga é mostrado na Figura 25. É possível observar várias seções e subseções neste painel, que podem ser encolhidas ou exibidas, o que facilita a navegação no painel.

Na seção de “Seleção de dados”, mostrada em detalhes na Figura 26, é possível observar uma lista na subseção “Dados criados”. Esta lista representa os dados resultantes da agregação de dados coletados. Esta agregação é realizada para homogeneizar a escala temporal, diminuir a quantidade de dados faltosos e o efeito de *outliers*. Para isso, os dados são consolidados em médias ou medianas a cada intervalo de agregação.



Figura 25 – Painel para previsão de dados.

Previsão de dados

Entidade selecionada

Tipo UNTRD COD\_ID IT\_ET58031\*72

+ Seleção de dados

+ Dados criados

+ Novo conjunto de dados

+ Modelos

+ Modelos criados

+ Novo modelo

+ Métricas de treino

Fonte: o autor.

Figura 26 – Seção “Seleção de dados” do painel de previsão de dados.

Previsão de dados

Entidade selecionada

Tipo UNTRD COD\_ID IT\_ET58031\*\*3

+ Seleção de dados

+ Dados criados

ID: 169434 - De 02/08/2021 a 15/12/2021, agregado a cada 2880 min. por mediana - 64 seqüências para intervalos de previsão de 10080 min de entrada e 2880 de saída.

ID: 206224 - De 02/08/2021 a 15/12/2021, agregado a cada 60 min. por mediana - 2722 seqüências para intervalos de previsão de 10080 min de entrada e 60 de saída.

ID: 485948 - De 02/08/2021 a 15/12/2021, agregado a cada 30 min. por mediana - 4341 seqüências para intervalos de previsão de 10080 min de entrada e 60 de saída.

Apagar

+ Novo conjunto de dados

Intervalo de tempo dos dados

Início 02/08/2021

Fim 15/12/2021

Agregação

Tipo Mediana

Intervalo (horas) 1

Intervalos para a previsão

De entrada (dias) 7

De saída (horas) 0,5

Variáveis do conjunto de dados

Selecione novas variáveis

Potência ativa total T

Criar

+ Modelos

Fonte: o autor.

Para criar os dados agregados, utiliza-se a subseção “Novo conjunto de dados”, em que é possível escolher o intervalo de tempo para as amostras do conjunto, o modo de agregação, por média ou mediana, e o intervalo de agregação, e os intervalos de entrada e saída para a previsão. Os intervalos de entrada e saída para previsão são parâmetros utilizados para calcular o número de sequências válidas para o treinamento de modelos. Esta informação é exibida na lista dos conjuntos de dados criados, conforme Figura 26.

É válido destacar que, nesta versão da aplicação, o intervalo de previsão de saída deve ser o mesmo do intervalo de agregação. Isto ocorre porque os modelos implementados realizam previsões para um intervalo de agregação, e então predizem para intervalos maiores de maneira iterativa. A criação de modelos com previsão *multi-step* direta, representa um aspecto interessante para melhorias futuras.

Além disso, os modelos utilizam apenas uma variável de entrada e saída. Esta limitação não está relacionada aos modelos, mas sim ao modo como as métricas de erro foram organizadas. É interessante que seja possível utilizar mais variáveis de entrada, como a temperatura ambiente. No entanto, conforme apresentado na Seção 2.4.1, para horizontes de previsão curtos, a influência da temperatura é menor.

A seção “Modelos” têm suas subseções mostrada em detalhes na Figura 27. A primeira subseção, “Modelos criados”, apresenta uma lista de modelos e alguns de seus atributos: tipo de camada recorrente, número de neurônios nas camadas recorrentes, tipo de camada de agrupamento, número de épocas máximo para treinamento, data e horário de criação e a ID do conjunto de dados utilizado para treinamento. Os demais parâmetros dos modelos são exibidos quando o usuário passa o *mouse* em cima dos modelos na lista.

A subseção “Novo modelo”, também mostrada na Figura 27, apresenta os vários parâmetros que podem ser escolhidos para os modelos. Além dos parâmetros exibidos na lista, temos o número de camadas recorrentes, a taxa de aprendizado, o tamanho do *batch*, o percentual de dados para a validação, além do tipo de normalização e seus parâmetros. O tamanho do *batch*, equivale ao número de amostras utilizadas em cada atualização dos pesos, no treinamento dos modelos.

Assim, os modelos implementados dinamicamente pelo usuário, consistem em uma ou mais camadas recorrentes, dos tipos recorrente simples, LSTM ou GRU, com ou sem camada de agrupamento, dos tipos GlobalMaxPooling1D ou GlobalAveragePooling1D. Estas camadas de agrupamento simplesmente calculam o máximo e a média das saídas da última camada recorrente, para todos os passos numa sequência de entrada.

Figura 27 – Seção “Modelos” do painel de previsão de dados.

The screenshot displays the 'Modelos' section of a data prediction dashboard. It is divided into three main panels:

- Modelos criados:** A table listing created models with columns for Tipo, Estado, Neurônios, Camada de agrupamento, Épocas, Criado em, and ID dos dados. The table contains three rows of LSTM models with 'Treino finalizado' status.
- Novo modelo:** A form for creating a new model, divided into three sections:
  - Parâmetros do modelo:** Tipo (LSTM), Neurônios na camada escondida (16), Número de camadas recorrentes (1), Camada de agrupamento (Nenhuma).
  - Pré-processamento de dados:** Tipo de normalização (Min-max), Novo mínimo (-1), Novo máximo (1).
  - Parâmetros do treino:** Épocas de treinamento (20), Taxa de aprendizado (0,002), Tamanho de batch (64), Percentual de validação (0,2).
- Métricas de treino:** A section for displaying training metrics, currently empty.

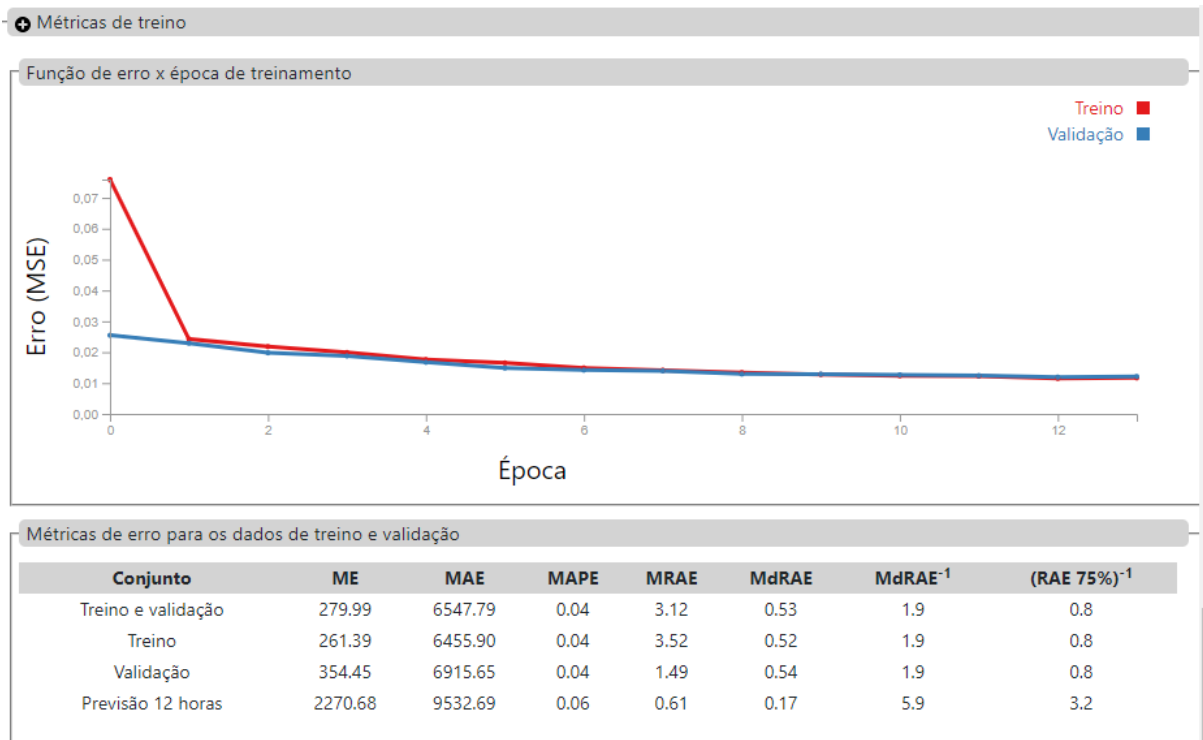
Fonte: o autor.

Após escolher os parâmetros, e selecionar um conjunto de dados, o modelo pode ser criado e treinado. Então é possível acompanhar o treino e, ao final, visualizar as métricas alcançadas pelo modelo, na subseção “Métricas de treino”, cuja parte superior é mostrada na Figura 28.

Nesta figura, são apresentados dois elementos: um gráfico de linha com a função de erro a cada época do treinamento, para os conjuntos de treino e validação; e uma tabela com várias métricas de erro para os conjuntos de dados. Caso o modelo selecionado esteja em treinamento, o gráfico da função de erro é atualizado a cada época.

As métricas mostradas na tabela foram definidas na Seção 2.4.2, com exceção de  $MdRAE^{-1}$  e  $(RAE\ 75\%)^{-1}$ . A métrica  $MdRAE^{-1}$  é o inverso da mediana do erro absoluto relativo. Assim, um valor de 0,54 para  $MdRAE$  indica que 50% das amostras do conjunto de validação tiveram uma previsão, no mínimo,  $\frac{1}{0,54} \approx 1,85$  vezes melhor com o modelo de rede neural em relação ao modelo de referência. De maneira análoga,  $(RAE\ 75\%)^{-1}$  indica o desempenho mínimo para 75% das amostras do conjunto avaliado.

Figura 28 – Parte superior da subseção “Métricas de treino”, do painel de previsão de dados.

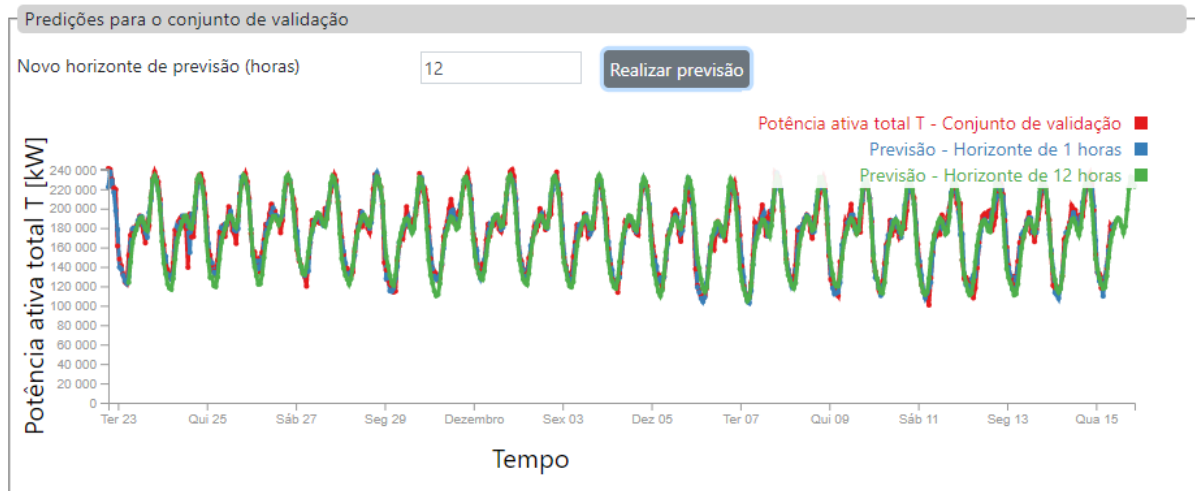


Fonte: o autor.

Nesta versão da aplicação, o algoritmo de referência utilizado equivale a repetir o último valor de cada subsequência utilizada para a previsão. Esta abordagem, apesar de relativamente eficiente num horizonte de previsão pequeno, é claramente inapropriada para horizontes de previsão de algumas horas. Este algoritmo de referência foi utilizado para possibilitar verificar as funcionalidades de previsão, uma vez que é comumente utilizado na literatura [101]. No entanto, deve ser trocado por um algoritmo mais eficiente caso o foco seja a seleção de modelos de previsão melhores.

Além disso, a tabela com as métricas de erro também inclui os dados para outros horizontes de previsão. Estes dados são inseridos na tabela conforme o usuário requisita novas previsões para o conjunto de validação, com outros horizontes. Para isso, basta utilizar o campo “Novo horizonte de previsão”, mostrado na Figura 29. Esta figura mostra um gráfico de linha com o conjunto de validação, e as previsões com diferentes horizontes temporais, realizadas para este conjunto.

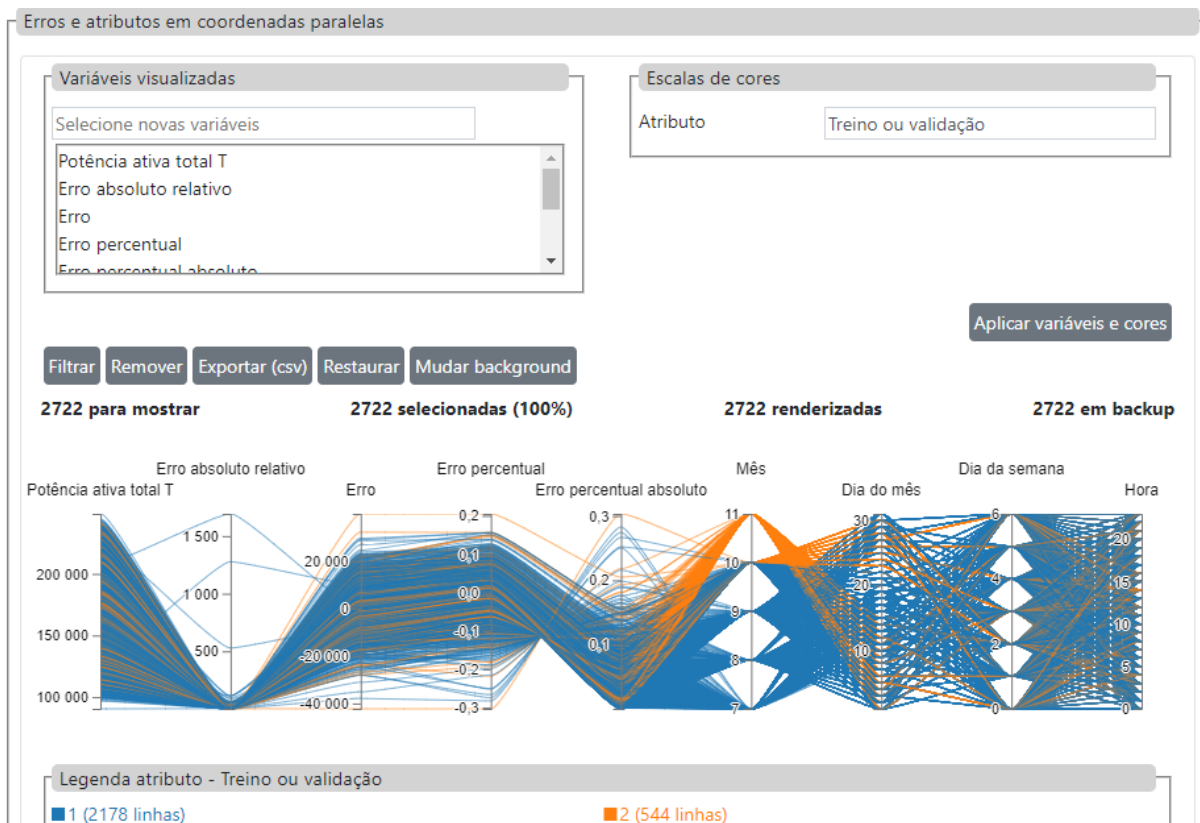
Figura 29 – Gráfico de linha com previsões para o conjunto de validação.



Fonte: o autor.

