

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Walquíria do Nascimento Silva

Geração de perfis de carga sintéticas em redes elétricas inteligentes: uma abordagem baseada em dados com redes generativas

Juiz de Fora

2023

Walquíria do Nascimento Silva

Geração de perfis de carga sintéticas em redes elétricas inteligentes: uma abordagem baseada em dados com redes generativas

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Engenharia Elétrica. Área de concentração: Sistemas de Energia Elétrica.

Orientador: Prof. Dr. Bruno Henriques Dias

Coorientador: Prof. Dr. Leonardo Willer de Oliveira

Coorientador: Prof. Dr. Madson Cortes de Almeida

Juiz de Fora

2023

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF com os dados fornecidos pelo(a) autor(a)

Silva, Walquiria N..

Geração de perfis de carga sintéticas em redes elétricas inteligentes : uma abordagem baseada em dados com redes generativas / Walquíria do Nascimento Silva. – 2023.
126 f. : il.

Orientador: Prof. Dr. Bruno Henriques Dias

Coorientador: Prof. Dr. Leonardo Willer de Oliveira

Coorientador: Prof. Dr. Madson Cortes de Almeida

Tese de Doutorado – Universidade Federal de Juiz de Fora, Faculdade de Engenharia.
Programa de Pós-Graduação em Engenharia Elétrica, 2023.

1. Curvas de Carga Sintética. 2. Rede Elétrica Inteligente. 3. Redes Generativas I.
Dias, Bruno H., orient. II. de Oliveira, Leonardo W. e Almeida, Madson C., coorient.
III. Título.

Walquíria do Nascimento Silva

Geração de perfis de carga sintéticos em redes elétricas inteligentes: uma abordagem baseada em dados com redes generativas

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Engenharia Elétrica. Área de concentração: Sistemas de Energia Elétrica

Aprovada em 12 de dezembro de 2023.

BANCA EXAMINADORA

Prof. Dr. Bruno Henriques Dias - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. Madson Cortes de Almeida - Coorientador

Universidade Estadual de Campinas

Prof. Dr. Leonardo Willer de Oliveira - Coorientador

Universidade Federal de Juiz de Fora

Prof. Dr. Sérgio Filipe Carvalho Ramos

Instituto Superior de Engenharia do Porto, Portugal

Prof. Dr. Vitor Hugo Ferreira

Universidade Federal Fluminense

Prof. Dr. Igor Delgado de Melo

Universidade Federal de Juiz de Fora

Prof. Dr. Moisés Vidal Ribeiro

Universidade Federal de Juiz de Fora

Juiz de Fora, 22/11/2023.



Documento assinado eletronicamente por **Bruno Henriques Dias, Professor(a)**, em 12/12/2023, às 17:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Willer de Oliveira, Professor(a)**, em 12/12/2023, às 17:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Igor Delgado de Melo, Professor(a)**, em 12/12/2023, às 17:23, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Sérgio Filipe Carvalho Ramos, Usuário Externo**, em 12/12/2023, às 17:25, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Vitor Hugo Ferreira, Usuário Externo**, em 12/12/2023, às 17:25, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Moises Vidal Ribeiro, Professor(a)**, em 12/12/2023, às 17:26, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1590138** e o código CRC **CA826A0D**.

*Dedico este trabalho à força suprema,
que me rege, me guia e me protege.*

AGRADECIMENTOS

Primeiramente eu agradeço a Deus, que me proveu de todas as ferramentas necessárias para alcançar meus objetivos até aqui.

Agradeço aos professores, Bruno Dias, Leonardo Willer e Madson Cortes de Almeida, pela orientação, apoio e conhecimentos compartilhados ao longo deste trabalho, que foram fundamentais para o meu crescimento acadêmico e profissional.

Agradeço aos professores colaboradores Sérgio Ramos, Igor Delgado, Vitor Hugo e Moisés Vidal pelas contribuições e sugestões, que enriqueceram esta pesquisa.

Também sou grata aos meus colegas que me auxiliaram durante o doutoramento, Luís Henrique, Letícia Henrique, Caio Amorim, Paulo Vitor Barbosa, Glaucus Lima e Jéssica Dohler, que compartilharam o trabalho colaborativo ao longo desses anos e que foram de grande importância para o desenvolvimento deste trabalho.

Um agradecimento especial ao Luís Henrique pelo auxílio e paciência em compartilhar conhecimento e orientação, principalmente na execução desse projeto. Além disso, sou grata ao grupo de pesquisa e aos colegas do *Smart4grids* pelo suporte e troca de experiências.

Agradeço a Universidade Federal de Juiz de Fora (UFJF), a Universidade Estadual de Campinas (UNICAMP), ao Programa de Pós-Graduação em Engenharia Elétrica (PPEE-UFJF) e a Pró-Reitoria de Infraestrutura e Gestão (PROINFA-UFJF) pelo suporte institucional, bem como pelos recursos e infraestrutura essenciais para a realização desta pesquisa.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro fornecido por meio do fomento do projeto 142583/2020-0, que viabilizou a execução desta pesquisa.

Agradeço aos meus pais, Antônio e Jurema, às minhas irmãs, sobrinhos e amigos, que estiveram ao meu lado, me apoiando e incentivando em todos os momentos. Por fim, agradeço a todos os colaboradores, professores, funcionários e pessoas que, direta ou indiretamente, contribuíram para a elaboração deste trabalho. Muito obrigada a todos!

"A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original".

Albert Einstein

RESUMO

O desenvolvimento de métodos sustentáveis e economicamente viáveis tem se tornado fundamental para aprimoramento das políticas de eficiência energética, redução de custos e mitigação dos impactos ambientais. Isso tem ocorrido devido a fatores como a transição energética, a digitalização dos sistemas elétricos e a difusão de recursos energéticos distribuídos (REDs), que vem transformando o paradigma do setor elétrico. Neste contexto, a análise de dados de redes elétricas surge como uma estratégia para identificar as oportunidades de melhorias e otimização dos recursos alocados nestas redes. Dessa forma, a análise de dados tem se tornado fundamental na operação e no planejamento das redes elétricas, bem como para todo o sistema elétrico de potência. Com base nessa perspectiva, ao longo da pesquisa conduzida nesta tese de doutoramento, foram identificados desafios que impactam a análise mais detalhada da dinâmica das redes elétricas, como a escassez de dados disponíveis, questões de segurança e privacidade dos dados, limitando a análise diagnóstica das redes e do desenvolvimento de aplicações voltadas para a eficiência energética e sustentabilidade. Para mitigar tais fatores, dados sintéticos emerge como uma abordagem para suprir estas lacunas. À vista disso, esta tese se propôs a investigar a aplicação de análise de dados em um sistema de distribuição de energia elétrica de um campus universitário, visando a geração de perfis sintéticos de curvas de carga. No âmbito desta perspectiva, adotou-se o modelo *Non-linear Independent Components Estimation* (NICE) com a inclusão de redes neurais com camadas convolucionais. Como resultado da metodologia adotada, obteve-se a representação dos perfis sintéticos. Essa abordagem permite uma representação dos padrões de consumo elétrico de um campus universitário, possibilitando avaliar o potencial dos modelos generativos na caracterização e quantificação dos perfis de carga. Portanto, a contribuição desta tese encontra-se na aplicação de técnicas de análise de dados e implementação de modelos generativos, como o NICE com camadas convolucionais. Estas abordagens têm por objetivo gerar perfis sintéticos de curvas de carga a partir de dados medidos em uma rede elétrica universitária, atuando como ambiente para testes e validação da metodologia proposta.

Palavras-chave: análise de dados; curva de carga sintética; NICE.

ABSTRACT

Developing sustainable and economically viable methods has become fundamental for improving energy efficiency policies, reducing costs, and mitigating environmental impacts. This has occurred due to factors such as the energy transition, the digitalization of electrical systems, and the diffusion of distributed energy resources (DERs), which have been transforming the paradigm of the electrical sector. In this context, data analysis from electrical networks emerges as a strategy to identify improvement opportunities and optimize resources allocated to these networks. In this way, data analysis has become fundamental in the operation and planning of electrical networks and the entire electrical power system. Based on this perspective, throughout the research conducted in this doctoral thesis, challenges were identified that impact the more detailed analysis of electrical network dynamics, such as the scarcity of available data, security issues, and data privacy, limiting the diagnostic analysis of networks and the development of applications aimed at energy efficiency and sustainability. To mitigate such factors, synthetic data emerges to fill these gaps. Given this, this thesis proposed to investigate the application of data analysis in an electrical energy distribution system on a university campus, aiming to generate synthetic profiles of load curves. Within the scope of this perspective, the Non-linear Independent Components Estimation (NICE) model was adopted with the inclusion of neural networks with convolutional layers. As a result of the methodology adopted, the representation of synthetic profiles was obtained. This approach allows a representation of the electrical consumption patterns of a university campus, making it possible to evaluate the potential of generative models in characterizing and quantifying load profiles. Therefore, the contribution of this thesis lies in the application of data analysis techniques and the implementation of generative models, such as NICE with convolutional layers. These approaches aim to generate synthetic profiles of load curves from data measured in a university electrical grid, acting as an environment for testing and validating the proposed methodology.

Keywords: data analysis, NICE, synthetic load curve.

LISTA DE FIGURAS

Figura 1 – Visão geral: <i>Smart campus</i>	18
Figura 2 – Processo KDD.	33
Figura 3 – Estrutura típica do modelo NICE.	37
Figura 4 – Exemplo do processo do fluxo de normalização.	38
Figura 5 – Visão geral da metodologia proposta.	46
Figura 6 – Elementos envolvidos nos cálculos de $S(i)$ objeto i pertence a A. 53	
Figura 7 – Estrutura do modelo NICE.	56
Figura 8 – Diagrama simplificado do Laboratório Vivo UNICAMP.	64
Figura 9 – Diagrama simplificado da rede de distribuição da UNICAMP.	65
Figura 10 – Medidores digitais.	67
Figura 11 – Exemplificação do perfil de curva de carga processados.	69
Figura 12 – Avaliação do número de <i>clusters</i> (k) (30 segundos).	72
Figura 13 – Avaliação do número de <i>clusters</i> (k) (15 minutos).	72
Figura 14 – Estimativa dos centroides para cada grupo.	74
Figura 15 – Arquitetura implementada do modelo NICE.	75
Figura 16 – Hiperparâmetros da função m (rede neural).	77
Figura 17 – Perfis de carga gerados - <i>Cluster 1</i> (Cenário 3 <i>clusters</i>).	79
Figura 18 – Perfis de carga gerados - <i>Cluster 2</i> (Cenário 3 <i>clusters</i>).	80
Figura 19 – Perfis de carga gerados - <i>Cluster 3</i> (Cenário 3 <i>clusters</i>).	80
Figura 20 – Perfis de carga gerados- <i>Cluster 1</i> (Cenário 4 <i>clusters</i>).	81
Figura 21 – Perfis de carga gerados - <i>Cluster 2</i> (Cenário 4 <i>clusters</i>).	81
Figura 22 – Perfis de carga gerados - <i>Cluster 3</i> (Cenário 4 <i>clusters</i>).	82
Figura 23 – Perfis de carga gerados- <i>Cluster 4</i> (Cenário 4 <i>clusters</i>).	82
Figura 24 – Estimativa PDF - <i>Cluster 1</i> (Cenário 3 <i>clusters</i>).	83
Figura 25 – Estimativa PDF - <i>Cluster 2</i> (Cenário 3 <i>clusters</i>).	84
Figura 26 – Estimativa PDF - <i>Cluster 3</i> (Cenário 3 <i>clusters</i>).	84
Figura 27 – Análise KL, WD e JS (Cenário 3 <i>clusters</i>).	85
Figura 28 – Estimativa PDF- <i>Cluster 1</i> (Cenário 4 <i>clusters</i>).	87
Figura 29 – Estimativa PDF - <i>Cluster 2</i> (Cenário 4 <i>clusters</i>).	87
Figura 30 – Estimativa PDF - <i>Cluster 3</i> (Cenário 4 <i>clusters</i>).	88
Figura 31 – Estimativa PDF - <i>Cluster 4</i> (Cenário 4 <i>clusters</i>).	88

Figura 32 – Análise KL, WD e JS (Cenário 4 clusters).	89
Figura 33 – Análise KL, WD, JS - <i>Cluster 1</i> (Cenário 3 clusters).	93
Figura 34 – Análise KL, WD, JS - <i>Cluster 2</i> (Cenário 3 clusters).	94
Figura 35 – Análise KL, WD, JS - <i>Cluster 3</i> (Cenário 3 clusters).	95
Figura 36 – Análise KL, WD, JS- <i>Cluster 1</i> (Cenário 4 clusters).	98
Figura 37 – Análise KL, WD, JS - <i>Cluster 2</i> (Cenário 4 clusters).	99
Figura 38 – Análise KL, WD, JS - <i>Cluster 3</i> (Cenário 4 clusters).	100
Figura 39 – Análise KL, WD, JS - <i>Cluster 4</i> (Cenário 4 clusters).	101
Figura 40 – Hiperparâmetros da função m (rede neural).	104

LISTA DE TABELAS

Tabela 1 – Grandezas aferidas pelos medidores digitais.	68
Tabela 2 – Valores numéricos: KL, WD, JS (Cenário 3 <i>Clusters</i>).	86
Tabela 3 – Valores numéricos: KL, WD, JS (Cenário 4 <i>Clusters</i>).	90
Tabela 4 – NICE denso: Valores KL, WD, JS (Cenário 3 <i>clusters</i>).	96
Tabela 5 – NICE convolucional: Valores KL, WD, JS (Cenário 3 <i>clusters</i>).	97
Tabela 6 – NICE denso: Valores KL, WD, JS (Cenário 4 <i>clusters</i>).	102
Tabela 7 – NICE convolucional: Valores KL, WD, JS (Cenário 4 <i>clusters</i>).	102
Tabela 8 – NICE denso: Valores KL, WD, JS (Cenário 3 <i>clusters</i>).	105
Tabela 9 – NICE denso: Valores KL, WD, JS (Cenário 4 <i>clusters</i>).	106

LISTA DE ABREVIATURAS E SIGLAS

DBI	Davies–Bouldin Index
DGM	Deep Generative Models
GD	Geração Distribuída
EE	Eficiência Energética
EMS	Energy Management System
ECVE	Estação de Carregamento de Veículo Elétrico
GAN	Generative Adversarial Network
GMM	Gaussian Mixture Model
IA	Inteligência Artificial
IoT	Internet of Things
JS	Jensen-Shannon
KDD	Knowledge Discovery in Databases
KDE	Kernel Density Estimation
KL	Kullback-Leibler
KNN	K-nearest neighbors
ML	Machine Learning
NICE	Non-linear Independent Component Estimation
PRD	Programas de Resposta da Demanda
RD	Resposta da Demanda
RMSE	Root Mean Squared Error
RNA	Redes Neurais Artificiais
RED	Recursos Energéticos Distribuído
ReLU	Rectified Linear Unit
RNC	Rede Neural Convolutacional
REI	Redes Elétricas Inteligente
PDF	Probability Density Function
SAE	Sistemas de Armazenamento de Energia
SC	Silhouette Coefficient
SFV	Sistema Fotovoltaico
TIC	Tecnologias da Informação e Comunicação
TL	Transfer Learning
UFJF	Universidade Federal de Juiz de Fora
UNICAMP	Universidade Estadual de Campinas

VAE	Variational Autoencoder
VE	Veículo Elétrico
WCSS	Within-Cluster Sum of Square
WD	Wasserstein Distance

SUMÁRIO

1	INTRODUÇÃO	16
1.1	REVISÃO BIBLIOGRÁFICA	20
1.2	PRINCIPAIS CONTRIBUIÇÕES	25
1.3	OBJETIVOS	26
1.3.1	Objetivo Geral	26
1.3.2	Objetivos Específicos	27
1.4	PUBLICAÇÕES REFERENTES À PESQUISA	27
1.4.1	Publicações Decorrentes da Pesquisa	28
1.4.2	Publicação Relacionados à Pesquisa	28
1.5	ESTRUTURA DA TESE	29
2	FUNDAMENTOS TEÓRICOS	31
2.1	ANÁLISE DE DADOS E PROCESSO KDD	31
2.1.1	Processo KDD	32
2.2	ESTIMATIVA DE COMPONENTES INDEPENDENTES NÃO LI- NEARES	36
2.2.1	Estruturação matemática do modelo NICE	39
3	METODOLOGIA PROPOSTA	45
3.1	VISÃO GERAL DA METODOLOGIA	45
3.2	SELEÇÃO DE DADOS	46
3.3	PRÉ-PROCESSAMENTO	47
3.4	TRANSFORMAÇÃO DOS DADOS	49
3.5	CLUSTERIZAÇÃO	50
3.5.1	Coeficiente de Silhueta	52
3.5.2	Índice de <i>Davies-Bouldin</i>	54
3.5.3	Método do Cotovelo	55
3.6	GERAÇÃO DE PERFIL DE CURVA DE CARGA SINTÉTICA	56
3.7	AVALIAÇÃO DO MODELO GENERATIVO	57
3.7.1	Função de densidade de probabilidade	58
3.7.2	Divergência de Kullback-Leibler (KL)	59
3.7.3	Divergência Jensen-Shannon (JS)	60
3.7.4	Distância <i>Wasserstein</i> (WD)	61

4	SIMULAÇÃO E RESULTADOS	63
4.1	CONTEXTUALIZAÇÃO DO AMBIENTE DE SIMULAÇÃO	63
4.2	PROCESSAMENTO E TRATAMENTO DOS DADOS	68
4.3	CLUSTERIZAÇÃO DOS DADOS	70
4.4	ARQUITETURA IMPLEMENTADA	74
4.5	GERAÇÃO E AVALIAÇÃO	78
4.6	ANÁLISE DA GRANULARIDADE TEMPORAL	92
4.7	TESTES ADICIONAIS E VARIAÇÕES DE MODELOS	103
4.8	ANÁLISES GERAIS	107
5	CONSIDERAÇÕES FINAIS	110
5.1	SUGESTÕES DE DESENVOLVIMENTOS FUTUROS	114
	REFERÊNCIAS	116

1 INTRODUÇÃO

A digitalização tem impulsionado a inovação nos sistemas elétricos de distribuição, trazendo desafios e mudanças significativas para o setor elétrico. Uma das principais transformações nesse sentido é a integração de infraestruturas de tecnologias da informação e comunicação (TIC) e medição avançada na rede elétrica (SINGH et al., 2022), promovendo a expansão das redes elétricas inteligentes, ou *smart grids* (BUTT; ZULQARNAIN; BUTT, 2021).

As redes inteligentes são caracterizadas pela integração de diferentes tecnologias emergentes, como fontes de geração distribuída (GD), estações de carregamento de veículos elétricos (VE), sistemas de armazenamento de energia (SAE), gerenciamento da demanda, como programas de resposta a demanda (PRD) e ações de eficiência energética (EE) (ESCOBAR et al., 2021; PANDA et al., 2022). Adicionalmente, essas redes incorporam recursos avançados de monitoramento, dispositivos de internet das coisas (do inglês, *internet of things (IoT)*), controle e automação, que possibilitam uma gestão mais eficiente da rede elétrica (RABIE et al., 2020).

A implementação das redes inteligentes visa maximizar a utilização dos recursos integrados na infraestrutura da rede elétrica, proporcionando uma série de benefícios. Entre eles, destacam-se a eficiência operacional, a redução das perdas técnicas e comerciais, a diminuição da emissão de gases de efeito estufa provenientes da geração de energia elétrica por fontes poluentes, a melhoria do custo-benefício e do desempenho da operação da rede, o uso racional da energia elétrica e a redução de custos, além de possibilitar uma maior interação com os consumidores (SALEH; RABIE; ABO-AL-EZ, 2016; DILEEP, 2020).

A digitalização das redes inteligentes, a disseminação em larga escala dos recursos energéticos distribuídos (REDs) e o aumento constante na demanda de energia introduzem uma complexidade substancial. Diante desse cenário, torna-se imperativo adotar estratégias eficientes de gerenciamento de ativos e processos, especialmente nas camadas arquitetônicas envolvidas, como na camada de sensores e medidores (PANDA; DAS, 2021).

Essas estratégias são essenciais não apenas para garantir a operacionalidade,

estabilidade, confiabilidade, segurança e sustentabilidade energética das redes, mas também para promover a interoperabilidade entre os dispositivos envolvidos, como destacado por (BUTT; ZULQARNAIN; BUTT, 2021). Para alcançar esses benefícios, tem sido realizada a integração das TICs às redes elétricas, garantindo a conectividade entre os equipamentos e a troca de informações em tempo real (RABIE et al., 2020).

Ademais, é fundamental que os processos de gerenciamento de dados e informações sejam eficientes e seguros, garantindo a privacidade e a confidencialidade dos dados dos consumidores e a proteção contra possíveis ataques cibernéticos. Para gerenciar os recursos de redes elétricas inteligentes, é comum a implementação de um sistema de gerenciamento de energia (do inglês, *energy management system (EMS)*) (RATHOR; SAXENA, 2020).

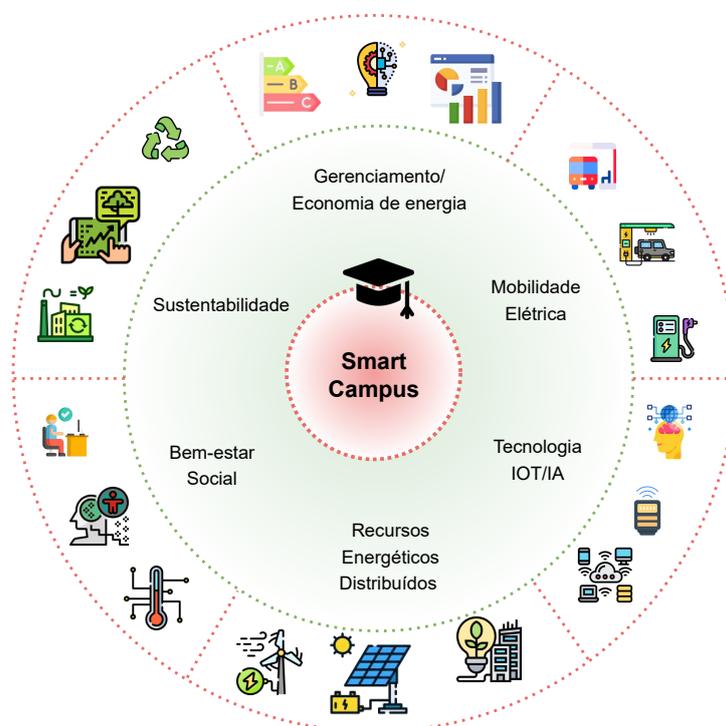
Os EMSs têm se mostrado essenciais para garantir a eficiência, estabilidade e confiabilidade das microrredes e redes elétricas inteligentes. Diante da crescente complexidade dessas redes, esses sistemas desempenham um papel fundamental ao permitir funcionalidades como o monitoramento, medição e o controle de geração e consumo de energia, além da otimização do fluxo de energia (SHUFIAN; MOHAMMAD, 2022). Ao utilizar dados em tempo real e algoritmos avançados, o EMS auxilia na tomada de decisões, permitindo a identificação de áreas onde a eficiência pode ser melhorada. Com isso, é possível reduzir os custos operacionais, além de garantir estabilidade e confiabilidade da rede elétrica (JAVED et al., 2021).

A difusão das redes elétricas inteligentes e o uso de EMS são temas amplamente discutidos, especialmente em relação às microrredes (BAYRAM; USTUN, 2017; RATHOR; SAXENA, 2020; SHARMA; SOOD, 2022). Nesse contexto, os *campi* universitários têm se destacado como uma aplicação promissora para a implementação de microrredes inteligentes (ZHANG et al., 2015). Os *campi* podem ser considerados cidades em escala reduzida (JAVED et al., 2021; GARCIA et al., 2022), uma vez que abrangem uma variedade de edifícios com diferentes finalidades, como departamentos administrativos, salas de aula, laboratórios, restaurantes/refeitórios, instalações recreativas e esportivas, além de dormitórios e/ou alojamentos, em alguns casos (PAPADOPOULOS et al., 2016; MUQEET et al., 2021).

As atividades realizadas nas instalações universitárias demandam frequentemente

uma quantidade significativa de energia elétrica. No entanto, nem sempre o consumo energético é eficiente e adequadamente gerenciado, o que torna necessário promover ações mais sustentáveis e efetivas. Diante desse cenário, surge a necessidade de implementar medidas voltadas para a economia e o aprimoramento da eficiência energética, visando promover a sustentabilidade nas instalações universitárias (SILVA et al., 2022). Uma solução viável para alcançar esse objetivo é a implementação de um *smart campus*, incorporando tecnologias inteligentes para monitorar, controlar e otimizar o consumo de energia. A Figura 1 ilustra o modelo de *smart campus*.

Figura 1 – Visão geral: *Smart campus*.



Fonte: Elaborado pela autora (2023), baseado em (HENRIQUE et al., 2023).

A Figura 1 exemplifica os elementos esperados em um *smart campus*, integrando diferentes aspectos fundamentais, tais como sustentabilidade, bem-estar, gerenciamento inteligente, economia, tecnologia, mobilidade elétrica e utilização de REDs.

Através da implementação de uma microrrede inteligente em um *campus* universitário, é possível alcançar diversos benefícios, como a identificação e a mitigação de

desperdícios de energia, o ajuste do consumo conforme as necessidades reais e a adoção de fontes de energia renováveis (HENRIQUE et al., 2023). Essa abordagem também traz vantagens adicionais, servindo como um modelo de negócio para implementação de microrredes locais e gestão de energia das edificações, além de ser um instrumento para estudos, integração, qualificação e teste de novas concepções e tecnologias (PAPADOPOULOS et al., 2016; JAVED et al., 2021).

Para maximizar os benefícios mencionados, é viável explorar o monitoramento dos dados das edificações do *campus*, visando desenvolver modelos de previsão e metodologias eficientes de gerenciamento de energia (MUQEET et al., 2021). No entanto, a coleta de dados ainda enfrenta desafios técnicos, regulatórios e econômicos que devem ser superados para promover uma adoção mais ampla desses sistemas (KUMARI; TANWAR, 2020; PONNUSAMY et al., 2021).

Diante dessas dificuldades, métodos baseados em dados sintéticos (BENDAOU; FARAH; Ben Ahmed, 2021) e técnicas de aprendizado por transferência (do inglês, *transfer learning (TL)*) têm sido amplamente difundidos como soluções para superar os desafios relacionados à gestão do consumo de energia e ao uso de dados (GAO et al., 2020). Nesse contexto, a abordagem comumente adotada de geração de perfis sintéticos de curvas carga permite a criação de conjuntos de dados novos a partir de informações existentes, possibilitando a análise de diferentes estratégias de gerenciamento de energia e a identificação de possíveis oportunidades de melhoria.

Essas abordagens, baseadas em *big data* e geração de perfis sintéticos de curvas carga, permitem o uso de modelos de previsão de energia, resultando no desenvolvimento de estratégias mais eficientes de gerenciamento de energia em edifícios (GAO et al., 2020). Com o auxílio dessas tecnologias, as edificações podem monitorar seu consumo de energia com maior precisão, identificando oportunidades para redução de custos (PÉREZ-CHACÓN et al., 2018). Consequentemente, esses avanços tecnológicos, contribuem significativamente para a otimização do uso de energia e para a promoção de práticas mais sustentáveis (KUMARI; TANWAR, 2020).

Considerando o contexto abordado, realiza-se uma breve revisão da literatura na Seção 1.1 para explorar os trabalhos relacionadas à análise de dados de energia em edificações, microrredes e redes inteligentes, com ênfase no desenvolvimento de modelos

de geração de perfis de curvas de carga sintéticas. O objetivo principal é fornecer uma visão abrangente das principais técnicas e metodologias utilizadas nessa vertente, ao mesmo tempo, em que são identificadas as lacunas e desafios que ainda precisam ser abordados na área de pesquisa.

1.1 REVISÃO BIBLIOGRÁFICA

A análise do perfil de carga é um fator determinante para a operação eficiente de redes elétricas inteligentes, especialmente no que diz respeito ao aprimoramento, gestão e planejamento do sistema de energia em edificações (AHMAD; ZHANG; YAN, 2020). Para atingir esse objetivo, diferentes estudos têm sido desenvolvidos, abordando modelos de alta performance para análise da demanda de energia elétrica (TOKTAROVA et al., 2019), considerando a complexidade da rede elétrica e a natureza detalhada dos diferentes perfis de carga que compõem o sistema (KAUR et al., 2020).

Técnicas avançadas, como inteligência artificial (IA), aprendizado de máquina (do inglês, *machine learning (ML)*) e modelos estatísticos, são amplamente utilizadas para identificar, avaliar e analisar o desempenho de dados em redes elétricas inteligentes (PACHECO et al., 2022). Com o auxílio dessas ferramentas, é possível desenvolver perfis de curvas de carga sintéticas com alta precisão.

Os modelos permitem o auxílio da operação, tornando o sistema de energia mais eficiente, garantindo o fornecimento de energia às edificações de forma confiável (TIGHTIZ; YOO, 2022). Além disso, as informações obtidas por meio desses modelos, também, podem ser utilizadas para o planejamento energético, preparando o sistema de energia para futuras demandas, alocação e dimensionamento otimizado de REDs (GRASSO et al., 2022), e permitindo ações corretivas e uma estruturação estratégica do sistema de forma mais precisa (RAZA; KHOSRAVI, 2015).

No âmbito da utilização de dados sintéticos em redes elétricas inteligentes, os modelos generativos profundos, como as redes adversárias generativas (do inglês, *generative adversarial networks (GAN)*), autoencoder variacional (do inglês, *variational autoencoders (VAEs)*) e modelos generativos profundos baseados em fluxo, têm se destacado recentemente na literatura. Esses modelos visam gerar perfis sintéticos de carga com alta precisão, buscando reproduzir perfis os mais realistas possíveis. Além

disso, a abordagem de geração de dados sintéticos apresenta uma vantagem adicional ao mitigar preocupações relacionadas à segurança e privacidade dos dados provenientes de medidores digitais (WANG; HONG, 2020).

Nesse contexto, um dos modelos mais difundidos na literatura para gerar perfis sintéticos de carga é o GAN e suas variações, conforme discutido por Wang e Hong (2020). Nesse artigo, os autores propuseram uma nova abordagem para gerar perfis realistas de carga elétrica em edifícios. A análise dessa abordagem foi realizada utilizando um banco de dados do projeto *Building Data Genome Project*, sendo as curvas geradas avaliadas pela média, o desvio padrão e a distribuição dos principais parâmetros dos perfis sintéticos de carga em comparação aos perfis reais. Os resultados alcançados evidenciaram que os perfis de carga gerados pela rede GAN conseguem capturar tanto a tendência geral quanto as variações aleatórias em comparação com os perfis de carga elétrica considerados.

Considerando a arquitetura condicional das redes adversárias generativas nomeadas de cGAN, Bendaoud, Farah e Ahmed (2021) utilizaram-na para estimar as perdas de carga diárias em um cenário de curto prazo, em que diferentes metodologias foram incorporadas ao GAN, como o *Deep Convolutional GAN*, *Least Squares GAN* e *Wasserstein GAN*. Os perfis sintéticos de carga gerados foram testados em um conjunto de dados de um ano e comparados com os perfis de carga reais por meio de métricas estatísticas. A precisão e eficácia ao longo do período previsto foram destacadas, onde os modelos GAN propostos forneceram previsões com uma média de erro percentual médio absoluto de 4,99%.

Também, considerando a abordagem cGAN, Pinceti, Sankar e Kosut (2021) utilizaram o modelo para gerar perfis de carga. O objetivo desse projeto foi aprender os padrões de um conjunto de dados reais de perfis de carga semanais amostrados por hora e gerar perfis sintéticos exclusivos sob demanda, com base na estação e no tipo de carga necessária. Foram realizados testes extensivos do modelo generativo para verificar se os dados sintéticos capturam totalmente as características das cargas reais e se podem ser usados para sistemas de energia *downstream* e/ou aplicações de aprendizado de máquina.

Além disso, Zhang et al. (2022) integraram a estimativa de média móvel (cGAN-MA) ao método cGAN para gerar perfis de carga. O modelo cGAN-MA utiliza infor-

mações como a forma da carga, clima e tempo. O desempenho do modelo foi validado em termos de similaridade, variabilidade e diversidade, sendo comparado com outros métodos tradicionais de geração de perfis de carga, como o GAN e o VAE. Os resultados mostram que o modelo proposto apresenta um bom desempenho em termos de diversidade, similaridade e variabilidade dos perfis de carga gerados. Assim, o modelo CGAN-MA conseguiu simular perfis de carga sob diferentes condições meteorológicas e de tempo, auxiliando no gerenciamento da operação de redes.

Hu et al. (2022) propuseram um modelo chamado MultiLoad-GAN, baseado na relação codificação-decodificação, que gera simultaneamente um grupo de perfis sintéticos de carga. A principal inovação do MultiLoad-GAN é capturar as correlações espaço-temporais entre as cargas atendidas por um mesmo transformador de distribuição, gerando uma quantidade significativa de curvas de carga correlacionadas para serem utilizadas em estudos de microrredes e sistemas de distribuição. O estudo também apresenta duas métricas para avaliar o realismo das curvas geradas, utilizando estatísticas computacionais e um classificador de aprendizado profundo. Os resultados das simulações mostram que o MultiLoad-GAN gera perfis de carga mais realistas do que as abordagens existentes, especialmente em termos de características de grupo.

Baasch, Rousseau e Evins (2021) aplicaram uma rede GAN com séries temporais denominada Time-GAN para a geração de perfis sintéticos de carga temporais. O objetivo foi avaliar a capacidade de geração de dados sintéticos, comparando o número de dados gerados com uma porcentagem menor. Para avaliar o desempenho do modelo, utilizaram a divergência de *Jensen-Shannon* e o erro de validação da previsão preditiva. Os resultados numéricos mostraram-se competitivos com outras metodologias de GAN consideradas na literatura, mesmo utilizando apenas 1% a 2% do tamanho dos dados residenciais e comerciais, fornecendo informações importantes para reduzir a escassez de dados nesse domínio.

Song, Li e Lu (2022) propuseram uma abordagem de dois estágios para a super-resolução de perfis de carga. No primeiro estágio, um modelo baseado em GAN é utilizado para restaurar os componentes de alta frequência de perfis de carga de baixa resolução, considerando a relação entre consumo e condições climáticas. Métricas de avaliação específicas para a forma de carga foram desenvolvidas para avaliar a realidade das cargas geradas. Os resultados indicaram um desempenho satisfatório na restauração

de componentes de alta frequência, demonstrando a eficácia do modelo em conjuntos de dados temporais reduzidos e melhorando o desempenho em tarefas subsequentes que requerem perfis de carga aprimorados

Li et al. (2022) propuseram um modelo designado como *Wasserstein deep convolutional generative adversarial network* para gerar séries temporais sintéticas de consumo de energia. Para avaliar o modelo proposto, os dados gerados foram comparados com os dados reais, usando métricas como a autocorrelação e a raiz quadrática média. O gerador de dados proposto foi avaliado utilizando dados da Irlanda, e os resultados mostram que o modelo implementado, pode gerar perfis de carga realistas com similaridade satisfatória em termos de distribuição de dados, padrões e desempenho.

Grasso et al. (2022) utilizaram a arquitetura GAN, mais especificamente a metodologia de convolução profunda Wasserstein com penalidade de gradiente, para criar perfis de carga artificial com propriedades estocásticas significativas. Os perfis são gerados a partir de *clusters* de dados de setores residenciais, capturando as propriedades estatísticas dos distintos agrupamentos. A precisão dos perfis gerados é avaliada comparando propriedades estatísticas, como média, desvio padrão e variância, com as do conjunto de dados original. Os resultados mostram que os perfis gerados pela GAN são estatisticamente semelhantes aos perfis reais, podendo ser úteis para lidar com dados de carga e geração de energia incompletos ou irregulares, fornecendo uma representação mais precisa e utilizável desses dados.

Considerando a metodologia do codificador autoencoder variacional, o trabalho de Pan et al. (2019) apresentou uma abordagem baseada em dados para a geração de perfis de carga de veículos elétricos, a fim de avaliar a penetração na análise de redes de distribuição. Este artigo verifica a validade dos perfis de carga gerados a partir de quatro aspectos: distribuição de probabilidade de cargas, correlação temporal de cargas, duração das cargas e volatilidade das cargas. Os resultados da simulação mostram que os perfis de carga veículos elétricos gerados conservam a correlação temporal e a natureza de distribuição de probabilidade dos perfis de carga originais.

Já utilizando-se a arquitetura de rede neural baseada em autoencoder variacional condicional, Wang, Tindemans e Palensky (2022) desenvolveram um modelo generativo para geração de perfis sintéticos de carga representativos de clientes industriais e

comerciais, utilizando a seleção aleatória de tipologias de cargas puras, geradores puros e cargas/geradores mistos. Para validar a capacidade de geração do modelo proposto, foram realizadas inspeções visuais, comparando os perfis de carga diária junto à análise estatística da distribuição cumulativa dos dados históricos e gerados em diferentes escalas de tempo. Os resultados experimentais obtidos demonstram que o modelo conseguiu gerar perfis sintéticos de carga, considerando o desempenho de testes estatísticos.

Usando o conceito de aprendizado de transferência, Salazar et al. (2020) apresentou em seu estudo uma estrutura para criar perfis de carga diários adaptados às características sociodemográficas das áreas atendidas por transformadores de distribuição de média e baixa tensão. O objetivo é reduzir o número de dispositivos de medição necessários na rede elétrica de distribuição. A estrutura utiliza o aprendizado de transferência com uma abordagem de adaptação de domínio, transferindo padrões de consumo de um transformador para outro. Para a avaliação da técnica proposta, o algoritmo *Robust Boosted* é selecionado. O estudo demonstra a aplicação da estrutura em dados reais de dois municípios na Holanda, destacando a importância das técnicas de classificação e das características relevantes identificadas para o processo de geração de perfis de carga.

Considerando-se os modelos baseados em fluxo, têm-se o trabalho de Ge et al. (2020), que propuseram o uso do modelo de estimativa de componentes independentes não lineares (do inglês, *nonlinear independent component estimation (NICE)*) para gerar perfis de carga diários realistas. O desempenho do método proposto foi testado considerando um conjunto de dados reais de edifícios. Para avaliar a similaridade entre os perfis de carga diários gerados e as amostras reais, foram utilizados indicadores de distribuição de probabilidade, correlação espaço-temporal e volatilidade dos perfis de carga. Os resultados da simulação mostraram que o modelo NICE pode capturar com precisão a correlação espaço-temporal das amostras reais, resultando em perfis sintéticos de carga semelhantes aos reais.

Também, com base no modelo NICE, uma abordagem foi proposta por Hu et al. (2021) para a previsão de cenários de energia eólica, representando diferentes comportamentos com base em dados históricos de potência eólica e previsões pontuais determinísticas. O objetivo da pesquisa foi capturar as características de flutuação e as correlações temporais da energia eólica. A estrutura proposta foi projetada para aprender padrões latentes das curvas históricas de energia eólica e utilizar técnicas de otimização

para gerar um amplo conjunto de cenários futuros. Os resultados das simulações indicam que a abordagem proposta supera técnicas populares de referência, como o GAN e VAE, sendo os modelos avaliados por meio de métricas como a largura média do intervalo de previsão e a porcentagem de cobertura do intervalo de previsão.

1.2 PRINCIPAIS CONTRIBUIÇÕES

Considerando a abordagem discutida anteriormente em relação à expansão de redes elétricas inteligentes e a análise de dados, a motivação do presente trabalho visa superar as lacunas encontradas na literatura e na temática abordada, tal como: a falta de infraestrutura de medição, a escassez de dados, a garantia de privacidade e segurança.

Nesse sentido, a presente tese explora o potencial da análise de dados e o uso de modelos generativos aplicados à rede elétrica de distribuição. Isso se dá por meio da análise e geração de perfis sintéticos de curva de carga baseados em dados reais, isto é, dados medidos por meio de medidores digitais instalados na rede elétrica.

Dessa forma, as contribuições referem-se à aplicação original de um método inovador para geração de perfis de curva de carga sintética a partir de dados medidos. Essa abordagem envolve a modificação do modelo NICE existente na literatura por meio da utilização de redes neurais com camadas convolucionais.

Portanto, este estudo inclui contribuições primárias e secundárias, embora as considerações secundárias não se destaquem por inovação, elas apresentam relevância no contexto abordado. As contribuições primárias desta tese são as seguintes:

- Desenvolvimento de perfis sintéticos de curvas de carga: aplicação de métodos de geração de dados sintéticos para geração de perfis sintéticos de curva de carga baseado em dados medidos (reais).
- Modelo NICE para geração de perfis de curva de carga: adoção do modelo NICE com camadas convolucionais para geração de perfis sintéticos de curva de cargas.
- Granularidade temporal: análise do impacto da granularidade temporal na modelagem de geração de perfis sintéticos de curva de cargas, considerando intervalos de tempo distintos, visando aprimorar a precisão e a utilidade das técnicas propostas

para abordar diferentes cenários tanto no planejamento quanto na operação das redes elétricas de distribuição.

As considerações secundárias desta tese incluem:

- Metodologia de análise de curva de carga: detalhamento da metodologia de preparação, processamento, tratamento, mineração e avaliação de dados de perfis de curva de carga diárias medidas baseado no processo de conhecimento de banco de dados.
- Representatividade dos dados: utilização de banco de dados reais (medidos) com grande volume de dados, obtidos a partir de medições em tempo real de uma microrrede universitária, permitindo análises mais precisas e representativas dos dados.

No geral, esta tese fornece uma abordagem aprimorada na análise e geração de perfis de curvas de carga. Até o momento da pesquisa, não foram encontradas aplicações do modelo NICE com camadas convolucionais para gerar perfis de curva de carga diária unidimensionais em pesquisas anteriores. Para validar o método proposto, é feita uma comparação entre os resultados obtidos do modelo NICE convolucional com os resultados obtidos pela aplicação do do modelo NICE com camadas densas, o qual já foi apresentado anteriormente na literatura para essa aplicação.

1.3 OBJETIVOS

Os objetivos deste trabalho estão dispostos em objetivo geral e específicos, apresentados a seguir.

1.3.1 Objetivo Geral

O objetivo geral desta tese é desenvolver uma abordagem abrangente para a geração de curvas de carga sintéticas, fundamentada na análise de dados reais. Esta pesquisa visa aprimorar a análise e a geração de perfis de curvas de carga, destacando-se

pela inovação na aplicação do modelo NICE com camadas convolucionais. A validação do método proposto será realizada por meio da comparação entre os resultados obtidos pelo modelo NICE convolucional e os resultados provenientes da aplicação do modelo NICE com camadas densas.

1.3.2 Objetivos Específicos

Esta tese tem como objetivos específicos:

- Abordar o processo de descoberta de conhecimento em bases de dados (do inglês, *knowledge discovery in databases (KDD)*) para realizar a análise exploratória de dados de perfis de carga diários.
- Utilizar técnicas de clusterização para agrupar os dados de consumo de energia elétrica em diferentes grupos de perfis de carga, com o objetivo de identificar padrões de consumo de energia elétrica em edificações e em redes elétricas universitárias.
- Implementar metodologias para a geração de perfis unidimensionais de curvas de carga sintéticas, baseando-se na aplicação do modelo NICE com camadas densas e convolucionais. Esta abordagem busca explorar e comparar os resultados obtidos pela utilização de diferentes arquiteturas do modelo NICE.
- Analisar a granularidade temporal dos dados de consumo de energia elétrica, considerando diferentes intervalos de tempo (30 segundos, 1 minuto, 5, 10, 15 minutos e 1 hora), visando compreender o impacto da granularidade na identificação de padrões de consumo e na precisão da geração de curvas de cargas sintéticas.

1.4 PUBLICAÇÕES REFERENTES À PESQUISA

As contribuições referentes à pesquisa deste doutoramento foram publicadas em congressos nacionais e internacionais, bem como em revistas na área de conhecimento da temática abordada neste projeto, sendo estas subdivididas em publicações decorrentes da pesquisa e publicações relacionadas à pesquisa.

1.4.1 Publicações Decorrentes da Pesquisa

- **Avaliação técnica e econômica da alocação de geração distribuída fotovoltaica em universidade pública: um estudo de caso.** In: 2021 14th IEEE International Conference on Industry Applications (INDUSCON), 2021, pp. 46-52, doi: 10.1109/INDUSCON51756.2021.9529490.
- **Análise preditiva do consumo de energia elétrica baseado em séries temporais: um estudo de caso.** Anais do LIV Simpósio Brasileiro de Pesquisa Operacional. O papel da PO na transição energética, 2022, Juiz de Fora. Anais eletrônicos. Campinas, Galoá, 2022.
- **Energy management system: a case study for a campus microgrid,"**2022 IEEE PES GTD Latin America, La Paz, Bolivia, 2022, pp. 1-6, doi: 10.1109/IEEE-PESGTDLatinAmeri53482.2022.10038294.
- **Optimal siting and sizing of distributed energy resources in a smart campus.** Electric Power Systems Research, Elsevier, Volume 217, page. 109095, 2022, ISSN 0378-7796, <https://doi.org/10.1016/j.epsr.2022.109095>.
- **Minimization of electrical losses in distribution transformers on a university campus,** aprovado para a conferência 15th IEEE International Conference on Industry Applications (INDUSCON) 2023.
- **Generating realistic load profiles in smart grids: an approach based on non-linear independent component estimation (NICE) and convolutional layers,** Applied Energy, Volume 351, 1 December 2023, 121902

1.4.2 Publicação Relacionados à Pesquisa

- **Market models and optimization techniques to support the decision-making on demand response for prosumers.** Electric Power Systems Research, Elsevier, Volume. 210, 108059, 2022, <https://doi.org/10.1016/j.epsr.2022.108059>.

- **Residential energy consumption forecasting using deep learning models**, Applied Energy, Elsevier, Volume 350, 2023, 121705, ISSN 0306-2619, <https://doi.org/10.1016/j.apenergy.2023.121705>.
- **Otimização dos parâmetros de controle para operação de microturbinas a gás**. In: Simpósio Brasileiro de Automação Inteligente-SBAI. 2021. doi: <https://doi.org/10.20906/sbai.v1i1.2660>.
- **Distributed generation hosting capacity evaluation for distribution systems through an optimal power flow** Workshop on Communication Networks and Power Systems (WCNPS), Brasilia, Brazil, 2023, pp. 1-6, doi: 10.1109/WCNPS60622.2023.10344817.

1.5 ESTRUTURA DA TESE

Esta tese está organizada em cinco capítulos. No Capítulo 1, é apresentada a contextualização da temática abordada, com ênfase na análise de consumo de energia em edificações, microrredes e redes inteligentes. Além disso, foi realizada uma revisão bibliográfica abrangendo as principais abordagens utilizadas para geração de perfis de curvas de carga sintéticas. O capítulo, também, descreve a motivação, a contribuição e os objetivos desta tese.

No Capítulo 2 é apresentada a fundamentação teórica, explorando-se os conceitos e etapas do processo de KDD, utilizada como base para a descoberta de conhecimento em bancos de dados. O modelo NICE também é descrito, abordando a modelagem matemática para desenvolver os perfis de curvas de carga sintéticas. Isso engloba a formulação que viabiliza a extração e compreensão das características dos dados medidos, possibilitando a geração de perfis sintéticos que sejam significativos.

No Capítulo 3, a metodologia proposta nesta tese é descrita detalhadamente. São apresentados os procedimentos metodológicos, incluindo a coleta e preparação dos dados, a aplicação do método de clusterização e o desenvolvimento dos perfis sintéticos de curvas de cargas. Essa descrição passo a passo permite uma compreensão clara do processo de análise de dados proposto e das diretrizes adotadas para alcançar os resultados pleiteados.

No Capítulo 4, inicialmente é apresentado o ambiente de teste, incluindo as variáveis consideradas nesta aplicação. Em seguida, são apresentados os resultados obtidos por meio das simulações computacionais da metodologia proposta. A análise e discussão dos resultados destacam as principais conclusões derivadas da abordagem proposta.

No Capítulo 5, é realizada uma conclusão que sintetiza os principais resultados alcançados ao longo desta pesquisa. São discutidas as contribuições alcançadas, destacando-se sua relevância e impacto e as perspectivas de trabalhos futuros.

2 FUNDAMENTOS TEÓRICOS

Neste capítulo, apresenta-se a fundamentação teórica que embasa esta pesquisa. Inicialmente, apresenta-se a análise de dados na Seção 2.1, considerando o processo de descoberta de conhecimento em bases de dados. Em seguida, na Seção 2.2, a apresentação do modelo NICE para a geração de perfis de curvas de carga sintéticas, considerando a estimação das componentes independentes não lineares.

2.1 ANÁLISE DE DADOS E PROCESSO KDD

Com a expansão das tecnologias conectadas, como a internet das coisas, os sistemas de sensoriamento e a ampla adoção de dispositivos digitais, observa-se um aumento substancial na geração de dados, sejam eles estruturados ou não estruturados, caracterizado a era do *big data* (MICHALIK; ŠTOFA; ZOLOTOVÁ, 2014; MARJANI et al., 2017). Essa enorme quantidade de dados apresenta um desafio significativo para a análise e extração de informações úteis, exigindo a utilização de ferramentas de análise de dados para compreender o volume coletado (AHMED et al., 2017; SESTINO et al., 2020).

Para lidar com esse desafio, as técnicas do processo de descoberta em bancos de dados, amplamente conhecida KDD (FRAWLEY; PIATETSKY-SHAPIRO; MATHEUS, 1992), vêm sendo adotadas (RAMOS et al., 2015a; CEMBRANEL et al., 2019; FAN et al., 2021). À medida que a capacidade de coletar e armazenar dados se expande, tornando-se mais acessível e econômica, abrem-se perspectivas promissoras para aproveitar todo o potencial de dados históricos (SESTINO et al., 2020).

O objetivo principal do processo KDD é extrair conhecimento útil e significativo a partir de grandes conjuntos de dados, utilizando técnicas automatizadas ou semi-automatizadas para identificar padrões, tendências, anomalias e relações entre os dados (CARO et al., 2020). Dessa forma, a análise de dados possibilita obter informações úteis para construção de modelos que possam diligenciar as complexidades e os processos subjacentes aos dados, explorando os padrões e tendências presentes nesses conjuntos de dados (MOHAMMADPOOR; TORABI, 2020; DEEPA et al., 2022). Com base nessas premissas, a Subseção 2.1.1 apresenta os processos básicos do KDD.

2.1.1 Processo KDD

O processo KDD representa uma abordagem estruturada para a aquisição de conhecimento útil a partir de dados. Esse método abrange várias etapas, como seleção, pré-processamento, transformação, mineração e interpretação dos dados. A finalidade central do KDD é extrair conhecimento significativo dos dados, conforme definido por Fayyad, Piatetsky-Shapiro e Smyth (1996) como o processo não trivial de identificar padrões válidos, inovadores, potencialmente úteis e, em última instância, compreensíveis em dados.

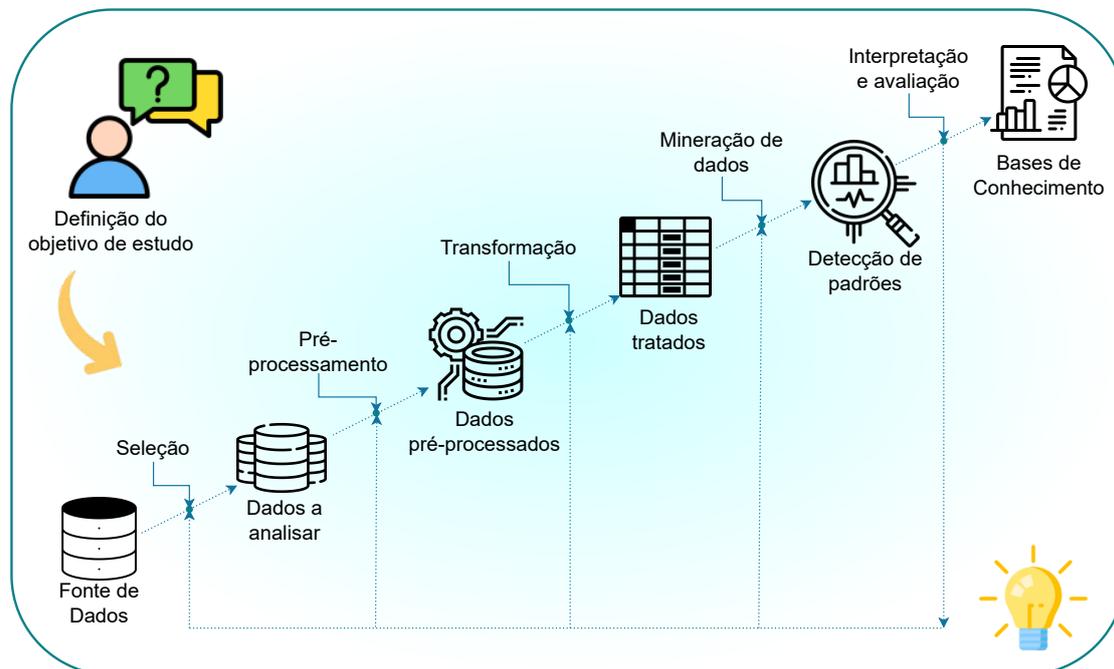
O objetivo principal do processo KDD é derivar conhecimento que possa ser aplicado para aprimorar o desempenho e os resultados em diversas áreas, auxiliando na tomada de decisões e em outras funcionalidades relevantes, como destacado por Cembranel et al. (2019). Isso inclui a incorporação de conhecimentos prévios relevantes e uma interpretação adequada dos resultados da mineração (FRAWLEY; PIATETSKY-SHAPIRO; MATHEUS, 1992).

A Figura 2 apresenta uma visão geral do processo de KDD, ilustrando as etapas fundamentais desde a seleção dos dados até a interpretação dos resultados obtidos.

Como ilustrado na Figura 2, o processo KDD constitui um ciclo iterativo que engloba diversas etapas interdependentes e interconectadas, desde a seleção e pré-processamento dos dados até a interpretação dos resultados obtidos. O propósito fundamental desse processo consiste em obter conhecimento relevante e compreensível a partir de conjuntos de dados. Cada uma das etapas representadas na figura é detalhadamente descrita conforme as referências (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996; CARO et al., 2020; SICSÚ; SAMARTINI; BARTH, 2023):

- a) **Definição do objetivo:** nesta etapa, é necessário compreender o domínio em questão (área específica de estudo) e utilizar o conhecimento prévio para estabelecer claramente os objetivos que serão buscados durante todo o processo de descoberta do conhecimento;
- b) **Seleção dos dados:** nesta etapa, realiza-se a seleção do conjunto de dados que será utilizado ao longo de todo o processo, considerando a escolha dos atributos

Figura 2 – Processo KDD.



Fonte: Elaborado pela autora (2023).

relevantes para a descoberta a ser realizada, alinhada com o objetivo da tarefa a ser desempenhada;

c) **Pré-processamento dos dados:** nesta etapa, os conjuntos de dados são submetidos a uma série de procedimentos visando à correção e padronização dos mesmos, desempenhando um papel fundamental ao preparar os dados em um formato apropriado para uso na etapa subsequente. Essa etapa, geralmente, pode ser subdividida em subetapas, tais como:

- **Limpeza dos dados:** esta subetapa abrange a remoção de dados duplicados ou redundantes, a correção de dados corrompidos ou inconsistentes. Além disso, pode ser necessário remover *outliers*, os quais são dados discrepantes contidos na base de dados, assegurando que a base de dados esteja completa, precisa e íntegra, contribuindo para a obtenção de resultados de qualidade na etapa de mineração de dados;

- **Tratamento de dados ausentes:** nessa etapa realiza-se o tratamento dos dados faltantes. Há duas maneiras principais de tratar dados ausentes, a exclusão ou a imputação de dados. A exclusão dos dados refere-se ao ato de remover informações ou registros de um conjunto de dados. A imputação é substituir os valores ausentes por valores estimados. Há diferentes métodos de imputação de dados, tal como uso de métricas estatísticas (média, mediana ou moda), modelos preditivos, *k-nearest neighbor* (KNN), estimativa baseada na distribuição multivariada do conjunto, entre outras. Essas abordagens oferecem maneiras diferentes de lidar com dados faltantes, dependendo dos requisitos e das características do conjunto de dados;
 - **Enriquecimento dos dados:** já nesta subetapa, visa-se melhorar a qualidade e a abrangência dos dados existentes, por meio da incorporação de informações adicionais, visando impulsionar a descoberta de conhecimento. Geralmente, neste estágio, utiliza-se bases externas para complementar e enriquecer a análise o conjunto dos dados;
- d) **Transformação dos dados:** nesta etapa, realiza-se o processo de conversão dos dados brutos em um formato apropriado para a aplicação de técnicas de mineração de dados. A transformação pode envolver diversas operações, tais como:
- **Normalização dos dados:** consiste em ajustar os valores dos atributos para uma escala específica ou padronizada, tornando os atributos comparáveis e evitando que valores mais elevados se sobressaiam nos resultados. Geralmente os dados são normalizados para valores entre 0 e 1, ou -1 e 1;
 - **Conversão de valores simbólicos para valores numéricos:** nessa etapa converte-se os valores em representações numéricas adequadas para poderem ser processados pelas técnicas de mineração de dados. Pois, frequentemente, os dados contêm atributos categóricos ou simbólicos que não podem ser diretamente utilizados, tendo-se, assim, a necessidade de convertê-los em um formato apropriado para análise;
 - **Discretização e composição de atributos:** em alguns casos, é vantajoso converter atributos contínuos em categorias discretas para simplificar o processo de mineração de dados. Isso pode ser feito por meio da discretização dos

valores em intervalos específicos. A composição de atributos pode servir para combinar características relacionadas, gerando outros atributos que podem fornecer informações mais significativas para a análise.

Finalizada a etapa de transformação dos dados, é necessário armazená-los adequadamente para serem posteriormente acessados e utilizados pelos algoritmos de mineração de dados;

e) **Mineração de dados:** nesta etapa, um conjunto ou subconjunto de dados transformados é selecionado para ser fornecido ao algoritmo de mineração de dados. O principal objetivo é identificar padrões e gerar resultados que serão analisados na próxima etapa do processo. O tipo de algoritmo utilizado depende dos objetivos a serem alcançados, tais como:

- **Classificação:** atribui rótulos ou categorias aos itens com base em seus atributos;
- **Agrupamento ou clusterização:** agrupa os itens similares sem rótulos pré-determinados, identificando agrupamento ideal;
- **Previsão de valores:** visa estimar quantidades numéricas com base em dados históricos ou tendências.

Na etapa de mineração de dados, outras abordagens e técnicas podem ser utilizadas, tais como: aprendizado de transferência, regressão logística, análise de componentes principais, KNN, entre outras. Normalmente, essas abordagens são exploradas para extrair informações úteis, auxiliando na tomada de decisão por meio de iterações repetidas de diferentes algoritmos. Isso permite identificar padrões que sejam mais relevantes para análise proposta;

f) **Interpretação e avaliação:** na fase final do processo de KDD, os padrões descobertos pela mineração de dados são analisados. A avaliação dos padrões obtidos pode ser realizada por meio da aplicação de métricas de avaliação, subjetivas e visuais, proporcionando uma compreensão abrangente dos resultados obtidos, sendo estas caracterizadas por:

- **Medidas de avaliação:** fornecem uma avaliação objetiva da qualidade e confiabilidade dos padrões. Medidas como precisão, erro médio, erro médio quadrático, entre outras, são utilizadas para determinar a eficácia dos modelos gerados durante a etapa anterior;
- **Medidas visuais:** são ferramentas que propiciam a identificação visual de padrões e percepções. Técnicas de visualização, incluindo a elaboração de *dashboards* interativos, gráficos e representações visuais, são empregadas com o propósito de facilitar a compreensão dos resultados, bem como a identificação de tendências e oportunidades;
- **Medidas subjetivas:** são relacionadas a perspectiva dos usuários na interpretação e uso dos resultados, contribuindo para uma análise mais completa e contextualizada;
- **Aprimoramento e ajustes:** durante essa etapa, pode ser necessário realizar ajustes nas etapas anteriores do processo, aprimorando os modelos ou experimentando diferentes técnicas para obter resultados ainda mais precisos e úteis. Pode ser necessário aprimorar a qualidade do conjunto de dados, adotando abordagens distintas de mineração ou ajustar os parâmetros de configuração dos algoritmos utilizados.

Finalizada a etapa de interpretação e avaliação, novos conhecimentos são gerados, podendo esses serem aplicados para tomada de decisões e no desenvolvimento de estratégias, explorando diferentes perspectivas em relação ao cenário considerado.

O processo KDD é caracterizado por um ciclo, conforme observado ao longo da descrição apresentada nesta seção. À medida que novos dados são coletados, torna-se necessário repetir as etapas a fim de atualizar e aprimorar os resultados obtidos, evidenciando-se como um processo iterativo em constante evolução.

2.2 ESTIMATIVA DE COMPONENTES INDEPENDENTES NÃO LINEARES

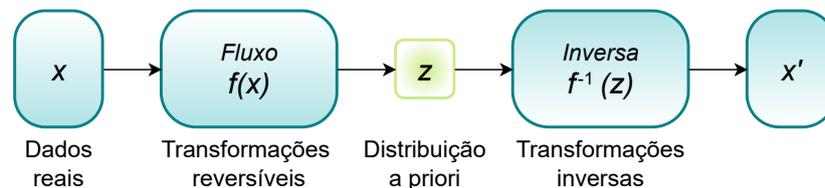
O modelo de estimativa de componentes independentes não lineares é um modelo generativo que faz parte da classe de modelos baseados em fluxo de normalização, cujo

objetivo é aprender a distribuição de probabilidade das amostras reais e, posteriormente, gerar novos dados de amostras (DINH; KRUEGER; BENGIO, 2014). Essa abordagem pertence à categoria de modelos generativos, semelhante às GANs e VAEs, mas se diferencia conceitualmente (RUTHOTTO; HABER, 2021).

Enquanto as GANs não impõem uma distribuição *a priori* (ALQAHTANI; KAVAKLI-THORNE; KUMAR, 2021) e os VAEs, aproximam-se da distribuição a priori através da inferência variacional (ODAIBO, 2019), o modelo NICE modela explicitamente a distribuição de probabilidade dos dados, baseando-se na distribuição real dos dados, a partir da estimação das componentes independentes não lineares (HU et al., 2021). Isso se dá por uma de série de transformações reversíveis, que modela a distribuição dos dados durante a codificação e decodificação dos dados (ZHANG; ZHANG, 2020a).

As transformações são projetadas de forma que a função de transformação é aplicável tanto no sentido direto (dos dados originais para um espaço de variáveis latentes), quanto no sentido inverso (do espaço de variáveis latentes de volta aos dados originais). Essa propriedade permite a reconstrução dos dados originais sem perda significativa de informação (KOBYZEV; PRINCE; BRUBAKER, 2021). A Figura 3 ilustra a estrutura típica do modelo NICE.

Figura 3 – Estrutura típica do modelo NICE.



Fonte: Elaborado pela autora (2023).

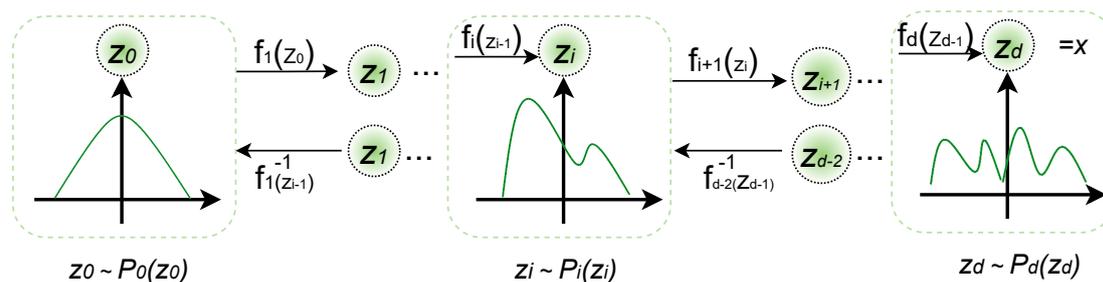
Na Figura 3 é ilustrada a estrutura típica do modelo NICE, onde x representa os dados reais, Fluxo ($f(x)$) refere-se à série de transformações reversíveis, z é a variável latente da distribuição a priori e $f^{-1}(z)$ denota as transformações inversas que reconstroem os dados a partir das variáveis latentes, gerando-se os dados sintéticos representado por x' . Em suma, no processo de codificação, os dados reais (x) são mapeados para o espaço variável latente (z), aprendendo uma série de transformações reversíveis (DINH;

KRUEGER; BENGIO, 2014; DINH; SOHL-DICKSTEIN; BENGIO, 2016).

A propriedade de transformações reversíveis refere-se à capacidade do modelo de mapear bidirecionalmente entre o espaço de entrada e o espaço de saída, assinalando ser possível gerar amostras sintéticas a partir de um espaço latente e, também, reconstruir a entrada original a partir de uma amostra gerada (saída). Essa reversibilidade permite uma maior flexibilidade e controle sobre o processo de geração e reconstrução de dados (DINH; SOHL-DICKSTEIN; BENGIO, 2016; KOBYZEV; PRINCE; BRUBAKER, 2021). Por meio dessas transformações, é possível mapear a distribuição de probabilidade das amostras reais para uma distribuição a priori.

No processo de decodificação, as amostras aleatórias provenientes da distribuição a priori são obtidas por amostragem do espaço variável latente e, em seguida, são mapeadas de volta ao espaço amostral usando as transformações inversas e diferenciais (KOBYZEV; PRINCE; BRUBAKER, 2021), conforme exemplificado na Figura 4.

Figura 4 – Exemplo do processo do fluxo de normalização.



Fonte: Elaborado pela autora (2023).

Na Figura 4 é ilustrado o processo de fluxo de normalização dos dados no modelo NICE. Esse processo envolve mapear a distribuição de probabilidade das amostras reais para uma distribuição a priori por meio de uma sequência de funções reversíveis (MÜLLER et al., 2019). Essas funções são formuladas matematicamente como transformações não lineares dos dados, sendo implementadas como camadas de acoplamento aditivo ou multiplicativo (HU et al., 2021).

Em cada camada de acoplamento, as variáveis de entrada são particionadas em duas partes: $f_{1:i}$ e $f_{i+1:d}$, de modo a manter uma parte dos dados inalterada, enquanto, a outra parte é transformada por uma função não linear dependente da parte inalterada.

A transformação é aplicada de forma que $f_{1:i}$ permaneça inalterado, enquanto $f_{i+1:d}$ é transformada por uma função não linear dependente de $f_{1:i}$ (DINH; SOHL-DICKSTEIN; BENGIO, 2016; ZHANG; ZHANG, 2020b). Essa abordagem permite a reversibilidade das transformações, uma vez que é possível reconstruir os dados originais a partir das variáveis latentes (KOBYZEV; PRINCE; BRUBAKER, 2021).

Ainda, na Figura 4, a variável d refere-se à quantidade de dimensões ou características presentes nos dados reais que estão sendo processados, sendo usada para indicar a dimensão total dos dados, determinada pela natureza dos dados de entrada, e pode variar conforme com o problema específico em que o modelo NICE está sendo aplicado. Já o termo i é usado para especificar o ponto de divisão entre as duas partes dos dados.