Por fim, abaixo deste gráfico, está uma visualização em coordenadas paralelas com os conjuntos de treino e validação. Além dos atributos originais, são mostradas métricas de erro. E assim como realizado na funcionalidade apresentada na Seção 4.2.4, para os dados coletados, também estão disponíveis eixos como hora, dia da semana, mês etc. Desta maneira, o usuário pode explorar como o modelo errou em relação aos demais atributos.

Figura 30 – Erros e atributos em coordenadas paralelas.



Fonte: o autor.

# 5 Resultados e discussões

Na primeira subseção deste capítulo, 5.1, são apresentados exemplos de aplicação das funcionalidades implementadas. Na Subseção 5.1.1 são mostrados alguns aspectos relacionados ao monitoramento das entidades: desempenho na recepção de dados em 5.1.1.1; e visualização de eventos em 5.1.1.2. Na Subseção 5.1.2 são apresentadas análises de dados estáticos e coletados com as coordenadas paralelas. Para os dados dinâmicos, atributos temporais, como dias da semana e horas, são explorados.

Na Seção 5.1.3, apresenta-se como as funcionalidades implementadas relacionadas à previsão de carga podem ser utilizadas. Neste contexto, são mostradas algumas métricas dos modelos de previsão de carga. Entretanto, é importante destacar que este trabalho não tem como objetivo a otimização extrema de um algoritmo de previsão. Por outro lado, ilustra-se como as coordenadas paralelas podem ser utilizadas para avaliar o desempenho de modelos de previsão, em conjunto com as várias métricas de erro calculadas.

Na Subseção 5.2 são apresentadas discussões sobre os resultados obtidos, com base na literatura e nos casos de uso apresentados na Subseção 5.1. Destacam-se alguns diferenciais da aplicação implementada, em relação aos sistemas presentes na literatura, bem como algumas limitações.

## 5.1 Casos de uso

### 5.1.1 Monitoramento

#### 5.1.1.1 Escalabilidade

Considerando um sistema de *smart grid* com milhares ou milhões de entidades monitoradas, manter as aplicações com desempenho adequado se torna um desafio. Neste contexto, simulou-se dois cenários: primeiramente com cerca de 28 mil transformadores de distribuição monitorados; e com cerca de 1,1 milhão de unidades consumidoras, além dos transformadores presentes no primeiro cenário. Em ambos os cenários o número total de entidades virtuais corresponde ao dobro dos elementos da rede de distribuição, uma vez que cada elemento da rede possui um dispositivo de comunicação associado.

Com um programa simulando dispositivos, isto é, enviando mensagens da mesma forma que estes enviariam à aplicação, verificou-se o limite máximo para a taxa de mensagens

processadas pela aplicação. Em ambos os cenários, a taxa de envio de pacotes foi mantida por 10 segundos. Isto é semelhante ao que aconteceria em caso de eventos simultâneos em vários elementos da rede de distribuição. Por exemplo, caso falte energia elétrica para muitos elementos monitorados, os dispositivos enviariam alertas, que chegariam com diferentes atrasos à aplicação.

Entre o VPlant e o *middleware* de mensagens, neste caso o programa *NodeStatus*, testou-se o envio de mensagens por UDP e *WebSockets*. A primeira alternativa se mostrou ineficiente, pois resultava em perdas de pacotes constantes. Para uma taxa de 2000 pacotes por segundo, as perdas ultrapassaram 10%. Por *WebSockets*, a aplicação suportou até 3000 pacotes por segundo. Isso equivale à 30 mil pacotes durante os 10 segundos do teste. Este valor é maior que o número de elementos monitorados neste cenário, o que demonstra que a aplicação estaria adequada neste contexto.

No segundo cenário, a taxa máxima utilizando *Websockets* chegou a 2000 pacotes por segundo. Isso sugere que o desempenho da aplicação, com relação a este aspecto, varia pouco com o número de entidades monitoradas. Por outro lado, os 20 mil pacotes recebidos representam apenas uma pequena parcela dos 1,1 milhão de entidades monitoradas neste cenário, o que representam uma limitação considerável. Tal limitação provavelmente poderia ser contornada com o uso de múltiplas instâncias da aplicação. Cada instância gerenciaria uma parte das entidades virtuais.

É interessante destacar que a em ambos os cenários a aplicação têm desempenho adequado para utilização da interface gráfica, estando a limitação relacionada ao *back-end*. Assim, o mapa e a árvore de relacionamentos apresentaram os 1,1 milhão de entidades de maneira satisfatória, como mostra a Figura 17.

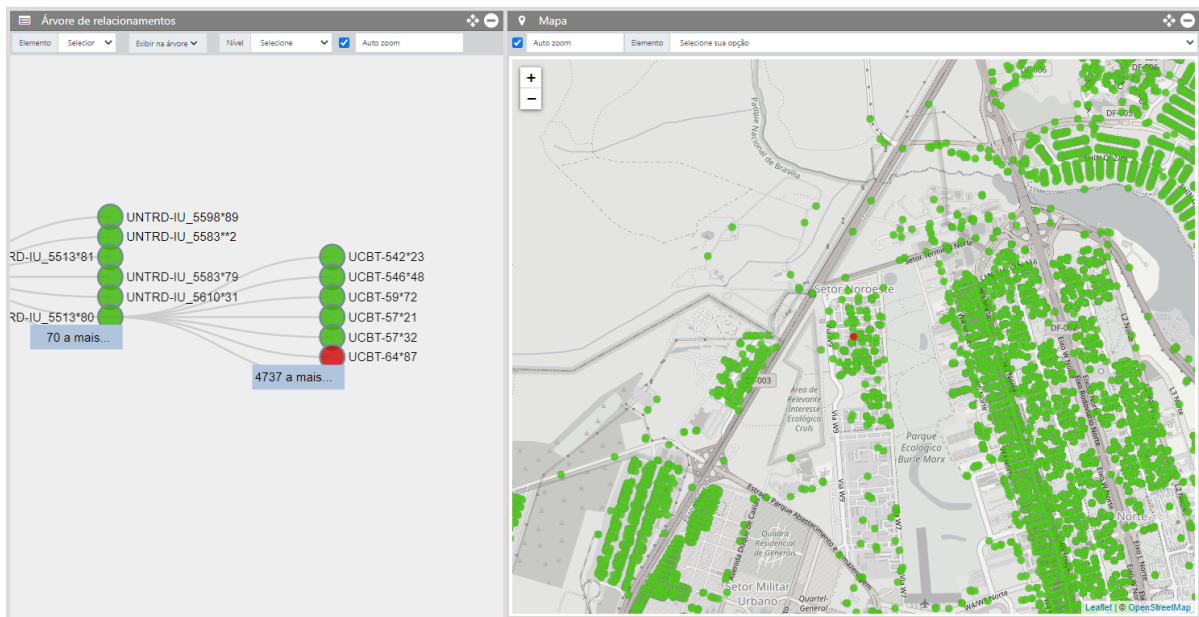
### 5.1.1.2 Visualização de estados

Com o intuito de verificar a utilidade da ferramenta para monitorar os estados das entidades, simulou-se alguns cenários variando o número de alarmes e o nível das entidades na hierarquia da rede. Primeiramente, em um cenário com um único evento, a sobreposição que ocorre no mapa poderia inviabilizar a visualização da ocorrência. No entanto, a função de *zoom* automático permitiu visualizar o evento de maneira eficiente, como mostra a Figura 31, em que os elementos em vermelho representam uma entidade em estado de alarme. Neste caso, tanto o mapa, quanto a árvore de relacionamentos, mostram que se trata de um evento isolado.

No segundo cenário, 17 alarmes foram enviados, referentes a um transformador de distribuição e às unidades consumidoras a jusante deste transformador. Na Figura 32, são

mostrados a árvore de relacionamentos e o mapa após a ocorrência destes eventos. Considerando que todas as UCs conectadas ao transformador na árvore estão em alarme, o usuário da ferramenta pode concluir que todos os eventos gerados provavelmente possuem uma única origem, relacionada ao transformador de distribuição. Para verificar esta suposição, o usuário poderia clicar no elemento que exibe a quantidade das demais UCs, a jusante do transformador, e visualizar todas as unidades.

Figura 31 – Evento em uma única entidade.



Fonte: o autor.

Figura 32 – Eventos em um transformador de distribuição e em 16 unidades consumidoras a jusante do transformador.



Fonte: o autor.



No terceiro cenário, foram enviados 112 alarmes para transformadores de distribuição, e cerca de 1500 alarmes para unidades consumidoras, que representam todas as entidades para um determinado circuito de média de tensão. O resultado exibido pela aplicação após os eventos é mostrado na Figura 33. Apesar da quantidade de eventos, as visualizações ainda conseguem exibir os eventos de forma satisfatória. A árvore de relacionamentos mais uma vez mostra que os eventos têm uma possível causa comum, desta vez relacionada ao circuito de média tensão.

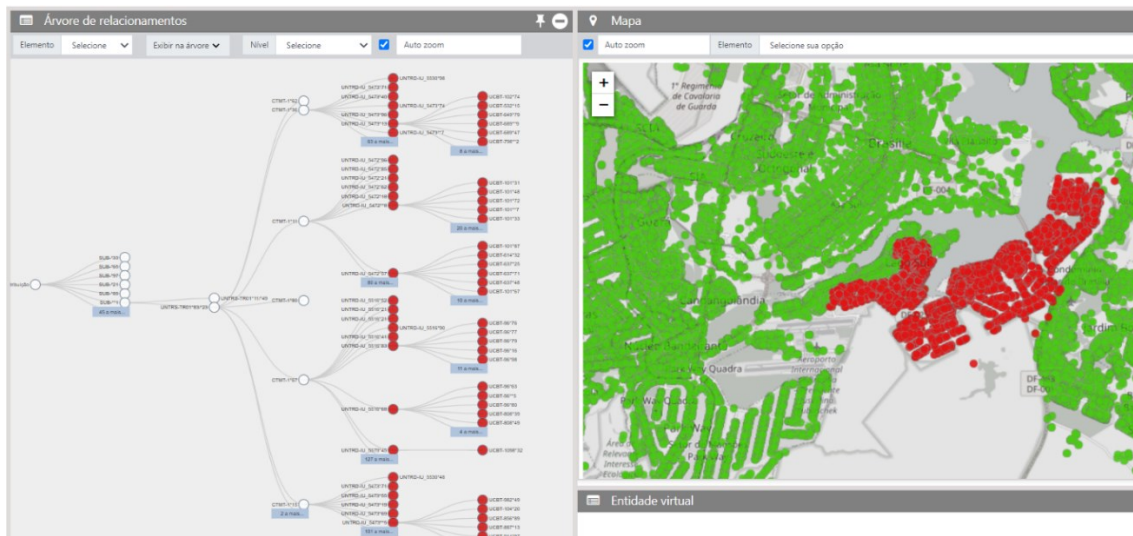
Figura 33 – Eventos em 112 transformadores de distribuição e cerca de 1500 unidades consumidoras, relacionados a um circuito de média tensão.



Fonte: o autor.

No quarto cenário, simulou-se alarmes em todos os transformadores de distribuição e unidades consumidoras a jusante de uma subestação, que consistem em cerca de 10 mil entidades. Na Figura 34 é possível observar novamente que a árvore auxilia na observação da origem comum dos eventos: a subestação. Além disso, a quantidade de entidades exibidas na árvore mostra, de forma visual, a gravidade e extensão dos eventos.

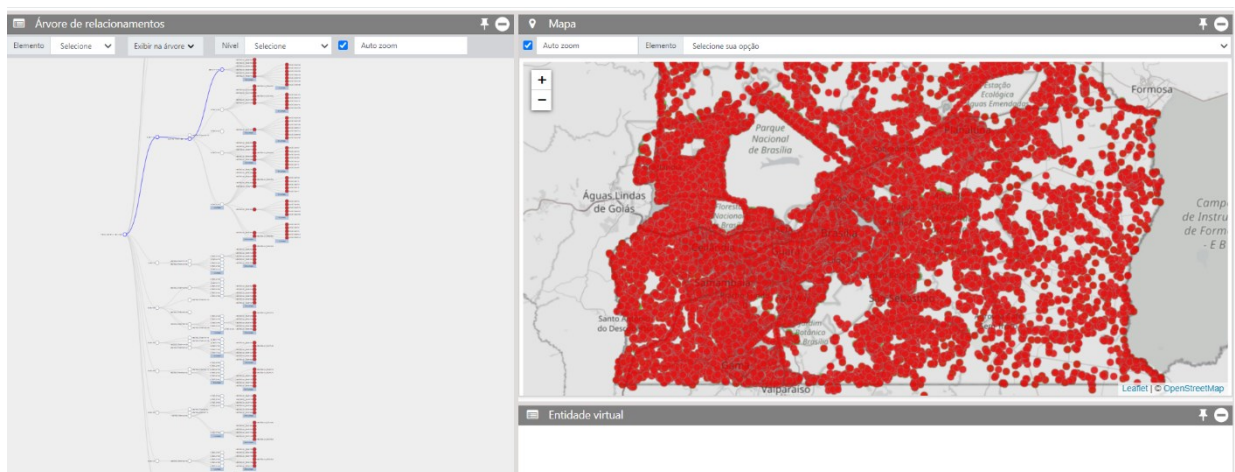
Figura 34 – Eventos em cerca de 10 mil entidades a jusante de uma subestação.



Fonte: o autor.

Por fim, simulou-se alarmes para todos os transformadores de distribuição. O resultado é mostrado na Figura 35. Desta vez, ambas as visualizações mostram claramente a gravidade do cenário. Destaca-se que em todos os cenários, apenas transformadores de distribuição e unidades consumidoras são elementos monitorados. Assim, mesmo sem informações das demais entidades, a visualização em árvore permitiu supor informações a respeito de algumas entidades não monitoradas.

Figura 35 – Eventos em quase 30 mil transformadores de distribuição.



Fonte: o autor.

## **5.1.2 Análises de dados em coordenadas paralelas**

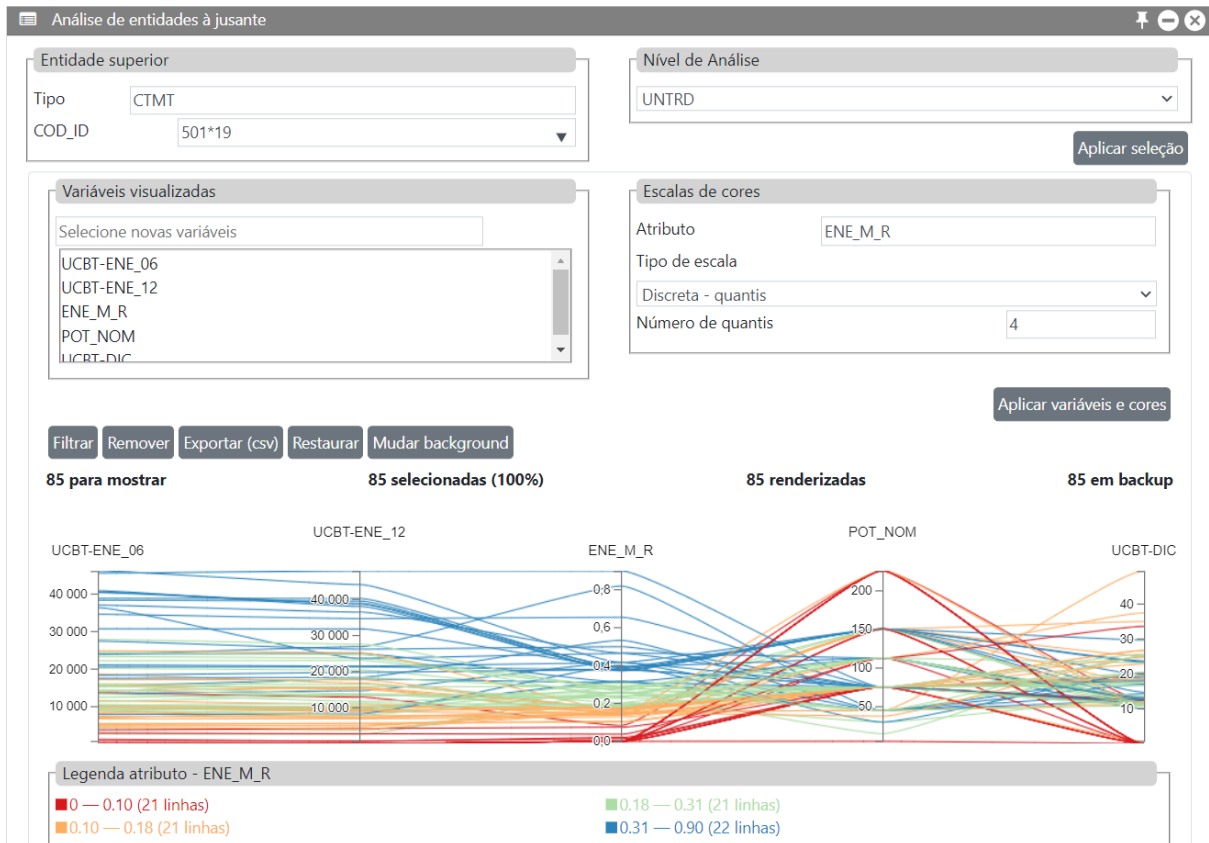
### **5.1.2.1 Dados estáticos**

Para analisar dados das BDGDs, considerando a grande quantidade de instâncias e atributos destas bases [146], é interessante aplicar visualizações adequadas para dados com alta dimensionalidade. Neste sentido, apresenta-se nesta subseção um exemplo de uso das coordenadas paralelas, aplicadas à análise de dados de uma BDGD. Como apresentado na Seção 4.2.4, esta funcionalidade foi agregada à ferramenta desenvolvida neste trabalho.

No caso de uso apresentado, supõe-se que se deseja diminuir os índices de perdas não técnicas em uma região, que é abastecida por determinado circuito. Considera-se ainda que, em unidades de tipo similar, uma carga instalada elevada, acompanhada de um baixo consumo de energia, seja um perfil suspeito. Também se supõe que a concessionária notou índices de qualidade piores em localidades com alto índice de perdas não técnicas. Outra suposição, é de que os meses com perdas não técnicas mais elevadas são junho e dezembro. Além disso, considera-se que os transformadores sobrecarregados, também indicam uma região interessante para análise.

Desta maneira, selecionando-se o circuito de interesse na árvore, e clicando no item “Análise de entidades a jusante com coordenadas paralelas”, a ferramenta exhibe o painel mostrado na Figura 36. Mais uma vez, os códigos mostrados estão mascarados com números aleatórios para preservar os dados.

Figura 36 – Painel com coordenadas paralelas e dados da BDGD.

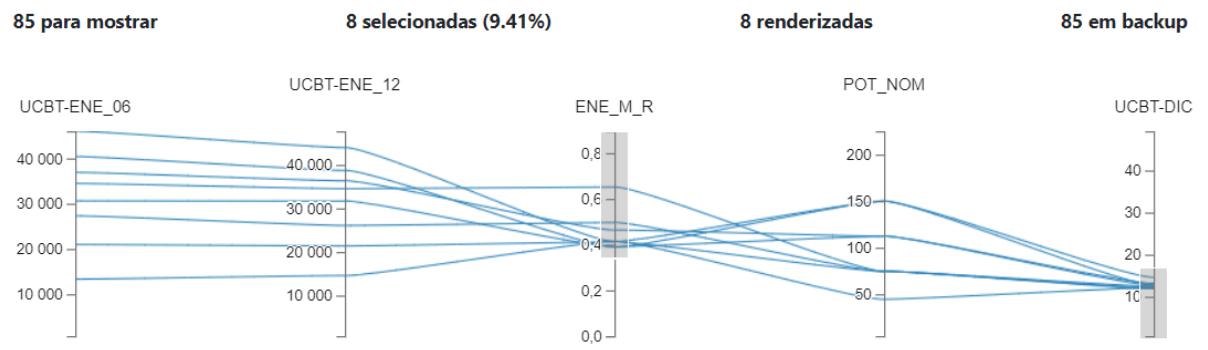


Fonte: o autor.

A seção “Nível de análise”, no canto superior direito da Figura 36, mostra que os dados mostrados nas coordenadas representam os atributos das unidades transformadoras de distribuição (UNTRDs). Os atributos mostrados na figura foram selecionados conforme as suposições apresentadas, e representam: a soma da energia consumida pelas unidades consumidoras a jusante do transformador, em junho (ENE\_06) e dezembro (ENE\_12); a razão entre a soma da potência média consumida por estas unidades, e a potência nominal de cada transformador (ENE\_M\_R), a potência nominal do transformador (POT\_NOM); e média do indicador DIC para as UCs a jusante dos transformadores.

A escala de cores foi ajustada para representar os quartis do atributo ENE\_M\_R, que representa o volume de cargas médio em relação à potência do transformador. Assim, as linhas em azul representam os 25% dos transformadores mais sobrecarregados. Selecionando-se o intervalo correspondente a este grupo, no eixo ENE\_M\_R, bem como os valores de DIC menores que 15, obteve-se o resultado mostrado na Figura 37. É possível observar que restaram apenas 8 transformadores selecionados, que podem ser exportados para um arquivo csv.

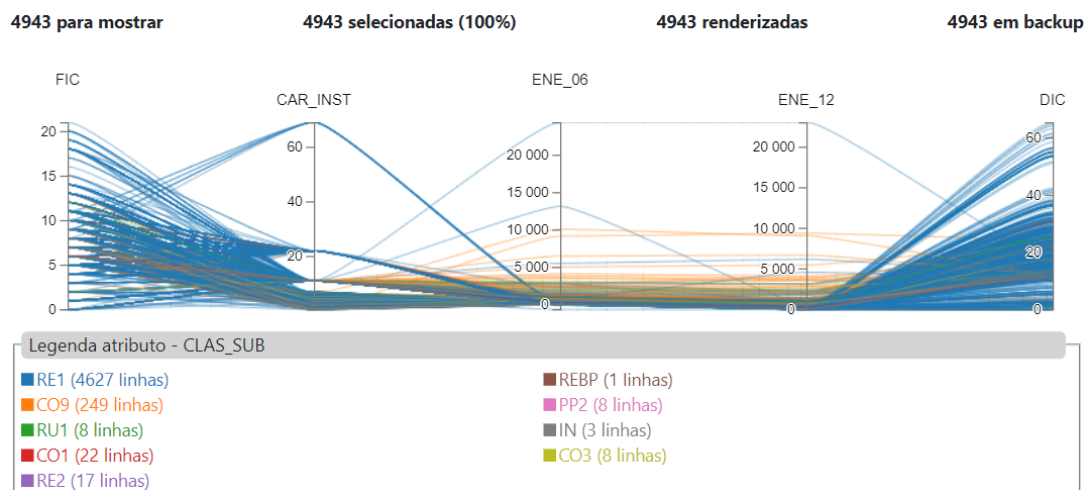
Figura 37 – Coordenadas paralelas com eixos selecionados.



Fonte: o autor.

Caso o usuário prefira fazer uma análise com unidades consumidoras de baixa tensão (UCBTs) diretamente, basta que selecione o respectivo nível na seção “Nível de análise”. O resultado é mostrado na Figura 38.

Figura 38 – Coordenadas paralelas com dados de unidades consumidoras.

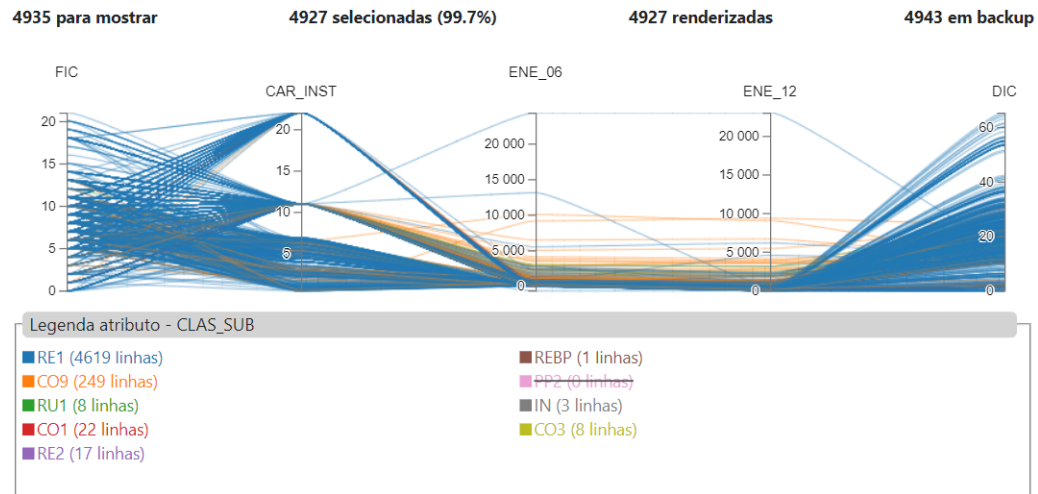


Fonte: o autor.

Nesta figura já é possível notar *outliers*, com elevada carga instalada, neste caso, acima de 60 kW. Seleciona-se então tais unidades com o uso do *brush*, no eixo da carga instalada, e salva-se a seleção em um arquivo csv.

Então, para seguir com a análise, retira-se esta seleção da visualização, o que permite os eixos se ajustarem aos dados restantes, como mostra Figura 39. Nesta figura, o limite máximo visualizado no eixo de carga instalada passou a ser de 20 kW. Além disso, como pode ser visto na legenda do gráfico de coordenadas paralelas, as linhas correspondentes à categoria poder público distrital ou estadual (PP2), foi retirada, uma vez que não fazem parte do perfil suspeito.

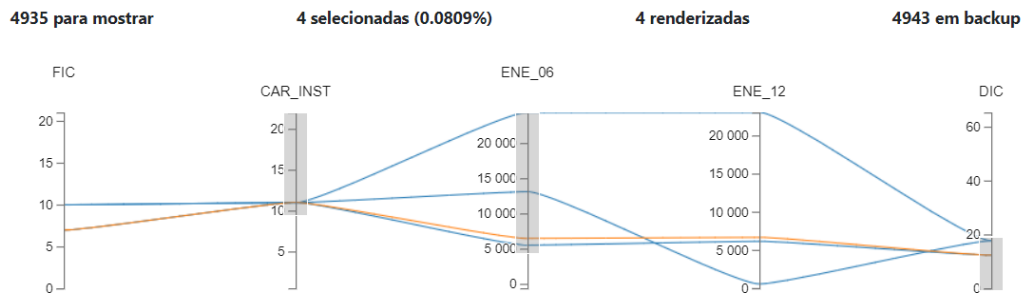
Figura 39 – Eixo de carga instalada com novo intervalo.



Fonte: o autor.

Pode-se então selecionar as linhas com carga instalada maior que 10 kW, energia consumida maior que 5000 kWh, e DIC maior que 10, conforme mostra a Figura 40. Segundo o que foi suposto, estas unidades seriam suspeitas, com relação a perdas não técnicas. É interessante destacar que, as 4 linhas selecionadas, das 4943 no total, mostram que o *layout* e as interações implementadas viabilizam uma rápida análise, uma vez que possibilitam a seleção das instâncias de interesse com facilidade.

Figura 40 – Dados selecionados para análise com coordenadas paralelas.



Fonte: o autor.

### 5.1.2.2 Dados coletados

Nesta subseção, apresenta-se uma análise sucinta de dados coletados por dispositivos que monitoram transformadores de distribuição. Como mencionado nas seções 4.1.2 e 4.1.3, estes dispositivos foram elaborados por outros pesquisadores do projeto que este trabalho faz parte. A Figura 41 mostra os dados de um transformador em um *layout* de coordenadas paralelas. São exibidas mais de 12 mil amostras, com os atributos de potência ativa e fator de potência na fase A, temperatura ambiente, dia da semana e hora das medidas coletadas. A escala de cores foi configurada para representar os quartis do atributo de potência ativa na fase A.

Figura 41 – Dados de um transformador em coordenadas paralelas.

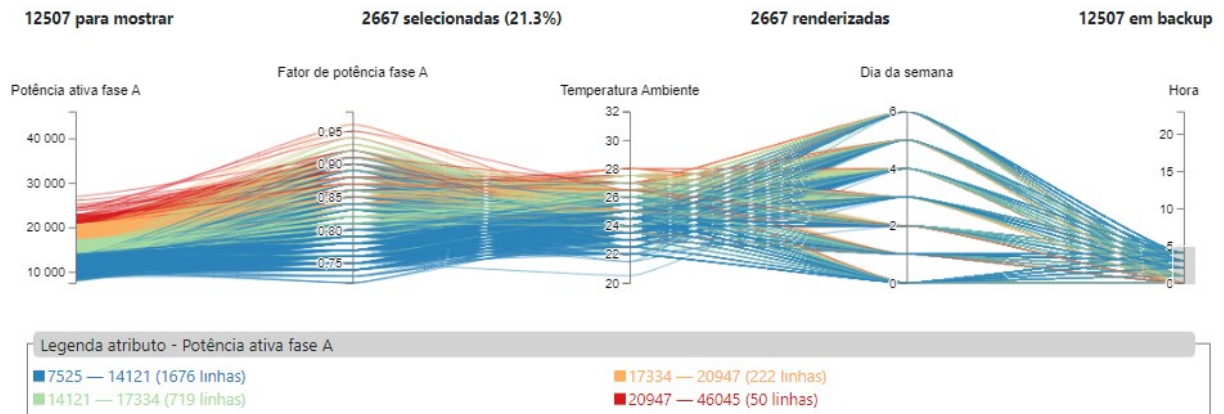


Fonte: o autor.

É possível visualizar que os atributos potência ativa, fator de potência, temperatura ambiente e horas, possuem uma concentração de linhas azuis e verdes na parte inferior dos eixos. Por outro lado, possuem uma concentração de linhas laranjas e vermelhas na parte superior dos eixos. Como mostra a legenda na parte inferior da Figura 41, o primeiro grupo de cores representa potências ativas de 7 a 17 MW, enquanto o segundo representa potências de 17 a 46 MW.

Assim, é possível observar que fatores de potência mais baixos estariam relacionados a um baixo consumo. No entanto, como as coordenadas paralelas exibem vários outros atributos, fica evidente que os fatores de potência baixos podem estar relacionados a outros atributos. O eixo das horas indica que os equipamentos ligados na rede de 0 a 5 horas, possuem baixo fator de potência, como destacado na Figura 42. Como exibido na legenda abaixo da visualização, neste intervalo de horas, as linhas laranjas e vermelhas são apenas 272, enquanto as cores representando um menor fator de potência, azul e verde, ultrapassam 2300 linhas.

Figura 42 – Dados de 0 a 5 horas em coordenadas paralelas.



Fonte: o autor.

### 5.1.3 Previsão de carga

Para ilustrar como as funcionalidades relacionadas à previsão de carga da aplicação podem utilizadas, são apresentados resultados referentes a dados de dois transformadores de distribuição reais.

#### 5.1.3.1 Caso 1: *outliers* de erro absoluto relativo e análise de erros percentuais

Para o primeiro transformador, os dados coletados foram agregados a cada uma hora, por mediana. Considerando sete dias de intervalo de entrada em cada previsão, puderam ser formadas 2722 sequências de entrada, com 168 amostras cada. A Figura 43 mostra o painel de previsão com as características do conjunto de dados e do modelo de previsão selecionados. Ambos são mostrados em destaque com fundo cinza escuro.

É importante mencionar que algumas características dos conjuntos de dados não estão visíveis na Figura 43, como o atributo selecionado para previsão. Neste caso, o atributo utilizado foi a potência ativa total fornecida pelo transformador. As demais características são mostradas ao se posicionar o *mouse* em cima do conjunto. O mesmo ocorre com o modelo de previsão. A escolha dos parâmetros do conjunto de dados considerou o número de sequências de entrada disponíveis para treinar o modelo.



Figura 43 – Painel com o conjunto de dados e o modelo selecionados para o primeiro transformador.