Ao passar pelas camadas de acoplamento, os dados são gradualmente transformados e normalizados, aproximando a distribuição das amostras reais da distribuição a priori desejada. Esse processo de normalização permite que as variáveis latentes sigam uma distribuição mais simples e conhecida, facilitando a geração de amostras sintéticas e outras operações no espaço latente. Portanto, baseado no princípio básico de estruturação do modelo NICE, é possível formular matematicamente tais processos. A modelagem apresentada na subseção 2.2.1 é fundamentada em estudos relevantes da literatura, tais como em (DINH; KRUEGER; BENGIO, 2014; DINH; SOHL-DICKSTEIN; BENGIO, 2016; ODAIBO, 2019; GE et al., 2020; ZHANG; ZHANG, 2020b; HU et al., 2021).

2.2.1 Estruturação matemática do modelo NICE

Considere x como os dados reais e $f(x)$ como uma transformação reversível que mapeia o espaço de dados de entrada para o espaço da variável latente z . Essa relação pode ser expressa como $z = f(x)$. Assim, a distribuição de z é denotada por $P_Z(z)$, é assumida como uma composição de componentes independentes que podem ser fatoradas, conforme a equação

$$P_Z(z) = \prod_{d \in D} P_{Z_d}(z_d), \quad (2.1)$$

onde $P_Z(z)$ representa a nova distribuição fatorizada no espaço latente, \prod é o operador de multiplicação sucessiva, d representa a dimensão dos dados de entrada, D é a di-

mensionalidade da variável latente e $P_{Z_d}(z_d)$ denota a distribuição da d -ésima variável latente.

Uma vez obtida a distribuição de probabilidade da variável latente z , é possível realizar a amostragem a partir dessa, utilizando o método da substituição. Nesse método, considera-se a mudança de variáveis $z = f(x)$, onde f é uma função invertível e z tem a mesma dimensão de x . Dessa forma é possível obter os novos dados, estabelecendo-se a distribuição de probabilidade $P_X(x)$ em relação à amostra original x em que

$$P_X(x) = P_Z(f(x)) \left| \det \frac{\partial f(x)}{\partial x} \right|, \quad (2.2)$$

onde $P_Z(f(x))$ representa a distribuição de probabilidade da variável latente z condicionada à função de transformação $f(x)$, e $\left| \det \frac{\partial f(x)}{\partial x} \right|$ corresponde ao determinante da matriz Jacobiana da função $f(x)$ em relação a x .

É importante ressaltar que, para delinear a estrutura do modelo NICE, a função f deve ser reversível, e o determinante da matriz Jacobiana de $f(x)$ deve ser trivialmente calculado (GE et al., 2020; HU et al., 2021), permitindo a amostragem de $P_X(x)$ da seguinte forma:

$$z \sim P_X(x); \quad (2.3)$$

$$x = f^{-1}(z). \quad (2.4)$$

Para maximizar a distribuição de probabilidade da amostra, visando otimizar a função de transformação $f(x)$, é aplicada a função logarítmica à Equação 2.2, resultando em

$$\log P_X(x) = \log P_Z(f(x)) + \log \left| \det \frac{\partial f(x)}{\partial x} \right|. \quad (2.5)$$

Desse modo, durante o treinamento do modelo NICE a função de transformação é otimizada para estimar a distribuição conjunta dos dados observados e dos componentes latentes. Isso é feito por meio da maximização da verossimilhança dos dados observados (ZHANG; ZHANG, 2020b). Após o treinamento, o modelo consegue gerar novas

amostras aplicando a função de transformação inversa a amostras de um espaço de características latentes.

Tipicamente, a função densidade de probabilidade $P_Z(z)$ é escolhida como uma distribuição normal multivariada ($\mathcal{N}(\mathbf{0}, \mathbf{I})$). Assim, na modelagem da distribuição $P_X(x)$ para a distribuição $P_Z(z)$, o mapeamento bijetivo desempenha um papel fundamental. No entanto, a transformação entre essas as distribuições é desconhecido (DINH; KRUEGER; BENGIO, 2014; ZHANG; ZHANG, 2020b). Dessa forma, é comum aplicar-se a função bijetiva parametrizada designada como $f_\theta(x)$.

A função $f_\theta(x)$ é definida de forma bijetiva, isto é, ela é tanto injetiva (cada valor de z é mapeado para um valor exclusivo de x , quanto sobrejetiva, todos os valores de x podem ser obtidos a partir de z através de $f_\theta(x)$). A propriedade bijetiva garante que a transformação seja reversível, e que seja possível calcular a probabilidade de densidade de P_X (ZHANG; ZHANG, 2020b).

Considerando tais aspectos e realizando-se uma mudança de variável, conforme mostrado nas Equações 2.2 e 2.5, pode-se relacionar a densidade de probabilidade modelada P_X com a distribuição de probabilidade P_Z por meio do mapeamento da função (f_θ) como apresentado em

$$P_X(x; \theta) = P_Z f_\theta(x) \left| \det \frac{\partial f_\theta(x)}{\partial x} \right|; \quad (2.6)$$

$$\log P_X(x; \theta) = \log P_Z f_\theta(x) + \log \left| \det \frac{\partial f_\theta(x)}{\partial x} \right|, \quad (2.7)$$

onde a distribuição modelada P_X pode ser considerada como uma função do parâmetro θ designada como a função de verossimilhança $L(\theta)$. O logaritmo de $L(\theta)$ é chamada de função log-verossimilhança. Dessa forma, ao fazer o uso do método de estimativa de máxima verossimilhança, pode-se treinar o modelo para escolher o valor apropriado de θ de modo que a verossimilhança seja maximizada como mostrado

$$\max_{\theta} \log P_X(x; \theta) = \log P_Z f_\theta(x) + \log \left| \det \frac{\partial f_\theta(x)}{\partial x} \right|. \quad (2.8)$$

A função bijetiva parametrizada f_θ é escolhida por ser uma composição de uma sequência de transformações reversíveis, isto é, $f_\theta = f_1 \circ f_2 \circ \dots \circ f_d$, de modo que o mapeamento de x para z , assim como o mapeamento inverso, podem ser representados da seguinte forma:

$$\mathbf{x} \xrightarrow{f_1} \mathbf{z}_1 \xrightarrow{f_2} \mathbf{z}_2 \cdots \mathbf{z}_{d-1} \xrightarrow{f_d} \mathbf{z}; \quad (2.9)$$

$$\mathbf{z} \xrightarrow{f_d^{-1}} \mathbf{z}_{d-1} \xrightarrow{f_{d-1}^{-1}} \mathbf{z}_{d-2} \cdots \mathbf{z}_1 \xrightarrow{f_1^{-1}} \mathbf{x}. \quad (2.10)$$

Essas sequências de transformação são denominadas fluxo de normalização, onde a variável z_i refere-se às representações intermediárias ou latentes, as quais são obtidas ao longo do processo de transformação após a aplicação da função de transformação f_i . Com base no projeto do fluxo de normalização, o logaritmo do determinante da matriz Jacobiana pode ser escrito, valendo-se da regra da cadeia, em que $z_0 = x$ e $h_k = z$:

$$\log \left| \det \left(\frac{\partial f_\theta(x)}{\partial x} \right) \right| = \sum_{i=1}^d \log \left| \det \left(\frac{\partial z_i}{\partial z_{i-1}} \right) \right|. \quad (2.11)$$

Para simplificar o cálculo do logaritmo do determinante do Jacobiano na Equação 2.11, cada transformação reversível f_i é implementada como uma camada de acoplamento. Essa abordagem permite que partes da entrada sejam passadas diretamente para a saída sem modificação.

As camadas de acoplamento dividem o vetor de entrada em duas partes, e aplicam transformações específicas a cada parte. Essas transformações podem ser aditivas ou multiplicativas, dependendo da variante do modelo. A ideia é preservar uma parte do vetor de entrada diretamente na saída, enquanto a outra parte é transformada de acordo com uma função específica.

Portanto, considerando uma camada de acoplamento aditiva, o vetor de dados observados x é dividido em duas partes distintas: x_1 e x_2 . A transformação é definida pelas equações:

$$y_1 = x_1; \quad (2.12)$$

$$y_2 = x_2 + m(x_1), \quad (2.13)$$

onde y_1 e y_2 são as variáveis chamadas de camada de acoplamento aditivo. A função m é uma função arbitrária escolhida pelo modelo, podendo ser qualquer função complexa, tal como uma rede neural. Este bloco de construção tem o determinante Jacobiano unitário para qualquer função m .

A reversibilidade da transformação das equações da camada de acoplamento pode ser expressa pela transformação inversa, permitindo recuperar o vetor de entrada original. Supondo que $x_1 = x_{1:d}$ sejam os primeiros (d) elementos de x e x_2 seja o elemento $(d + 1)$ a D de x . Dessa forma, a matriz Jacobiana (J) pode ser dada como uma transformação invertível usando a matriz de blocos conforme apresentada na equação:

$$\mathbf{J} = \begin{bmatrix} I & \frac{\partial x_1}{\partial y_2} \\ 0 & \frac{\partial x_2}{\partial y_2} \end{bmatrix}, \quad (2.14)$$

onde I é a matriz identidade de dimensão quadrada d ($I_{d \times d}$) e $\frac{\partial y_2}{\partial x_1}$ e $\frac{\partial y_2}{\partial x_2}$ são as matrizes jacobianas parciais das equações da camada de acoplamento em relação às variáveis x_1 e x_2 , respectivamente. A matriz Jacobiana dessa transformação será uma matriz triangular, onde todos os elementos diagonais são iguais a 1, o que implica que o determinante é igual a 1, tornando-a trivialmente invertível.

A reversibilidade da transformação dessas equações é dada em função inversa da camada de acoplamento, permitindo obter novamente o vetor de entrada original x a partir dos vetores de saída y_1 e y_2 , sendo essas mudanças definidas nas equações:

$$x_1 = y_1; \quad (2.15)$$

$$x_2 = y_2 - m(y_1). \quad (2.16)$$

Com as camadas de acoplamento, o modelo NICE torna as transformações reversíveis, simplificando o cálculo dos determinantes jacobianos e tornando o modelo mais eficiente para o cálculo do log-determinante da matriz jacobiana. Isso é essencial para o objetivo de treinamento do modelo.

Além disso, para melhorar a capacidade do modelo NICE de aprender a distribuição de probabilidade das amostras reais e gerar novos dados de amostra, é comum

adicionar ruídos aos dados de entrada durante o treinamento. Esses ruídos são amostras aleatórias seguindo uma distribuição normal, sendo utilizados como dados de entrada para o modelo NICE, aumentando a robustez do modelo, permitindo capturar melhor a variabilidade e complexidade dos dados de entrada (DINH; KRUEGER; BENGIO, 2014).

A composição das camadas de acoplamento ao longo do fluxo de normalização completa a modelagem matemática do modelo NICE. Cada camada de acoplamento recebe o vetor de saída da camada anterior e gera um novo vetor de saída, o qual é passado para a próxima camada. Esse processo se repete até que todas as camadas de acoplamento tenham sido aplicadas.

Portanto, a modelagem matemática do modelo NICE envolve a definição das camadas de acoplamento, das transformações reversíveis dentro de cada camada e a composição dessas camadas ao longo do fluxo de normalização. Essas escolhas podem variar conforme com o problema em questão e as necessidades do modelo.

3 METODOLOGIA PROPOSTA

Neste capítulo, descreve-se a metodologia proposta para gerar perfis de carga sintéticos, fundamentada no processo KDD, detalhada nas Seções 3.1 a 3.7. Em cada etapa, destacam-se as técnicas e abordagens específicas utilizadas para garantir a eficácia no processo. Embora fundamentada no processo KDD, cada etapa da metodologia é direcionada para o propósito específico de utilizar o modelo generativo NICE para atingir os objetivos almejados nesta tese.

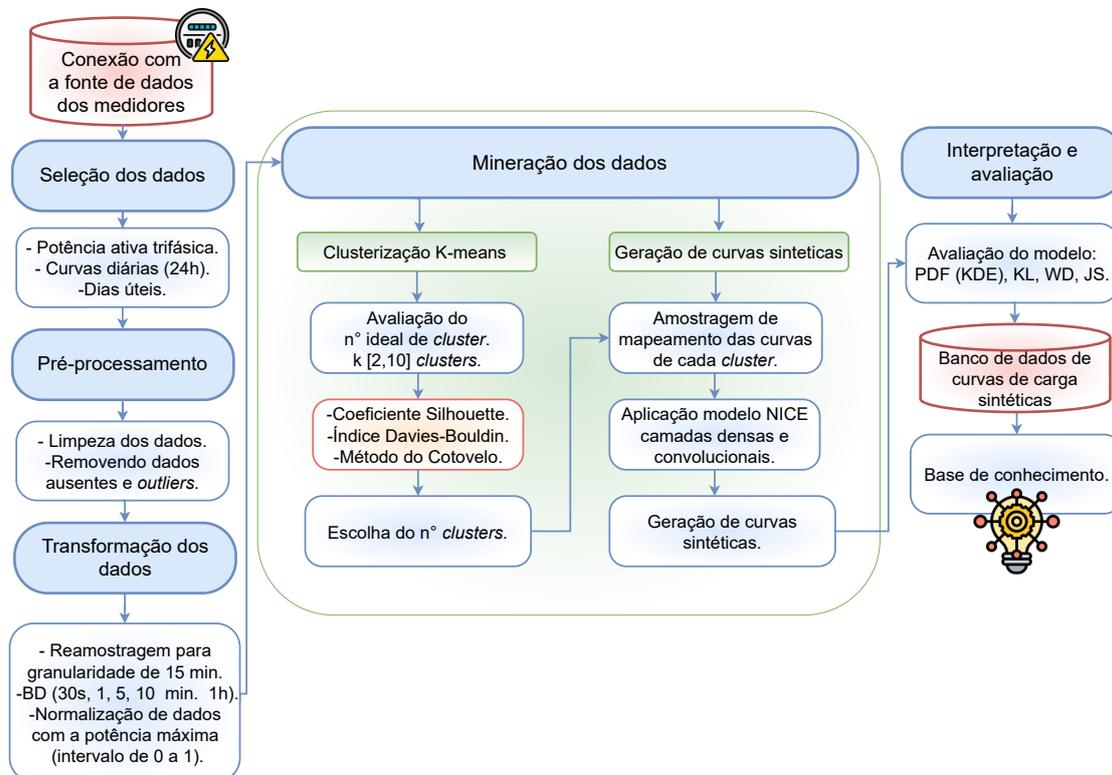
3.1 VISÃO GERAL DA METODOLOGIA

Esta seção fornece uma visão abrangente da metodologia proposta, que incorpora técnicas avançadas de análise de dados e a implementação do modelo generativo NICE. O foco principal é a geração de perfis sintéticos de curvas de carga a partir de dados medidos em uma rede elétrica. O panorama geral destaca a interconexão das principais etapas, incluindo a seleção de dados, o pré-processamento, a transformação dos dados, a clusterização, a aplicação do modelo generativo e a avaliação dos resultados. A Figura 5 fornece um fluxograma visual que esquematiza as etapas adotadas nesta metodologia.

Como pode-se observar a Figura 5 apresenta um panorama das etapas envolvidas, desde a obtenção dos dados brutos dos medidores digitais até a criação dos perfis sintéticos de curva de carga. Cada uma dessas etapas são discutidas em detalhes nas seções seguintes.

Essas etapas combinadas permitem gerar perfis sintéticos de curva de carga, que podem ser utilizados para análises e simulações em estudos de demanda de energia, operação e planejamento de redes elétricas, entre outros. É importante ressaltar que a metodologia proposta pode ser adaptada e refinada em conformidade com as necessidades e características específicas do contexto em que for aplicada.

Figura 5 – Visão geral da metodologia proposta.



Fonte: Elaborado pela autora (2023).

3.2 SELEÇÃO DE DADOS

A primeira etapa do processo na metodologia proposta é estabelecer a conexão com a fonte de dados dos medidores digitais. Os dados dos medidores são coletados de forma bruta e enviados para um banco de dados central. Nesta etapa, realiza-se a conexão com esse banco de dados, adquirindo assim os dados necessários para as etapas subsequentes da análise. A variável selecionada para análise é a potência ativa trifásica. Os dados coletados são representados em forma de curvas diárias, registrando a potência ativa trifásica a cada 24 horas.

Após a obtenção dos dados, estes são separados em dois grupos distintos: dias úteis e dias não úteis. Para fins de análise, para esta pesquisa apenas se selecionam os dias úteis. Essa seleção visa focar nas variações de consumo que ocorrem em dias de maior atividade e demanda energética, fornecendo uma visão mais precisa do comportamento

do consumo durante períodos típicos de uso.

Ao escolher exclusivamente os dias úteis, evitam-se distorções decorrentes de eventos atípicos, como feriados ou fins de semana, que podem exercer uma influência significativa nos padrões de consumo. Essa abordagem visa garantir a precisão e a confiabilidade dos perfis sintéticos de curva de carga, possibilitando uma representação mais fiel do comportamento de consumo durante os dias úteis, quando se observa uma ocupação e atividade mais intensas, resultando em uma demanda mais representativa por energia elétrica (WANG; HONG, 2020).

Almeja-se que a análise dos dados possibilite uma melhor compreensão em relação aos hábitos e as tendências de consumo energético durante os dias de maior atividade econômica e social. Essa seleção é fundamentada no fato de que o consumo de energia em edifícios é influenciado por diversos fatores, tais como: ocupação, clima, horário de funcionamento, atividades específicas, presença e uso de equipamentos e dispositivos, eficiência energética, políticas e práticas de conservação de energia, bem como características estruturais do edifício.

3.3 PRÉ-PROCESSAMENTO

A etapa de pré-processamento dos dados é caracterizada pelo processamento dos dados brutos selecionados na etapa anterior, submetendo-os a um conjunto de procedimentos a fim de prepará-los para serem utilizados nas etapas subsequentes. Esses procedimentos incluem a limpeza e tratamento de possíveis anomalias e irregularidades nos dados, garantindo que os mesmos estejam em um formato compatível com as técnicas de análise e aprendizado a serem utilizadas.

Dessa forma, ao iniciar a análise do conjunto de dados, deve-se identificar quais são as incidências proeminentes, tais como dados faltantes, incoerentes ou com ruídos de medição. Ao detectar a principal questão relacionada aos dados, é necessário decidir entre a exclusão ou a imputação de dados. No processo de imputação, uma abordagem específica é adotada, aplicando-a quando a porcentagem de dados ausentes é baixa, geralmente em torno de 10% a 15%.

Neste trabalho para a completude dos dados, utiliza-se à técnica de imputação

com base no valor médio das amostras em um intervalo de tempo específico, selecionado de acordo com a granularidade temporal dos dados analisados. Essa estratégia preserva a tendência temporal da curva de carga diária. Entretanto, é fundamental destacar que, caso a porcentagem de dados faltantes ultrapasse o limite tolerável estipulado, considera-se a exclusão do medidor associado a um alto percentual de falha na coleta dos dados.

Essa medida é implementada para assegurar a confiabilidade dos resultados, visto que uma proporção considerável de dados ausentes pode prejudicar a precisão e a validade da análise. O objetivo dessas ações é atenuar os efeitos dos dados incompletos, garantindo a integridade e a qualidade das informações obtidas. Vale ressaltar que a abordagem de completude dos dados deve ser adotada apenas quando o impacto nos resultados é mínimo e não compromete a integridade das análises disponíveis para o estudo.

Além disso, nesta etapa, é aplicado um método de detecção e remoção de valores atípicos com base na teoria do quartil, conforme descrito por Favero e Belfiore (2017), que define que os quartis representam uma medida estatística, dividindo o conjunto de dados em quatro partes iguais; caracterizando os percentis de 25% (primeiro quartil, Q_1), 50% (mediana, Q_2) e 75% (terceiro quartil, Q_3) dos dados.

De acordo com esse método, um ponto de dados é considerado atípico se estiver acima de 1,5 vezes o intervalo interquartil (IQR) abaixo do primeiro quartil (Q_1) ou acima do terceiro quartil (Q_3), como demonstrado em:

$$\begin{cases} x_i > Q_1 - 1,5 \cdot \text{IQR}, \\ x_i < Q_3 + 1,5 \cdot \text{IQR}. \end{cases} \quad (3.1)$$

A variável x_i representa o valor em análise e o IQR é calculado como a diferença entre o terceiro quartil e o primeiro quartil ($\text{IQR} = Q_3 - Q_1$). O primeiro quartil é o valor que divide a primeira metade do conjunto de dados, isto é, 25% dos dados são menores ou iguais a esse valor. O terceiro quartil é o valor que divide a segunda metade do conjunto de dados, indicando que 75% dos dados são menores ou iguais a esse valor.

Essa técnica estatística permite identificar e descartar valores que se afastam significativamente da distribuição típica dos dados, visando garantir a qualidade e a consistência dos perfis sintéticos de curva de carga a serem gerados. A remoção desses

valores atípicos contribui para evitar distorções indesejadas e assegurar que os modelos de aprendizado subsequentes sejam alimentados com dados mais representativos.

3.4 TRANSFORMAÇÃO DOS DADOS

Na etapa de transformação é realizada a normalização e a reamostragem dos dados, considerando diferentes intervalos de tempo. A normalização é efetuada para ajustar o perfil de curva de carga diária de cada um dos medidores analisados. Nesta tese, a variável utilizada para a normalização dos dados é a potência ativa trifásica máxima, conforme a seguinte relação:

$$W_{nom_{ij}} = \frac{W_{ij}}{\max(W_i)}, \quad (3.2)$$

em que $W_{nom_{ij}}$ representa o valor normalizado da potência ativa trifásica do medidor i no dia j , caracterizado por unidade (p.u), W_{ij} é a potência ativa trifásica diária em kW, $\max(W_i)$ representa o valor máximo da potência ativa trifásica em kW para o medidor i ao longo dos dias analisados.

O objetivo da normalização dos dados é garantir que os perfis de curva de carga de diferentes edifícios e com comportamentos semelhantes, apresentem curvas características com magnitudes comparáveis na faixa de 0 a 1. Desse modo, ao considerar a variação máxima de potência ao longo do tempo, para cada um das localidades monitoradas, torna-se possível a comparação entre os perfis de curvas obtidos, contribuindo para o aprimoramento e a eficiência dos modelos generativos.

No que diz respeito à reamostragem dos dados, é comum que a amostragem dos medidores digitais sejam registrados em intervalos de 30 segundos ou a cada 1 minuto. No entanto, com o propósito de obter uma representação mais abrangente das variações, procede-se à reamostragem desses dados. Essa abordagem pode ser conduzida considerando as diversas granularidades temporais, possibilitando a identificação de comportamentos e variações em intervalos de tempo mais substanciais e significativos.

Desse modo, leva-se em consideração a granularidade temporal dos dados de consumo de energia elétrica. A granularidade refere-se ao intervalo de tempo no qual os

dados são registrados e representa uma escolha relevante para a análise de padrões de consumo e geração de perfis de curva de carga sintética.

Inicialmente, optou-se por empregar uma granularidade de 15 minutos na reamostragem dos dados. Essa escolha possibilita uma representação mais realista do consumo de energia dos edifícios, capturando variações mais detalhadas ao longo do dia. Essa granularidade é especialmente pertinente para técnicas que utilizam perfis de curva de carga sintética, como em aplicações que realizem o cálculo de fluxo de potência em sistemas de distribuição e estimação de estados (PRIMADIANTO; LU, 2016).

Nesse contexto, tem-se que as diferentes granularidades temporais podem ser adequadas para aplicações específicas na geração de perfis de curva de carga sintética. Por exemplo, uma granularidade menor como 30 segundos ou 1 minuto pode ser mais apropriada para análises em tempo real, em que há necessidade de capturar variações mais detalhadas do consumo de energia. Por outro lado, granularidades maiores, como 10 minutos ou 15 minutos, podem ser mais adequadas para análises de maior prazo que visem a identificação de tendências de consumo ao longo do dia.

A análise de granularidade temporal não apenas permite uma compreensão mais detalhada dos padrões de consumo em diferentes escalas temporais, mas também é uma abordagem para lidar com os grandes volumes de dados. À medida que se opta por análises com granularidades maiores, como reamostragem a cada 10, 15, 30 minutos ou 1 hora, estima-se que haverá uma redução significativa no volume de dados processados. Isso se apresenta como uma estratégia viável na otimização de recursos computacionais, simplificando o processo de análise. Entretanto, é fundamental avaliar criteriosamente os resultados obtidos de modo a garantir que não ocorram perdas significativas de informações no processo de agregação.

3.5 CLUSTERIZAÇÃO

A clusterização é uma técnica amplamente utilizada em diversas áreas para identificar os grupos naturais nos dados, agrupando os objetos semelhantes tomando por base as características e propriedades (RAMOS et al., 2012; HAN; PEI; TONG, 2022). No que se referem aos perfis de curvas de carga, a clusterização possibilita lidar com a variabilidade dos dados, permitindo reduzir a quantidade considerável de perfis de carga

em conjuntos menores que compartilham características semelhantes (KONTOGIANNIS et al., 2022; MICHALAKOPOULOS et al., 2023).

Apesar da ampla variedade de algoritmos de agrupamento disponíveis, não existe um algoritmo único capaz de identificar todas as formas de agrupamentos e estruturas presentes nos dados. Agrupamentos eficazes são caracterizados pela alta similaridade entre os objetos dentro do mesmo grupo e, ao mesmo tempo, por uma baixa ou significativa diferença de similaridade entre objetos pertencentes a diferentes classes. Portanto, ao lidar com um conjunto de dados específico, surgem questões fundamentais relacionadas ao processo de agrupamento, como a escolha do método mais adequado para gerar a partição ideal dos dados e a determinação do número apropriado de *clusters* a serem aplicados (RAMOS; VALE, 2008; RAMOS et al., 2012; RAMOS et al., 2015b; MICHALAKOPOULOS et al., 2023).

Um dos algoritmos de agrupamentos mais populares é o *K-means*, empregado em 1967 por James MacQueen (MACQUEEN et al., 1967), amplamente difundido na literatura para solução de problemas de agrupamento de dados em diferentes domínios da ciência (IKOTUN et al., 2023), devido à simplicidade de implementação e baixa complexidade computacional e eficácia (JAIN, 2010; EZUGWU, 2020; AHMED; SERAJ; ISLAM, 2020; IKOTUN et al., 2023).

Embora, o método de agrupamento *K-means* seja um dos mais populares, este algoritmo ainda possui alguns desafios que precisam ser superados, tais como: limitações de escalabilidade, sensibilidade à seleção inicial de centroides e tratamento de valores atípicos. Além disso, requer conhecimento prévio do número desejado de *clusters* e tende a convergir para soluções locais em vez da ótima (EZUGWU, 2020).

Apesar dessas limitações, o algoritmo de agrupamento *K-means* ainda é muito empregado para resolver problemas de agrupamento, devido às suas particularidades em relação aos demais modelos de clusterização presentes na literatura (IKOTUN et al., 2023). Matematicamente, o agrupamento *K-means* é representado por:

$$\arg \min_S \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2, \quad (3.3)$$

em que o argumento mínimo ($\arg \min_S$) do conjunto S , k representa o número esperado

de *clusters* produzidos pelo processo de clusterização. Os membros do conjunto de dados são representados como x_1 a x_n , e os conjuntos *clusters* são representados como S_1 a S_n . Sendo $k \leq n$, μ_i representa o centroide do *cluster* S_i , que é a média dos membros de S_i .

O objetivo do *K-means* é agrupar os dados de forma que as amostras sejam divididas em grupos de igual variância, minimizando a distância entre a soma dos erros quadrados de todos os agrupamentos k (MACQUEEN et al., 1967; JAIN, 2010). No entanto, escolher o número ideal de *clusters* é um desafio (EZUGWU, 2020). Para resolver essa questão, medidas de avaliação do número de *cluster* são adotadas.

Para abordar a definição do número ideal de *clusters* será utilizadas medidas de eficácia de agrupamento, tal como o coeficiente de silhueta (do inglês, *silhouette coefficient (SC)*), o índice *Davies-Bouldin* (do inglês, *davies–bouldin index (DBI)*) e o método do cotovelo (do inglês, *elbow method*). Nesta pesquisa, realiza-se a análise k variando na faixa de 2 a 10 *clusters* e calculando essas medidas para cada uma possibilidades analisadas.

Essa abordagem permite estimar o impacto de diferentes números de *clusters* de acordo com esses indicadores. Através da análise comparativa dessas medidas, pode-se selecionar o valor de k mais adequado às características intrínsecas do conjunto de dados analisados. Ressalta-se, que essa estimação de k também deve estar alinhada com o conhecimento prévio sobre o contexto da aplicação.

Além disso, ressalta-se que a análise de clusterização não se restringirá ao uso exclusivo do algoritmo *K-means*. Outros métodos de clusterização são testados e avaliados, visando à identificação da abordagem mais adequada para o cenário em questão. Essa abordagem diversificada permitirá uma análise mais abrangente e a seleção do método de clusterização mais apropriado para as características intrínsecas dos dados sob investigação.

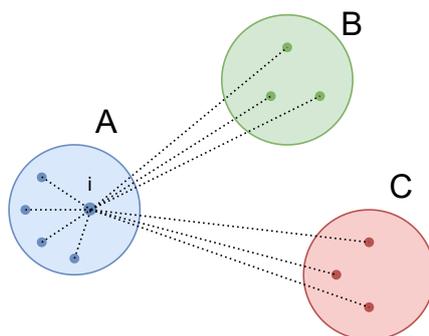
3.5.1 Coeficiente de Silhueta

O coeficiente de silhueta é uma medida de validação da clusterização, relacionando o grau de adequação de cada objeto ao *cluster* em comparação com os demais. Matematicamente, esse coeficiente é analisado conforme o estudo proposto Rousseeuw (1987), elucidando que para a análise da silhueta, é necessário analisar as partições

obtidas da técnica de agrupamento e as proximidades dos objetos.

Para calcular o SC, primeiramente, deve-se considerar que o número de *clusters*, denotado como k , é maior que 1. Assim, para cada objeto i no conjunto de dados, determina-se um valor específico $S(i)$. Suponha que ao escolher aleatoriamente um objeto i do conjunto de dados, e esse objeto pertença ao *cluster* denominado A , como ilustrado na Figura 6.

Figura 6 – Elementos envolvidos nos cálculos de $S(i)$ objeto i pertence a A .



Fonte: Elaborado pela autora (2023).

A Figura 6 fornece uma representação ilustrativa desse conjunto, exemplificando a média das distâncias entre o objeto i e os outros objetos no *cluster* A . Dessa forma, para cada objeto no *cluster* A , calcula-se $A(i)$, a média da dissimilaridade de i em relação a todos os outros objetos em A .

Em seguida, para qualquer *cluster* diferente de A , por exemplo o *cluster* C , calcula-se a distância $d(i,C)$, que representa a média das distâncias entre o objeto i e os objetos contidos em C . Essa consideração é aplicada para todos os objetos do conjuntos $C \neq A$, de modo a identificar o menor desses valores, usando a seguinte relação matemática:

$$B(i) = \min(d(i,C)). \quad (3.4)$$

O *cluster* B , que corresponde a esse valor mínimo, é chamado de *cluster* de vizinhança do objeto i e representa a segunda melhor opção de agrupamento para o

objeto. Portanto, a alocação do objeto i aos respectivos *clusters* é determinada pela seguinte equação:

$$S(i) = \begin{cases} \frac{1-A(i)}{B(i)}, & \text{se } A(i) < B(i), \\ 0, & \text{se } A(i) = B(i), \\ \frac{B(i)}{A(i)-1}, & \text{se } A(i) > B(i). \end{cases} \quad (3.5)$$

Nessa relação matemática, $S(i)$ representa o SC para o objeto, $A(i)$ é a média da dissimilaridade de i em relação a todos os outros objetos no *cluster* A , e $B(i)$ é o valor mínimo da média da dissimilaridade de i em relação aos objetos de outros *clusters* C .

O coeficiente silhueta $S(i)$ pode variar de

$$-1 \leq S(i) \leq 1. \quad (3.6)$$

Para cada objeto i , os valores próximos a 1 indicam uma alocação dos objetos de forma mais concisa aos *clusters*. Já os valores próximos a -1 sugerem uma possível atribuição incorreta aos *clusters*, e os valores próximos a 0 indicam que o objeto está entre os *clusters*.

3.5.2 Índice de *Davies-Bouldin*

O DBI é uma estimativa para avaliar a similaridade *intracluster* e as diferenças entre *clusters*, cuja separação entre *clusters* é calculada com base no *medóide* (KARO; MAULANAADHINUGRAHA; HUDA, 2017). Matematicamente, o índice é representado por uma função que visa minimizar a variância *intracluster* e maximizar a distância entre *clusters* (ROS; RIAD; GUILLAUME, 2023). O índice DBI é definido de acordo com Davies e Bouldin (1979), que de forma simplificada, conforme (ROS; RIAD; GUILLAUME, 2023) é dada por

$$DB(k) = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left(\frac{S_i + S_j}{d(\bar{x}_i, \bar{x}_j)} \right), \quad (3.7)$$

em que $DB(k)$ é DBI para o número de *cluster* k analisado, os índices i e j representam os diferentes conjuntos, (S_i) representa a distância média entre o centro do *cluster* e todos os seus elementos e $d(\bar{x}_i, \bar{x}_j)$ é a distância média entre os elementos.

Desse modo, tem-se que o DBI é calculado como a média da razão máxima entre a distância dentro do *cluster* e a distância entre *clusters* para cada conjunto. Quanto menor o DBI, melhor é a qualidade do *cluster* do conjunto de dados analisados.

3.5.3 Método do Cotovelo

O método do cotovelo é uma abordagem heurística que auxilia na análise da consistência de agrupamentos em conjuntos de dados, cujo objetivo é determinar o número mais apropriado de *clusters* para o conjunto de dados analisados ao se considerar diferentes números de agrupamento (BAHRI et al., 2023).