Entidade selecionada

Tipo: UNTRD COD\_ID: 59\_ET5803145\*90

Seleção de dados

Dados criados

ID: 169434 - De 02/08/2021 a 15/12/2021, agregado a cada 2880 min. por mediana - 64 sequências para intervalos de previsão de 10080 min de entrada e 2880 de saída.  
 ID: 206224 - De 02/08/2021 a 15/12/2021, agregado a cada 60 min. por mediana - 2722 sequências para intervalos de previsão de 10080 min de entrada e 60 de saída.  
 ID: 485948 - De 02/08/2021 a 15/12/2021, agregado a cada 30 min. por mediana - 4341 sequências para intervalos de previsão de 10080 min de entrada e 30 de saída.  
 ID: 801994 - De 02/08/2021 a 15/12/2021, agregado a cada 60 min. por mediana - 2722 sequências para intervalos de previsão de 10080 min de entrada e 60 de saída.  
 ID: 840455 - De 02/08/2021 a 15/12/2021, agregado a cada 60 min. por mediana - 2170 sequências para intervalos de previsão de 10080 min de entrada e 60 de saída.

Novo conjunto de dados

Modelos

Modelos criados

Tipo	Estado	Neurônios	Camada de agrupamento	Épocas	Criado em	ID dos dados
LSTM	Treino finalizado	32	Nenhuma	20	25/4/22, 21:14	206224
LSTM	Treino finalizado	16	Nenhuma	20	25/4/22, 22:23	206224
GRU	Treino finalizado	16	Nenhuma	20	25/4/22, 23:05	206224
LSTM	Treino finalizado	32	Nenhuma	20	25/4/22, 21:14	840455

Visualizar métricas de treino Apagar

Fonte: o autor.

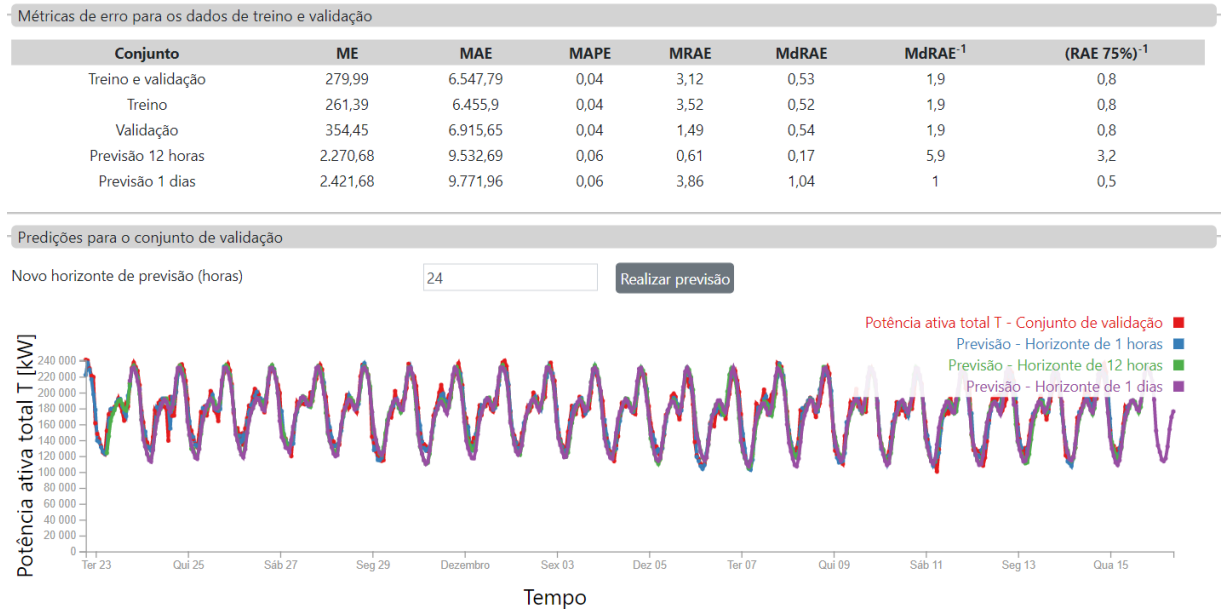
A escolha dos parâmetros do modelo é resultado de uma busca limitada entre os parâmetros disponíveis na aplicação. Assim, como não é objetivo deste trabalho encontrar os melhores modelos de previsão, limita-se a apresentar os parâmetros utilizados: duas camadas recorrentes do tipo LSTM, com 32 neurônios, sem camada de agrupamento e uma camada densa.

Para o treinamento foram utilizados os seguintes parâmetros: taxa de aprendizado de  $2 \times 10^{-3}$ ; tamanho de *batch* de 16; e percentual de validação igual a 20% dos dados. A normalização *min-max* foi aplicada aos dados, resultando nos valores de mínimo e máximo de  $-1$  e  $1$ , respectivamente, para o conjunto de treino. É válido mencionar que os parâmetros desta normalização e do conjunto de treino, são armazenados para permitir reescalar outros dados, como o conjunto de validação.

A Figura 44 mostra as métricas do modelo de previsão para vários conjuntos de dados e, na parte inferior, um gráfico com estes conjuntos em função do tempo. Na tabela, há métricas para os conjuntos de treino e validação. É importante destacar que as métricas do conjunto de validação estão relacionadas a erros considerando apenas uma iteração de previsão. Neste caso, isto equivale a um horizonte de uma hora.

Além dos conjuntos de validação e treino, são mostrados dados para previsões com horizontes de 12 horas e um dia. Estes dados são resultado de previsões recursivas. Isto é, para obter a uma previsão com dados do dia anterior, são realizadas 24 previsões, com 23 delas utilizadas também como entrada para o modelo.

Figura 44 – Métricas de erro e gráfico de previsões em função do tempo na aplicação VPlant, para o primeiro transformador.



Fonte: o autor.

É possível notar que o ME é consideravelmente maior para os horizontes de previsão de 12 horas e 1 dia, ultrapassando 2200 kW. Esta métrica positiva indica que o modelo, em média, subestima o valor real. Para os conjuntos de treino e validação, o MAE foi de cerca de 6400 e 6900 kW respectivamente. Para os horizontes de 12 horas e um dia, esta métrica foi ligeiramente maior, cerca de 9500 e 9700 kW. Comportamento análogo pode ser observado para o MAPE: 4% para os conjuntos de treino e validação; e 6% para os horizontes de 12 horas e 1 dia.

Entre as métricas de erro relativo, destaca-se que, apesar de MRAE ser maior que 1 para quase todos os conjuntos, a métrica MdRAE é menor que 1 para quatro dos cinco conjuntos. Como explicado na Seção 2.4.2.3, MRAE maior que 1 indica que, em média, o modelo erra mais que o modelo de referência.

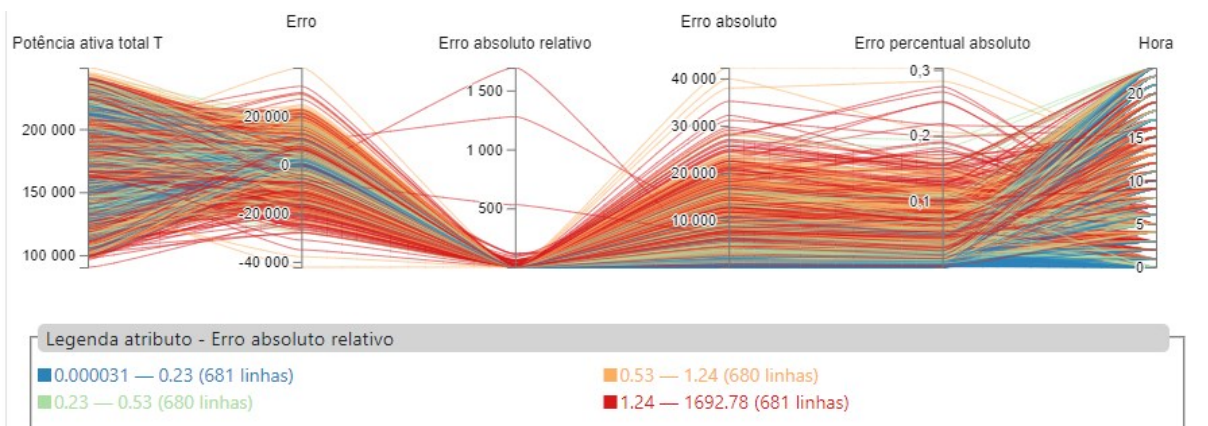
Por outro lado, como explicado na Seção 4.2.5, se MdRAE é menor que 1, 50% das amostras do conjunto têm uma previsão melhor que o modelo de referência. Neste sentido,  $MdRAE^{-1} = 1,9$  para o conjunto de validação, indica um desempenho que é, no mínimo, quase duas vezes melhor, para 50% dos dados, em relação ao modelo de referência.

Para o horizonte de um dia, é interessante destacar que  $MdRAE^{-1} = 1,04$  indica que o modelo tem desempenho similar ao modelo de referência. E como o modelo de referência

utilizado consiste em repetir o último dado conhecido, utilizar o valor do dia anterior, para o mesmo horário, como previsão, resulta em uma métrica semelhante.

A Figura 45 mostra uma visualização em coordenadas paralelas de algumas métricas de erro e outros atributos para os conjuntos de treino e validação. As cores nesta figura representam os quartis do erro relativo absoluto. É possível observar três linhas com erros relativos acima de 500, que representam *outliers* para este atributo. Assim, a distribuição dos erros relativos possui *outliers* consideráveis. Neste cenário, a mediana do erro absoluto relativo, MdRAE, é uma medida mais representativa destes erros, do que a média, MRAE.

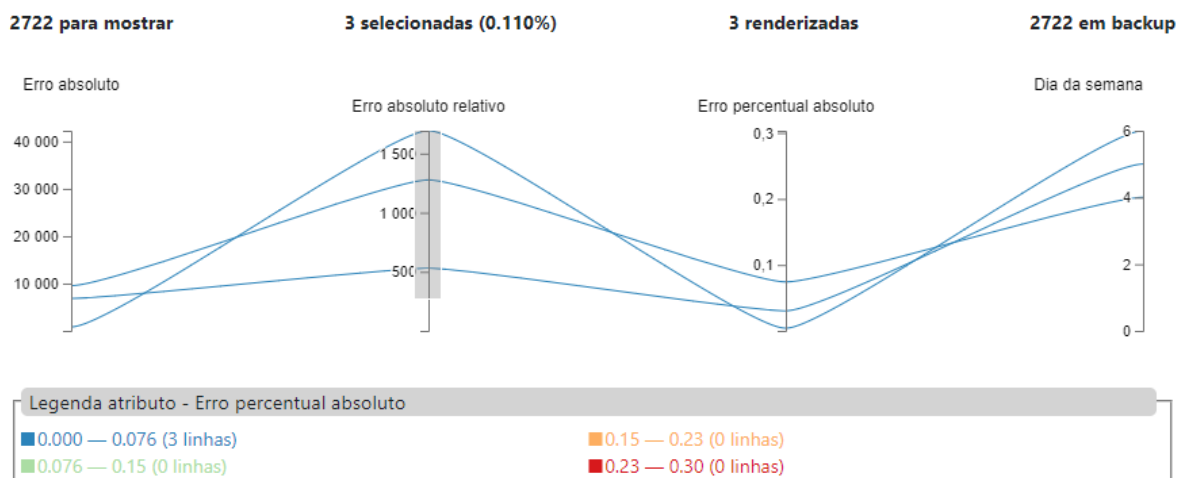
Figura 45 – Coordenadas paralelas mostrando quartis do erro absoluto relativo para o primeiro transformador.



Fonte: o autor.

A Figura 46 mostra estes três *outliers* selecionados. É possível observar pelo gráfico e pela legenda na parte inferior, que estas três amostras possuem erro percentual absoluto menor que 7,6%. Assim, apesar de terem erros relativos elevados, têm erros percentuais baixos.

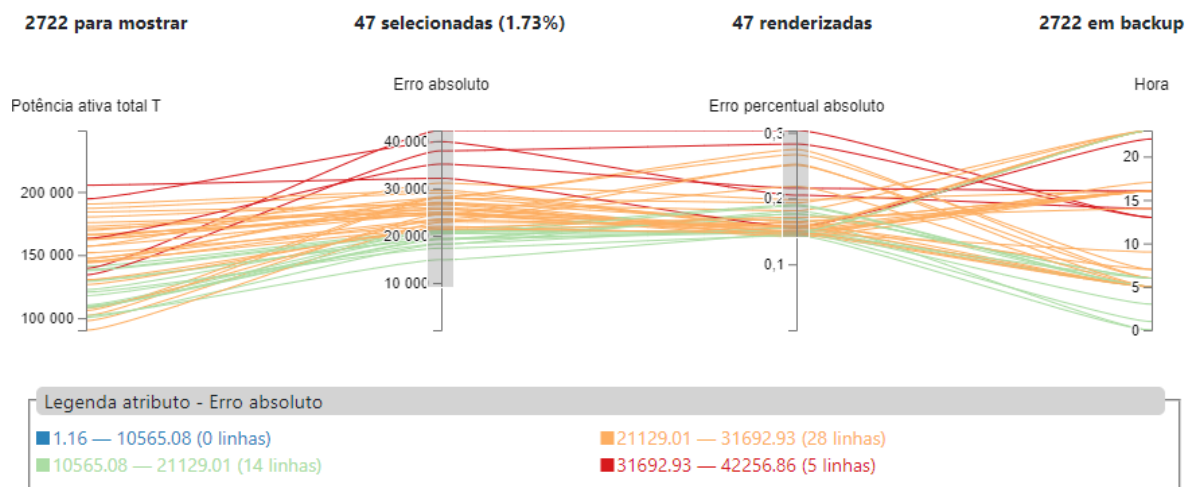
Figura 46 – Coordenadas paralelas mostrando três *outliers* do erro absoluto relativo e classe de cor por erro percentual para o primeiro transformador.



Fonte: o autor.

Selecionando-se as amostras com erros absolutos acima 10 MW, e erros percentuais absolutos acima de 15%, como mostrado na Figura 47, é possível verificar que estas amostras se concentram em determinados intervalos do dia: entre 5 e 10 horas, e próximo a 15 ou 23 horas. A escala de cores está relacionada aos valores de erro absoluto, de tal forma que as amostras em verde possuem os menores valores para este erro. Desta forma, é possível notar que entre zero e cinco horas, há amostras com elevados erros percentuais, isto é, acima de 15%, mas que possuem baixos valores de erro absoluto, em relação às outras amostras visualizadas.

Figura 47 – Coordenadas paralelas mostrando amostras com mais de 15% de erro percentual absoluto e classes de cor por erro absoluto para o primeiro transformador.



Fonte: o autor.

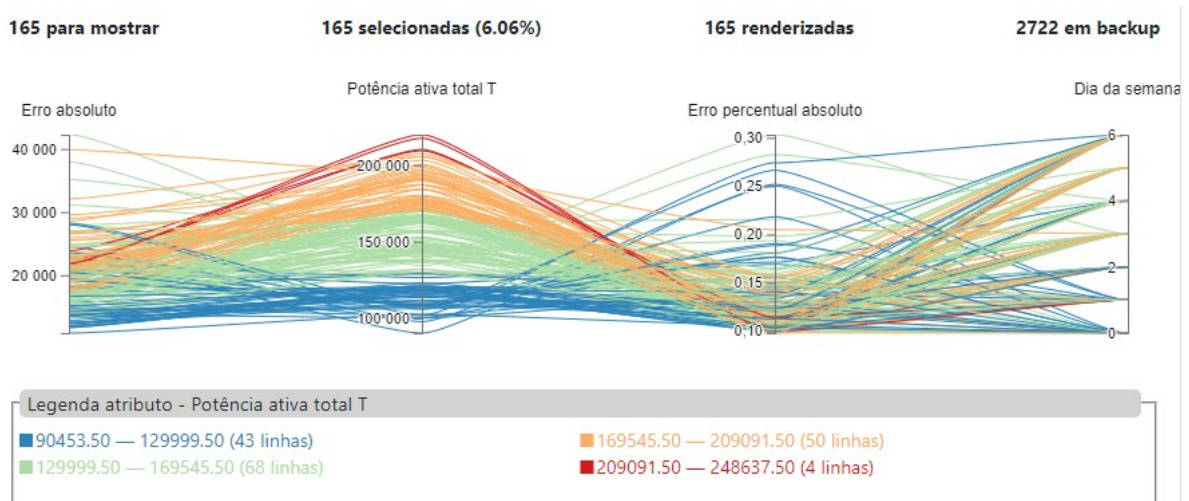
A concentração de linhas com elevados erros absolutos e percentuais absolutos em determinados intervalos do dia, pode indicar a necessidade de utilização de modelos de previsão específicos para tais intervalos. Além disso, conhecer esta tendência de erros ao longo do dia, permite que o modelo seja usado de modo seletivo em horários favoráveis, e indica que é necessário cautela ao utilizar as previsões do modelo nos horários com erros mais elevados.