Para cada valor de k proposto é calculado a soma quadrada dentro do *cluster* (do inglês, *Within-Cluster Sum of Square*) (WCSS), realizando-se a soma quadrada das distâncias entre os pontos e o centroides do *cluster*, conforme apresentado em (WEN et al., 2023). Matematicamente, a função custo WCSS é expressa por:

$$WCSS = \sum_{i=1}^k \sum_{x \in S_i} d(g_i, x)^2, \quad (3.8)$$

em que d é a distância euclidiana entre dois pontos, k é o número de *clusters*, x é o ponto de amostra, S_i é o conjunto de dados no i -ésimo *cluster* e g_i é o i -ésimo centro do *cluster*.

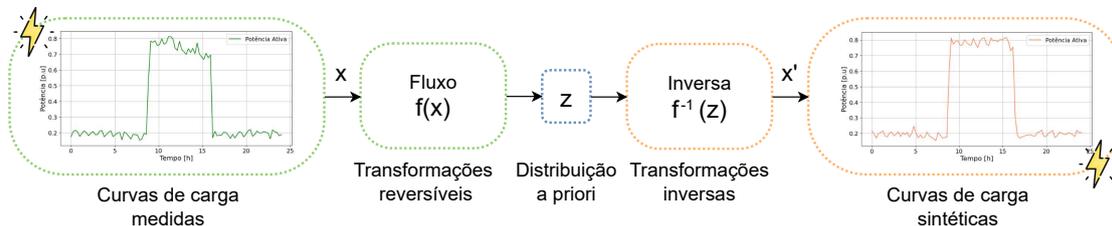
Ao calcular o WCSS em relação a k , observa-se que, à medida que o número de *clusters* aumenta, o valor do WCSS inicialmente diminui, como apresentado em (MARUTHO et al., 2018). Embora o WCSS seja mais alto quando k é igual a 1, a tendência é que este valor diminui-se ao aumentar gradualmente o valor de k e, dessa forma, a curva assume o formato característico de um cotovelo. A partir desse ponto, o gráfico permanece quase paralelo ao eixo x , e o valor de k associado a esse ponto representa o número ideal de *clusters*. Após esse ponto específico, a curva não demonstra uma redução na função de custo tão acentuada.

3.6 GERAÇÃO DE PERFIL DE CURVA DE CARGA SINTÉTICA

A geração de perfis sintéticos de curva de carga, que possuam características estatísticas semelhantes aos perfis de cargas medidas, representam um desafio significativo no campo da análise de dados. Para abordar essa questão, nesta tese é proposta a utilização do modelo NICE. A escolha desse modelo é motivada pela sua eficiência na modelagem de funções de densidade complexas e de alta dimensão contidas nos conjuntos de dados.

Conforme discutido na Seção 2.2, o modelo NICE é composto por uma sequência de transformações reversíveis aplicadas a um vetor de entrada, garantindo a preservação da reversibilidade dos dados. Essa abordagem se mostra adequada para a geração de perfis sintéticos de curva de carga, uma vez que consegue capturar as características estatísticas presentes nos perfis de cargas medidas. Ao modelar o fluxo de normalização, o modelo NICE estima a distribuição dos dados de entrada e gera amostras sintéticas que apresentam similaridade estatística com os dados medidos (GE et al., 2020). A Figura 7 ilustra a estrutura do modelo NICE considerada neste estudo.

Figura 7 – Estrutura do modelo NICE.



Fonte: Elaborado pela autora (2023).

Na Figura 7, x representa as curvas de carga medidas, provenientes dos medidores digitais, enquanto o Fluxo ($f(x)$) refere-se à sequência de transformações reversíveis aplicadas a esses dados. A variável z representa as variáveis latentes da distribuição a priori. Por meio das transformações inversas ($f^{-1}(z)$), os dados são reconstruídos a partir das variáveis latentes, gerando as curvas de carga sintéticas representadas por x' .

Desse modo, considerando a modelagem matemática do modelo NICE apresentada na Subseção 2.2.1, na etapa de fluxo de normalização é necessário designar a função

complexa m , responsável por transformar os dados de entrada em um vetor de saída.

Nesta tese, propõe-se a incorporação de redes neurais com camadas convolucionais como função complexa m no modelo NICE. As redes neurais convolucionais são conhecidas pela capacidade de aprender representações hierárquicas dos dados, especialmente, em tarefas relacionadas a imagens e sinais unidimensionais (KRIZHEVSKY; SUTSKEVER; HINTON, 2012; KRIZHEVSKY; SUTSKEVER; HINTON, 2017), como os perfis de curva de carga diários.

Estima-se que ao utilizar redes neurais com camadas convolucionais na função m , o modelo NICE possa capturar as relações espaciais e temporais presentes nos dados dos perfis de curva carga, permitindo uma modelagem mais precisa e eficiente. Além disso, as transformações reversíveis aplicadas pela função garantem que o modelo NICE preserve a informação necessária para reconstruir os dados originais durante a etapa de geração de amostras sintéticas.

Essa abordagem representa uma inovação em relação aos métodos anteriores, que predominantemente utilizavam redes neurais com camadas densas para essa aplicação, tal como apresentado nos trabalhos de Ge et al. (2020) e Hu et al. (2021). A inserção de redes neurais convolucionais no modelo NICE tem como intuito a ampliação do potencial da modelagem, permitindo uma representação mais adequada dos perfis de carga diários ao longo do tempo.

Para validar a eficácia do método proposto, NICE com redes neurais camadas convolucionais, que neste trabalho é denominado como modelo NICE convolucional, realiza-se a avaliação comparativa da estimativa da função de densidade de probabilidade entre os perfis de curva de carga medidos e os gerados pelo modelo. O mesmo é aplicado para o modelo NICE com redes neurais camadas densas, denominado neste trabalho como NICE denso.

3.7 AVALIAÇÃO DO MODELO GENERATIVO

A avaliação da qualidade dos perfis de curva de carga sintéticos gerados pelo modelo NICE é etapa fundamental para garantir a confiabilidade das curvas obtidas. Para avaliar a similaridade entre os perfis de carga diários gerados e as amostras medidas,

é utilizada a análise da estimativa da função de densidade de probabilidade, fazendo o uso de diferentes indicadores. Entre as medidas propostas, destacam-se a divergência Kullback-Leibler (KL), a distância de Wasserstein (do inglês, *Wasserstein distance (WD)*) e a divergência Jensen-Shannon (JS). Essas medidas serão detalhadas em subseções subsequentes.

No contexto abordado nesta tese, a análise da função de densidade de probabilidade (do inglês, *probability density function (PDF)*) desempenha um papel significativo, pois essa analogia permite a comparação da estimativa das PDFs de curvas reais e sintéticas. Isso possibilita avaliar aspectos específicos para cada objetivo a ser considerado. Por exemplo, a divergência KL é indicada para avaliar a assimetria na divergência entre as distribuições, enfatizando as diferenças em uma direção particular na comparação.

Já a WD é particularmente relevante quando as distribuições apresentam topologias complexas, medindo o custo de transporte entre elas. E por fim, a divergência JS propicia uma avaliação equilibrada da similaridade entre as distribuições, sendo uma medida simétrica, que propicia uma análise da média entre as distribuições analisadas.

Através da avaliação das medidas consideradas, é possível analisar e comparar o desempenho dos modelos NICEs, propostos para a geração de perfis sintéticos de curva de carga. Essas abordagens proporcionam uma medida para verificar a capacidade dos modelos em capturar as características essenciais dos dados medidos. Com base nos resultados obtidos, é possível ajustar e aprimorar os modelos generativos, garantindo a qualidade dos perfis sintéticos gerados.

3.7.1 Função de densidade de probabilidade

Uma das medidas de avaliação do modelo NICE para comparação das curvas medidas e as curvas sintéticas é a estimativa da função de densidade de probabilidade, que é uma função que descreve a forma da curva de probabilidade de ocorrência de cada valor da variável aleatória contínua (FAVERO; BELFIORE, 2017). Essa estimativa é fundamental para compreender a distribuição dos dados, tanto dos medidos quanto dos gerados. No presente contexto, para estimativa da PDF, utilizou-se a estimativa de densidade de *Kernel* (do inglês, *Kernel density estimation (KDE)*).

O KDE é uma técnica estatística não paramétrica que permite inferir estimar

a função de densidade de probabilidade de uma variável aleatória. O diferencial do KDE está na capacidade de prescindir de suposições prévias acerca da forma funcional da distribuição subjacente, baseando-se na distribuição empírica dos próprios dados amostrais (HE et al., 2016; ZHANG; HE; YANG, 2021).

Dessa forma, a estimativa da PDF é adaptada aos dados específicos em análise, por meio de uma suavização intrínseca, gerando uma representação contínua e adaptável (MERINO; FARMER; JACOBS, 2023). Tal característica possibilita a análise comparativa entre os dados medidos e os dados gerados pelo modelo. A representação matemática do KDE segue os modelos descritos em (LIAO et al., 2015; ABBASI, 2022), sendo:

$$P(x) = \frac{1}{nh} \sum_{i=1}^n H\left(\frac{x-x_i}{h}\right), \quad (3.9)$$

em que $P(x)$ é a estimativa da PDF em x , n é o número total de amostras, h é a largura de banda de suavização (largura da janela), x_i são as observações no conjunto de dados e $H(u)$ é a função *Kernel* que suaviza os dados. Nesta aplicação, a função *Kernel* Gaussiano é usada para estimar a distribuição marginal de variáveis aleatórias, sendo descrito da seguinte forma:

$$H(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}. \quad (3.10)$$

O *Kernel* Gaussiano usa a variável u para estimar a probabilidade de que um ponto específico pertença a uma região em torno de um ponto de referência, e $H(u)$ é a função de densidade de probabilidade estimada para esse ponto.

3.7.2 Divergência de Kullback-Leibler (KL)

Uma medida comumente utilizada para a avaliação de curvas de cargas sintéticas é a divergência de KL, também conhecida como entropia relativa. Tal medida visa mensurar o quão a distribuição de probabilidade diverge de uma segunda distribuição. Isto é, mede a diferença entre a estimativa da distribuição de probabilidade dos dados medidos, representa por $P(x)$, e a estimativa da distribuição de probabilidade gerada pelo

modelo NICE, representada por $Q(x)$ (KULLBACK, 1997). Neste contexto, a divergência KL pode ser calculada de duas maneiras: $D_{KL}(P||Q)$ e $D_{KL}(Q||P)$. Matematicamente, tem-se

$$D_{KL}(P||Q) = \sum_{x \in X} P(x) \log \left(\frac{P(x)}{Q(x)} \right), \quad (3.11)$$

em que $P(x)$ representa a estimativa da distribuição de probabilidade dos dados medidos, enquanto $Q(x)$ representa a estimativa da distribuição de probabilidade dos dados gerados pelo modelo NICE. Ambas as distribuições estão definidas no mesmo espaço de probabilidade x .

A divergência KL é uma soma ponderada da diferença logarítmica das estimativas de probabilidades $P(x)$ e $Q(x)$, em que $D_{KL}(P||Q)$ mede o quão diferente é a distribuição $P(x)$ em relação a $Q(x)$, quantificando a informação de quanto $P(x)$ e $Q(x)$ diferem nesse resultado específico.

Já a medida $D_{KL}(Q||P)$ mensura o quanto a estimativa da distribuição $Q(x)$ difere da estimativa da distribuição $P(x)$ e quantifica a informação perdida quando $P(x)$ é utilizada para aproximar $Q(x)$. É importante ressaltar que a divergência de KL não é simétrica, e as fórmulas podem produzir resultados diferentes. A escolha entre as duas formas de cálculo depende do contexto do problema e das perspectivas objetivadas.

Ao aplicar a divergência de KL para avaliar os perfis sintéticos de curva de carga gerados pelo modelo NICE, busca-se um valor de $D_{KL}(P||Q)$ ou $D_{KL}(Q||P)$ o mais próximo possível de zero, em que um valor mais baixo indica uma maior semelhança entre as estimativas de distribuições de probabilidade, evidenciando uma qualidade superior na geração dos perfis de curva de cargas sintéticas.

3.7.3 Divergência Jensen-Shannon (JS)

A divergência *Jensen-Shannon* JS é uma medida estatística usada para avaliar a diferença entre duas estimativas de distribuições de probabilidade, sendo uma extensão da divergência KL (LIN, 1991). O intuito da aplicação da divergência JS é determinar o grau de similaridade entre a estimativa de duas distribuições de probabilidade, neste trabalho denotadas por $P(x)$ e $Q(x)$.

A divergência JS é definida como uma medida que combina ponderadamente as divergências KL entre duas distribuições, e a média ponderada dessas distribuições, atribuindo pesos iguais as distribuições, ocasionando que a medida seja simétrica (ENDRES; SCHINDELIN, 2003). A representação matemática da divergência JS é descrita em (LIN, 1991; ENDRES; SCHINDELIN, 2003) e simplificada de acordo com (GU et al., 2019), sendo:

$$JS(P||Q) = \frac{1}{2} \sum_{x \in X} \left(P(x) \log \left(\frac{2P(x)}{P(x) + Q(x)} \right) + Q(x) \log \left(\frac{2Q(x)}{P(x) + Q(x)} \right) \right). \quad (3.12)$$

Na fórmula acima, $P(x)$ e $Q(x)$, representa a estimativa de distribuição de probabilidade da carga medida e gerada, respectivamente, em x em todo o intervalo X . De maneira semelhante, na avaliação da divergência KL e nos resultados de JS, espera-se que os valores estejam contidos no intervalo $[0, 1]$. Quanto mais próximo de zero, maior a similaridade entre as distribuições de probabilidade, e quanto mais próximo de um, maior é a dissimilaridade entre as distribuições.

3.7.4 Distância *Wasserstein* (WD)

A distância de *Wasserstein*, também conhecida como *Earth Mover's Distance* (EMD), é uma medida utilizada para quantificar a dissimilaridade entre a estimativa de duas distribuições de probabilidade contínuas (YANG; ZHENG; ZHANG, 2022). A WD pode ser calculada tal como apresentado em (ARJOVSKY; CHINTALA; BOTTOU, 2017; PANARETOS; ZEMEL, 2019), estabelecendo que

$$WD(P, Q) = \inf_{j \sim \Pi(P, Q)} \mathbb{E}_{(x, y) \sim j} \|x - y\|, \quad (3.13)$$

em que, de modo análogo as outras medidas consideradas, $P(x)$ e $Q(x)$ representam as distribuições de probabilidade e $\Pi(P, Q)$ é o conjunto (j) de todas as distribuições conjuntas e marginais de P e Q . \mathbb{E} representa a esperança das distâncias entre os pares de variáveis aleatórias, x e y , extraídas de distribuições P e Q , respectivamente.

Portanto, a formulação de WD calcula a distância esperada entre os pares de variáveis aleatórias extraídas de distribuições conjuntas j , em que o intuito é encontrar a

menor distância sobre todas as distribuições conjuntas possíveis. Isto é, o valor mínimo da média das distâncias entre x e y . Dessa forma, para avaliação dessa medida, os valores menores indicam maior similaridade, enquanto valores maiores sugerem maior dissimilaridade (PANARETOS; ZEMEL, 2019).

4 SIMULAÇÃO E RESULTADOS

Neste capítulo, são apresentados os resultados obtidos por meio das simulações realizadas para avaliar o desempenho da pesquisa proposta. As simulações foram conduzidas de acordo com a metodologia discutida no Capítulo 3. Para a implementação dos modelos desenvolvidos, utilizou-se a linguagem de programação *Python* e a plataforma *Google Colab Pro*.

Inicialmente, na Seção 4.1, realiza-se a contextualização do ambiente de simulação, fornecendo uma visão geral do local de estudo e da importância para as simulações propostas. Em seguida, na Seção 4.2, os resultados do processamento e tratamento dos dados são discutidos, abrangendo as etapas de limpeza, transformação e normalização dos dados.

Na sequência, na Seção 4.3, são apresentados os resultados da técnica de clusterização e do modelo generativo; nesta etapa, o desempenho do modelo NICE é avaliado. Na Seção 4.4, descreve-se a arquitetura implementada dos modelos NICE denso e convolucional. Na Seção 4.5, apresentam-se os resultados dos perfis sintéticos de curva gerados por cada modelo e as medidas de avaliação.

Já na Seção 4.6, é realizada uma análise dos resultados considerando a variação da granularidade temporal, explorando os diferentes intervalos de tempo propostos para geração de perfil sintético de curva de carga. Na Seção 4.7, são apresentados os testes adicionais realizados para validar a metodologia proposta nesta tese, considerando a variação do modelo NICE denso. Por fim, na Seção 4.8, são discutidas as análises abrangentes realizadas ao longo deste capítulo.

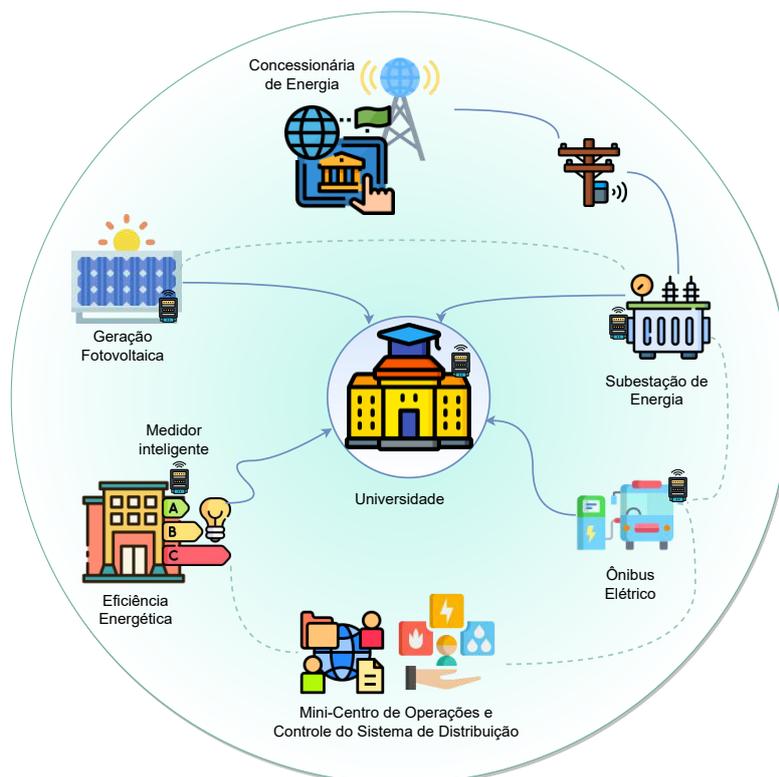
4.1 CONTEXTUALIZAÇÃO DO AMBIENTE DE SIMULAÇÃO

Esta seção oferece uma breve apresentação do ambiente de simulação, destacando elementos fundamentais para as simulações propostas. A compreensão detalhada deste contexto é essencial para a avaliação dos resultados, pois ele atua como cenário fundamental para a coleta de dados e aplicação da metodologia proposta.

Os dados utilizados incluem perfis de carga medidos por medidores digitais nos

alimentadores de distribuição da Universidade Estadual de Campinas (UNICAMP), parte da infraestrutura avançada de medição implementada pelo Projeto Campus Sustentável. Esse projeto visa tornar a instituição mais eficiente e sustentável energeticamente, conforme ilustrado na Figura 8.

Figura 8 – Diagrama simplificado do Laboratório Vivo UNICAMP.



Fonte: Elaborado pela autora (2023).

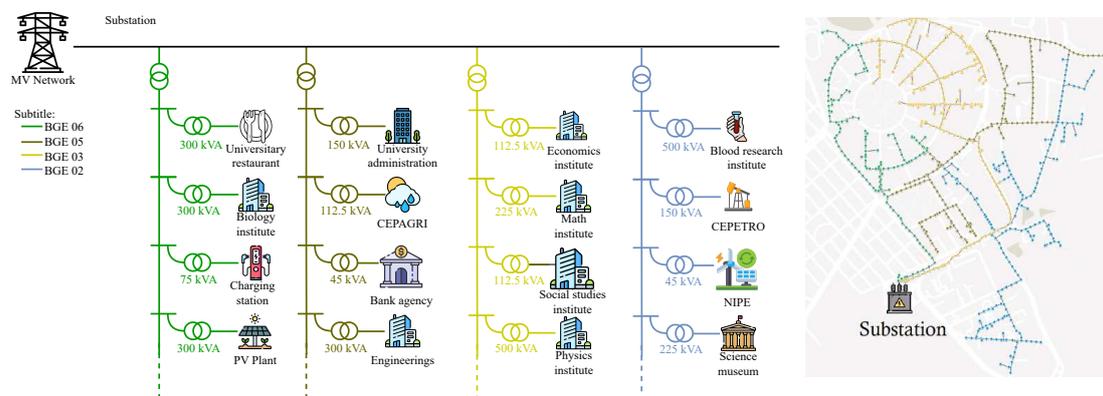
O Projeto Campus Sustentável, representado na Figura 8, engloba diferentes áreas, incluindo geração de energia renovável, mobilidade elétrica, eficiência energética, monitoramento em tempo real das redes elétricas e gerenciamento do consumo de energia com o propósito de aprimorar a infraestrutura do campus, conforme destacado por Silva *et al.* (SILVA, 2021).

O sistema de distribuição da universidade é composto por 26 km de redes de média tensão, tanto aéreas quanto subterrâneas, distribuídas em cinco alimentadores radiais de 11,9 kV nomeados como BGE02, BGE03, BGE04, BGE05 e BGE06. Esses

alimentadores possuem um total de 284 transformadores de serviço (11,9 kV/220 V). O sistema inclui mais de 3.000 pontos de iluminação pública e 30 km de redes de baixa tensão (220 V/127 V), tanto aéreas quanto subterrâneas (UNICAMP, 2023).

Destaca-se, ainda, a presença de cargas especiais no sistema, como um ônibus elétrico e sistemas fotovoltaicos (SFV). A Figura 9 ilustra a subestação de energia de alta tensão para média tensão (HV/MV) da UNICAMP, a topologia dos alimentadores e algumas cargas representativas em cada um deles.

Figura 9 – Diagrama simplificado da rede de distribuição da UNICAMP.



Fonte: Silva et al. (2023).

O fornecimento de energia nas instalações é realizado por diferentes alimentadores. O alimentador BGE02 tem como principal função suprir energia aos laboratórios de pesquisa. Já o alimentador BGE03 é responsável por abastecer a maioria das salas de aula e prédios administrativos. O BGE05, por sua vez, provê energia para os prédios administrativos, agências bancárias e edifícios relacionados a diferentes áreas da engenharia, como salas de aula e unidades administrativas.

O alimentador BGE06 concentra a maioria das cargas especiais, tais como uma usina fotovoltaica de 364 kWp e o posto de carregamento do ônibus elétrico. O alimentador BGE04 é exclusivamente dedicado ao fornecimento de energia para o hospital universitário. Essa carga especial não foi considerada neste estudo e, portanto, não está representada na Figura 9.

No sistema atual de distribuição de energia da UNICAMP, o monitoramento é

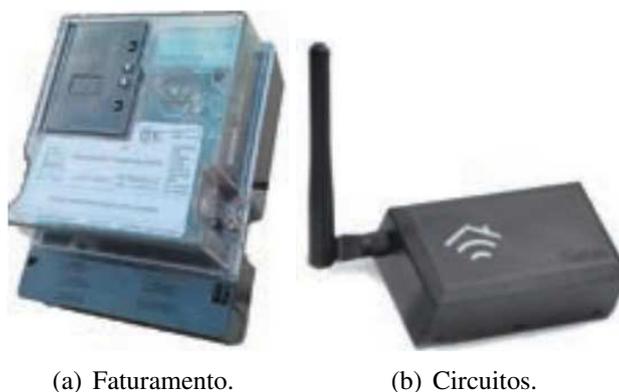
feito por meio de medidores digitais instalados na rede de baixa tensão, abrangendo todas as unidades consumidoras do *campus* universitário. Esses medidores coletam dados de consumo e operação das redes elétricas, sendo uma parceria com a Companhia Paulista de Força e Luz (CPFL Energia), como parte do projeto *Campus Sustentável* (SILVA, 2021). O projeto busca estabelecer um modelo de gestão e eficiência energética, podendo ser replicado em outras instituições de ensino superior no país. Para isso, são utilizadas tecnologias avançadas para acompanhar e gerenciar o consumo de energia eficientemente (CYPRIANO et al., 2021).

Os medidores digitais coletam dados elétricos, como consumo, corrente e tensão elétrica na rede de baixa tensão dos transformadores da Cidade Universitária Zeferino Vaz. Esses dados são analisados para implementar ações de eficiência energética. Além disso, os medidores permitem o faturamento remoto das unidades consumidoras e a gestão da geração de energia fotovoltaica. De acordo com Silva (2021), para realizar a medição das unidades consumidoras e permissionárias, foram utilizados 351 medidores, os quais estão divididos da seguinte forma:

- Medidores diretos: 61;
- Medidores indiretos com transformador de corrente: 182;
- Medidores indiretos com bobinas *rogowski*: 76;
- Medidores da rede de média tensão (alimentadores): 05;
- Medidores AC das plantas de geração fotovoltaica: 23;
- Medidores DC para geração fotovoltaica: 04;
- Medidores de circuito: 09.

Para a medição das unidades consumidoras (conjunto de transformadores) e permissionárias, utilizam-se os medidor digital de faturamento (ilustrado na Figura 10 (a)). Já para as medições das usinas fotovoltaicas, circuitos internos das edificações, e aplicações de medição e verificação, são utilizados medidores digitais de circuitos (CYPRIANO et al., 2021), Figura 10 (b).

Figura 10 – Medidores digitais.



Fonte: Cypriano et al. (2021), pg. 96.

Os dados coletados pelos medidores são transmitidos por três tecnologias distintas: *Wi-Fi*, *RF-Mesh* e *LoRaWan* (*Long Range Wan*). A rede *Wi-Fi* utiliza um canal exclusivo para soluções de IoT na rede já existente no campus. As tecnologias *RF-Mesh* e *LoRaWan* são aplicadas para análise do funcionamento do sistema elétrico das tecnologias. A conexão final desses equipamentos é estabelecida com a rede de comunicação interna do centro de computação da UNICAMP, via cabo *ethernet*. As aplicações são configuradas para acessar portas exclusivas em conformidade com a Lei Geral de Proteção de Dados (LGPD) da UNICAMP (CYPRIANO et al., 2021).

Os medidores instalados nas unidades consumidoras registram diversas grandezas elétricas, tais como: potência ativa, potência reativa, tensão, corrente e frequência, as quais são apresentadas resumidamente na Tabela 1.

Conforme apresentado na Tabela 1, os medidores digitais registram um total de 10 grandezas elétricas distintas, sendo que algumas medidas por fase. Esses medidores fornecem dados em tempo real e com alta granularidade, registrando informações a cada 30 segundos. Nesta tese, o atributo de interesse é a potências ativa trifásica, considerando-se um sistema trifásico equilibrado.

Os dados coletados pelos medidores são utilizados para desenvolver os perfis sintéticos de carga das curvas de demanda. Os perfis de carga abrangem uma variedade de áreas, como prédios administrativos, laboratórios/comerciais, iluminação pública e salas de aula. Nos casos em que transformadores alimentam mais de um prédio, os perfis

Tabela 1 – Grandezas aferidas pelos medidores digitais.

Grandeza	Unidade	Grandeza	Unidade
Data/Hora	-	Potência Reativa A	VAr
ID do Dispositivo	-	Potência Reativa B	VAr
Tensão A	V	Potência Reativa C	VAr
Tensão B	V	Frequência A	Hz
Tensão C	V	Frequência B	Hz
Corrente A	A	Frequência C	Hz
Corrente B	A	Ângulo A	rad
Corrente C	A	Ângulo B	rad
Potência Ativa A	W	Ângulo C	rad
Potência Ativa B	W	Corrente Neutro	A
Potência Ativa C	W	Temperatura	°C

Fonte: Elaborado pela autora (2023).

de carga correspondem à soma dos perfis de carga individuais de cada edifício.

Ressalta-se que, embora o sistema de medição do *campus* universitário abranje outras tipologias de carga, como de carregamento de ônibus elétrico e geração fotovoltaica, esses perfis de carga não foram incluídos na presente tese. O foco principal concentra-se nos dados de consumo de potência ativa trifásica, visando uma análise detalhada e específica desse aspecto.

4.2 PROCESSAMENTO E TRATAMENTO DOS DADOS

A primeira análise a ser apresentada é referente ao processamento e ao tratamento dos dados brutos coletados pelo sistema de medição do *campus* universitário da UNICAMP. Durante essa etapa, os dados passaram por diferentes técnicas, resultando em perfis de curva de carga prontos para as análises posteriores. O conjunto de dados considerado é composto por medições de potência ativa por fase, coletadas a cada 30 segundos, durante o período de análise de 01/06/2022 a 01/06/2023.

Os dados selecionados consideram, portanto, a variável de potência ativa trifásica, restringindo-se aos dias úteis. Essa seleção baseia-se na observação de que os dias úteis oferecem padrões de consumo mais consistentes e representativos das atividades do

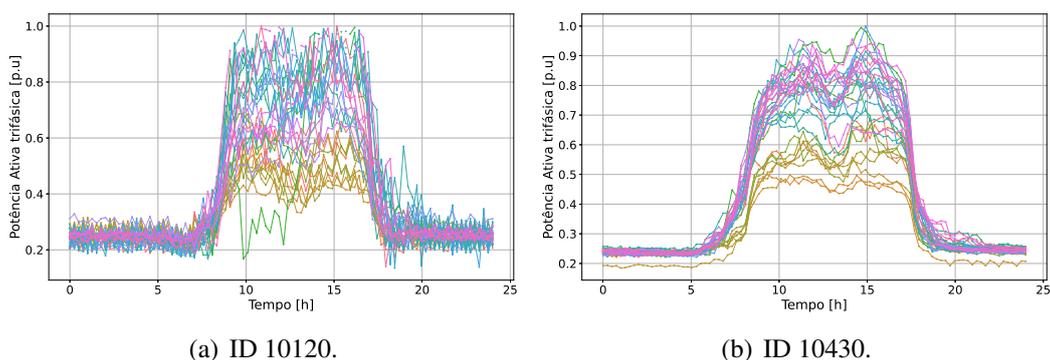
campus universitário. Em seguida, realizou-se o processo de limpeza dos dados, visando identificar e remover possíveis *outliers*, ruídos ou registros inconsistentes.

Finalizada a etapa de processamento dos dados, realiza-se a transformação dos mesmos, iniciando-se pela reamostragem e considerando a granularidade temporal de 15 minutos, o que resulta em perfis de curvas de carga diários compostos por 96 pontos (representando as 24 horas do dia). A escolha desse intervalo foi baseada em considerações práticas e na capacidade de capturar as variações relevantes do consumo de energia ao longo do tempo (JANGID; MATHRURIA; GUPTA, 2021).

Além disso, os dados foram normalizados considerando o intervalo de 0 a 1 em relação à potência ativa trifásica máxima de cada um dos medidores analisados. A etapa de normalização é fundamental para a padronização dos dados, pois isso facilita a comparação entre os diferentes perfis de carga, contribuindo para a eficácia e precisão dos processos de geração de curvas sintéticas.

O conjunto de dados utilizado nesta análise consiste em um total de 42.264 perfis de carga, abrangendo-se uma variedade de padrões de consumo encontrados nos edifícios da universidade. Essa diversidade foi capturada durante o processo de coleta de dados, que considerou os diferentes tipos de edifícios, atividades e horários de funcionamento. Essa abordagem, também, proporciona uma análise abrangente e representativa do consumo de energia no *campus* universitário. Para ilustrar os resultados obtidos, a Figura 11 apresenta exemplos de perfis de curva de carga diária de dois edifícios distintos do estudo de caso.

Figura 11 – Exemplificação do perfil de curva de carga processados.



Fonte: Elaborado pela autora (2023).

A representação gráfica apresentada na Figura 11 (a) e (b) ilustram a variação da potência ativa ao longo do tempo, isto é, perfis de carga diários expressos em unidades padronizadas por unidade (p.u.), discretizadas em intervalos de 15 minutos, representando o consumo de energia de edifícios específicos, tal como o ilustrado pelo ID 10120 e ID 10403, respectivamente.

4.3 CLUSTERIZAÇÃO DOS DADOS

No processo de clusterização os perfis de curvas de carga processados e tratados são agrupadas, considerando os diferentes medidores responsáveis pelo monitoramento da rede elétrica analisada. O objetivo dessa etapa é identificar e categorizar os perfis de consumo semelhantes entre os diversos medidores instalados no *campus*.

Embora existam diversas técnicas de clusterização disponíveis para análise de agrupamento, cada método pode identificar grupos distintos com diferentes associações de objetos quando aplicado a um conjunto de dados específico. Portanto, a escolha do método de clusterização mais apropriado para produzir a partição de dados ideal é uma decisão crítica (RAMOS; VALE, 2008).

Inicialmente, procedeu-se a uma análise exploratória dos dados, empregando o método *K-means*, com o propósito de realizar a clusterização dos dados. Além disso, o conjunto de dados foi submetido a outras metodologias, como o método *Two-Step*. Inicialmente, considerou-se a aplicação da clusterização hierárquica (Agglomerative Hierarchical Clustering) em combinação com o método *K-means*. Adicionalmente, foi conduzido um teste empregando a clusterização hierárquica com o método *Weigh the partitions* (WEACS) proposto por em (RAMOS et al., 2007).

Também, foi considerada a aplicação do método *K-means* em conjunto com o modelo *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) (WANG et al., 2015). Além disso, foi realizada uma análise empregando a abordagem *Single Link* no contexto do agrupamento hierárquico, conforme discutido por Tokuda (TOKUDA; COMIN; COSTA, 2022). Todas as análises foram conduzidas utilizando a biblioteca *Python scikit-learn* (Scikit-Learn Community, 2023).

Destaca-se que a escolha dessas abordagens metodológicas adicionais de clus-

terização visou a busca por resultados mais robustos e abrangentes, tendo em vista a utilização de métodos considerados na literatura para a clusterização de conjuntos de dados de grande volume. No entanto, os resultados obtidos por meio dessas técnicas alternativas mostraram-se inferiores aos centroides obtidos pelo método *K-means*.

Portanto, todas as análises subsequentes apresentadas referem-se exclusivamente ao agrupamento utilizando o método *K-means*, uma vez que este se mostrou adequado para a tarefa de agrupar os conjuntos de dados em questão. A eficácia desse método pode ser atribuída à sua habilidade em identificar centroides representativos e agrupar os dados de maneira coerente.

Para realizar a categorização das curvas de carga por meio do algoritmo *K-means*, primeiramente é necessário determinar o número ideal de grupos para o conjunto de dados analisado. Para isso, realizou-se um teste do número de grupos k na escala de 2 a 10, avaliando-se cada número de grupos por meio da análise do SC, DBI e o método do cotovelo.

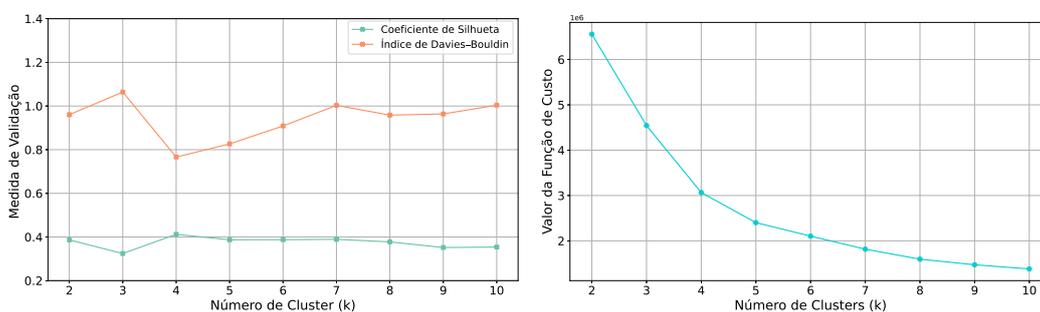
A realização do teste de categorização no processo de clusterização envolveu a avaliação das medidas sob duas abordagens de granularidade temporal. A primeira abordagem utiliza o banco de dados com uma granularidade temporal de 30 segundos, que corresponde a granularidade original dos dados coletados dos medidores.

A segunda abordagem envolve o banco de dados reamostrado com a granularidade de 15 minutos, sendo a principal granularidade analisada nesta pesquisa. Essa perspectiva tem como objetivo determinar se o uso da granularidade original dos dados coletados melhora a capacidade de separação entre os *clusters*, em comparação com a outra granularidade. Os resultados são representados nas Figuras 12 e 13 para análise dos métodos SC, DBI e método do cotovelo para os dados com granularidade de 30 segundos e 15 minutos, respectivamente.

Ao analisar a Figura 12, em relação às avaliações consideradas, e ao banco de dados com granularidade de 30 segundos, identifica-se que a configuração ideal para o número de *clusters* é igual a 4. Quando k é igual a 4, obtém-se os valores de SC atingindo um pico de 0,412, e de índice de DBI, com um mínimo de 0,766.

Além disso, analisando-se a Figura 13 (b), observa-se uma redução notável na função de custo à medida que o número de *clusters* (k) aumenta. No entanto, essa

Figura 12 – Avaliação do número de *clusters* (k) (30 segundos).

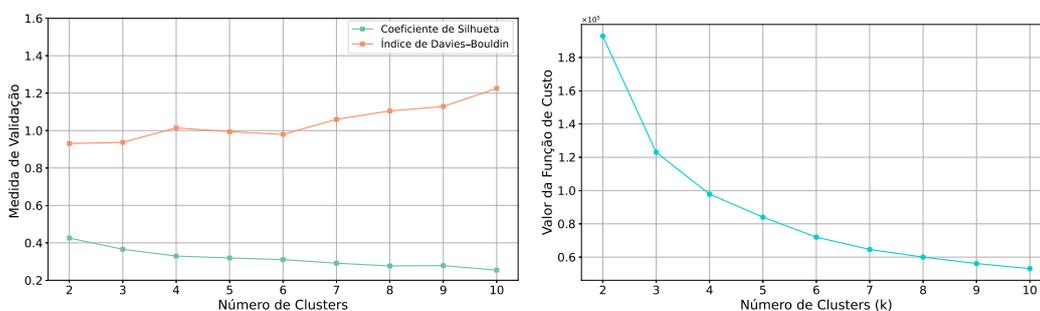


(a) SC e DBI-30 segundos.

(b) Método do cotovelo-30 segundos.

Fonte: Elaborado pela autora (2023).

Figura 13 – Avaliação do número de *clusters* (k) (15 minutos).



(a) SC e DBI- 15 minutos.

(b) Método do cotovelo- 15minutos.

Fonte: Elaborado pela autora (2023).

redução torna-se significativamente menos acentuada após atingir o valor mínimo para k , que, neste caso, foi identificado como 4. Esse ponto de inflexão no gráfico do método do cotovelo sugere que adicionar mais *clusters*, além desse ponto, não proporcionará uma melhoria significativa na separabilidade dos dados.

Por outro lado, ao analisar a Figura 13, em relação às avaliações consideradas e ao banco de dados com granularidade de 15 minutos, o SC sugere que a configuração com 2 ou 3 *clusters* é a mais adequada, uma vez que apresenta os valores mais alto entre todas as opções consideradas. Já considerando DBI, o método indica k igual a 3 *clusters* como a melhor escolha, uma vez que apresenta o valor mais baixo registrado nessa configuração. O método do cotovelo, por sua vez, sugere que a transição de 3 para

4 *clusters* é onde ocorre a maior redução na função de custo.

Dessa forma, ao considerar os resultados da clusterização provenientes da análise com a granularidade original do banco de dados e os reamostrados, não se pode afirmar categoricamente que uma é melhor do que a outra. A granularidade original de 30 segundos preserva melhor os dados, porém, com um custo computacional mais alto. Por outro lado, a granularidade de 15 minutos gera um *cluster* a menos, o que pode simplificar a interpretação, mas também pode ocasionar a perda de detalhes importantes. Assim, a decisão final sobre o número ideal de *clusters* depende do equilíbrio entre a qualidade desejada da clusterização e os recursos computacionais disponíveis.

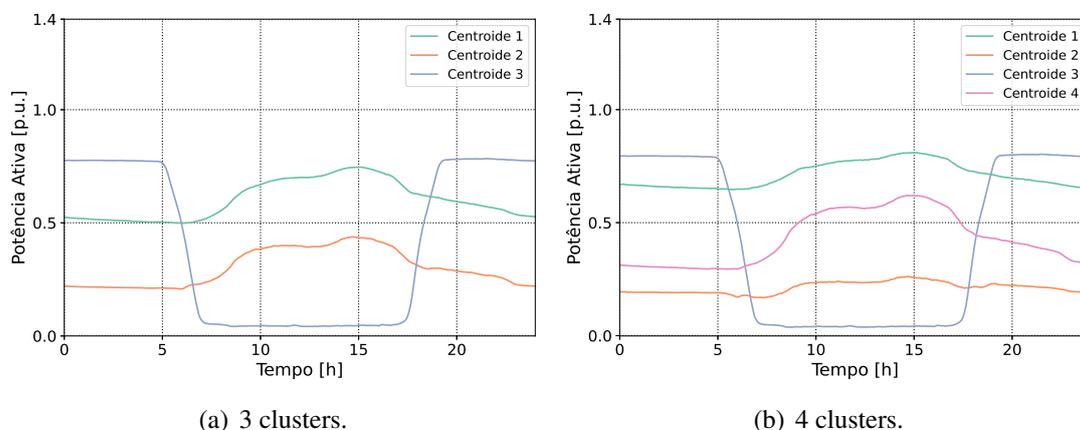
Outro ponto importante na escolha do número ideal de *clusters*, além dos valores obtidos pela avaliação, é o conhecimento do domínio e a interpretação dos resultados. Muitas vezes para ajuste mais apropriado, ao contexto analisado, é necessário avaliar detalhadamente e experimentalmente a quantidade de *clusters* que se faz necessária. Portanto, ao considerar a análise avaliativa, conclui-se que a escolha do número ideal de *clusters* recai nos valores de k igual a 3 ou 4, os quais são analisados em seus respectivos cenários para geração de perfis de curvas de carga sintéticas.

Com base no número de grupos, considerados 3 e 4 *clusters*, respectivamente, estima-se os centroides de cada grupo, conforme ilustrado na Figura 14. Após a conclusão do processo de agrupamento, os perfis de carga atribuídos a cada grupo podem ser utilizados como entrada para o modelo generativo.

Os centroides de *clusters* para os perfis de carga em um *campus* universitário podem ilustrar a diversidade das curvas de carga das diferentes áreas e atividades da localidade. Cada *cluster* pode representar um grupo específico, como salas de aula e laboratórios, setor administrativo e comercial, restaurante universitário e iluminação pública, entre outros. Essa divisão em *clusters* permite uma análise detalhada e a implementação de estratégias de gestão de energia direcionadas a cada grupo, considerando as características e demandas específicas da localidade.

No cenário de 3 *clusters*, as curvas de carga foram distribuídas da seguinte forma: o *cluster* 1 representou 33,88% das curvas (14.321 curvas), o *cluster* 2 representou 45,44% das curvas (19.205 curvas) e o *cluster* 3 representou 20,68% das curvas (8.738 curvas). No cenário de 4 *clusters*, as curvas de carga foram distribuídas da seguinte

Figura 14 – Estimativa dos centroides para cada grupo.



Fonte: Elaborado pela autora (2023).

forma: o *cluster* 1 foi composto por 15,53% das curvas (6.566 curvas), o *cluster* 2 por 23,32% das curvas (9.857 curvas), o *cluster* 3 por 19,43% (8.207 curvas) e o *cluster* 4 por 41,72% das curvas (17.634 curvas).

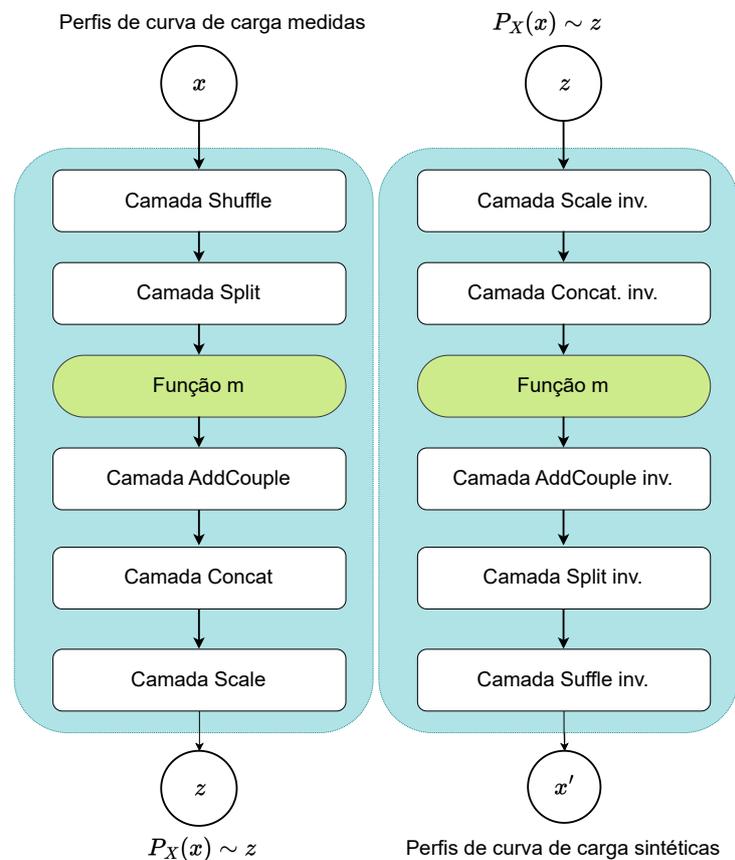
Esses valores foram calculados com base no total de 42.264 curvas no conjunto de dados. É importante ressaltar que as porcentagens indicam a proporção de cada *cluster* em relação ao total de curvas. Essa análise proporciona uma compreensão mais clara da distribuição dos perfis de carga nos diferentes *clusters*, auxiliando no entendimento da variabilidade dos dados.

4.4 ARQUITETURA IMPLEMENTADA

Esta seção descreve a simulação da arquitetura dos modelos NICE, considerando os hiperparâmetros utilizados. Essa análise baseia-se em dois modelos generativos para a geração de curvas de carga: um modelo NICE denso e um modelo NICE convolucional. A Figura 15 apresenta a arquitetura do modelo NICE implementado.

A arquitetura do modelo NICE é composta por diferentes camadas e operações, essenciais na aprendizagem do mapeamento direto entre o espaço de entrada (x) e o espaço latente (z), conforme ilustrado na Figura 15. Na parte esquerda da figura, uma sequência de funções reversíveis é aplicada, resultando na obtenção da função de

Figura 15 – Arquitetura implementada do modelo NICE.



Fonte: Elaborado pela autora (2023).

distribuição de probabilidade ($P_X(x) \sim z$).

Por outro lado, na parte direita da Figura 15, são aplicadas as funções inversas, por meio das transformações inversas das camadas, visando gerar os dados sintéticos (x'), conforme apresentado na Seção 2.2. O modelo proposto é implementado utilizando a biblioteca *keras* (CHOLLET, 2015), considerando as funções *Shuffle*, *SplitVector*, *ConcatVector*, *AddCouple* e *Scale* no desenvolvimento da arquitetura do modelo.

A camada *Shuffle* realiza o embaralhamento da dimensão de entrada, permitindo diferentes formas de rearranjo dos dados. Ela pode ser configurada para inverter diretamente a dimensão ou embaralhar aleatoriamente, evitando a dependência da ordem de entrada (LAUBE; ZELL, 2019).

A camada *SplitVector* divide a entrada em duas partes correspondendo às variáveis x_1 e x_2 nas Equações (2.12) e (2.13), e as intercala, permitindo a extração de diferentes características. Essa divisão e intercalação são essenciais para que o modelo aprenda as informações independentes em cada sub-vetor (KINGMA; DHARIWAL, 2018).

A função m ocasiona a transformação não linear em um dos sub-vetores obtidos pela camada *SplitVector*. Essa transformação pode ser realizada por meio de qualquer função arbitrária capaz de transformar os dados, conforme apresentado na Subseção 2.2.1. No presente estudo, foram utilizadas redes neurais com camadas densas e convolucionais, recorrendo à capacidade intrínseca dessas redes para aprender e representar relações complexas e não lineares nos dados (GOODFELLOW; BENGIO; COURVILLE, 2016). Essa capacidade das redes neurais contribui para a eficácia do modelo NICE na extração de características relevantes e na geração de dados sintéticos consistentes.

A camada *AddCouple* realiza uma operação de acoplamento aditivo entre os dois sub-vetores resultantes da etapa anterior. Essa operação combina as informações independentes aprendidas em cada sub-vetor, permitindo que o modelo capture as relações entre eles. A reversibilidade dessa operação é importante para preservar a informação original e possibilitar a geração de dados sintéticos a partir da representação latente.

A camada *ConcatVector* (nomeada de Concat, na Figura 15) mescla novamente as partes intercaladas dos sub-vetores obtidos na etapa de *SplitVector*, reconstituindo a entrada original. Essa etapa é fundamental para garantir que a estrutura e as características originais dos dados de entrada sejam preservadas na representação latente.

A camada *Scale* ajusta a escala dos elementos na representação latente, controlando a variância das características aprendidas pelo modelo. Isso pode ser útil para normalizar os dados ou melhorar a interpretabilidade da representação latente.

Para o treinamento dos modelos NICE propostos, realizou-se um particionamento do conjunto de dados, onde 70% foram destinados ao treinamento e 30% à validação. Esse procedimento permite avaliar o desempenho dos modelos em dados não utilizados durante o treinamento, fornecendo uma medida de generalização.

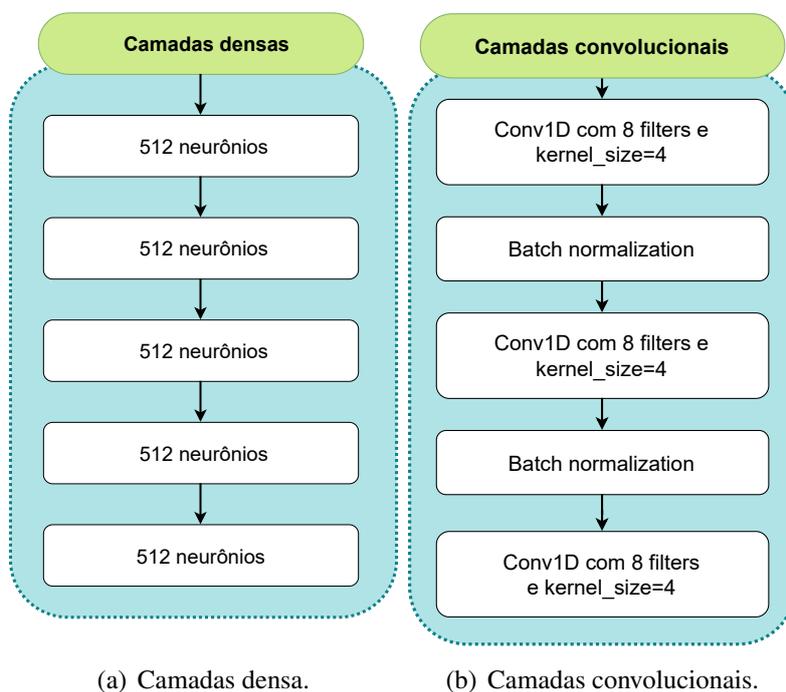
Para otimizar os modelos, foi utilizado o algoritmo de retropropagação do gradiente estocástico em conjunto com o otimizador Adam (KINGMA; BA, 2014). Esses métodos são amplamente empregados para ajustar os parâmetros das redes neurais

durante o processo de treinamento, buscando minimizar a função de perda (GOODFELLOW; BENGIO; COURVILLE, 2016).

No modelo NICE com camadas densas, a configuração da função m é implementada conforme proposto no trabalho (GE et al., 2020). No entanto, é importante destacar que foram realizados testes adicionais em relação ao modelo proposto na literatura. Esses novos testes foram conduzidos por meio dos ajustes dos hiperparâmetros da rede neural densa, sendo realizados experimentos na geração de curvas de carga sintéticas. As análises adicionais é apresentada na Seção 4.7.

Já para determinar os hiperparâmetros ideais do modelo NICE com camadas convolucionais, proposto nesta tese, conduziu-se uma avaliação manual de diversas arquiteturas quanto ao modelo. Foram investigadas diversas combinações de parâmetros, com a compilação e teste dos modelos para identificar as configurações mais apropriadas. A partir dessa análise, foi selecionada a configuração que demonstrou melhor desempenho. A Figura 16 apresenta a arquitetura final escolhida para cada modelo.

Figura 16 – Hiperparâmetros da função m (rede neural).



Fonte: Elaborado pela autora (2023).

A arquitetura do modelo NICE com camadas densas, representada pela função m , consiste em redes neurais com camadas totalmente conectadas, também conhecidas como *perceptron* multicamadas. Essas camadas possuem conexões entre todos os neurônios de uma camada e a próxima, permitindo a transmissão de informações em todas as direções, sendo eficaz na identificação de padrões complexos nos dados, especialmente quando existem relações não lineares entre as variáveis (BISHOP et al., 1995; HAYKIN, 2001).

A Figura 16 (a) ilustra a arquitetura proposta da função m de acordo com o proposto na pesquisa de Ge et al. (2020). Essa arquitetura consiste em cinco camadas densas, cada uma com 512 neurônios, e utiliza a função de ativação conhecida como unidade linear retificada (do inglês, *rectified linear unit (ReLU)*).

Em contrapartida, o modelo NICE convolucional, incorpora uma arquitetura redes neurais com camadas convolucionais que visa adaptar-se de maneira flexível às hierarquias espaciais das características dos dados (KRIZHEVSKY; SUTSKEVER; HINTON, 2012; YAMASHITA et al., 2018).

Essa abordagem tem como objetivo aprender de forma adaptativa, proporcionando uma compreensão mais eficaz das complexidades presentes nos dados. Esse modelo é composto por camadas de convolução, camadas de agrupamento (*pooling*) e camadas totalmente conectadas, empregando o algoritmo de retropropagação para ajustar seus parâmetros e capturar eficientemente os padrões espaciais presentes nos dados.

No modelo NICE convolucional, são utilizadas três camadas convolucionais com a função de ativação ReLU. Entre as duas primeiras camadas convolucionais, é aplicada a técnica de normalização em lote (*Batch Normalization*), que contribui para acelerar o processo de aprendizado e facilitar o treinamento dessas camadas. A normalização em lote promove a estabilidade e a convergência mais rápida do modelo (IOFFE; SZEGEDY, 2015).

4.5 GERAÇÃO E AVALIAÇÃO

Com base nos pressupostos anteriores, a fase de treinamento dos modelos NICE resulta na geração de um conjunto de curvas sintéticas de carga. Em seguida, esse conjunto de dados gerado é submetido a um processo de classificação por meio de

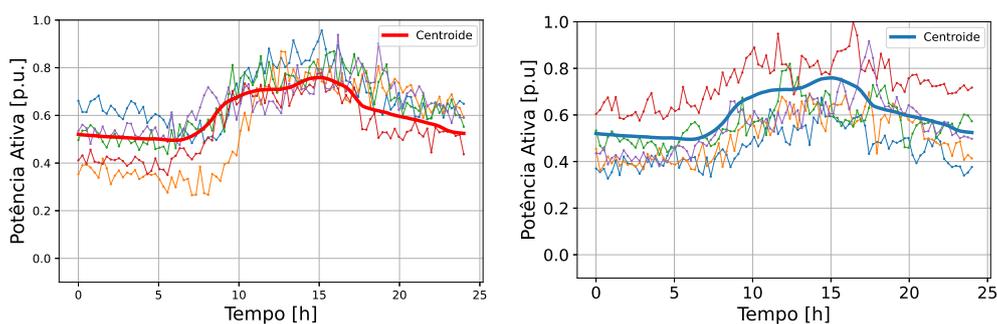
um modelo classificador baseado em redes convolucionais. Esse modelo classifica as curvas nos quatro *clusters* predefinidos na etapa anterior. A etapa de classificação é essencial para orientar análises subsequentes, possibilitando uma avaliação detalhada das características específicas de cada *cluster*.

A análise da geração de perfil de curvas de carga sintéticas, geradas pelo modelo NICE denso e convolucional, requer avaliar diferentes aspectos ao que se refere as curvas geradas por cada um dos modelos. Inicialmente, concentra-se na análise dos perfis de carga gerados em relação aos centroides dos *clusters*, visando uma compreensão das características de agrupamento dessas curvas.

Na sequência, realiza-se a análise da estimativa da função de densidade de probabilidade, com base no KDE. Essa análise propicia o entendimento de como os dados sintéticos se assemelham ou se diferenciam dos perfis de cargas medidas. Para avaliação dessa etapa, foram utilizadas a divergência KL, JS e WD.

Primeiramente, apresentam-se os resultados relativos à análise dos cenários de geração de perfil de carga sintética com três *clusters*, abordando-se, também, os resultados referentes ao cenário que considera a avaliação com quatro *clusters*. As Figuras 17, 18 e 19 exibem os perfis de carga gerados com base nos cenários de três *clusters*, representando o perfil de carga gerado para o respectivo *cluster* e modelo. As subfiguras (a) e (b), em cada um das figuras apresentadas, correspondem, respectivamente, aos perfis curvas de cargas geradas pelos modelos NICE denso e NICE convolucional.

Figura 17 – Perfis de carga gerados - *Cluster 1* (Cenário 3 *clusters*).

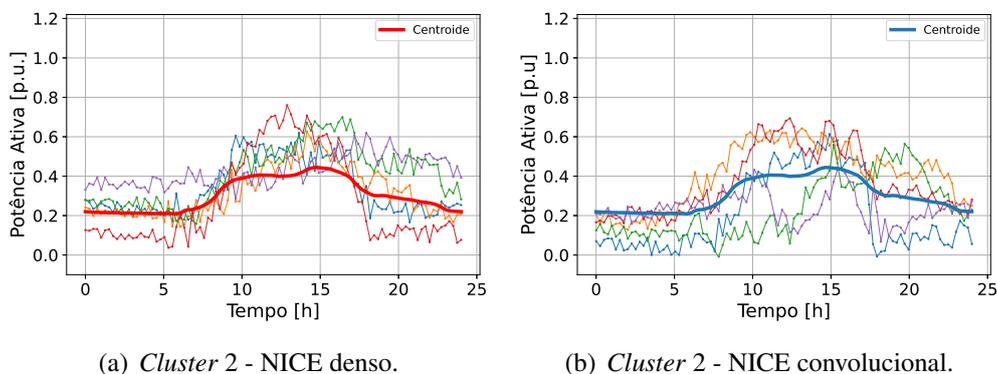


(a) *Cluster 1* - NICE denso.

(b) *Cluster 1* - NICE convolucional.

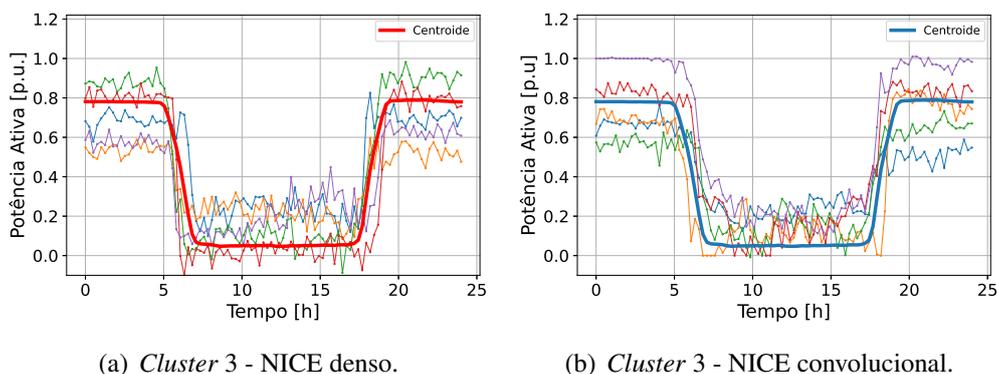
Fonte: Elaborado pela autora (2023).

Figura 18 – Perfis de carga gerados - *Cluster 2* (Cenário 3 *clusters*).



Fonte: Elaborado pela autora (2023).

Figura 19 – Perfis de carga gerados - *Cluster 3* (Cenário 3 *clusters*).



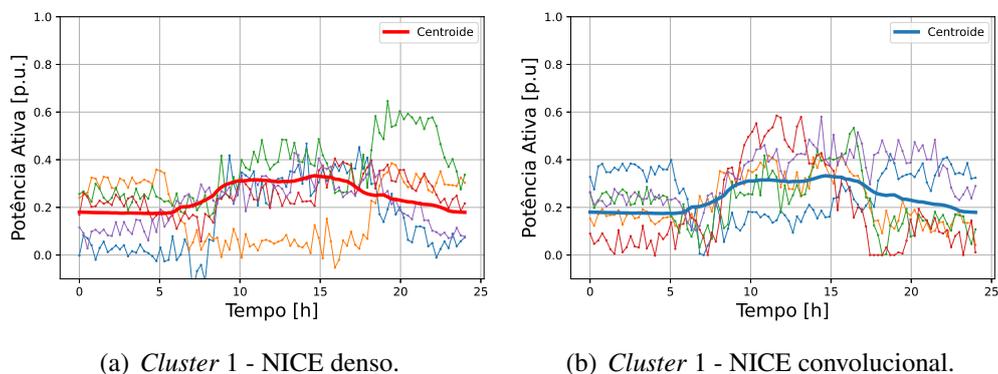
Fonte: Elaborado pela autora (2023).

De modo análogo, por sua vez, as Figuras 20, 21, 22 e 23, apresentam os perfis de carga gerados com base no cenário de quatro *clusters*. Cada subfigura (a) e (b) exibem o perfil de carga gerado para o respectivo *cluster* e modelo NICE denso e NICE convolucional.

Conforme ilustrado nas figuras que representam o cenário de três *clusters* (Figuras 17, 18 e 19), e no cenário de geração de curvas para os quatro *clusters* (Figuras 20, 21, 22 e 23), os perfis de curva de carga sintéticos foram gerados tanto pelo modelo NICE denso como NICE convolucional.

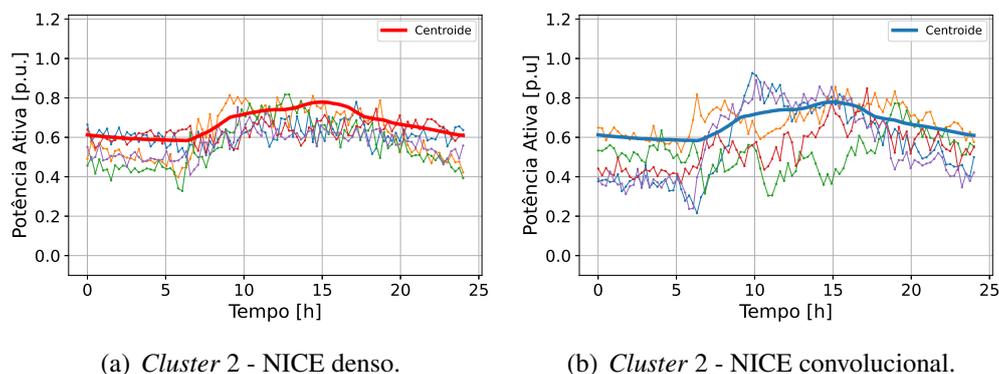
Ambos os modelos evidenciam um comportamento semelhante na geração das

Figura 20 – Perfis de carga gerados- *Cluster 1* (Cenário 4 *clusters*).



Fonte: Elaborado pela autora (2023).

Figura 21 – Perfis de carga gerados - *Cluster 2* (Cenário 4 *clusters*).



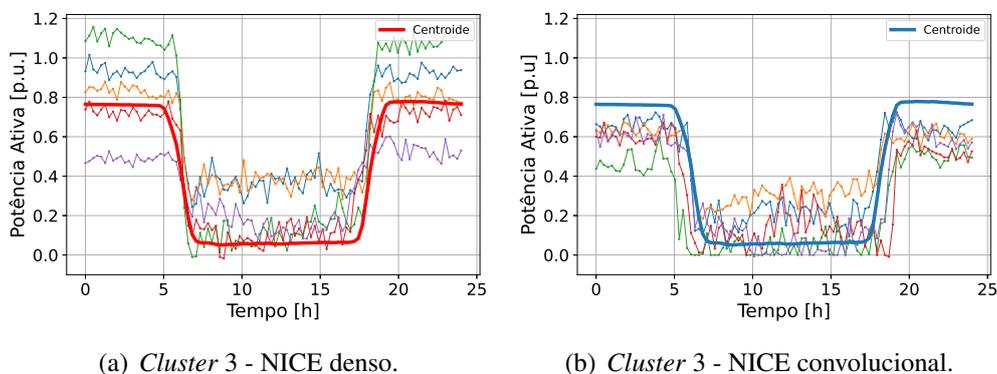
Fonte: Elaborado pela autora (2023).

curvas de carga, evidenciando a capacidade de capturar as características essenciais do conjunto de dados. No entanto, em alguns casos específicos, nota-se que as curvas geradas pelos modelos mostram maior semelhança na reprodução de curvas mais próximas ao centroide dos *clusters* correspondentes.

Além disso, é importante observar que o modelo NICE convolucional gerou curvas com uma variação maior em relação ao centroide, comparado com o modelo NICE denso. Essa diferença pode ser atribuída às características de capacidade de variabilidade ao longo do espaço temporal incorporadas pelo modelo convolucional.

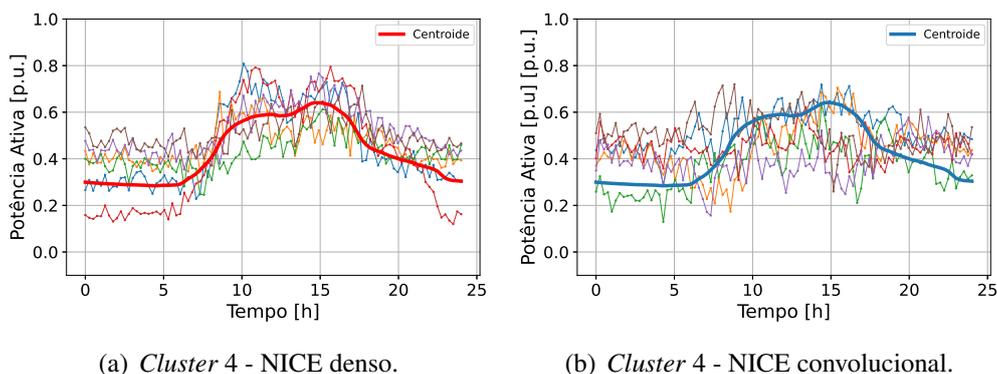
O centroide do *cluster* pode ser considerado uma representação média ou central

Figura 22 – Perfis de carga gerados - *Cluster 3* (Cenário 4 *clusters*).



Fonte: Elaborado pela autora (2023).

Figura 23 – Perfis de carga gerados- *Cluster 4* (Cenário 4 *clusters*).



Fonte: Elaborado pela autora (2023).

das curvas reais nesse grupo específico. Portanto, quando um modelo gera curvas de carga que se aproximam mais do centroide, isso sugere uma captura mais representativa das características principais desse *cluster* em particular.

Para evitar que o modelo generativo gere perfis de carga além da faixa normalizada, mais dados podem ser incluídos no processo de treinamento, especialmente com um conjunto de dados mais equilibrado contendo um número semelhante de perfis de carga para cada *cluster*. Alternativamente, técnicas de regularização podem ser empregadas durante o treinamento para evitar esse comportamento (TIAN; ZHANG, 2022).