Na Figura 48, são mostradas as amostras com mais de 10% de erro. Como observado na figura, estas amostras representam apenas cerca de 6% dos dados. As cores na figura indicam a faixa de potência ativa total em que se encontram as amostras, conforme legenda na parte inferior do gráfico. Considerando as linhas se cruzando entre os eixos de potência ativa e erro percentual absoluto, é possível observar que algumas amostras com maior potência, possuem menor erro percentual.

Além disso, ao contrário do conjunto de dados inteiro, que possui mais amostras com potência superior a 169 MW, representadas por linhas laranjas e vermelhas, o subconjunto de amostras com erro percentual maior que 10%, possui mais linhas nas classes de menor potência

ativa, representadas por linhas verdes e azuis. Assim, verifica-se que o modelo tem maiores erros percentuais para amostras de menor potência ativa.

Figura 48 – Coordenadas paralelas mostrando as amostras com mais de 10% de erro e classes de cor por potência ativa total para o primeiro transformador.



Fonte: o autor.

### 5.1.3.2 Caso 2: medidas escassas e intervalo de entrada dos modelos

Para o segundo transformador, que possui um menor número de medidas coletadas, serão apresentados os resultados de dois modelos de previsão: um com sete dias de entrada para cada previsão, e outro com apenas três dias. Devido a isto, o conjunto de sequências de entrada é diferente para os dois modelos.

Para ambos, o intervalo de agregação e de saída foi de uma hora, o tamanho de batch foi 16, a taxa de aprendizado foi  $2 \times 10^{-3}$  e o percentual de dados para validação foi de 20%. Os modelos possuem duas camadas recorrentes LSTM com 16 neurônios cada, e uma camada de saída densa. Novamente, a variável utilizada foi a potência ativa total.

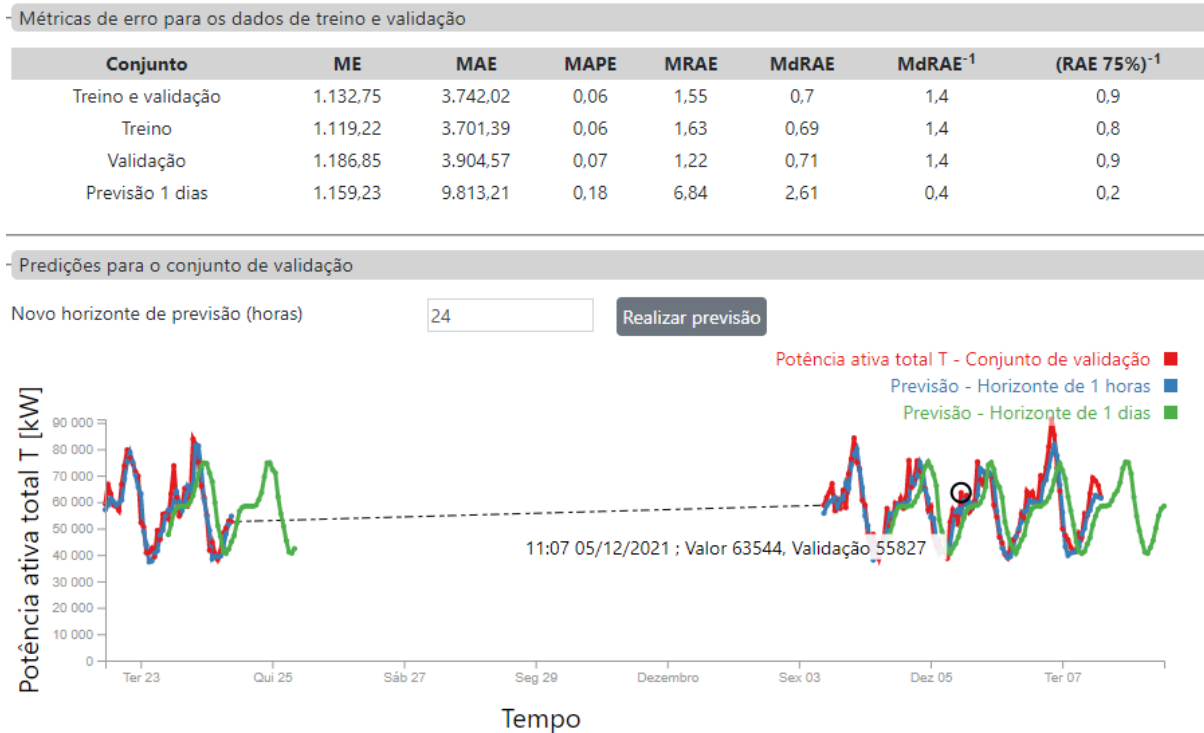
O modelo com entradas de 7 dias possui apenas 745 sequências disponíveis para o treino, enquanto o modelo com entradas de 3 dias possui 1378. Esta diferença ocorre, principalmente, porque para serem válidas como entradas para o primeiro modelo, as sequências de dados coletados devem ser interrompidas por no mínimo 7 dias e uma hora.

É válido citar que esta continuidade se refere às sequências agregadas. Deste modo, uma sequência só é interrompida caso não haja nenhuma medida num intervalo superior ao de agregação, neste caso, de uma hora.

A Figura 49 mostra as métricas de erro e o gráfico de previsões em função do tempo para o primeiro modelo, com entradas de 7 dias. É possível observar que há um atraso na reprodução dos picos de consumo para a previsão com horizonte de 1 dia. Além disso, mostra-

se o valor medido, de 63544 kW, e previsto, de 55827 kW, no dia 05/12/2021 às 11:07. O valor previsto se refere ao conjunto de validação, isto é, com um horizonte de previsão de uma hora.

Figura 49 – Métricas de erro e gráfico de previsões em função do tempo para o modelo com entradas de 7 dias.

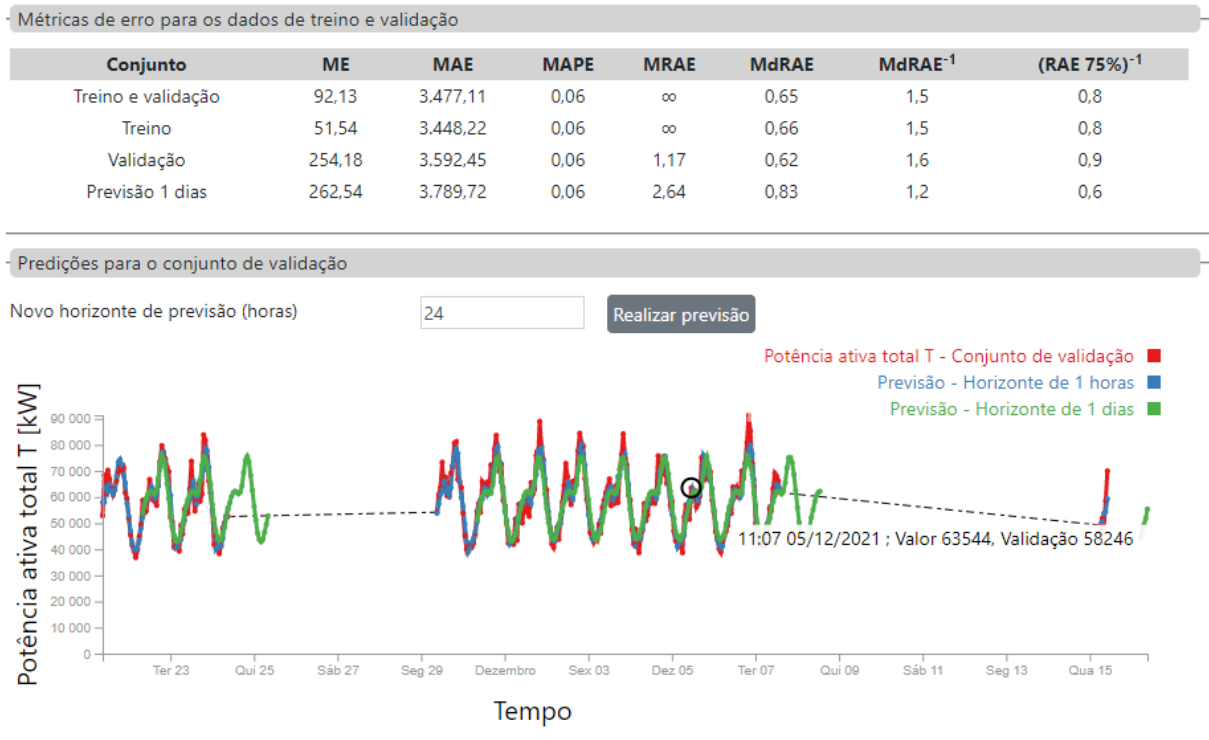


Fonte: o autor.

Ainda na Figura 49, observa-se que a métrica ME se manteve em cerca de 1100 kW para todos os conjuntos de dados. No entanto, MAE e MAPE aumentaram em cerca de três vezes para a previsão com horizonte de um dia, em relação aos conjuntos de treino e validação, atingindo 9813 kW e 18% respectivamente. Ainda para este horizonte, a métrica  $MdRAE^{-1} = 0,4$  indica que, para a mediana dos erros relativos, o modelo teria um erro duas vezes maior que o modelo de referência.

Na Figura 50, apresenta-se as métricas de erro e o gráfico de previsões em função do tempo para o segundo modelo, com entradas de 3 dias. Neste caso, observa-se que não há atrasos perceptíveis na reprodução dos picos de consumo. Além disso, mostra-se o valor medido, de 63544 kW, e previsto, de 58246 kW, no dia 05/12/2021 às 11:07. Novamente, o valor previsto se refere ao horizonte de previsão de uma hora.

Figura 50 – Métricas de erro e gráfico de previsões em função do tempo para o modelo com entradas de 3 dias.



Fonte: o autor.

Quanto às métricas de erro, destaca-se que o ME aumentou cerca de cinco vezes para o conjunto de validação e para a previsão de um dia, em relação ao conjunto de treino. No entanto, esta métrica foi mais baixa para todos os conjuntos de dados que o modelo anterior, com entradas de 7 dias. Para o modelo com entradas de 3 dias, o ME não ultrapassou 263 kW para previsão com horizonte de um dia.

Além disso, este modelo praticamente manteve MAE e MAPE para as previsões com horizonte de um dia, em relação aos conjuntos de treino e validação, atingindo  $MAE = 3790 \text{ kW}$  e  $MAPE = 6\%$ . Ainda para este horizonte, a métrica  $MdRAE = 0,83$  indica que, para a mediana dos erros relativos, o modelo teria um erro 17% menor que o modelo de referência.

Um fato interessante é que a métrica MRAE foi indicada como infinito para o conjunto de treino. Isto aconteceu porque o modelo de referência produziu, para uma amostra, um erro nulo. Esta situação reforça a importância da análise de várias métricas de erro para avaliar os modelos de previsão.

## 5.2 Discussões

No que se refere ao monitoramento de elementos do sistema de distribuição, a aplicação implementada neste trabalho apresenta uma limitação em relação à recepção de dados, conforme apresentado na Seção 5.1.1.1. Tal limitação poderia ser superada com a utilização de múltiplas instâncias da aplicação, que monitorariam partes da rede de distribuição.

Além disso, a separação de algumas funções da aplicação VPlant *back-end*, como o gerenciamento de entidades e os algoritmos de análise, poderia resultar em melhoria no desempenho do sistema. Neste sentido, outras aplicações poderiam ser criadas, para executarem funcionalidades específicas.

Outra opção, seria utilizar uma arquitetura semelhante à proposta de [121], em que se utiliza multiagentes independentes e distribuídos para representar as entidades da rede. No entanto, é importante mencionar que em [121], o autor avalia o desempenho do sistema proposto utilizando uma *microgrid*, com apenas algumas dezenas de entidades. Por outro lado, neste trabalho, os cenários apresentados envolveram mais de um milhão de entidades.

Este também é um diferencial em relação à referência [122], em que o número de entidades nos cenários simulados não ultrapassa 900. É válido destacar que as simulações apresentadas em [121] e [122], possuem objetivos diferentes das simulações apresentadas neste trabalho. Os autores destacam, por exemplo, diversos parâmetros dos sistemas de comunicação.

Em [123] e [126], são apresentadas aplicações semelhantes à proposta deste trabalho. Entretanto, em ambas as aplicações as visualizações se limitam às técnicas clássicas, como gráficos de linha e de barra. Por outro lado, este trabalho explora as técnicas de visualização em árvore e coordenadas paralelas.

A visualização em árvore permitiu explorar a hierarquia inerente aos dados da rede elétrica. Apesar de não refletir a complexidade e o dinamismo das conexões da rede em totalidade, é interessante observar que a visualização implementada pode ser adaptada para representar configurações não radiais. Para isso, podem ser utilizados recursos de visualização de ciclos [85], múltipla herança [86] ou pseudoelementos. Neste trabalho, não se pode implementar a complexidade e o dinamismo das conexões, uma vez que não se teve acesso a estas informações.

Já as coordenadas paralelas, são adequadas para o grande volume de dados de elevada dimensão, presente em sistemas de distribuição e *smart grids*. Em contexto semelhante, a referência [10] emprega várias técnicas de visualização, inclusive as coordenadas paralelas. No entanto, os autores utilizam as técnicas em ferramentas de análise genéricas, isto é, não



apresentam um sistema próprio para sistemas de distribuição e *smart grids*. Por isso, os *layouts* possuem baixo nível de iteração. Além disso, a visualização em coordenadas paralelas apresentada, consiste em uma versão simplificada, que possui apenas dois eixos ortogonais, como um plano cartesiano.

Por outro lado, na aplicação proposta neste trabalho, os *layouts* apresentados possuem diversas funcionalidades de iteração. Os *layouts* de coordenadas paralelas, por exemplo, possuem eixos graduados de forma independente e permitem a seleção de linhas em qualquer eixo.

Em [129] e [130] são empregadas diversas técnicas de visualização, mas em sistemas de transmissão de energia elétrica. Entre as técnicas utilizadas, destacam-se os mapas geográficos de calor com gráficos de pizza utilizados em [129], e *sparklines* utilizadas em [130]. Desta forma, estas referências apresentam possibilidades que podem ser empregadas no contexto deste trabalho. No entanto, é importante destacar que os sistemas de distribuição possuem características peculiares em relação aos de transmissão, como um elevado grau de capilaridade.

No que se refere a algoritmos de previsão, os *outliers* de erro relativo e a concentração de erros percentuais elevados em horários específicos, mostrados na Seção 5.1.3.1, bem como as várias métricas dos modelos comparados na Seção 5.1.3.2, reforçam que o uso de métricas isoladas, como ocorre em muitos trabalhos citados, pode não ser uma forma de avaliação eficaz para os modelos de previsão.

Alguns trabalhos utilizam somente a métrica MAPE [139, 15], enquanto outros utilizam somente o RMSE [14, 16]. Em [18], os autores comparam os modelos com ambas as métricas, MAPE e RMSE. Desta forma, a disponibilização de diversas métricas de erro pela aplicação, consiste em um diferencial deste trabalho.

Ainda neste sentido, a análise em nível de amostras realizada por meio das coordenadas paralelas, é um artefato importante na avaliação de modelos de previsão. Esta abordagem pode ser utilizada em diferentes tipos de modelos e contextos de aplicação. Em [123], por exemplo, os autores utilizam modelos de previsão voltados para a previsão de geração de energia fotovoltaica, e comparam os modelos apenas com base na métrica MAE. Neste contexto, as coordenadas paralelas poderiam ser utilizadas para analisar relações entre diversas variáveis relacionadas à variação da energia gerada.

Além disso, para modelos de caixa preta, esta abordagem representa uma possibilidade de avaliação detalhada dos resultados, o que contribui para mitigar problemas com entradas peculiares.

Como destacado anteriormente, ressalta-se que não é objetivo deste trabalho encontrar um modelo excepcional para a previsão de carga. Além disso, a busca por modelos de previsão melhores pode ser realizada pelo usuário da ferramenta, que pode configurar diversos parâmetros, tanto para o pré-processamento dos dados, quanto para a criação e treinamento dos modelos.

# 6 Conclusão e trabalhos futuros

## 6.1 Conclusão

Com relação aos objetivos propostos, a capacidade do sistema implementado gerenciar e monitorar entidades virtuais foi avaliada. Para isso, considerou-se um cenário com mais de um milhão de elementos da rede de distribuição, entre transformadores e unidades consumidoras. Desta forma, a solução proposta representa uma base para aplicação em grandes sistemas de *smart grids*.

O objetivo específico de exemplificar como técnicas de visualização de informação podem ser aplicadas no contexto de redes de distribuição e SGs foi alcançado. Explorando a hierarquia inerente aos dados da rede de distribuição, verificou-se que a visualização em árvore é capaz de auxiliar o usuário na compreensão de eventos da rede. O mapa, capaz de mostrar mais de um milhão de entidades, também foi útil nesta tarefa. Além disso, aplicou-se a técnica de coordenadas paralelas em diversas análises de dados neste contexto. É válido destacar que todas as análises foram feitas por meio de um navegador de internet comum.

Com relação aos objetivos específicos de explorar técnicas de inteligência artificial e a utilizar algoritmos de previsão de carga, para demonstrar como a aplicação implementada pode ser utilizada para gerar e apresentar informações úteis, utilizou-se redes neurais recorrentes para realizar a previsão de carga. Neste sentido, a aplicação permite que o usuário explore diversos parâmetros dos modelos e visualize diferentes métricas de erros.

Por fim, verificou-se o potencial de aplicação de técnicas de visualização de informação para avaliar os algoritmos de previsão, com a utilização de coordenadas paralelas. Neste sentido, esta técnica permitiu visualizar com facilidade, aspectos importantes em relação ao desempenho dos modelos, que não são compreendidos apenas com as métricas baseadas em médias ou medianas de erros.

Com as coordenadas paralelas, foram visualizados *outliers* do erro relativo, que impactam severamente na média destes erros. Além disso, visualizou-se os horários em que ocorrem os maiores erros absolutos e percentuais absolutos. Estes são apenas exemplos de como a técnica pode ser utilizada para avaliar os modelos. Assim, não esgotam as possibilidades de utilização neste contexto, uma vez que esta permite visualizar simultaneamente várias medidas de erro e atributos, em nível de amostras.

Neste sentido, a utilização “exótica” de coordenadas paralelas para avaliar o desempenho das redes neurais utilizadas para previsão de carga, é uma importante contribuição deste trabalho. Além de representar um caso de uso interessante para esta técnica de visualização, considerando a atual relevância dos modelos de ANNs, esta contribuição é um exemplo de como as técnicas de visualização de informação podem otimizar importantes funcionalidades em aplicações de IoT e/ou de análises de dados. Também é interessante destacar que as coordenadas paralelas podem ser utilizadas para avaliar outros modelos de previsão, e inclusive, para comparar vários modelos.

Desta forma, o objetivo principal deste trabalho foi alcançado, uma vez que os vários conceitos e técnicas aplicados – entidades virtuais, algoritmos de inteligência artificial e técnicas de visualização – permitiram criar um sistema capaz de monitorar entidades do sistema de distribuição em grande escala, visualizando os eventos, bem como as informações coletadas e sintetizadas, de maneira eficaz. Assim, o sistema proposto pode ser considerado base para grandes sistemas de *smart grids*, e representa uma importante contribuição deste trabalho.

## 6.2 Trabalhos futuros

Considerando a limitação de desempenho na recepção de dados relacionada ao VPlant *back-end*, a redistribuição das funcionalidades de gerenciamento de entidades virtuais e execução de algoritmos para outras aplicações, constitui um cenário para trabalhos futuros. Neste mesmo sentido, utilizar múltiplas instâncias da aplicação de maneira sincronizada, é outra abordagem que pode gerar melhorias de desempenho.

A utilização de coordenadas paralelas representa apenas um exemplo de como técnicas de visualização podem ser utilizadas no contexto de sistemas de distribuição e SGs. Neste sentido, a aplicação de técnicas, como mapas de calor [151], para visualizar o consumo típico, representa outra oportunidade para trabalhos futuros. Além disso, pode-se implementar melhorias no *layout* de coordenadas paralelas, como a escolha dos eixos a serem visualizados de forma automática, baseada, por exemplo, em aspectos relacionados à entropia da informação. Esta técnica de visualização pode ainda ser utilizada para finalidades específicas, como encontrar equipamentos mais susceptíveis a defeitos.

A visualização em árvore também pode ser otimizada para representar, com melhor fidelidade, o dinamismo e a complexidade dos sistemas de distribuição. Isto é, podem ser representados trechos não radiais com a utilização de *links* extras ou pseudoelementos. Os nós

da árvore também poderiam representar informações a respeito das entidades a jusante, por meio de pequenos gráficos de pizza ou barra.

De forma análoga, os modelos de previsão utilizados representam apenas um exemplo de como a aplicação pode gerar informações úteis. Desta forma, é interessante implementar outros algoritmos, como por exemplo, de avaliação de perdas não técnicas, análises de fluxo de carga e otimização do reestabelecimento automático energia. Estas funcionalidades poderiam ser integradas à visualização em árvore e ao mapa, explorando a hierarquia, as conexões e os aspectos espaciais.

Além disso, é interessante otimizar os algoritmos de previsão empregados, seja com o uso de arquiteturas de redes neurais mais robustas, ou com etapas de pré-processamento, como indicado na Seção 2.4.3. A utilização de modelos com previsão *multi-step* direta, representa uma abordagem interessante para obter melhores resultados, em previsões com horizontes maiores.

Outro aspecto interessante para trabalhos futuros, consiste em permitir a utilização de múltiplas variáveis de entrada e saída nos modelos. Também é interessante utilizar outros modelos de previsão, visando comparar com os modelos já empregados.

Com relação a avaliação dos modelos, a utilização de métricas interessantes no contexto de previsão de carga, como erro máximo diário, ou erro no ponto de máximo consumo, conforme apresentado em 2.4.2.4, representa outra oportunidade de melhoria. Neste contexto, outras técnicas de visualização, como histogramas, poderiam ser utilizadas em conjunto com as coordenadas paralelas, com o intuito de otimizar ainda mais a avaliação dos modelos.

# Referências

- [1] D. Reinsel, J. Gantz and J. Rydning, "The Digitization of the World from Edge to Core. Data Age 2025," IDC, 2018.
- [2] J. A. S. Maia, "Arquitetura base para soluções de Internet das Coisas: aplicações de telemetria e computação na ponta com uso de Microsoft Azure nos modelos de IaaS, PaaS e SaaS," Dissertação (Mestrado em Sistemas Mecatrônicos) – Universidade de Brasília, Brasília, 2020.
- [3] Banco Nacional de Desenvolvimento Econômico e Social - BNDES, "Internet das Coisas: Um plano de ação para o Brasil. Cartilha de Cidades," Janeiro 2018. [Online]. Available: <https://www.bndes.gov.br/wps/wcm/connect/site/db27849e-dd37-4fbd-9046-6fda14b53ad0/produto-13-cartilha-das-cidades-publicada.pdf?MOD=AJPERES&CVID=m7tz8bf>. [Acesso em 22 Dezembro 2021].
- [4] Banco Nacional de Desenvolvimento Econômico e Social - BNDES, "Internet das Coisas: Um plano de ação para o Brasil. Produto Produto 9A: Relatório Final do Estudo," Janeiro 2018. [Online]. Available: <https://www.bndes.gov.br/wps/wcm/connect/site/d22e7598-55f5-4ed5-b9e5-543d1e5c6dec/produto-9A-relatorio-final-estudo-de-iot.pdf?MOD=AJPERES&CVID=m5WVild>. [Acesso em 22 Dezembro 2021].
- [5] R. M. González, F. D. Wattjes, M. Gibescu, W. Vermeiden, J. G. Slootweg and W. L. Kling, "Applied Internet of Things Architecture to Unlock the Value of Smart Microgrids," *IEEE Internet of Things Journal*, vol. 5(6), pp. 5326-5336, 2018. doi: <https://doi.org/10.1109/JIOT.2018.2869389>.
- [6] M. Jeffin, G. Madhu, A. Rao, G. Singh and C. Vyjayanthi, "Internet of Things Enabled Power Theft Detection and Smart Meter Monitoring System," in *2020 International Conference on Communication and Signal Processing (ICCSP)*, 2020. doi: <https://doi.org/10.1109/ICCSP48568.2020.9182144>.
- [7] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 2nd ed., O'Reilly Media, Inc, 2019.

- [8] M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," *IEEE Communications Surveys Tutorials*, vol. 20(4), pp. 2923-2960, 2018. doi: <https://doi.org/10.1109/COMST.2018.2844341>.
- [9] D. A. Keim, "Information visualization and visual data mining," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8(1), pp. 1-8, 2002. doi: <https://doi.org/10.1109/2945.981847>.
- [10] D. V. Nga, O. H. See, D. N. Quang, C. Y. Xuen and L. L. Chee, "Visualization Techniques in Smart Grid," *Smart Grid and Renewable Energy*, vol. 3(3), pp. 175-185, 2012. doi: <https://doi.org/10.4236/sgre.2012.33025>.
- [11] A. J. P. Soares, "Avaliação da propagação do sinal LoRa e desenvolvimento de um método para auxiliar o planejamento de redes IoT usando otimização do modelo de Hata," Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Uberlândia, Uberlândia, 2019. doi: <https://doi.org/10.14393/ufu.di.2019.1267>.
- [12] W. D. C. Nunes, "Proposta e desenvolvimento de um sistema de gerência de dispositivos de rede restritos baseados em LoRaWAN e LwM2M," Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica – Universidade Federal de Uberlândia, Uberlândia, 2020. doi: <https://doi.org/10.14393/ufu.di.2020.254>.
- [13] L. J. Duarte, "Metodologia de avaliação da condição de operação de transformadores de distribuição por meio de análise da degradação térmica, integridade operacional e confiabilidade em ambiente de IoT," Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Uberlândia, Uberlândia, 2021. doi: <https://doi.org/10.14393/ufu.di.2021.239>.
- [14] D. Smith, K. Jaskie, J. Cadigan, J. Marvin e A. Spanias, "Machine Learning for Fast Short-Term Energy Load Forecasting," *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, vol. 1, pp. 433-436, 2020. doi: <https://doi.org/10.1109/ICPS48405.2020.9274781>.
- [15] U. N. Khan, "Short-Term Load Forecasting by Using Artificial Neural Networks," Thesis (Master in Electrical Engineering) - Graduate School of Natural and Applied Sciences - Bahçeşehir University, Istanbul, 2018. doi: <https://doi.org/http://hdl.handle.net/123456789/1257>.

- [16] Q. Jiang, J.-X. Zhu, M. Li and H.-Y. Qing, "Electricity Power Load Forecast via Long Short-Term Memory Recurrent Neural Networks," in *2018 4th Annual International Conference on Network and Information Systems for Computers (ICNISC)*, 2018. doi: <https://doi.org/10.1109/ICNISC.2018.00060>.
- [17] Hasan-Al-Shaikh, M. A. Rahman and A. Zubair, "Electric Load Forecasting with Hourly Precision Using Long Short-Term Memory Networks," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2019. doi: <https://doi.org/10.1109/ECACE.2019.8679244>.
- [18] X. Guo, Y. Gao, Y. Li, D. Zheng and D. Shan, "Short-term household load forecasting based on Long- and Short-term Time-series network," *Energy Reports*, vol. 7, pp. 58-64, 2021. doi: <https://doi.org/10.1016/j.egy.2021.02.023>.
- [19] International Energy Agency (IEA), "Key world energy statistics 2021," September 2021. [Online]. Available: <https://www.iea.org/reports/key-world-energy-statistics-2021>. [Accessed 10 12 2021].
- [20] International Energy Agency (IEA), "Global Energy & CO2 Status Report 2019 - Report extract - Electricity," 2019. [Online]. Available: <https://www.iea.org/reports/global-energy-co2-status-report-2019/electricity>. [Acesso em 11 12 2021].
- [21] S.-H. Yoo, "Electricity consumption and economic growth: evidence from Korea," *Energy Policy*, vol. 33(12), pp. 1627-1632, 2005. doi: <https://doi.org/10.1016/j.enpol.2004.02.002>.
- [22] K. Ashton, "That "Internet of Things" Thing," *RFID Journal*, vol. 22, pp. 97-114, 2009.
- [23] S. O. Muhanji, A. E. Flint e A. M. Farid, eIoT: The Development of the Energy Internet of Things in Energy Infrastructure, Cham, Switzerland: Springer, 2019. doi: <https://doi.org/10.1007/978-3-030-10427-6>.
- [24] L. L. A. Gomes e R. F. Cameira, "Internet of Things: Conceitos e Aplicações Baseadas em uma Revisão Bibliográfica," em *Anais do XXXVIII Encontro Nacional de Engenharia de Produção - Enegep*, Maceió, 2018. doi: [https://doi.org/10.14488/enegep2018\\_tn\\_sto\\_258\\_486\\_35576](https://doi.org/10.14488/enegep2018_tn_sto_258_486_35576).
- [25] B. P. Santos, L. A. M. Silva, C. S. F. S. Celes, J. B. B. Neto, B. S. Peres, M. A. M. Vieira, L. F. M. Vieira, O. N. Goussevskaia e A. A. F. Loureiro, "Internet das Coisas:



- da Teoria à Prática,” em *XXXIV Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos - SBRC. Livro Texto Minicursos.*, Salvador, 2016.
- [26] Banco Nacional de Desenvolvimento Econômico e Social - BNDES, “Internet das Coisas: Um plano de ação para o Brasil. Produto 9B: Síntese do relatório final do estudo,” Janeiro 2018. [Online]. Available: <https://www.bndes.gov.br/wps/wcm/connect/site/445c4dd8-069b-47c1-b191-767caee4a5ae/produto-9B-relatorio-final-sintese-do-estudo-de-Iot-atualizado.pdf?MOD=AJPERES&CVID=m87qOtl>. [Acesso em 22 Dezembro 2021].
- [27] L. Atzori, A. Iera e G. Morabito, “The Internet of Things: A Survey,” *Computer Networks*, vol. 54(15), pp. 2787-2805, 2010. doi: <https://doi.org/10.1016/j.comnet.2010.05.010>.
- [28] J. Gubbi, R. Buyya, S. Marusic e M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Future Generation Computer Systems*, vol. 29(7), pp. 1645-1660, 2013. doi: <https://doi.org/10.1016/j.future.2013.01.010>.
- [29] P. Sethi and S. R. Sarangi, "Internet of Things: Architectures, Protocols, and Applications," *Journal of Electrical and Computer Engineering*, pp. 2090-0147, 2017. doi: <https://doi.org/doi.org/10.1155/2017/9324035>.
- [30] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys Tutorials*, vol. 17(4), pp. 2347-2376, 2015. doi: <https://doi.org/10.1109/COMST.2015.2444095>.
- [31] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang and W. Zhao, "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications," *IEEE Internet of Things Journal*, vol. 4(5), pp. 1125-1142, 2017. doi: <https://doi.org/10.1109/JIOT.2017.2683200>.
- [32] M. Lombardi, F. Pascale and D. Santaniello, "Internet of Things: A General Overview between Architectures, Protocols and Applications," *Information*, vol. 12(2), 2021. doi: <https://doi.org/10.3390/info12020087>.
- [33] M. R. Abdmeziem, D. Tandjaoui and I. Romdhani, "Architecting the Internet of Things: State of the art," *Studies in Systems, Decision and Control*, pp. 55-75, 2015. doi: [https://doi.org/10.1007/978-3-319-22168-7\\_3](https://doi.org/10.1007/978-3-319-22168-7_3).