A capacidade de capturar as principais características do conjunto de dados é um

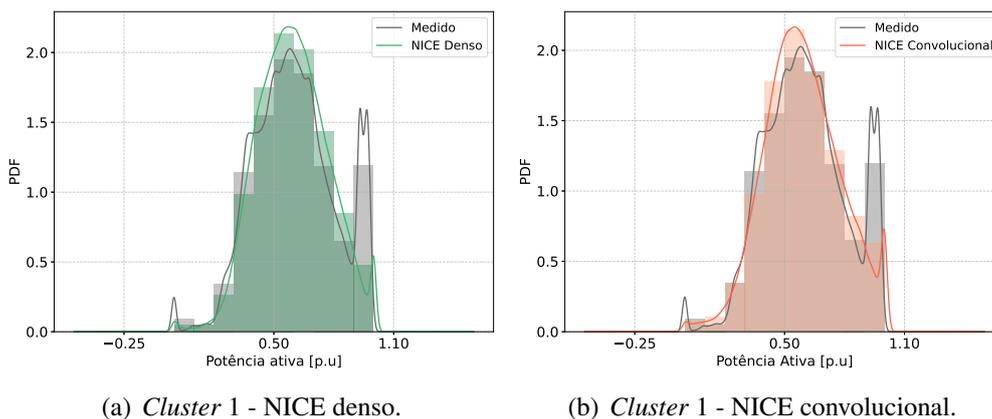
dos indicativos de que o modelo está aprendendo os padrões subjacentes, sendo capaz de gerar curvas de carga sintéticas mais próximas das curvas medidas.

Por exemplo, no contexto de estudos de planejamento energético, é fundamental ter modelos precisos que consigam gerar curvas de carga sintéticas mais realistas. Esses modelos podem auxiliar na simulação de cenários futuros, na otimização do dimensionamento de infraestruturas elétricas e no desenvolvimento de estratégias de gerenciamento de demanda.

No entanto, é importante destacar que a avaliação da qualidade da geração das curvas de carga sintéticas não se limita apenas à proximidade com o centroide do *cluster*. Outros critérios de avaliação, como a análise da distribuição dos dados, a variabilidade e a consistência em relação aos dados medidos, também devem ser considerados. Portanto, após a análise dos perfis das curvas geradas, procede-se à avaliação da estimativa PDFs das curvas medidas e das curvas geradas, considerando KDE em conjunto com o histograma da distribuição dos dados medidos.

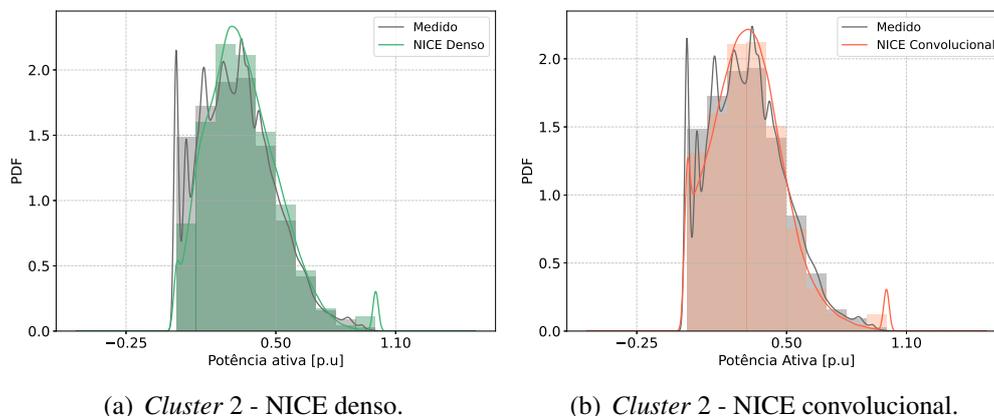
Nas Figuras 24, 25 e 26, são apresentadas as estimativas PDFs das curvas de carga medidas em comparação com as geradas pelos modelos NICE denso e NICE convolucional, para cenários de três *clusters*. Estes gráficos oferecem uma visão comparativa dos modelos em relação a cada *cluster* analisado, proporcionando a análise visual sobre a concordância ou discrepância entre as distribuições de carga sintética e medidas.

Figura 24 – Estimativa PDF - *Cluster 1* (Cenário 3 *clusters*).



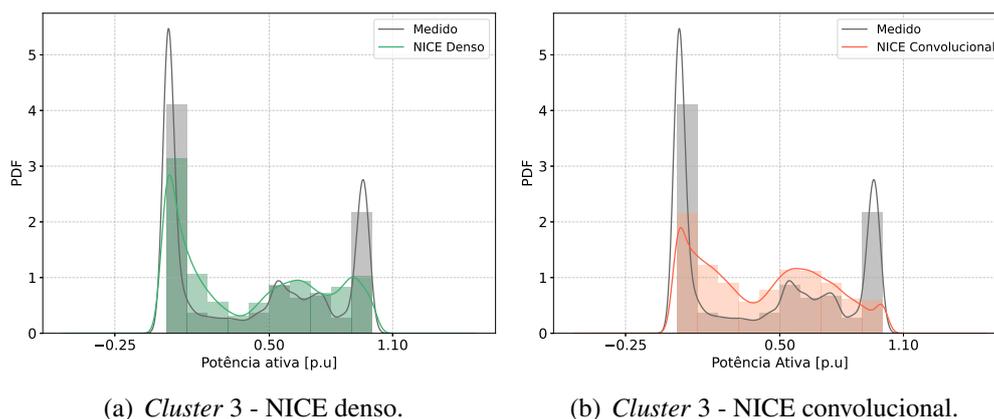
Fonte: Elaborado pela autora (2023).

Figura 25 – Estimativa PDF - *Cluster 2* (Cenário 3 clusters).



Fonte: Elaborado pela autora (2023).

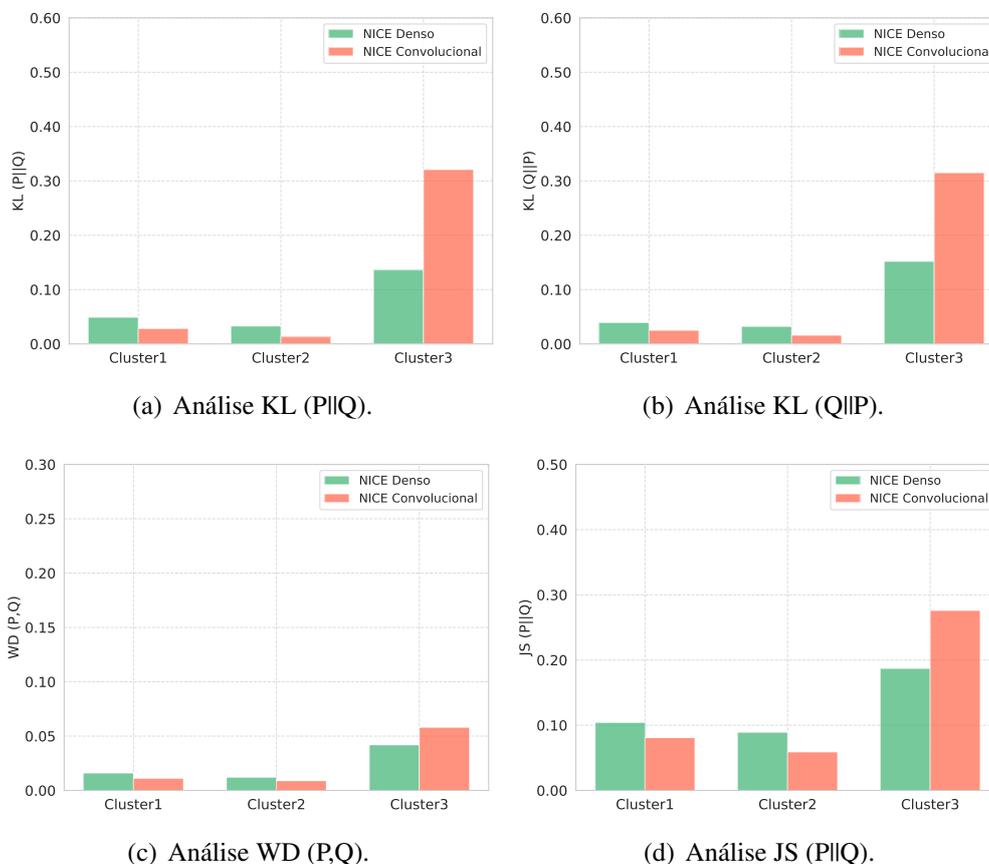
Figura 26 – Estimativa PDF - *Cluster 3* (Cenário 3 clusters).



Fonte: Elaborado pela autora (2023).

A partir da análise PDFs, são avaliadas as medidas KL, WD, e JS, as quais estão representadas graficamente na Figura 27, traçando-se um comparativo dos valores numéricos obtidos em cada um dos modelos considerados.

Observando a Figura 27, para a divergência KL ($P||Q$) e KL ($Q||P$), nota-se que, no que diz respeito ao *Cluster 1* e *Cluster 2*, o modelo NICE convolucional demonstrou valores inferiores em comparação ao modelo NICE denso. Isso sugere que, nesse cenário específico, o modelo convolucional se ajustou melhor às curvas de carga medidas,

Figura 27 – Análise KL, WD e JS (Cenário 3 *clusters*).

Fonte: Elaborado pela autora (2023).

indicando uma maior similaridade entre as estimativas das distribuições de probabilidade.

Já analisando a medida WD, também, o modelo NICE convolucional exibiu uma menor distância em relação ao modelo NICE denso para o *Cluster 1* e *Cluster 2*, em relação às PDFs das curvas medidas e geradas. Isso implica que as curvas de carga geradas pelo modelo NICE convolucional são mais similares às curvas de carga medidas, indicando uma melhor correspondência na forma das distribuições de probabilidade.

No que se refere à divergência JS, a análise numérica revela resultados semelhantes aos observados nas medidas anteriores. Para os *Cluster 1* e *Cluster 2*, o modelo NICE convolucional exibe uma menor divergência em relação ao modelo NICE denso. Isso sugere que a distribuição de probabilidade gerada pelo modelo convolucional está mais

próxima da distribuição real das curvas de carga.

Ao se avaliar as medidas em conjunto, destaca-se o desempenho superior do modelo NICE denso em relação ao modelo NICE convolucional no que diz respeito ao *Cluster 3*. Este *cluster* representa a centroide que caracteriza a forma de onda da iluminação pública. Tal superioridade pode ser atribuída, na maioria das vezes, às características inerentes das redes neurais densas, que conseguem capturar e representar as nuances complexas das curvas características inerentes a essas curvas.

Para uma análise mais detalhada dos valores numéricos apresentados na Figura 27, a Tabela 2 apresenta os resultados numéricos referentes às medidas de avaliação relacionadas à análise da estimativa das PDFs para os modelos NICE.

Tabela 2 – Valores numéricos: KL, WD, JS (Cenário 3 *Clusters*).

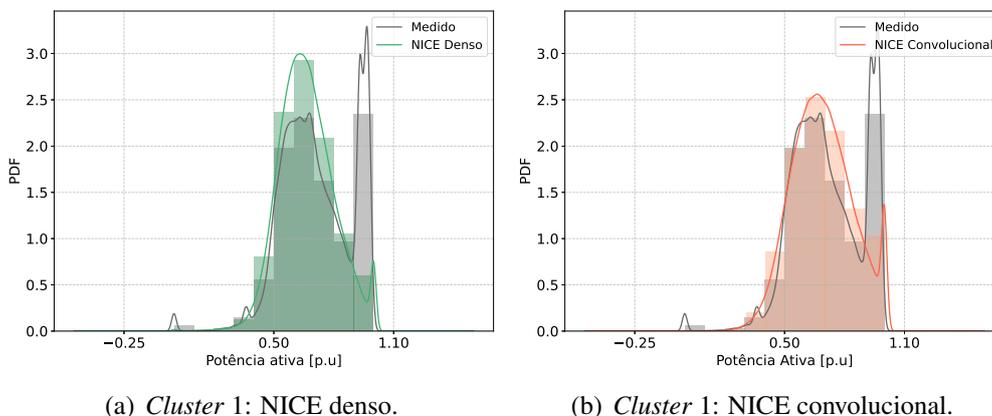
		<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>
NICE denso	<i>KL (P Q)</i>	0,049	0,033	0,136
	<i>KL (Q P)</i>	0,039	0,032	0,152
	<i>WD</i>	0,016	0,012	0,042
	<i>JS (P Q)</i>	0,104	0,089	0,187
NICE convolucional	<i>KL (P Q)</i>	0,028	0,013	0,321
	<i>KL (Q P)</i>	0,025	0,016	0,315
	<i>WD</i>	0,011	0,009	0,058
	<i>JS (P Q)</i>	0,081	0,059	0,276

Fonte: Elaborado pela autora (2023).

Já em relação aos cenários com quatro *clusters*, nas Figuras 28, 29, 30 e 31 são apresentadas as representações gráficas em relação aos modelos NICE denso e NICE convolucional, utilizando o KDE em conjunto com o histograma para a estimativa das PDFs dos dados medidos e gerados para cada um dos *clusters* considerados.

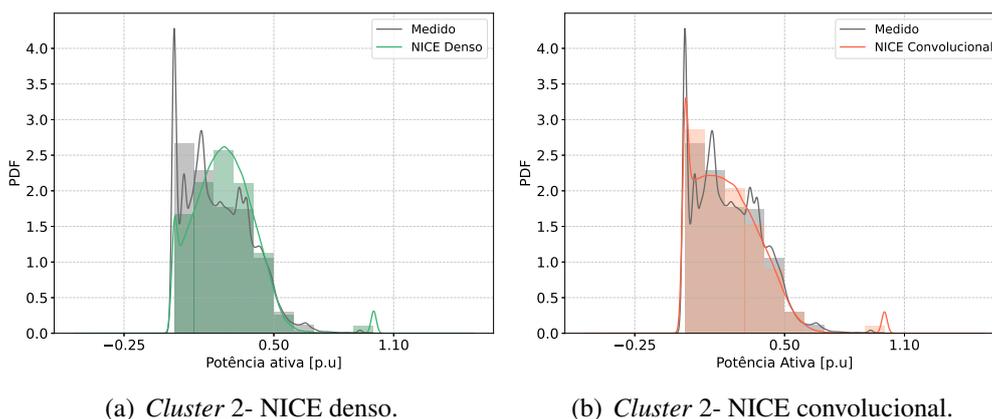
A partir da análise da estimativa das PDFs, são avaliadas a divergência KL, WD, e JS, as quais estão representadas graficamente na Figura 32, traçando-se um comparativo dos valores numéricos obtidos em cada um dos modelos considerados.

Figura 28 – Estimativa PDF- *Cluster 1* (Cenário 4 clusters).



Fonte: Elaborado pela autora (2023).

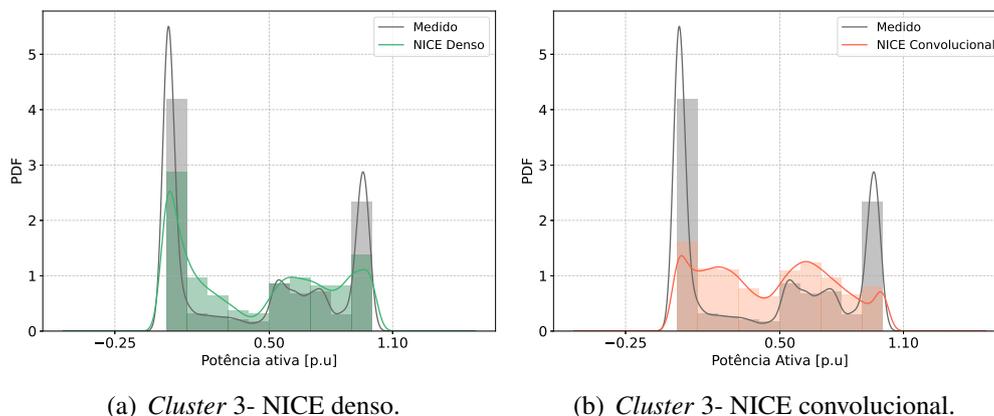
Figura 29 – Estimativa PDF - *Cluster 2* (Cenário 4 clusters).



Fonte: Elaborado pela autora (2023).

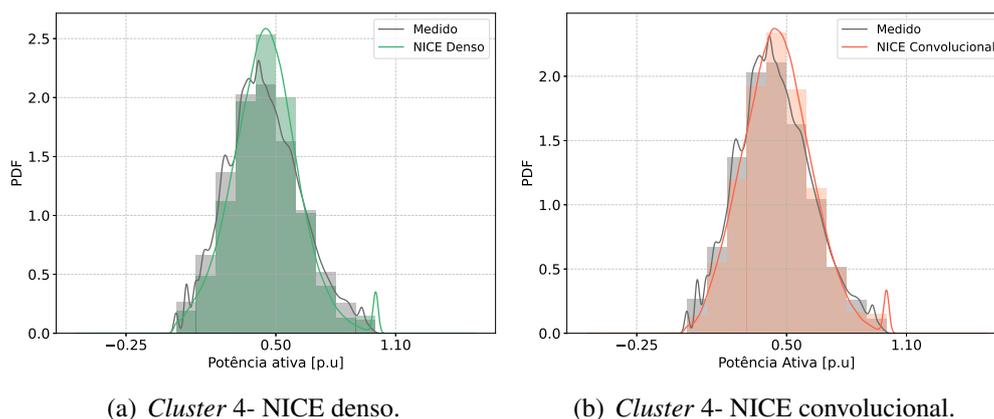
Analisando-se a Figura 32, para as medidas $KL(P||Q)$ e $KL(Q||P)$, $WD(P,Q)$, $JS(P||Q)$, é possível observar que ambos os modelos considerados apresentam valores baixos em relação às medidas consideradas. Isso evidencia a capacidade de ambos os modelos em reproduzir curvas de cargas sintéticas satisfatoriamente, destacando a semelhança entre as estimativas da distribuição real ($P(x)$) e a estimativa da distribuição gerada pelo modelo ($Q(x)$). No entanto, é possível observar uma tendência de superioridade do modelo NICE denso em relação ao modelo NICE convolutacional, considerando $KL(P||Q)$, $KL(Q||P)$ e $WD(P,Q)$.

Figura 30 – Estimativa PDF - *Cluster 3* (Cenário 4 clusters).



Fonte: Elaborado pela autora (2023).

Figura 31 – Estimativa PDF - *Cluster 4* (Cenário 4 clusters).

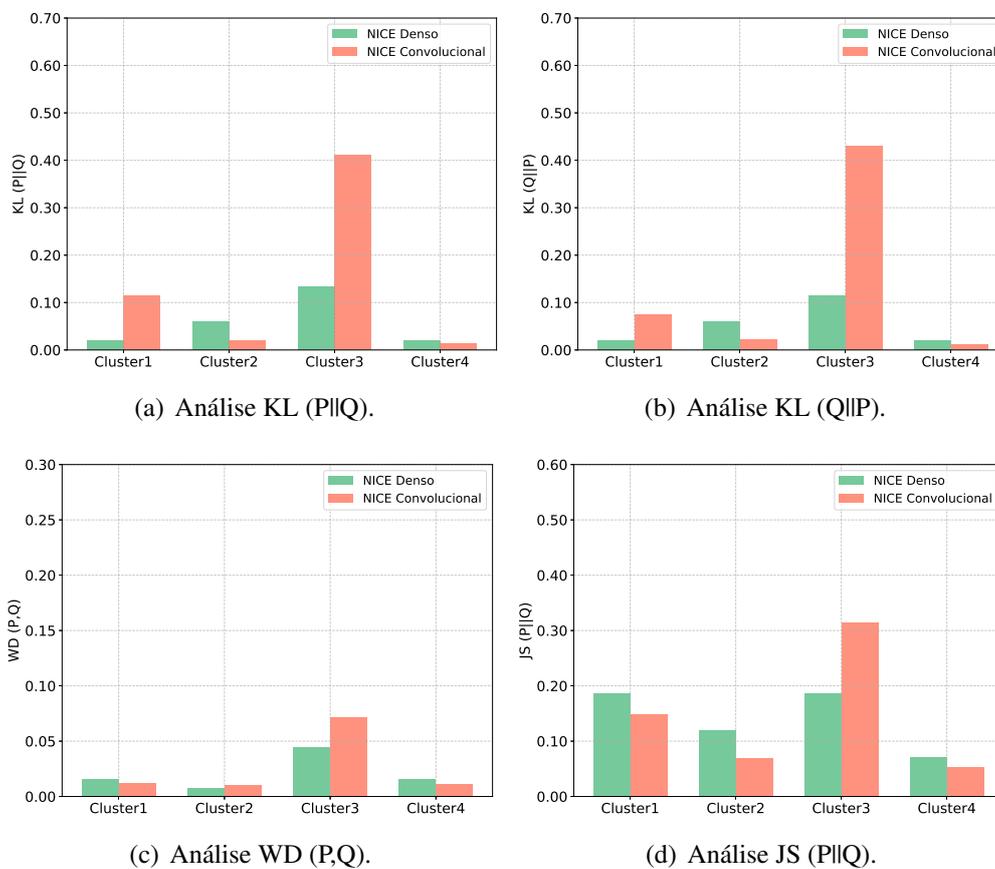


Fonte: Elaborado pela autora (2023).

Já considerando a divergência JS ($P||Q$), observa-se uma superioridade do modelo NICE convolucional na minimização da divergência entre a distribuição real e a distribuição gerada pelo modelo. Para este caso em questão, o modelo convolucional se mostrou mais adequado para reproduzir as distribuições de probabilidade de maneira mais equilibrada.

As diferenças nos resultados das medidas podem indicar que cada modelo tem suas próprias vantagens, desvantagens e particularidades na modelagem das curvas de

Figura 32 – Análise KL, WD e JS (Cenário 4 clusters).



Fonte: Elaborado pela autora (2023).

carga. Sendo assim, a escolha entre os modelos dependerá dos objetivos específicos pleiteados.

Para uma análise mais detalhada dos valores numéricos apresentados na Figura 32, a Tabela 3 apresenta os resultados numéricos referentes às medidas de avaliação relacionadas à análise da estimativa das PDFs para os modelos NICE.

Tabela 3 – Valores numéricos: KL, WD, JS (Cenário 4 *Clusters*).

		<i>Cluster 1</i>	<i>Cluster 2</i>	<i>Cluster 3</i>	<i>Cluster 4</i>
NICE denso	<i>KL (P Q)</i>	0,021	0,059	0,133	0,021
	<i>KL (Q P)</i>	0,020	0,059	0,155	0,020
	<i>WD</i>	0,016	0,008	0,045	0,016
	<i>JS (P Q)</i>	0,186	0,119	0,187	0,072
NICE convolucional	<i>KL (P Q)</i>	0,115	0,021	0,411	0,013
	<i>KL (Q P)</i>	0,075	0,022	0,431	0,012
	<i>WD</i>	0,012	0,010	0,072	0,011
	<i>JS (P Q)</i>	0,148	0,069	0,315	0,054

Fonte: Elaborado pela autora (2023).

Outro fator importante a ser considerado na análise é a divergência em relação aos resultados obtidos no cenário de três *clusters* e quatro *clusters*. A comparação dos resultados entre esses cenários possibilita averiguar o impacto do número de *clusters* na avaliação dos modelos propostos.

Para o cenário com 3 *clusters*, observa-se que o modelo NICE denso e convolucional apresentaram resultados similares, mas, para os *clusters* 1 e 2, os índices obtidos em relação as medidas considerados, foram menores para o modelo NICE convolucional, mostrando-se eficaz na representação das curvas de cargas sintéticas. Em contraste, o modelo NICE denso para o *cluster* 3 apresentou resultado mais satisfatório, mostrando a capacidade desse modelo em capturar nuances complexas nas distribuições de probabilidade das curvas de carga.

Ao expandir o cenário para 4 *clusters*, observa-se uma disparidade mais acentuada nos resultados. Os modelos exibiram um desempenho variado em cada medida e em diferentes *clusters*. Na maioria das análises consideradas, o modelo NICE denso foi mais satisfatório em relação ao modelo NICE convolucional, refletindo a capacidade de representar a estimativa das distribuições de probabilidade subjacentes.

A divergência nos resultados entre os cenários de 3 *clusters* e 4 *clusters* pode ser atribuída a diferentes fatores, incluindo a complexidade crescente dos *clusters*, a variação *intracluster* ampliada e a distribuição de dados entre os *clusters* e a sensibilidade das medidas às variações nos *clusters*. Essa análise ressalta a influência crítica do número de

clusters, destacando a necessidade de escolher a configuração do modelo mais apropriada, levando-se em conta o banco de dados analisado e os objetivos do projeto.

De modo, geral, ao analisar as distribuições de probabilidade das curvas de carga, observou-se uma notável semelhança entre a estimativa das PDFs dos dados medidos e dos dados sintéticos, para ambos os modelos NICE considerados. Essa semelhança indica que ambos os modelos conseguem gerar perfis sintéticos de carga que reproduzem em termos gerais as características dos dados medidos (reais).

Essa conclusão é válida tanto para a análise considerando o cenário de três *clusters*, quanto para a análise considerando o cenário de quatro. Em ambas as situações, os modelos NICEs demonstraram eficácia na geração de curvas de carga sintéticas, apresentando distribuições de probabilidade semelhantes às dos dados reais.

Essa capacidade dos modelos NICEs de reproduzir com precisão as características das curvas de carga medidas é de grande importância, possibilitando a utilização dos perfis sintéticos de carga como representações confiáveis dos dados reais. Esses perfis sintéticos podem ser aplicados em simulações e análises de cenários futuros, visando auxiliar a tomada de decisões, embasadas em informações precisas e realistas.

No geral, os resultados obtidos evidenciam um desempenho adequado dos modelos NICEs na geração de perfis sintéticos de carga. Cada modelo apresenta suas particularidades na geração de características específicas de diferentes curvas de carga. Essa capacidade de capturar as características distintas das curvas de carga contribui para a utilidade e relevância dos modelos no contexto analisado.

Ressalta-se que ao aplicar o modelo NICE para a geração de perfis de curvas de cargas sintéticas, observa-se a necessidade de realizar ajustes nos hiperparâmetros da rede neural densa e convolucional, conforme o banco de dados é modificado. Essa prática é comum e esperada em modelos de aprendizado de máquina, uma vez que os hiperparâmetros devem se adaptar às características e complexidade dos dados para obter resultados mais precisos e representativos.

4.6 ANÁLISE DA GRANULARIDADE TEMPORAL

A granularidade temporal, no contexto da geração de curvas de carga, refere-se ao intervalo de tempo entre as amostras medidas. Neste estudo, foram considerados diferentes intervalos de granularidade temporal, como 30 segundos, 1 minuto, 5 minutos, 10 minutos, 15 minutos e 1 hora. A seleção da granularidade adequada é um fator importante na aplicação prática, uma vez que pode impactar, tanto a compreensão da demanda em tempo real, quanto a precisão das previsões de carga para o planejamento das redes de distribuição elétrica.

Do ponto de vista operacional, uma granularidade temporal menor, como 30 segundos, 1 minuto ou 5 minutos, pode ser necessária para obter um entendimento mais detalhado da demanda em tempo real. Isso permite ajustes mais rápidos na rede elétrica em resposta as variações da demanda. Por outro lado, no contexto do planejamento da rede, uma granularidade maior, como 10 minutos ou 15 minutos, pode ser suficiente, uma vez que as variações de carga em períodos mais longos podem ser previstas com maior precisão.

No âmbito dos modelos propostos, NICE denso e convolucional, a escolha da granularidade pode influenciar nos valores ótimos dos hiperparâmetros do modelo. Por exemplo, para granularidades menores, pode ser necessário aumentar o número de camadas e neurônios nos modelos, aumentando a complexidade da arquitetura do modelo e, conseqüentemente, o tempo de treinamento necessário para execução do algoritmo.

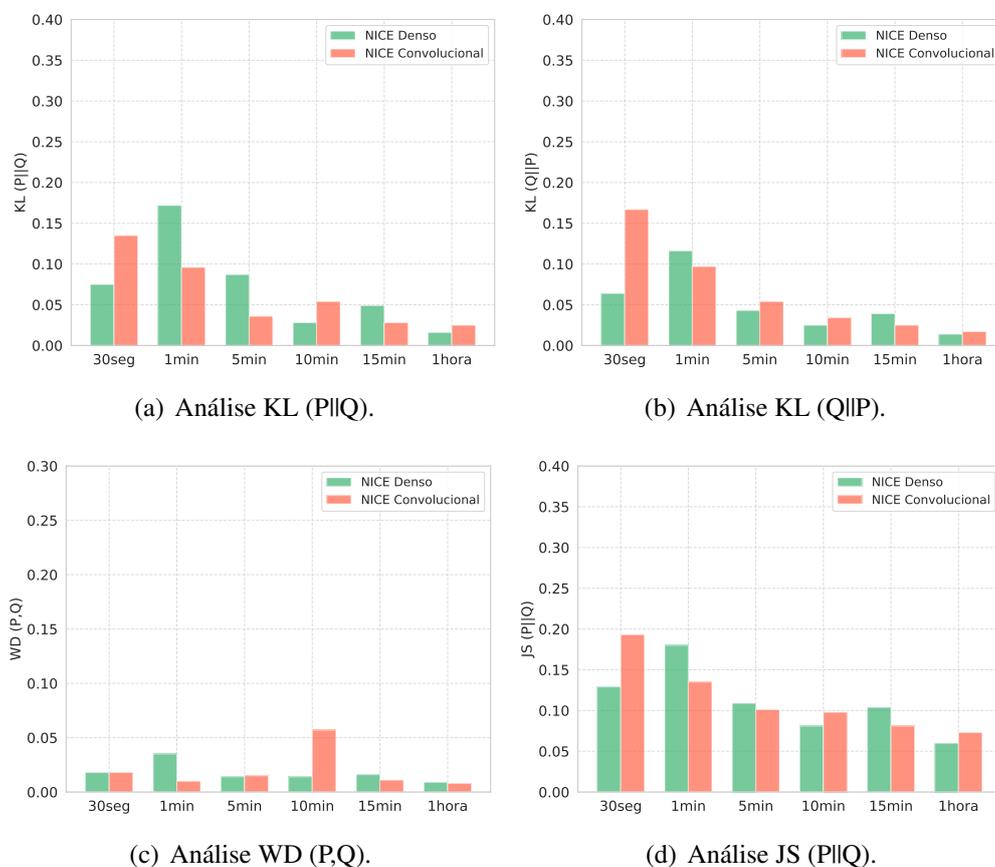
Ressalta-se, que cada cenário abordado requer uma adaptação dos hiperparâmetros do modelo conforme a granularidade escolhida, considerando a relação entre o tempo de processamento e a precisão na geração dos dados sintéticos. Além disso, o tempo necessário para gerar as curvas de carga sintéticas também é uma consideração importante, uma vez que uma granularidade mais fina pode resultar em um aumento significativo no tempo de simulação computacional.

Dessa forma, para realizar a geração de perfis de curvas sintéticas considerando as distintas granularidades temporais, optou-se por manter a mesma configuração de hiperparâmetros para os modelos NICE denso e convolucional, conforme descrito em detalhes na Seção 4.4. Essa abordagem padronizada possibilita avaliar a capacidade de

generalização dos modelos diante das diversas granularidades temporais em questão.

Para uma análise simplificada dos resultados das simulações com diferentes granularidades temporais, apresentam-se figuras de divergência, incluindo KL, WD e JS. Essas medidas são utilizadas para avaliar a similaridade entre as estimativas das distribuições de probabilidade nos modelos NICE com camadas densas e convolucionais, em cenários de três e quatro *clusters*. As Figuras 33, 34 e 35 ilustram a análise de KL, WD e JS no cenário de três *clusters*.

Figura 33 – Análise KL, WD, JS - *Cluster 1* (Cenário 3 *clusters*).



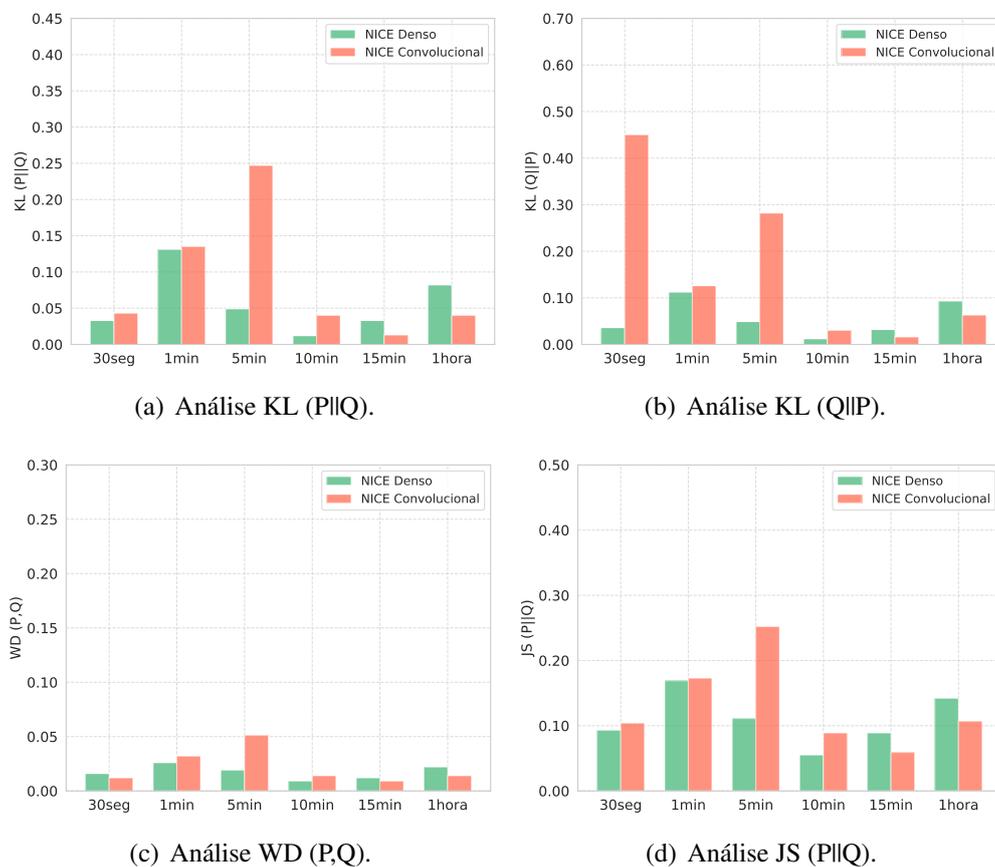
Fonte: Elaborado pela autora (2023).