- [34] L. D. Xu, W. He and S. Li, "Internet of Things in Industries: A Survey," *IEEE Transactions on Industrial Informatics*, vol. 10(4), pp. 2233-2243, 2014. doi: <https://doi.org/10.1109/TII.2014.2300753>.
- [35] A. R. Al-Ali, "Internet of Things Role in the Renewable Energy Resources," em *Energy Procedia, 3rd International Conference on Power and Energy Systems Engineering*, Kitakyushu, Japan, 2016. doi: <https://doi.org/10.1016/j.egypro.2016.10.144>.
- [36] Banco Nacional de Desenvolvimento Econômico e Social - BNDES, "Internet das Coisas: Um plano de ação para o Brasil," 2018. [Online]. Available: <https://www.bndes.gov.br/wps/portal/site/home/conhecimento/pesquisaedados/estudos/estudo-internet-das-coisas-iot/estudo-internet-das-coisas-um-plano-de-acao-para-o-brasil>. [Acesso em 22 Dezembro 2021].
- [37] Banco Nacional de Desenvolvimento Econômico e Social - BNDES, "Internet das Coisas: Um plano de ação para o Brasil. Produto 6 – Relatório final da priorização de verticais e horizontais," Novembro 2017. [Online]. Available: [https://www.bndes.gov.br/wps/wcm/connect/site/9cc660de-bb54-408a-a7e9-322f5dbc3f03/Produto+6\\_Relat%C3%B3rio\\_Final\\_Prioriza%C3%A7%C3%A3o\\_v1\\_atualizado.pdf?MOD=AJPERES&CVID=m0SUIr](https://www.bndes.gov.br/wps/wcm/connect/site/9cc660de-bb54-408a-a7e9-322f5dbc3f03/Produto+6_Relat%C3%B3rio_Final_Prioriza%C3%A7%C3%A3o_v1_atualizado.pdf?MOD=AJPERES&CVID=m0SUIr). [Acesso em 22 Dezembro 2021].
- [38] Brasil, "Decreto nº 9.854, de 25 de junho de 2019. Institui o Plano Nacional de Internet das Coisas e dispõe sobre a Câmara de Gestão e Acompanhamento do Desenvolvimento de Sistemas de Comunicação Máquina a Máquina e Internet das Coisas," Presidência da República, Secretaria-Geral, Subchefia para Assuntos Jurídicos, 2019. [Online]. Available: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2019-2022/2019/decreto/D9854.htm](http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D9854.htm). [Acesso em 11 Junho 2022].
- [39] Brasil, "Lei nº 14.108, de 16 de dezembro de 2020. Diário Oficial da União: edição 241, seção 1, p. 3," 17 Dezembro 2020. [Online]. Available: <https://www.in.gov.br/en/web/dou/-/lei-n-14.108-de-16-de-dezembro-de-2020-294616158>. [Acesso em 11 Junho 2022].
- [40] A. Ghasempour, "Internet of Things in Smart Grid: Architecture, Applications, Services, Key Technologies, and Challenges," *Inventions*, vol. 4(1), 2019. doi: <https://doi.org/10.3390/inventions4010022>.

- [41] A. Ghasempour, "Optimum Packet Service and Arrival Rates in Advanced Metering Infrastructure Architecture of Smart Grid," *2016 IEEE Green Technologies Conference (GreenTech)*, pp. 1-5, 2016. doi: <https://doi.org/10.1109/GreenTech.2016.8>.
- [42] K. Siozios, "Mastering the Challenges of Changing Energy Systems: The Smart-Grid Concept," em *IoT for Smart Grids*, Springer International Publishing, 2019. doi: [https://doi.org/10.1007/978-3-030-03640-9\\_1](https://doi.org/10.1007/978-3-030-03640-9_1).
- [43] F. Al-Turjman and M. Abujubbeh, "IoT-enabled smart grid via SM: An overview," *Future Generation Computer Systems*, vol. 96, pp. 579-590, 2019. doi: <https://doi.org/10.1016/j.future.2019.02.012>.
- [44] F. Zavoda, "The key role of intelligent electronic devices (IED) in advanced Distribution Automation (ADA)," in *2008 China International Conference on Electricity Distribution*, 2008. doi: <https://doi.org/10.1109/CICED.2008.5211637>.
- [45] Y. Saleem, N. Crespi, M. H. Rehmani and R. Copeland, "Internet of Things-Aided Smart Grid: Technologies, Architectures, Applications, Prototypes, and Future Research Directions," *IEEE Access*, vol. 7, pp. 62962-63003, 2019. doi: <https://doi.org/10.1109/ACCESS.2019.2913984>.
- [46] Grupo de Trabalho de Redes Elétricas Inteligentes, "Relatório. Smart Grid," 2010. [Online]. Available: <https://www.gov.br/mme/pt-br/assuntos/secretarias/energia-eletrica/relatorio-smart-grid-1/documentos/relatorio-smart-grid.pdf/view>. [Acesso em 22 Dezembro 2021].
- [47] F. F. d. S. Luz e P. H. B. d. Souza, "Telemedição de clientes do grupo B no âmbito de concessão da Companhia Paulista de Força e Luz - CPFL," Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica) – Universidade Federal de Itajubá, Itajubá, 2016.
- [48] M. Lamagna, D. Groppi, M. M. Nezhad and G. Piras, "A Comprehensive Review on Digital Twins for Smart Energy Management System," *International Journal of Energy Production and Management*, vol. 6(4), pp. 323-334, 2021. doi: <https://doi.org/10.2495/EQ-V6-N4-323-334>.
- [49] M. Grieves and J. Vickers, "Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems," in *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, Springer International Publishing, 2017, pp. 85-113. doi: [https://doi.org/10.1007/978-3-319-38756-7\\_4](https://doi.org/10.1007/978-3-319-38756-7_4).

- [50] Z. Jiang, H. Lv, Y. Li and Y. Guo, "A novel application architecture of digital twin in smart grid," *Journal of Ambient Intelligence and Humanized Computing*, 2021. doi: <https://doi.org/10.1007/s12652-021-03329-z>.
- [51] M. Grieves, "Digital Twin: Manufacturing Excellence through Virtual Factory Replication," Digital Twin Institute, March 2015. [Online]. Available: [https://www.researchgate.net/publication/275211047\\_Digital\\_Twin\\_Manufacturing\\_Excellence\\_through\\_Virtual\\_Factory\\_Replication](https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication). [Accessed 03 Janeiro 2022].
- [52] T. Cioara, I. Anghel, M. Antal, I. Salomie, C. Antal and A. G. Ioan, "An Overview of Digital Twins Application Domains in Smart Energy Grid," *arXiv:2104.07904*, 2021.
- [53] S. Y. Teng, M. Touš, W. D. Leong, B. S. How, H. L. Lam and V. Máša, "Recent Advances on Industrial Data-Driven Energy Savings: Digital Twins and Infrastructures," *Renewable and Sustainable Energy Reviews*, vol. 135, 2021. doi: <https://doi.org/10.1016/j.rser.2020.110208>.
- [54] D. d. O. Ferreira, A. P. Pinheiro, L. J. Duarte, M. Pereira e C. Barroso, "Visualização de Dados com Coordenadas Paralelas para Sistemas de Distribuição em Aplicação Web de Alto Desempenho," em *Anais da XIV Conferência Brasileira sobre Qualidade da Energia Elétrica.*, 2021. doi: <https://doi.org/10.17648/cbqee-2021-130670>.
- [55] M.-A. Sanchez-Hidalgo and M.-D. Cano, "A survey on visual data representation for smart grids control and monitoring," *Sustainable Energy, Grids and Networks*, vol. 16, pp. 351-369, 2018. doi: <https://doi.org/10.1016/j.segan.2018.09.007>.
- [56] N. Kagan, C. C. B. d. Oliveira e E. J. Robba, *Introdução aos Sistemas de Distribuição de Energia Elétrica*, 2ª ed., Blucher, 2010.
- [57] N. V. d. Melo, "Uma técnica de análise de causa raiz para sistemas elétricos de potência," Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-graduação em Ciência da Computação, Centro de Engenharia Elétrica e Informática – Universidade Federal de Campina Grande, Campina Grande, 2015.
- [58] L. H. S. d. Toledo, *Estudo de caso de um centro de controle de sistema elétrico de potência: uma abordagem tridimensional*, Belo Horizonte: Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica, Escola de Engenharia – Universidade Federal de Minas Gerais, 2017.
- [59] J. T. Pinho e M. A. Galdino, *Manual de Engenharia para Sistemas Fotovoltaicos*, 2ª ed., Rio de Janeiro: CEPTEL - CRESESB, 2014.

- [60] R. D. d. S. e. Silva, “Nota Técnica Nº 69. Contextualização do Setor Elétrico Brasileiro e o Planejamento da Infraestrutura no Longo Prazo,” Instituto de Pesquisa Econômica Aplicada - Ipea, 2020.
- [61] D. M. Leme, M. A. d. Cunha, T. A. Pitocco e W. C. Rizzardi, “Sistema de Proteção da Rede de Distribuição de Energia Elétrica,” Monografia (Graduação em Engenharia Elétrica) – Universidade São Francisco., Itatiba, 2013.
- [62] A. C. Santana, “Aula 13. Transformadores,” [Online]. Available: [http://professor.ufop.br/sites/default/files/adrielle/files/aula\\_13.pdf](http://professor.ufop.br/sites/default/files/adrielle/files/aula_13.pdf). [Acesso em 15 Janeiro 2022].
- [63] R. P. Buratti, “Reguladores de tensão em redes elétricas com alta penetração de cargas não lineares,” Dissertação (Mestrado em Engenharia Elétrica) – Universidade Estadual Paulista (Unesp), Ilha Solteira, 2016.
- [64] J. Johansson and C. Forsell, "Evaluation of Parallel Coordinates: Overview, Categorization and Guidelines for Future Research," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22(1), pp. 579-588, 2016. doi: <https://doi.org/10.1109/TVCG.2015.2466992>.
- [65] M. Behrisch, M. Blumenschein, N. W. Kim, L. Shao, M. El-Assady, J. Fuchs, D. Seebacher, A. Diehl, U. Brandes, H. Pfister, T. Schreck, D. Weiskopf and D. A. Keim, "Quality metrics for information visualization," *Computer Graphics Forum*, vol. 37(3), pp. 625-662, 2018. doi: <https://doi.org/10.1111/cgf.13446>.
- [66] D. F. Andrews, "Plots of High-Dimensional Data," *Biometrics*, vol. 28(1), pp. 125-136, 1972. doi: <https://doi.org/10.2307/2528964>.
- [67] J. M. Chambers, W. S. Cleveland, B. Kleiner and P. A. Tukey, *Graphical Methods for Data Analysis*, Boston: Duxbury Press, 1983.
- [68] N. H. Müller, B. Liebold, D. Pietschmann, a. P. Ohler and P. Rosenthal, "Hierarchy visualization designs and their impact on perception and," in *ACHI 2017: The Tenth International Conference on Advances in Computer-Human Interactions*, Nice, France, 2017.
- [69] L. Woodburn, Y. Yang and K. Marriott, "Interactive Visualisation of Hierarchical Quantitative Data: An Evaluation," in *2019 IEEE Visualization Conference (VIS)*, Vancouver, Canada, 2020. doi: <https://doi.org/10.1109/VISUAL.2019.8933545>.

- [70] S. B. Azhar and M. J. Rissanen, "Evaluation of Parallel Coordinates for Interactive Alarm Filtering," in *2011 15th International Conference on Information Visualisation*, London, 2011. doi: <https://doi.org/10.1109/IV.2011.30>.
- [71] Y.-H. Fua, M. O. Ward and E. A. Rundensteiner, "Hierarchical Parallel Coordinates for Exploration of Large Datasets," in *Proceedings Visualization '99 (Cat. No.99CB37067)*, Worcester, 1999. doi: <https://doi.org/10.1109/VISUAL.1999.809866>.
- [72] M. Nöllenburg, "Geographic Visualization," *Human-Centered Visualization Environments. Lecture Notes in Computer Science*, vol. 4417, pp. 253-290, 2007. doi: [https://doi.org/10.1007/978-3-540-71949-6\\_6](https://doi.org/10.1007/978-3-540-71949-6_6).
- [73] A. Inselberg, "The plane with parallel coordinates," *The Visual Computer*, p. 69–91, 1985. doi: <https://doi.org/10.1007/BF01898350>.
- [74] E. J. Wegman, "Hyperdimensional Data Analysis Using Parallel," *Journal of the American Statistical Association*, vol. 85(411), pp. 664-675, 1990. doi: <https://doi.org/10.1080/01621459.1990.10474926>.
- [75] H. D. Griffin, "Graphic Computation of Tau as a Coefficient of Disarray," *Journal of the American Statistical Association*, vol. 53(282), pp. 441-447, 1958.
- [76] J. A. Hartigan, *Clustering Algorithms*, New York: John Wiley & Sons, 1975.
- [77] P. Diaconis and J. H. Friedman, "M and N Plots," in *Recent Advances in Statistics*, New York, Academic Press, 1983, pp. 425-447. doi: <https://doi.org/10.1016/B978-0-12-589320-6.50024-1>.
- [78] J. Heinrich and D. Weiskopf, "State of the Art of Parallel Coordinates," in *Eurographics 2013 - State of the Art Reports*, 2013. doi: <https://doi.org/10.2312/conf/EG2013/stars/095-116>.
- [79] J. Li and J. J. v. W. Jean-Bernard Martens, "Judging Correlation from Scatterplots and Parallel Coordinate Plots," *Information Visualization*, vol. 9(1), pp. 13-30, 1 Maio 2008. doi: <https://doi.org/10.1057/ivs.2008.13>.
- [80] M. Ward, G. Grinstein and D. Keim, *Interactive Data Visualization: Foundations, Techniques, and Applications*, A K Peters, 2010. doi: <https://doi.org/10.1201/b10683>.
- [81] R. M. Tarawaneh, P. Keller and A. Ebert, "A General Introduction To Graph Visualization Techniques," in *Visualization of Large and Unstructured Data Sets:*

- Applications in Geospatial Planning, Modeling and Engineering. Proceedings of IRTG 1131 Workshop 2011*, 2012. doi: <https://doi.org/10.4230/OASIS.VLUDS.2011.151>.
- [82] A. L. Souza, "Teoria dos grafos e aplicações," Dissertação (Mestrado Profissional em Matemática) – Programa de Pós-graduação em Matemática – Universidade Federal do Amazonas, Manaus, 2013.
- [83] C.-h. Chen, W. Härdle and A. Unwin, *Handbook of Data Visualization*, 2008. doi: <https://doi.org/10.1007/978-3-540-33037-0>.
- [84] D. K. Nogueira, "Introdução à Teoria dos Grafos: Proposta para o Ensino Médio," Dissertação (Mestrado Profissional em Matemática) – Programa de Mestrado Profissional em Matemática em Rede Nacional – Universidade de Brasília, Brasília, 2015. doi: <https://doi.org/10.26512/2015.07.D.19363>.
- [85] B. Lee, C. S. Parr, C. Plaisant, B. B. Bederson, V. D. Veksler, W. D. Gray and C. Kotfila, "TreePlus: Interactive Exploration of Networks with Enhanced Tree Layouts," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12(6), pp. 1414-1426, 2006. doi: <https://doi.org/10.1109/TVCG.2006.106>.
- [86] R. Schmuecker, "Multiple Parent Nodes D3.js," 27 April 2021. [Online]. Available: <http://bl.ocks.org/robschmuecker/6afc2ecb05b191359862>. [Acesso em 13 January 2022].
- [87] H.-J. Schulz, "Treevis.net: A Tree Visualization Reference," *IEEE Computer Graphics and Applications*, vol. 31(6), pp. 11-15, 2011. doi: <https://doi.org/10.1109/MCG.2011.103>.
- [88] S. Zhao, M. J. McGuffin and M. H. Chignell, "Elastic Hierarchies: Combining Treemaps and Node-Link Diagrams," in *IEEE Symposium on Information Visualization. INFOVIS 2005*, 2005. doi: <https://doi.org/10.1109/INFVIS.2005.1532129>.
- [89] B. Johnson and B. Shneiderman, "Tree-maps: a space-filling approach to the visualization of hierarchical information structures," in *Proceeding Visualization '91*, 1991. doi: <https://doi.org/10.1109/VISUAL.1991.175815>.
- [90] J. Stasko and E. Zhang, "Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations," in *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, 2000. doi: <https://doi.org/10.1109/INFVIS.2000.885091>.