Analisando a Figura 33, um conjunto de resultados para o *Cluster 1*. É possível notar que, em ambos os modelos, que à medida que a granularidade temporal aumenta, as divergência KL ($P||Q$) e KL ($P||Q$), bem como WD (P,Q) e JS ($P||Q$), tendem a

diminuir. Esse declínio indica uma maior semelhança entre os dados gerados pelo modelo e os dados medidos.

Além disso, é importante notar que, em ambos os modelos, os valores numéricos se aproximam, o que demonstra um comportamento semelhante. Em algumas granularidades temporais, como 30 segundos, 10 minutos e 1 hora, o modelo NICE denso apresentou um desempenho superior. No entanto, para as demais granularidades temporais analisadas, o modelo NICE convolucional obteve valores inferiores, mais próximos de zero, sugerindo uma maior semelhança entre as estimativas das distribuições de probabilidade dos dados medidos e dos dados gerados.

Figura 34 – Análise KL, WD, JS - *Cluster 2* (Cenário 3 clusters).

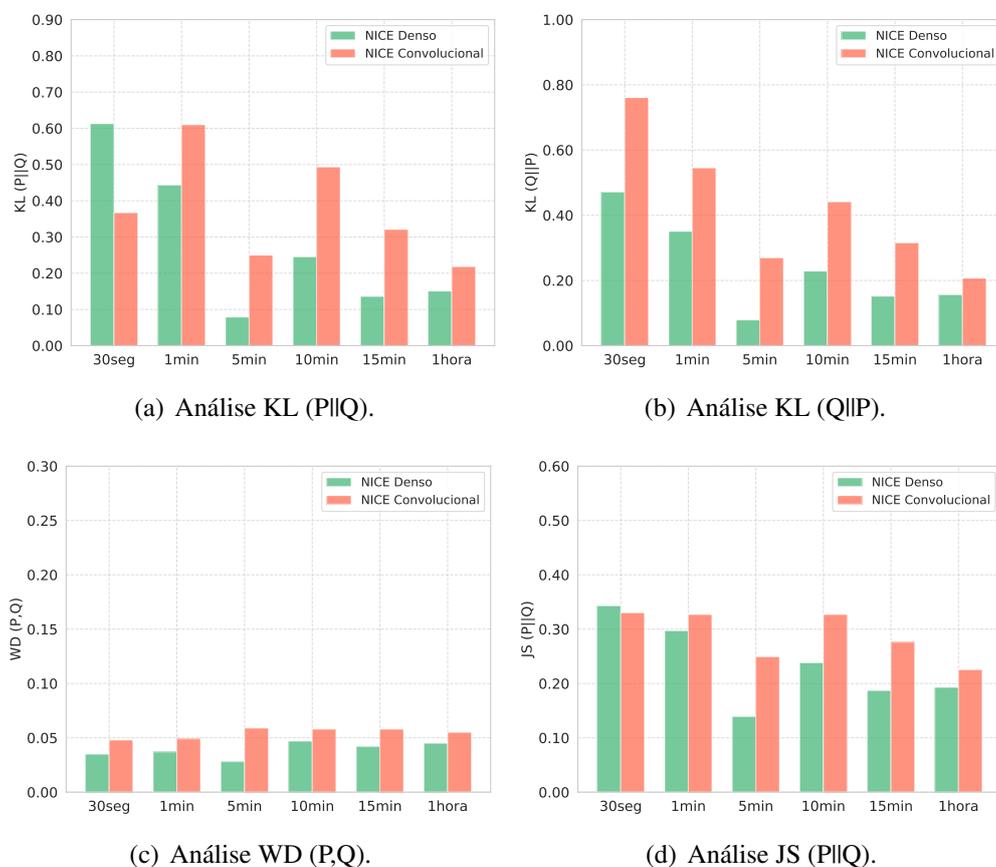


Fonte: Elaborado pela autora (2023).

Analisando a Figura 34, observa-se que, de modo análogo aos resultados obtidos

pelo conjunto de medida do *Cluster 1*, ao analisar os valores das medidas avaliativas no *Cluster 2*, à medida que a granularidade temporal aumenta, as medidas avaliadas tendem a diminuir. Esse declínio indica uma maior semelhança entre os dados gerados pelo modelo e os dados medidos. Ao comparar os valores obtidos pelos dois modelos, pode-se observar que ambos obtiveram valores baixos. No entanto, para granularidades menores, o modelo NICE denso obteve valores menores, enquanto para granularidades maiores, como 15 minutos e 1 hora, o modelo NICE convolucional obteve valores menores.

Figura 35 – Análise KL, WD, JS - *Cluster 3* (Cenário 3 clusters).



Fonte: Elaborado pela autora (2023).

Analisando os resultados para o *Cluster 3* na Figura 35, é possível observar que, em ambos os modelos considerados, os valores obtidos em relação às medidas se mantiveram em uma escala mediana. No entanto, observa-se que as avaliações

apresentaram valores mais elevados ao considerar granularidades temporais mais finas, como 30 segundos e 1 minuto

Além disso, também se destaca que o modelo NICE denso demonstrou um desempenho superior em comparação com o modelo convolucional. Como discutido previamente em seções anteriores, este *cluster* se relaciona ao centroide da curva representativa da iluminação pública, no qual as redes neurais com camadas densas conseguiram capturar as características intrínsecas e nuances dessas curvas.

Ressalta-se que a diferença de desempenho entre as avaliações dos modelos propostos está intrinsecamente ligada à forma como cada medida avalia as discrepâncias e semelhanças entre as estimativas das distribuições. Cada medida é adaptada a contextos específicos, destacando aspectos relevantes, como assimetria, topologia e simetria.

Os valores numéricos relacionadas aos resultados obtidos para a granularidade temporal no cenário de três *clusters* (Figuras 33, 34, 35) são apresentadas na Tabela 4 para o modelo NICE denso e na Tabela 5 para o modelo NICE convolucional.

Tabela 4 – NICE denso: Valores KL, WD, JS (Cenário 3 *clusters*).

<i>Cluster</i>	<i>Medida</i>	30seg	1min	5min	10min	15min	1hora
1	<i>KL (P Q)</i>	0,075	0,172	0,087	0,028	0,049	0,016
	<i>KL (Q P)</i>	0,064	0,116	0,043	0,025	0,039	0,014
	<i>WD</i>	0,018	0,035	0,014	0,014	0,016	0,009
	<i>JS (P Q)</i>	0,129	0,180	0,109	0,081	0,104	0,060
2	<i>KL (P Q)</i>	0,033	0,131	0,049	0,012	0,033	0,082
	<i>KL (Q P)</i>	0,036	0,112	0,049	0,012	0,032	0,093
	<i>WD</i>	0,016	0,026	0,019	0,009	0,012	0,022
	<i>JS (P Q)</i>	0,093	0,169	0,111	0,055	0,089	0,142
3	<i>KL (P Q)</i>	0,613	0,443	0,079	0,245	0,136	0,151
	<i>KL (Q P)</i>	0,471	0,351	0,079	0,229	0,152	0,156
	<i>WD</i>	0,035	0,037	0,028	0,047	0,042	0,045
	<i>JS (P Q)</i>	0,343	0,297	0,139	0,238	0,187	0,193

Fonte: Elaborado pela autora (2023).

Tabela 5 – NICE convolucional: Valores KL, WD, JS (Cenário 3 *clusters*).

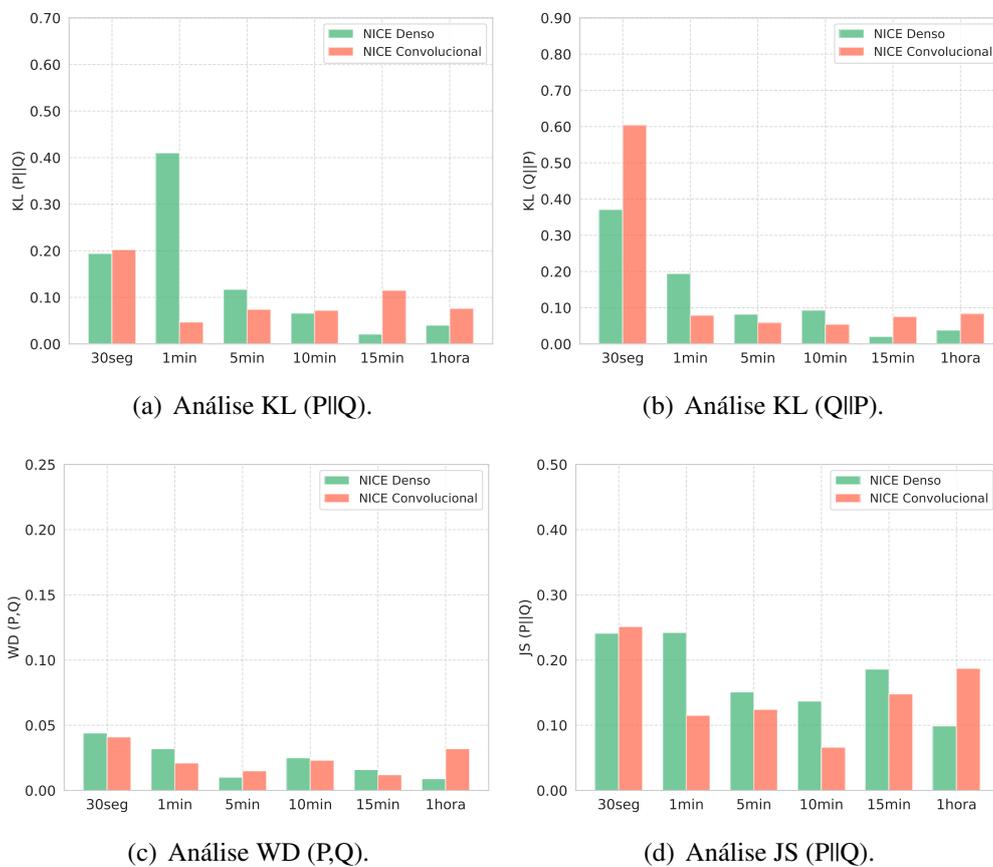
<i>Cluster</i>	<i>Medida</i>	30seg	1min	5min	10min	15min	1hora
1	<i>KL (P Q)</i>	0,135	0,096	0,036	0,054	0,028	0,025
	<i>KL (Q P)</i>	0,167	0,097	0,054	0,034	0,025	0,017
	<i>WD</i>	0,018	0,010	0,015	0,057	0,011	0,008
	<i>JS (P Q)</i>	0,193	0,135	0,101	0,098	0,081	0,073
2	<i>KL (P Q)</i>	0,043	0,135	0,247	0,040	0,013	0,040
	<i>KL (Q P)</i>	0,450	0,126	0,282	0,030	0,016	0,063
	<i>WD</i>	0,012	0,032	0,051	0,014	0,009	0,014
	<i>JS (P Q)</i>	0,104	0,173	0,252	0,089	0,059	0,107
3	<i>KL (P Q)</i>	0,367	0,610	0,250	0,493	0,321	0,218
	<i>KL (Q P)</i>	0,761	0,545	0,269	0,441	0,315	0,207
	<i>WD</i>	0,048	0,049	0,059	0,058	0,058	0,055
	<i>JS (P Q)</i>	0,330	0,327	0,249	0,327	0,276	0,225

Fonte: Elaborado pela autora (2023).

Ao analisar a Figura 36 no contexto da avaliação do *Cluster 1* em um cenário com quatro *clusters*, observa-se que ambos os modelos propostos apresentaram valores baixos em relação às medidas consideradas. Isso sugere um desempenho satisfatório, uma vez que valores menores indicam uma maior similaridade entre as estimativas das PDFs analisadas.

Outro fator, também constatado nas análises anteriores, é que, à medida que se consideram granularidades mais elevadas, os valores das medidas tendem a diminuir, o que demonstra que o modelo está conseguindo capturar de forma mais precisa as discrepâncias e semelhanças entre as curvas geradas e as curvas medidas. Para esse *cluster* em questão, na maioria das granularidades temporais consideradas, o modelo NICE convolucional apresentou valores mais baixos em comparação com o modelo NICE denso.

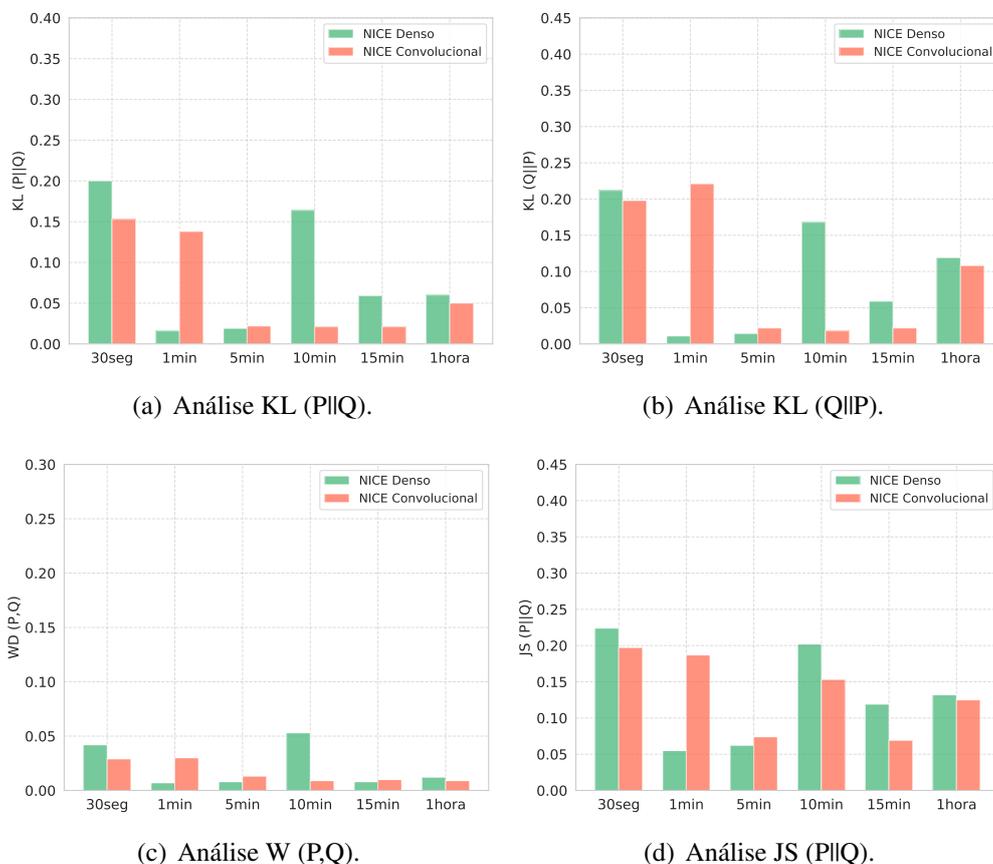
Figura 36 – Análise KL, WD, JS- *Cluster 1* (Cenário 4 *clusters*).



Fonte: Elaborado pela autora (2023).

A análise da Figura 37 revela que ambos os modelos propostos exibem valores das medidas consistentemente baixos. No entanto, ao realizar-se uma análise numérica comparativa entre os modelos em relação às medidas avaliadas, observa-se que o modelo NICE convolucional demonstra valores ainda mais baixos em comparação ao modelo NICE denso. Esse resultado enfatiza a eficácia do modelo convolucional em relação à sua contraparte densa na representação das características específicas das curvas referente ao *cluster* considerado e na obtenção de um desempenho ainda mais otimizado.

Figura 37 – Análise KL, WD, JS - *Cluster 2* (Cenário 4 *clusters*).

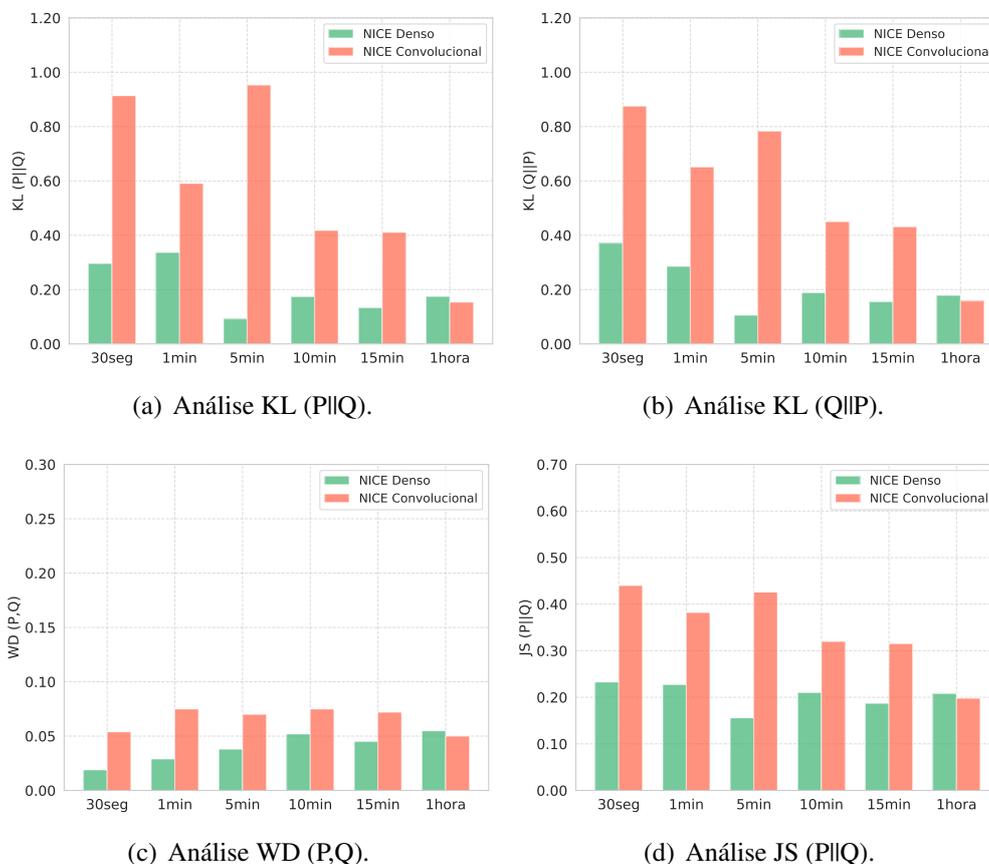


Fonte: Elaborado pela autora (2023).

Já para o *Cluster 3*, analisando-se a Figura 38, observa-se que para granularidades de 30 segundos e 1 minuto, ambos os modelos apresentam valores mais altos nas medidas, sugerindo um desempenho inferior em capturar as nuances das curvas de iluminação pública. Isso pode ser atribuído à complexidade das curvas com degraus distintos durante o dia e a noite.

Outro fator que deve ser considerado em relação a granularidade temporal é que à medida que a granularidade aumenta, observa-se uma melhoria no desempenho de ambos os modelos. Os valores das medidas diminuem, indicando que os modelos são mais eficazes em representar as diferenças e semelhanças entre as curvas geradas e as curvas reais.

Figura 38 – Análise KL, WD, JS - *Cluster 3* (Cenário 4 *clusters*).

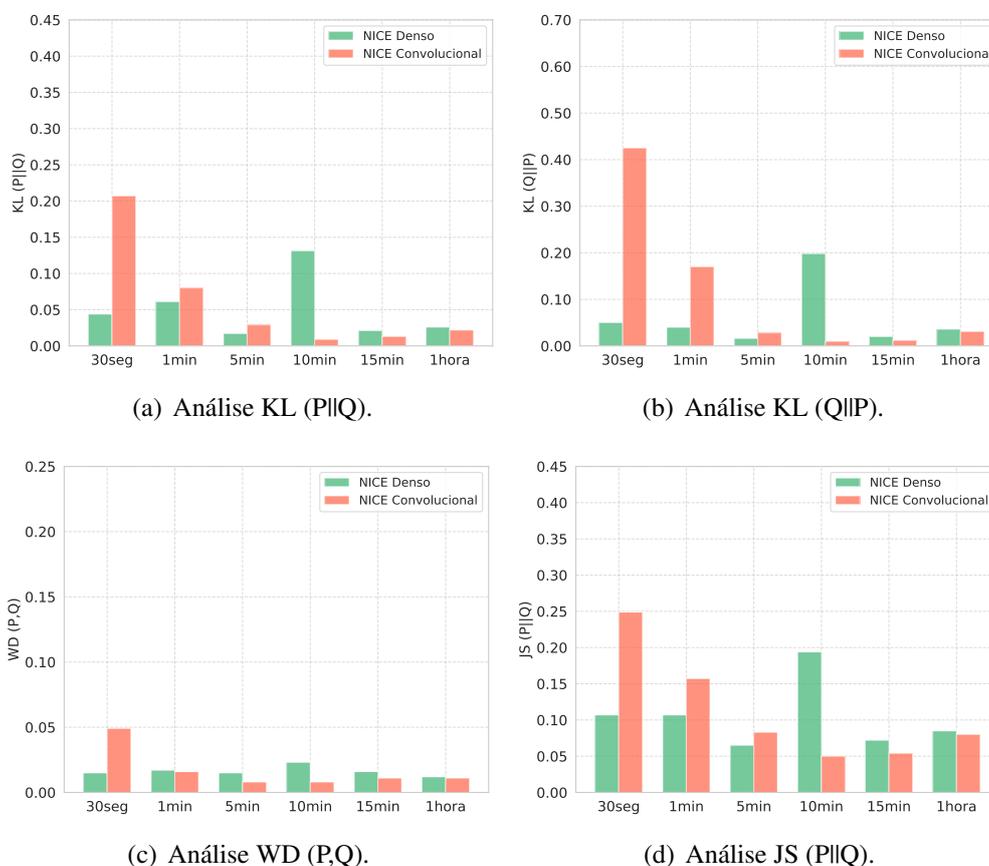


Fonte: Elaborado pela autora (2023).

Em comparação ao modelos propostos, tem-se que o modelo NICE denso demonstra um desempenho superior em comparação com o modelo NICE convolucional em todas as granularidades consideradas para este *cluster* em questão. Isso sugere que as redes neurais com camadas densas são mais eficazes em capturar as características intrínsecas e nuances das curvas de iluminação pública com degraus distintos, como já discutido nas análises similares.

Analisando-se a Figura 39, tem-se que análise dos dados e das medidas consideradas, indicam que ambos os modelos têm um desempenho satisfatório na representação das curvas no *Cluster 4*, independentemente da granularidade. No entanto, o modelo NICE denso tende a ter uma vantagem sobressalente, para granularidades mais finas.

Figura 39 – Análise KL, WD, JS - *Cluster 4* (Cenário 4 *clusters*).



Fonte: Elaborado pela autora (2023).

Além disso, o modelo NICE convolucional destaca-se com valores inferiores para granularidades maiores, indicando uma melhor capacidade de alinhamento das distribuições.

Já os resultados numéricos relacionados a análise da granularidade temporal do cenário quatro *clusters* apresentados nas Figuras 36, 37, 38 e 39 estão detalhados na Tabela 6 para o modelo NICE denso e na Tabela 7 para o modelo NICE convolucional.

Tabela 6 – NICE denso: Valores KL, WD, JS (Cenário 4 *clusters*).

<i>Cluster</i>	<i>Medida</i>	30seg	1min	5min	10min	15min	1hora
1	<i>KL (P Q)</i>	0,194	0,410	0,117	0,066	0,021	0,040
	<i>KL (Q P)</i>	0,371	0,194	0,082	0,093	0,020	0,038
	<i>WD</i>	0,044	0,032	0,010	0,025	0,016	0,009
	<i>JS (P Q)</i>	0,241	0,242	0,151	0,137	0,186	0,099
2	<i>KL (P Q)</i>	0,200	0,016	0,019	0,164	0,059	0,060
	<i>KL (Q P)</i>	0,212	0,011	0,014	0,168	0,059	0,119
	<i>WD</i>	0,042	0,007	0,008	0,053	0,008	0,012
	<i>JS (P Q)</i>	0,224	0,055	0,062	0,202	0,119	0,132
3	<i>KL (P Q)</i>	0,296	0,337	0,092	0,174	0,133	0,175
	<i>KL (Q P)</i>	0,371	0,286	0,106	0,188	0,155	0,179
	<i>WD</i>	0,019	0,029	0,038	0,052	0,045	0,055
	<i>JS (P Q)</i>	0,233	0,227	0,156	0,210	0,187	0,208
4	<i>KL (P Q)</i>	0,044	0,061	0,017	0,131	0,021	0,026
	<i>KL (Q P)</i>	0,050	0,040	0,016	0,198	0,020	0,036
	<i>WD</i>	0,015	0,017	0,015	0,023	0,016	0,012
	<i>JS (P Q)</i>	0,107	0,107	0,065	0,194	0,072	0,085

Fonte: Elaborado pela autora (2023).

Tabela 7 – NICE convolucional: Valores KL, WD, JS (Cenário 4 *clusters*).

<i>Cluster</i>	<i>Medida</i>	30seg	1min	5min	10min	15min	1hora
1	<i>KL (P Q)</i>	0,202	0,047	0,074	0,072	0,115	0,076
	<i>KL (Q P)</i>	0,603	0,079	0,058	0,054	0,075	0,083
	<i>WD</i>	0,041	0,021	0,015	0,023	0,012	0,032
	<i>JS (P Q)</i>	0,251	0,115	0,124	0,066	0,148	0,187
2	<i>KL (P Q)</i>	0,153	0,138	0,022	0,021	0,021	0,050
	<i>KL (Q P)</i>	0,198	0,221	0,022	0,018	0,022	0,108
	<i>WD</i>	0,029	0,030	0,013	0,009	0,010	0,009
	<i>JS (P Q)</i>	0,197	0,187	0,074	0,153	0,069	0,125
3	<i>KL (P Q)</i>	0,914	0,591	0,953	0,418	0,411	0,154
	<i>KL (Q P)</i>	0,875	0,651	0,783	0,450	0,431	0,158
	<i>WD</i>	0,054	0,075	0,070	0,075	0,072	0,050
	<i>JS (P Q)</i>	0,440	0,382	0,426	0,320	0,315	0,198
4	<i>KL (P Q)</i>	0,207	0,080	0,029	0,009	0,013	0,022
	<i>KL (Q P)</i>	0,425	0,170	0,028	0,010	0,012	0,031
	<i>WD</i>	0,049	0,016	0,008	0,008	0,011	0,011
	<i>JS (P Q)</i>	0,249	0,157	0,083	0,050	0,054	0,080

Fonte: Elaborado pela autora (2023).

4.7 TESTES ADICIONAIS E VARIAÇÕES DE MODELOS

Nesta seção, apresentam-se os testes adicionais realizados para validar a metodologia proposta nesta tese, cujo cerne foi o desenvolvimento do modelo NICE com a incorporação de redes neurais convolucionais para geração de perfil de curvas de carga sintéticas, em comparativo à metodologia difundida na literatura.

Com o objetivo de validar a metodologia proposta, procedeu-se a uma comparação entre os resultados obtidos do modelo NICE convolucional e o NICE denso, conforme previamente apresentados na Seção 4.5. No modelo NICE com camadas densas analisado nessas respectivas seções, a função m consistia em cinco camadas, cada uma com 512 neurônios, ilustrados na Seção 4.4 na Figura 16 (a).

Além da análise do modelo NICE denso, proposto por (GE et al., 2020), foram realizados testes adicionais para avaliar o impacto das variações nos hiperparâmetros das redes neurais densas do modelo NICE denso (função m). Isso incluiu a modificação do número de camadas e neurônios em cada camada, adotando uma abordagem progressiva que iniciou-se com uma configuração simples composta por apenas 2 neurônios em uma única camada.

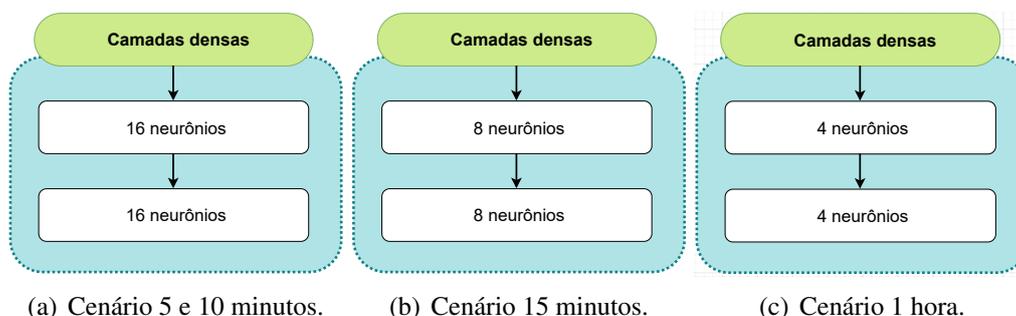
A complexidade da arquitetura da rede foi sistematicamente incrementada, aumentando o número de neurônios e adicionando camadas quando necessário, visando garantir um desempenho adequado na geração de perfis de curvas sintéticas. Esses testes foram conduzidos em relação à granularidade temporal dos dados de consumo de energia elétrica, abrangendo diferentes cenários de granularidade.

Os resultados apresentados nesta seção correspondem às configurações de hiperparâmetros que possibilitaram ao modelo gerar curvas de carga sintéticas para todos os *clusters* considerados. É importante destacar que, em diversas instâncias, o modelo enfrentou dificuldades em gerar curvas de carga satisfatórias com o número mínimo de camadas e neurônios. Em tais casos, a adição de camadas e neurônios adicionais foi necessária para capacitar o modelo a desempenhar eficazmente a geração de curvas de carga para todos os *clusters*, de modo a assegurar uma representação precisa e robusta dos dados de consumo de energia elétrica.

As diversas configurações da função m do modelo, representadas pelas arquitetu-

ras de redes neurais densas na Figura 40, englobaram uma análise das granularidades temporais, as quais foram variadas para abranger diferentes intervalos analisados neste trabalho.

Figura 40 – Hiperparâmetros da função m (rede neural).



Fonte: Elaborado pela autora (2023).

Para a configuração dos hiperparâmetros da função m na rede neural, foi adotada uma abordagem que considerou a granularidade temporal de 5 e 10 minutos, conforme ilustrado na Figura 40 (a). Nesses cenários, a configuração selecionada compreendeu o uso de 2 camadas densas e 16 neurônios em cada camada. No cenário de 15 minutos, como apresentado na Figura 40 (b), optou-se por uma configuração similar, mantendo-se 2 camadas densas, com 8 neurônios em cada camada. De modo análogo, para o cenário de 1 hora, ilustrado na Figura 40 (c), a configuração básica obtida foi de 2 camadas densas, com 4 neurônios em cada camada.

Entretanto, no contexto dos cenários de granularidade de 30 segundos e 1 minuto, observou-se que o modelo não convergiu ao empregar um número reduzido de camadas. Por conseguinte, optou-se por manter a configuração prévia apresentada na Figura 16 (a). Essa decisão foi respaldada pelo reconhecimento de que, em determinados casos, a simplificação da arquitetura neural poderia comprometer a capacidade do modelo de aprender representações adequadas para os dados. Portanto, a configuração foi mantida de modo a garantir a convergência do modelo e a qualidade das saídas geradas nos cenários mais exigentes em termos de granularidade temporal.

Os resultados obtidos referentes à aplicação dos modelos são apresentados na Figura 40, considerando a estimativa da PDF. Para cada cenário de 3 *clusters* e 4 *clusters*,

apresentam-se os valores das medidas KL, WD, e JS, nas Tabelas 8 e 9, respectivamente.

Tabela 8 – NICE denso: Valores KL, WD, JS (Cenário 3 *clusters*).

<i>Cluster</i>	<i>Medida</i>	30seg	1min	5min	10min	15min	1hora
1	<i>KL (P Q)</i>	0,075	0,172	0,035	0,034	0,053	0,067
	<i>KL (Q P)</i>	0,064	0,116	0,029	0,031	0,053	0,100
	<i>WD</i>	0,018	0,035	0,015	0,022	0,090	0,021
	<i>JS (P Q)</i>	0,129	0,180	0,189	0,083	0,156	0,140
2	<i>KL (P Q)</i>	0,033	0,131	0,170	0,152	0,068	0,097
	<i>KL (Q P)</i>	0,036	0,112	0,156	0,173	0,077	0,099
	<i>WD</i>	0,016	0,026	0,033	0,035	0,058	0,026
	<i>JS (P Q)</i>	0,093	0,169	0,197	0,199	0,099	0,158
3	<i>KL (P Q)</i>	0,613	0,443	0,141	0,322	0,134	0,201
	<i>KL (Q P)</i>	0,471	0,351	0,142	0,311	0,162	0,210
	<i>WD</i>	0,035	0,037	0,043	0,047	0,055	0,052
	<i>JS (P Q)</i>	0,343	0,297	0,186	0,267	0,057	0,224

Fonte: Elaborado pela autora (2023).

Durante os testes de variação dos hiperparâmetros das redes densas do modelo NICE, observou-se que os resultados em relação às medidas de avaliação foram, em geral, satisfatórios. No entanto, emergiram observações notáveis ao efetuar modificações no número de camadas e neurônios nas redes.

Ao reduzir o número de camadas e neurônios, constatou-se que os resultados tornaram-se propensos a ruídos na geração das curvas. Essa manifestação se refletiu sob a forma de picos não normalizados nas curvas geradas, que extrapolavam a faixa normalizada. Tais picos representam discrepâncias notáveis em relação aos dados medidos.