- [91] J. B. Kruskal and J. M. Landwehr, "Icicle Plots: Better Displays for Hierarchical Clustering," *The American Statistician*, vol. 37(2), pp. 162-168, 1983. doi: <https://doi.org/10.1080/00031305.1983.10482733>.
- [92] G. G. Robertson, J. D. Mackinlay and S. K. Card, "Cone Trees: Animated 3D Visualizations of Hierarchical Information," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New Orleans, Louisiana, USA, 1991. doi: <https://doi.org/10.1145/108844.108883>.
- [93] T. Hong, "Short Term Electric Load Forecasting," NC State University, Raleigh, North Carolina, 2010.
- [94] L. Burg, G. Gürses-Tran, R. Madlener e A. Monti, "Comparative Analysis of Load Forecasting Models for Varying Time Horizons and Load Aggregation Levels," *Energies*, vol. 14 (21): 7128, 2021. doi: <https://doi.org/10.3390/en14217128>.
- [95] A. Botchkarev, "Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 14, pp. 45-76, 2019. doi: <https://doi.org/10.28945/4184>.
- [96] L. M. S. Saúde, "Análise comparativa entre os métodos auto-regressivo, integrado de médias móveis e rede neural artificial para previsão de séries temporais," Universidade Federal de Itajubá, Itajubá, 2018.
- [97] D. C. Montgomery, C. L. Jennings and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, John Wiley & Sons, 2008.
- [98] J. Fischer, "What the Mape Is FALSELY Blamed For, Its TRUE Weaknesses and BETTER Alternatives!," STATWORX, 16 August 2019. [Online]. Available: <https://www.statworx.com/at/blog/what-the-mape-is-falsely-blamed-for-its-true-weaknesses-and-better-alternatives/>. [Accessed 22 12 2021].
- [99] C. J. Willmott and K. Matsuura, "Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance," *Climate Research*, vol. 30, pp. 79-82, 2005.
- [100] A. Jierula, S. Wang, T.-M. OH and P. Wang, "Study on Accuracy Metrics for Evaluating the Predictions of Damage Locations in Deep Piles Using Artificial Neural Networks with Acoustic Emission Data," *Applied Sciences*, vol. 11(5): 2314, 5 March 2021. doi: <https://doi.org/10.3390/app11052314>.



- [101] J. G. D. Gooijer and R. J. Hyndman, "25 years of time series forecasting," *International Journal of Forecasting*, vol. 22(3), pp. 443-473, 2006.
- [102] R. J. Hyndman, "Another Look at Forecast-Accuracy Metrics for Intermittent Demand," *Foresight: The International Journal of Applied Forecasting*, vol. 4, pp. 43-46, 2006.
- [103] S. Makridakis, "Accuracy measures: theoretical and practical concerns," *International Journal of Forecasting*, vol. 9, pp. 527-529, 1993.
- [104] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22(4), pp. 679-688, 2006.
- [105] M. Kadowaki, T. Ohishi, S. S. Filho e W. d. S. Lima, "Modelo de Previsão da Demanda de Carga de Curtíssimo Prazo para o Período da Ponta," em *XXXVI SBPO - O Impacto da Pesquisa Operacional nas Novas Tendências Multidisciplinares*, São João del-Rei, 2004.
- [106] K. Aceli, A. C. Lorena, J. Gama e A. C. P. d. L. F. d. Carvalho, *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina*, Rio de Janeiro: LTC, 2011.
- [107] V. A. Padilha e A. C. P. d. L. F. d. Carvalho, *Mineração de Dados em Python*, São Paulo: Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo - ICMC-USP, 2017.
- [108] M. I. V. Furtado, *Redes Neurais Artificiais: uma abordagem para sala de aula*, Ponta Grossa: Atena Editora, 2019. doi: <https://doi.org/10.22533/at.ed.262191504>.
- [109] L. V. Fausett, *Fundamentals of neural networks: architectures, algorithms, and applications*, New Delhi: Pearson, 1993.
- [110] S. Sun, Z. Cao, H. Zhu and J. Zhao, "A Survey of Optimization Methods From a Machine Learning Perspective," *IEEE Transactions on Cybernetics*, vol. 50(8), pp. 3668-3681, 2020. doi: <https://doi.org/10.1109/TCYB.2019.2950779>.
- [111] K. Li and J. Malik, "Learning to Optimize Neural Nets," *arXiv*, November 2017. doi: <https://doi.org/10.48550/arXiv.1703.00441>.
- [112] I. N. d. Silva, D. H. Spatti e R. A. Flauzino, *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas. Curso Prático, 2ª ed.*, São Paulo: Artliber Editora Ltda, 2019.
- [113] F. Sultana, A. Sufian and P. Dutta, "Advancements in Image Classification using Convolutional Neural Network," in *2018 Fourth International Conference on Research*

- in Computational Intelligence and Communication Networks (ICRCICN)*, 2018. doi: <https://doi.org/10.1109/ICRCICN.2018.8718718>.
- [114] K. Fukushima, "Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position," *Biological Cybernetics*, vol. 36, pp. 193-202, 1980. doi: <https://doi.org/10.1007/BF00344251>.
- [115] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano and K. J. Lang, "Phoneme Recognition Using Time-Delay Neural Networks," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37(3), pp. 328-339, 1989. doi: <https://doi.org/10.1109/29.21701>.
- [116] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, vol. 1(4), pp. 541-551, 1989. doi: <https://doi.org/10.1162/neco.1989.1.4.541>.
- [117] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9(8), pp. 1735-1780, 1997. doi: <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [118] K. Cho, B. v. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk e Y. Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *arXiv:1406.1078v3*, 2014. doi: <https://doi.org/10.3115/v1/D14-1179>.
- [119] V. Liberatore and A. Al-Hammouri, "Smart Grid Communication and Co-Simulation," in *IEEE 2011 EnergyTech*, 2011. doi: <https://doi.org/10.1109/EnergyTech.2011.5948542>.
- [120] C. Shum, W.-H. Lau, T. Mao, H. S.-H. Chung, K.-F. Tsang, N. C.-F. Tse and L. L. Lai, "Co-Simulation of Distributed Smart Grid Software Using Direct-Execution Simulation," *IEEE Access*, vol. 6, pp. 20531-20544, 2018. doi: <https://doi.org/10.1109/ACCESS.2018.2824341>.
- [121] R. S. Netto, "Framework em Tempo Real para Sistema de Gerenciamento de Energia de uma Smart Microgrid Utilizando Sistema Multiagente," Tese (Doutorado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica – Universidade Federal de Itajubá, Itajubá, 2018.
- [122] J. B. Leite, "Desenvolvimento de um sistema de gerenciamento de energia (EMS - Energy Management System) para a rede elétrica inteligente (Smart Grid)," Tese

- (Doutorado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica, Faculdade de Engenharia de Ilha Solteira – Universidade Estadual Paulista “Júlio De Mesquita Filho”, Ilha Solteira, 2015.
- [123] E. O’Dwyer, I. Pan, R. Charlesworth, S. Butler and N. Shah, "Integration of an energy management tool and digital twin for coordination and control of multi-vector smart energy systems," *Sustainable Cities and Society*, vol. 62, November 2020. doi: <https://doi.org/10.1016/j.scs.2020.102412>.
- [124] P. Ghosh, S. Eisele, A. Dubey, M. Metelko, I. Madari, P. Volgyesi and G. Karsai, "Designing a decentralized fault-tolerant software framework for smart grids and its applications," *Journal of Systems Architecture*, October 2020. doi: <https://doi.org/10.1016/j.sysarc.2020.101759>.
- [125] A. Albarakati, B. Moussa, M. Debbabi, A. Youssef, B. L. Agba and M. Kassouf, "OpenStack-Based Evaluation Framework for Smart Grid Cyber Security," in *2018 CIEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2018. doi: <https://doi.org/10.1109/SmartGridComm.2018.8587420>.
- [126] A. A. Munshi and Y. A.-R. I. Mohamed, "Data Lake Lambda Architecture for Smart Grids Big Data Analytics," *IEEE Access*, vol. 6, pp. 40463-40471, 2018. doi: <https://doi.org/10.1109/ACCESS.2018.2858256>.
- [127] N. Marz and J. Warren, *Big Data: Principles and best practices of scalable realtime data systems*, Shelter Island, New York, USA: Manning, 2015.
- [128] Tableau Software LLC., "Tableau," 2021. [Online]. Available: <https://www.tableau.com/>. [Accessed 11 12 2021].
- [129] T. J. Overbye and J. D. Weber, "New methods for the visualization of electric power system information," in *IEEE Symposium on Information Visualization 2000. INFOVIS 2000*, 2000. doi: <https://doi.org/10.1109/INFVIS.2000.885101>.
- [130] T. J. Overbye and J. D. Weber, "Smart Grid Wide-Area Transmission System," *Engineering*, vol. 1(4), pp. 466-474, December 2015. doi: <https://doi.org/10.15302/J-ENG-2015098>.
- [131] D. B. d. Alencar, “Modelo híbrido baseado em séries temporais e redes neurais para previsão da geração de energia eólica,” Tese (Doutorado em Engenharia Elétrica) –

- Programa de Pós-graduação em Engenharia Elétrica, Instituto de Tecnologia, Campus Universitário de Guamá – Universidade Federal do Pará, Belém, 2018.
- [132] Y. Liu, L. Li and S. Zhou, "Ensemble Forecasting Frame Based on Deep Learning and Multi-Objective Optimization for Planning Solar Energy Management: A Case Study," *Frontiers in Energy Research*, vol. 9, 2021. doi: <https://doi.org/10.3389/fenrg.2021.764635>.
- [133] S. Mujeeb, N. Javaid, M. Akbar, R. Khalid, O. Nazeer and M. Khan, "Big Data Analytics for Price and Load Forecasting in Smart Grids," in *Advances on Broadband and Wireless Computing, Communication and Applications. BWCCA 2018. Lecture Notes on Data Engineering and Communications Technologies*, 2018. doi: [https://doi.org/10.1007/978-3-030-02613-4\\_7](https://doi.org/10.1007/978-3-030-02613-4_7).
- [134] D. Toffanin, "Generation of customer load profiles based on smart-metering time series, building-level data and aggregated measurements," Thesis (Master in Electrical Engineering) - Department of Information Engineering - University of Padua, Padua, 2016.
- [135] M. d. S. Marques, "Metodologia para modelagem de curvas típicas de demanda elétrica utilizando redes neurais artificiais considerando variáveis climáticas," Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica – Universidade Federal do Pampa, Alegrete, 2014.
- [136] V. A. d. Almeida, "Uma metodologia para tratamento de dados de curvas de carga baseada em técnicas de inteligência artificial," Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós-graduação em Engenharia Elétrica – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2017.
- [137] A. M. Pirbazari, M. Farmanbar, A. Chakravorty and C. Rong, "Improving Load Forecast Accuracy of Households Using Load Disaggregation Techniques," in *2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, 2020. doi: <https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics50389.2020.00140>.
- [138] S. Haben, S. Arora, G. Giasemidis, M. Voss and D. V. Greetham, "Review of low voltage load forecasting: Methods, applications, and recommendations," *Applied*

- Energy*, vol. 304, December 2021. doi: <https://doi.org/10.1016/j.apenergy.2021.117798>.
- [139] A. Piras, A. Germond, B. Buchenel, K. Imhof and Y. Jaccard, "Heterogeneous artificial neural network for short term electrical load forecasting," *IEEE Transactions on Power Systems*, vol. 11(1), pp. 397-402, 1996. doi: <https://doi.org/10.1109/59.486124>.
- [140] Google LLC, "TensorFlow.js: TensorFlow for JavaScript," 2021. [Online]. Available: <https://www.tensorflow.org/js>. [Accessed 11 12 2021].
- [141] M. Bostock, "D3.js: Data-Driven Documents," 2021. [Online]. Available: <https://d3js.org/>. [Acesso em 11 12 2021].
- [142] A. Feuer, "D3.js Tree Editor," 2015. [Online]. Available: <https://bl.ocks.org/adamfeuer/042bfa0dde0059e2b288>. [Accessed 11 12 2021].
- [143] K. Chang, "Nutrient Parallel Coordinates," 2012. [Online]. Available: <http://bl.ocks.org/syntagmatic/3150059>. [Accessed 11 12 2021].
- [144] R. Plummer, "Leaflet.glify: fully functional, ridiculously fast web gl renderer plugin for leaflet," January 2019. [Online]. Available: <https://github.com/robertleplummerjr/Leaflet.glify>. [Accessed 11 12 2021].
- [145] MongoDB, Inc, "MongoDB," 2021. [Online]. Available: <https://www.mongodb.com/>. [Accessed 11 12 2021].
- [146] Agência Nacional de Energia Elétrica (ANEEL), "Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional (PRODIST) Módulo 10 – Sistema de Informação Geográfica Regulatório," 7 Dezembro 2021. [Online]. Available: [https://www2.aneel.gov.br/cedoc/aren2021956\\_prod\\_dist\\_modulo\\_10\\_v4.pdf](https://www2.aneel.gov.br/cedoc/aren2021956_prod_dist_modulo_10_v4.pdf). [Acesso em 11 12 2021].
- [147] Apache Software Foundation, "Apache Kafka," 2017. [Online]. Available: <https://kafka.apache.org/>. [Accessed 11 12 2021].
- [148] Docker Inc, "Docker Engine overview," 2021. [Online]. Available: <https://docs.docker.com/engine/>. [Acesso em 11 12 2021].
- [149] J. Dooley, *Software Development and Professional Practice*, New York: Apress, 2011.
- [150] C.-L. Hor and P. A. Crossley, "Knowledge Extraction from Intelligent Electronic Devices," *Transactions on Rough Sets III. Lecture Notes in Computer Science*, vol. 3400, p. 82–111, 2005. doi: [https://doi.org/10.1007/11427834\\_4](https://doi.org/10.1007/11427834_4).

- [151] M. Bostock, "Electricity Usage, 2019," 28 February 2020. [Online]. Available: <https://observablehq.com/@mbostock/electric-usage-2019>. [Accessed 11 12 2021].

# Apêndice A: Lista de publicações

Abaixo lista-se as produções científicas que foram frutos deste trabalho até a sua apresentação para a banca.

- Publicação do artigo “Aplicação de Técnicas de Visualização de Informação para Seleção Unidades Consumidoras para Telemedicação” no “VIII Simpósio Brasileiro de Sistemas Elétricos (SBSE 2020)”. Disponível em:  
<<https://doi.org/10.48011/sbse.v1i1.2267>>  
<[https://www.sba.org.br/open\\_journal\\_systems/index.php/sbse/article/view/2267](https://www.sba.org.br/open_journal_systems/index.php/sbse/article/view/2267)>;
- Publicação do artigo “Visualização de Dados com Coordenadas Paralelas para Sistemas de Distribuição em Aplicação Web de Alto Desempenho” na “XIV Conferência Brasileira sobre Qualidade da Energia Elétrica”. Disponível em:  
<<https://doi.org/10.17648/cbqee-2021-130670>>  
<<https://proceedings.science/cbqee-2021/papers/visualizacao-de-dados-com-coordenadas-paralelas-para-sistemas-de-distribuicao-em-aplicacao-web-de-alto-desempenho>>.