Por outro lado, ao aumentar o número de neurônios nas camadas, notou-se uma redução significativa desses problemas de ruído. A ampliação da capacidade da rede neural densa, mediante um número maior de neurônios, pareceu aprimorar a habilidade do modelo NICE na geração de curvas de carga mais suaves e compatíveis com os dados medidos.

Essas observações sugerem que a configuração adequada dos hiperparâmetros das redes neurais densas desempenha um papel essencial na qualidade e consistência das

Tabela 9 – NICE denso: Valores KL, WD, JS (Cenário 4 clusters).

<i>Cluster</i>	<i>Medida</i>	30seg	1min	5min	10min	15min	1hora
1	<i>KL (P Q)</i>	0,202	0,047	0,059	0,055	0,064	0,067
	<i>KL (Q P)</i>	0,603	0,079	0,033	0,079	0,057	0,094
	<i>WD</i>	0,041	0,021	0,016	0,125	0,020	0,026
	<i>JS (P Q)</i>	0,251	0,115	0,093	0,125	0,119	0,138
2	<i>KL (P Q)</i>	0,153	0,138	0,018	0,167	0,032	0,079
	<i>KL (Q P)</i>	0,198	0,221	0,015	0,171	0,036	0,150
	<i>WD</i>	0,029	0,030	0,010	0,203	0,015	0,020
	<i>JS (P Q)</i>	0,197	0,187	0,062	0,204	0,091	0,152
3	<i>KL (P Q)</i>	0,914	0,591	0,332	0,180	0,319	0,180
	<i>KL (Q P)</i>	0,875	0,651	0,372	0,195	0,335	0,196
	<i>WD</i>	0,054	0,075	0,073	0,112	0,065	0,054
	<i>JS (P Q)</i>	0,440	0,382	0,289	0,212	0,280	0,214
4	<i>KL (P Q)</i>	0,207	0,080	0,025	0,124	0,018	0,126
	<i>KL (Q P)</i>	0,425	0,170	0,035	0,188	0,022	0,175
	<i>WD</i>	0,049	0,016	0,012	0,190	0,011	0,041
	<i>JS (P Q)</i>	0,249	0,157	0,084	0,199	0,069	0,187

Fonte: Elaborado pela autora (2023).

curvas de carga geradas pelo modelo NICE. A otimização criteriosa dos hiperparâmetros, adaptada ao contexto do problema, pode se tornar uma área de investigação pertinente em futuras pesquisas.

Diante desse fato, observou-se que a incorporação de múltiplas camadas à rede neural propiciou uma capacidade superior no aprendizado das nuances e padrões sutis nos dados relacionados as curvas características imputados no modelo. Desse modo, neste caso, constatou-se que o número de neurônios alocados em cada camada exerceu impacto direto na habilidade da rede para modelar relações complexas provenientes das curvas de carga.

Outro aspecto a ser considerado nesta análise é que a redução do número de neurônios em uma camada não demonstrou simplificação na complexidade da implementação do modelo. Isso decorre do fato de que, a função m consiste em apenas uma das características e funcionalidades específicas do modelo NICE, e a variação no número de neurônios não necessariamente afeta significativamente a sua capacidade de modelar

relações intrincadas nos dados. Em vez disso, outras partes da arquitetura do modelo e configurações de hiperparâmetros também desempenham um papel importante na modelagem geral dos dados.

Por outro lado, reduzir a complexidade do modelo ao diminuir o número de camadas nem sempre se traduz em ganhos substanciais no tempo de simulação computacional e no uso de recursos de computação. Embora menos camadas possam, teoricamente, levar a um modelo mais simples, outros fatores, como o número de neurônios em cada camada e a capacidade da rede em modelar padrões complexos, também influenciam o custo computacional. Assim, a otimização da arquitetura da rede neural deve ser abordada considerando uma variedade de fatores para garantir um equilíbrio adequado entre a complexidade do modelo e o uso eficiente dos recursos computacionais.

Essas constatações evidenciam que redes com um número menor de neurônios e camadas tenderam a gerar saídas mais suscetíveis a ruídos e desvios, enquanto um aumento no número de neurônios culminou na geração de curvas de carga mais suaves e consistentes com as curvas medidas. Constata-se que a seleção e configuração adequada dos hiperparâmetros da rede é fundamental para otimizar o desempenho do modelo.

Além disso, os resultados obtidos ressaltam a importância das considerações práticas na construção dos modelos generativos. Desse modo, a complexidade da arquitetura da rede deve ser adequada à natureza do problema em análise e à quantidade de dados analisados, bem como dos recursos computacionais disponíveis. Portanto, o ajuste dos hiperparâmetros é uma das principais etapas na construção dos modelos computacionais, e portanto, deve-se levar em conta a complexidade do problema, visando aprimorar a capacidade do modelo em propiciar resultados de qualidade e alta performance.

4.8 ANÁLISES GERAIS

Com base na análise abrangente realizada ao longo deste capítulo, que se concentrou na geração de curvas de carga sintéticas utilizando modelos NICE denso e convolucional em diferentes granularidades temporais, é possível extrair algumas conclusões. Na maioria dos cenários analisados, observa-se que ambos os modelos evidenciam um desempenho notável na representação das curvas de carga. A avaliação das medidas evidencia o desempenho dos modelos, destacando variações significativas conforme a

granularidade temporal, mostrando a adaptação dos modelos às características intrínsecas das curvas de carga analisadas.

No contexto da representação de *clusters* dedicados a curvas específicas, como as relacionadas à iluminação pública, a superioridade do desempenho do modelo NICE denso em comparação com o modelo NICE convolucional é evidenciada nas diferentes granularidades temporais avaliadas. Este fator pode ser atribuído à notável capacidade do modelo NICE denso em extrair e representar de maneira eficaz as características intrínsecas e nuances presentes nas curvas de iluminação pública. Essa eficácia é particularmente evidente ao lidar com curvas que apresentam transições abruptas ou degraus distintos, evidenciando a adaptabilidade e poder de representação das redes neurais densas em situações onde há variações mais complexas nas curvas de carga.

Ao considerar granularidades mais finas, como 30 segundos e 1 minuto, os valores das medidas tendem a aumentar. Essa observação reflete a complexidade das curvas, que podem incluir transições rápidas nas incidências dos valores medidos, aumentando a aleatoriedade da curva de carga. Em contrapartida, ao adotar granularidades mais elevadas, como 5 minutos, 10 minutos e 15 minutos, os valores das medidas tendem a diminuir. Isso indica que o modelo está capturando de forma mais precisa as discrepâncias e semelhanças entre as curvas geradas e as curvas medidas.

Um aspecto interessante observado é que, em determinadas circunstâncias, ao adotar uma granularidade de 1 hora, observa-se que há indícios de perda de informações pelo modelo ao analisar as medidas referentes à estimativa das PDFs. Esta observação está alinhada com a compreensão de que granularidades mais amplas podem resultar em uma representação menos detalhada dos padrões subjacentes nos dados.

Outro fator a ser considerado na análise dos resultados é a avaliação do desempenho geral dos modelos NICE em relação à dimensão temporal e à qualidade das gerações desses modelos nos diversos cenários considerados. Essa avaliação permitiu a avaliação do impacto da granularidade temporal ao que se refere à eficiência e eficácia dos modelos, auxiliando na escolha da granularidade mais apropriada em diferentes contextos de aplicabilidade da geração de curvas de cargas sintéticas.

Para otimizar a seleção da granularidade temporal e do modelo mais adequado, é essencial considerar não apenas a análise das medidas discutidas nesta seção, mas

também avaliar a qualidade das curvas geradas por cada modelo. É fundamental garantir que essas curvas atendam aos requisitos da aplicação. Caso necessário, recomenda-se realizar o ajuste dos hiperparâmetros dos modelos para alcançar os melhores resultados. Esta abordagem proporcionará um melhor suporte para a tomada de decisões embasadas em cenários de simulações e diversas necessidades de perfis de curva de carga sintéticas.

5 CONSIDERAÇÕES FINAIS

Esta pesquisa de doutorado teve como o objetivo principal o desenvolvimento de perfis sintéticos de carga que representassem com precisão os perfis medidos de curva de carga de consumo de energia. Isso foi alcançado por meio da aplicação de metodologias de análise de dados provenientes das redes elétricas inteligentes. Diante das lacunas identificadas, este trabalho buscou soluções para superar os desafios encontrados, tais como a escassez de dados para análise, preocupações com a segurança e privacidade dos dados, bem como os custos substanciais associados à digitalização em larga escala e aos recursos computacionais exigidos.

Para superar os desafios identificados ao longo da pesquisa desta tese, optou-se pela incorporação de dados sintéticos. Nesse contexto, foram empregados modelos de análise de dados baseado no processo KDD, incluindo a implementação de um modelo generativo profundo baseado em fluxo, para gerar perfis de curvas de carga sintéticas. Dessa forma, introduziu-se uma abordagem inovadora, utilizando o modelo NICE com inserção de camadas convolucionais, uma das principais contribuições desta pesquisa.

Até onde se sabe, esta contribuição representa a primeira abordagem do modelo proposto no contexto de geração de perfil de curvas sintéticas unidimensionais, usando dados medidos (reais) de curvas de carga. Estima-se que esta abordagem possa enriquecer a compreensão dos estudos relacionados aos padrões de consumo de energia, propiciando ações a serem incorporadas para melhorias na eficiência e sustentabilidade energética em diferentes aplicações do setor elétrico. Isso se torna especialmente relevante em um cenário de transição energética e digitalização.

A metodologia proposta foi validada por meio da comparação entre o modelo NICE com camadas convolucionais, proposto neste trabalho, e o modelo NICE com camadas densas presente na literatura. Essa validação envolveu a análise de dados em dois cenários: um com três *clusters* e outro com quatro *clusters*.

Durante a análise, os modelos propostos foram comparados com base na avaliação das curvas geradas por cada um deles, utilizando a estimativa da PDF dos dados medidos e gerados. Para avaliar a similaridade da estimativa da PDF de cada modelo em relação à distribuição dos dados medidos, foram consideradas as medidas de similaridade,

incluindo a divergência KL, JS e WD.

No cenário com três *clusters*, os modelos NICE com camadas densas e convolucionais demonstraram um desempenho semelhante em relação às medidas consideradas. No entanto, o modelo convolucional se destacou nos *Clusters* 1 e 2, enquanto o modelo denso teve um desempenho superior no *Cluster* 3. Ao expandir para a análise de quatro *Clusters*, observou-se uma maior diversificação nos resultados. O modelo NICE denso demonstrou ser mais eficaz na maioria das medidas, refletindo sua capacidade de representar com precisão as distribuições de probabilidade subjacentes.

A divergência nos resultados entre os cenários de três e quatro *clusters* pode ser atribuída a diversos fatores, como a crescente complexidade dos *clusters*, a ampla variação *intracluster*, a distribuição de dados entre os *clusters* e a sensibilidade das medidas às variações nos *clusters*. Essa análise destaca a influência crítica do número de *clusters* e a importância de escolher a configuração do modelo mais apropriada com base nos dados e nos objetivos do projeto.

Em geral, a validação sistemática dos perfis sintéticos de carga demonstrou a capacidade dos modelos em reproduzir com precisão a diversidade observada nos perfis de carga medidos. Notavelmente, foi observada uma similaridade no desempenho entre os modelos NICE densos e convolucionais em relação às medidas consideradas. Na maioria dos casos, ambos os modelos apresentaram valores reduzidos, indicando uma concordância entre as distribuições de probabilidade dos dados medidos e sintéticos.

Ademais, o modelo convolucional exibiu uma habilidade notável para mapear e replicar a variabilidade intrínseca presente nos perfis de carga de consumo, resultando na geração de perfis diversos. Por outro lado, o modelo denso demonstrou capacidade de capturar nuances complexas nas distribuições de probabilidade das curvas de carga.

Além disso, este trabalho conduziu uma análise sistemática da granularidade temporal das curvas de carga sintéticas. Foram considerados diferentes intervalos de granularidade, como 30 segundos, 1 minuto, 5 minutos, 10 minutos, 15 minutos e 1 hora. O propósito foi examinar o impacto da granularidade sobre a acurácia dos modelos, bem como sobre as medidas de divergência entre os perfis de carga medidos e os sintéticos, com o intuito de compreender o comportamento dos diferentes intervalos temporais na precisão das curvas de carga sintéticas.

Destaca-se a observação de que a granularidade temporal tem uma influência nas medidas consideradas. Constatou-se que as medidas exibem um desempenho inferior à medida que a granularidade diminui. Tal fato denota um aumento na disparidade entre as distribuições dos dados medidos e dos dados sintéticos, à medida que a granularidade se reduz. Dessa forma, para as granularidades analisadas, observou-se que quanto menor o intervalo temporal entre as amostras, maior foi a discrepância entre os dados medidos e sintéticos.

Essa conjuntura pode estar relacionada ao fato que uma granularidade temporal mais fina permite a captura de variações mais detalhadas nos perfis de carga, resultando, por conseguinte, em distribuições de probabilidade mais distintas. Esses resultados ressaltam a importância de escolher a granularidade temporal adequada ao utilizar o modelo NICE em diferentes aplicações e, portanto, equiparadas ao objetivo pleiteado. Além disso, deve-se atentar-se às configurações adequadas dos hiperparâmetros do modelo conforme a aplicabilidade pretendida para se obter resultados mais precisos e úteis.

Como resultado, esta pesquisa não só amplia o escopo da análise de dados em redes elétricas, mas também apresenta soluções práticas para os desafios enfrentados pelo setor elétrico, ressaltando a importância da análise de dados e do desenvolvimento de curvas de carga sintéticas como instrumentos para otimizar a gestão eficaz da energia.

A ideia central, é que ao adotar essas abordagens, seja possível impulsionar práticas que visam usar a energia de forma mais eficiente, promovendo a sustentabilidade e melhorando o controle energético. Com base nos resultados obtidos, a expectativa é que essas técnicas possam ser aplicadas não apenas em universidades, mas também em outros sistemas de gerenciamento de energia, fomentando a adoção de soluções mais eficientes e sustentáveis em todo o setor.

Em resumo, esta tese de doutorado apresentou uma metodologia inovadora para abordar os desafios críticos na análise de dados de consumo de energia. Três pontos principais da metodologia se destacam como contribuições significativas:

- **Geração de dados sintéticos:** a utilização de dados sintéticos é um recurso essencial para mitigar a escassez de dados reais. Estima-se que a metodologia proposta para a geração de curvas de cargas sintéticas desempenhe um papel

fundamental, especialmente, no contexto em testes e no desenvolvimento de pesquisas exploratórias, permitindo analisar padrões de consumo de energia com mais profundidade e precisão, além de promover aplicações significativas na perspectiva de gestão energética, entre outros domínios do sistema elétrico de potência;

- **Desenvolvimento modelo NICE:** abrangeu a incorporação de redes neurais com camadas convolucionais, que visou expandir a capacidade de detecção das características espaço-temporais, representando uma evolução da metodologia existente na literatura. Enquanto o modelo NICE com camadas densas se destacou na abordagem de nuances específicas dos dados, a introdução das camadas convolucionais trouxe uma característica distintiva, permitindo uma compreensão mais abrangente das variações temporais. A comparação entre esses modelos ilustrou a versatilidade de cada um em diferentes cenários, além de evidenciar sua capacidade de se complementarem em diferentes aplicações. Esse desenvolvimento destacou a relevância da abordagem inovadora, contribuindo de maneira significativa para a análise exploratória de modelos generativos no contexto desta pesquisa;
- **Análise do impacto da granularidade temporal:** a seleção criteriosa dos intervalos de tempo desempenha um papel de importância na geração de perfis de curva de carga sintética. Observou-se que a utilização de intervalos de tempo mais curtos, como 30 segundos e 1 minuto, permitiu a captura de variações altamente detalhadas nas curvas de carga, enquanto a adoção de intervalos mais longos, como 10 minutos e 15 minutos, tendeu a suavizar as variações de curto prazo. A metodologia evidenciou que a decisão relativa à granularidade temporal exerce uma influência direta na eficácia e eficiência na geração de curvas de carga sintéticas. Esse fato ressalta a magnitude da escolha apropriada no âmbito da análise de dados.

Os aspectos fundamentais delineados nesta tese representam avanços notáveis no campo da análise de dados nas redes elétricas inteligentes. Esses progressos têm como objetivo fornecer suporte efetivo para soluções práticas e estratégicas. A expectativa é que essas contribuições possam ser aplicadas com eficácia em várias aplicações, promovendo

assim a eficiência e sustentabilidade energética em meio às mudanças paradigmáticas que o setor elétrico está enfrentando.

5.1 SUGESTÕES DE DESENVOLVIMENTOS FUTUROS

Com base nos resultados e nas contribuições desta tese de doutoramento, identificam-se diferentes áreas promissoras para futuros desenvolvimentos e aprimoramentos. As sugestões a seguir destacam direções que podem ampliar o escopo desta pesquisa:

- **Aprimoramento dos modelos propostos:** aprimorar os modelos generativos propostos de modo a garantir que sejam aplicáveis em uma ampla gama de situações e cenários. É essencial avaliar continuamente o desempenho do modelo com base nos diferentes aspectos, como o tamanho do conjunto de dados de treinamento, a arquitetura, os hiperparâmetros e as técnicas de regularização. Além disso, encontrar métodos de treinamento mais eficientes e explorar novas arquiteturas pode propiciar modelos mais robustos e confiáveis;
- **Comparação de modelos de geração de curvas de carga:** realizar uma avaliação quantitativa da fidelidade na geração de curvas de carga por meio da implementação de modelos mais elementares, incluindo abordagens lineares, senoidais, aleatórias, de média móvel, exponenciais e de degrau. De modo a comparar estatisticamente as curvas de carga geradas pelos modelos mais simples e aquelas produzidas pelo modelo NICE proposto, evidenciando a eficácia comparativa em relação a abordagens mais simplificadas existentes na literatura.
- **Incorporação de características adicionais em perfis de carga:** inserir diferentes cenários de análise na geração de curvas de cargas sintética, considerando a inclusão de aspectos que se refere a sazonalidade e variações temporais específicas, como picos de consumo durante o verão ou inverno. Além disso, a integração de informações sobre sistemas fotovoltaicos e veículos elétricos nos perfis de carga pode ser um passo importante. Essas adições permitiriam uma representação mais abrangente e precisa dos padrões de consumo de energia;

- **Ampliação do escopo para outras grandezas elétricas:** estender a metodologia para outras grandezas elétricas de relevância. Isso inclui análise da potência reativa, do fator de potência, que é uma das variáveis que auxilia na análise da eficiência do sistema elétrico, e a potência aparente. Além disso, a análise do neutro em sistemas elétricos complexos pode ser explorada. Tal extensão permitiria uma compreensão mais completa do comportamento dos sistemas elétricos e abrangendo novas possibilidades de gestão em diversos contextos e setores.
- **Aprendizagem por transferência e diversificação de dados:** incorporar a metodologia deste estudo por meio da abordagem de aprendizagem por transferência e a diversificação de dados, envolvendo a integração de conjuntos de dados de diferentes localidades e redes elétricas, mesmo aquelas que não possuem infraestrutura de medição avançada. Essa abordagem pode considerar as variações sazonais em cada localidade e aplicá-las aos perfis de carga. Isso é particularmente útil para a gestão de energia em edifícios que atualmente não possuem dados em quantidade suficiente para treinar eficazmente o modelo generativo. Dessa forma, a pesquisa pode buscar soluções mais abrangentes e acessíveis para o gerenciamento de energia em diferentes contextos.

REFERÊNCIAS

- ABBASI, A. R. Comparison parametric and non-parametric methods in probabilistic load flow studies for power distribution networks. *Electrical Engineering*, Springer, v. 104, n. 6, p. 3943–3954, 2022.
- AHMAD, T.; ZHANG, H.; YAN, B. A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society*, Elsevier, v. 55, p. 102052, 2020.
- AHMED, E. et al. The role of big data analytics in internet of things. *Computer Networks*, v. 129, p. 459–471, 2017. ISSN 1389-1286. Special Issue on 5G Wireless Networks for IoT and Body Sensors. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1389128617302591>>.
- AHMED, M.; SERAJ, R.; ISLAM, S. M. S. The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, v. 9, n. 8, 2020. ISSN 2079-9292. Disponível em: <<https://www.mdpi.com/2079-9292/9/8/1295>>.
- ALQAHTANI, H.; KAVAKLI-THORNE, M.; KUMAR, G. Applications of generative adversarial networks (gans): An updated review. *Archives of Computational Methods in Engineering*, Springer, v. 28, p. 525–552, 2021.
- ARJOVSKY, M.; CHINTALA, S.; BOTTOU, L. Wasserstein generative adversarial networks. In: PMLR. *International conference on machine learning*. [S.l.], 2017. p. 214–223.
- BAASCH, G.; ROUSSEAU, G.; EVINS, R. A conditional generative adversarial network for energy use in multiple buildings using scarce data. *Energy and AI*, v. 5, p. 100087, 2021. ISSN 2666-5468.
- BAHRI, R. S. et al. Data mining techniques to uncovering customer segments: K-means clustering using the elbow method approach in medium-scale grocery. In: *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*. [S.l.: s.n.], 2023. p. 262–266.
- BAYRAM, I. S.; USTUN, T. S. A survey on behind the meter energy management systems in smart grid. *Renewable and Sustainable Energy Reviews*, v. 72, p. 1208–1232, 2017. ISSN 1364-0321. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1364032116306852>>.
- BENDAOU, N. M. M.; FARAH, N.; AHMED, S. B. Comparing generative adversarial networks architectures for electricity demand forecasting. *Energy and Buildings*, v. 247, p. 111152, 2021. ISSN 0378-7788.

- BENDAOUD, N. M. M.; FARAH, N.; Ben Ahmed, S. Comparing generative adversarial networks architectures for electricity demand forecasting. *Energy and Buildings*, v. 247, p. 111152, 2021. ISSN 0378-7788. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378778821004369>>.
- BISHOP, C. M. et al. *Neural networks for pattern recognition*. [S.l.]: Oxford university press, 1995.
- BUTT, O. M.; ZULQARNAIN, M.; BUTT, T. M. Recent advancement in smart grid technology: Future prospects in the electrical power network. *Ain Shams Engineering Journal*, Elsevier, v. 12, n. 1, p. 687–695, 2021.
- CARO, F. D. et al. A review of the enabling methodologies for knowledge discovery from smart grids data. *Energies*, v. 13, n. 24, 2020. ISSN 1996-1073. Disponível em: <<https://www.mdpi.com/1996-1073/13/24/6579>>.
- CEMBRANEL, S. S. et al. A short review on data mining techniques for electricity customers characterization. In: *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*. [S.l.: s.n.], 2019. p. 194–199.
- CHOLLET, F. *keras*. [S.l.]: GitHub, 2015. <<https://github.com/fchollet/keras>>.
- CYPRIANO, J. G. I. et al. Centro de operações e monitoramento do sistema elétrico da unicamp. In: _____. *Campus Sustentável: um modelo de inovação em gestão energética para a América Latina e o Caribe*. Campinas, São Paulo, Brasil: Editora da Unicamp, 2021. cap. 6. ISBN 978-65-86214-61-1.
- DAVIES, D. L.; BOULDIN, D. W. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1, n. 2, p. 224–227, 1979.
- DEEPA, N. et al. A survey on blockchain for big data: Approaches, opportunities, and future directions. *Future Generation Computer Systems*, v. 131, p. 209–226, 2022. ISSN 0167-739X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167739X22000243>>.
- DILEEP, G. A survey on smart grid technologies and applications. *Renewable Energy*, v. 146, p. 2589–2625, 2020. ISSN 0960-1481. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0960148119312790>>.
- DINH, L.; KRUEGER, D.; BENGIO, Y. Nice: Non-linear independent components estimation. *arXiv preprint arXiv:1410.8516*, 2014.
- DINH, L.; SOHL-DICKSTEIN, J.; BENGIO, S. Density estimation using real nvp. *arXiv preprint arXiv:1605.08803*, 2016.

- ENDRES, D.; SCHINDELIN, J. A new metric for probability distributions. *IEEE Transactions on Information Theory*, v. 49, n. 7, p. 1858–1860, 2003.
- ESCOBAR, J. J. M. et al. A comprehensive review on smart grids: Challenges and opportunities. *Sensors*, v. 21, n. 21, 2021. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/21/21/6978>>.
- EZUGWU, A. E. Nature-inspired metaheuristic techniques for automatic clustering: a survey and performance study. *SN Applied Sciences*, Springer, v. 2, p. 1–57, 2020.
- FAN, C. et al. A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in Energy Research*, Frontiers Media SA, v. 9, p. 652801, 2021.
- FAVERO, L. P.; BELFIORE, P. *Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®*. [S.l.]: Elsevier Brasil, 2017.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. *AI magazine*, v. 17, n. 3, p. 37–37, 1996.
- FRAWLEY, W. J.; PIATETSKY-SHAPIRO, G.; MATHEUS, C. J. Knowledge discovery in databases: An overview. *AI magazine*, v. 13, n. 3, p. 57–57, 1992.
- GAO, Y. et al. Deep learning and transfer learning models of energy consumption forecasting for a building with poor information data. *Energy and Buildings*, v. 223, p. 110156, 2020. ISSN 0378-7788. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378778819324843>>.
- GARCIA, Y. V. et al. Methodology to implement a microgrid in a university campus. *Applied Sciences*, v. 12, n. 9, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/9/4563>>.
- GE, L. et al. Modeling daily load profiles of distribution network for scenario generation using flow-based generative network. *IEEE Access*, v. 8, p. 77587–77597, 2020.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press, 2016.
- GRASSO, F. et al. Artificial load profiles and pv generation in renewable energy communities using generative adversarial networks. In: *2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON)*. [S.l.: s.n.], 2022. p. 709–714.
- GU, Y. et al. Gan-based model for residential load generation considering typical consumption patterns. In: *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. [S.l.: s.n.], 2019. p. 1–5.

HAN, J.; PEI, J.; TONG, H. *Data mining: concepts and techniques*. [S.l.]: Morgan kaufmann, 2022.

HAYKIN, S. *Redes neurais: princípios e prática*. [S.l.]: Bookman Editora, 2001.

HE, Y. et al. Short-term power load probability density forecasting based on quantile regression neural network and triangle kernel function. *Energy*, v. 114, p. 498–512, 2016. ISSN 0360-5442. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0360544216311264>>.

HENRIQUE, L. F. et al. Optimal siting and sizing of distributed energy resources in a smart campus. *Electric Power Systems Research*, v. 217, p. 109095, 2023. ISSN 0378-7796. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378779622011440>>.

HU, S. et al. Scenario forecasting for wind power using flow-based generative networks. *Energy Reports*, v. 7, p. 369–377, 2021. ISSN 2352-4847. 2021 The 4th International Conference on Electrical Engineering and Green Energy. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2352484721006387>>.

HU, Y. et al. *MultiLoad-GAN: A GAN-Based Synthetic Load Group Generation Method Considering Spatial-Temporal Correlations*. [S.l.]: arXiv, 2022.

IKOTUN, A. M. et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, v. 622, p. 178–210, 2023. ISSN 0020-0255. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0020025522014633>>.

IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: PMLR. *International conference on machine learning*. [S.l.], 2015. p. 448–456.

JAIN, A. K. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, v. 31, n. 8, p. 651–666, 2010. ISSN 0167-8655. Award winning papers from the 19th International Conference on Pattern Recognition (ICPR). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167865509002323>>.

JANGID, B.; MATHRURIA, P.; GUPTA, V. Load profile segmentation of various load categories using clustering. In: *2021 IEEE 2nd International Conference On Electrical Power and Energy Systems (ICEPES)*. [S.l.: s.n.], 2021. p. 1–6.

JAVED, H. et al. Optimal energy management of a campus microgrid considering financial and economic analysis with demand response strategies. *Energies*, MDPI, v. 14, n. 24, p. 8501, 2021.

KARO, I. M. K.; MAULANAADHINUGRAHA, K.; HUDA, A. F. A cluster validity for spatial clustering based on davies bouldin index and polygon dissimilarity function. In: *2017 Second International Conference on Informatics and Computing (ICIC)*. [S.l.: s.n.], 2017. p. 1–6.

KAUR, D. et al. Energy forecasting in smart grid systems: A review of the state-of-the-art techniques. *arXiv preprint arXiv:2011.12598*, 2020.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

KINGMA, D. P.; DHARIWAL, P. Glow: Generative flow with invertible 1x1 convolutions. *Advances in neural information processing systems*, v. 31, 2018.

KOBYZEV, I.; PRINCE, S. J.; BRUBAKER, M. A. Normalizing flows: An introduction and review of current methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 43, n. 11, p. 3964–3979, 2021.

KONTOGIANNIS, D. et al. Structural ensemble regression for cluster-based aggregate electricity demand forecasting. *Electricity*, v. 3, n. 4, p. 480–504, 2022. ISSN 2673-4826. Disponível em: <<https://www.mdpi.com/2673-4826/3/4/25>>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: PEREIRA, F. et al. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012. v. 25. Disponível em: <https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, ACM New York, NY, USA, v. 60, n. 6, p. 84–90, 2017.

KULLBACK, S. *Information theory and statistics*. [S.l.]: Courier Corporation, 1997.

KUMARI, A.; TANWAR, S. Secure data analytics for smart grid systems in a sustainable smart city: Challenges, solutions, and future directions. *Sustainable Computing: Informatics and Systems*, v. 28, p. 100427, 2020. ISSN 2210-5379. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2210537920301542>>.

LAUBE, K. A.; ZELL, A. Shufflenasnets: Efficient cnn models through modified efficient neural architecture search. In: IEEE. *2019 International Joint Conference on Neural Networks (IJCNN)*. [S.l.], 2019. p. 1–6.

LI, J. et al. Energy data generation with wasserstein deep convolutional generative adversarial networks. *Energy*, v. 257, p. 124694, 2022. ISSN 0360-5442.

- LIAO, G. et al. Wind power prediction errors model and algorithm based on non-parametric kernel density estimation. In: *2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*. [S.l.: s.n.], 2015. p. 1864–1868.
- LIN, J. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, v. 37, n. 1, p. 145–151, 1991.
- MACQUEEN, J. et al. Some methods for classification and analysis of multivariate observations. In: OAKLAND, CA, USA. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. [S.l.], 1967. v. 1, n. 14, p. 281–297.
- MARJANI, M. et al. Big iot data analytics: architecture, opportunities, and open research challenges. *iee access*, IEEE, v. 5, p. 5247–5261, 2017.
- MARUTHO, D. et al. The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. In: *2018 International Seminar on Application for Technology of Information and Communication*. [S.l.: s.n.], 2018. p. 533–538.
- MERINO, Z. D.; FARMER, J.; JACOBS, D. J. Probability density estimation through nonparametric adaptive partitioning and stitching. *Algorithms*, v. 16, n. 7, 2023. ISSN 1999-4893. Disponível em: <<https://www.mdpi.com/1999-4893/16/7/310>>.
- MICHALAKOPOULOS, V. et al. A machine learning-based framework for clustering residential electricity load profiles to enhance demand response programs. *arXiv preprint arXiv:2310.20367*, 2023.
- MICHALIK, P.; ŠTOFA, J.; ZOLOTOVÁ, I. Concept definition for big data architecture in the education system. In: *2014 IEEE 12th International Symposium on Applied Machine Intelligence and Informatics (SAMII)*. [S.l.: s.n.], 2014. p. 331–334.
- MOHAMMADPOOR, M.; TORABI, F. Big data analytics in oil and gas industry: An emerging trend. *Petroleum*, v. 6, n. 4, p. 321–328, 2020. ISSN 2405-6561. SI: Artificial Intelligence (AI), Knowledge-based Systems (KBS), and Machine Learning (ML). Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405656118301421>>.
- MÜLLER, T. et al. Neural importance sampling. *ACM Transactions on Graphics (ToG)*, ACM New York, NY, USA, v. 38, n. 5, p. 1–19, 2019.
- MUQEET, H. A. et al. An energy management system of campus microgrids: State-of-the-art and future challenges. *Energies*, v. 14, n. 20, 2021. ISSN 1996-1073. Disponível em: <<https://www.mdpi.com/1996-1073/14/20/6525>>.

ODAIBO, S. Tutorial: Deriving the standard variational autoencoder (vae) loss function. *arXiv preprint arXiv:1907.08956*, 2019.

PACHECO, F. et al. Generation of synthetic data for the analysis of the physical stability of tailing dams through artificial intelligence. *Mathematics*, v. 10, n. 23, 2022. ISSN 2227-7390. Disponível em: <<https://www.mdpi.com/2227-7390/10/23/4396>>.

PAN, Z. et al. Data-driven ev load profiles generation using a variational auto-encoder. *Energies*, v. 12, n. 5, 2019. ISSN 1996-1073.

PANARETOS, V. M.; ZEMEL, Y. Statistical aspects of wasserstein distances. *Annual review of statistics and its application*, Annual Reviews, v. 6, p. 405–431, 2019.

PANDA, D. K.; DAS, S. Smart grid architecture model for control, optimization and data analytics of future power networks with more renewable energy. *Journal of Cleaner Production*, v. 301, p. 126877, 2021. ISSN 0959-6526.

PANDA, S. et al. An insight into the integration of distributed energy resources and energy storage systems with smart distribution networks using demand-side management. *Applied Sciences*, v. 12, n. 17, 2022. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/12/17/8914>>.

PAPADOPOULOS, T. A. et al. Study of electricity load profiles in university campuses: The case study of democritus university of thrace. In: *Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016)*. [S.l.: s.n.], 2016. p. 1–8.

PINCETI, A.; SANKAR, L.; KOSUT, O. Synthetic time-series load data via conditional generative adversarial networks. In: *2021 IEEE Power & Energy Society General Meeting (PESGM)*. [S.l.: s.n.], 2021. p. 1–5.

PONNUSAMY, V. K. et al. A comprehensive review on sustainable aspects of big data analytics for the smart grid. *Sustainability*, MDPI, v. 13, n. 23, p. 13322, 2021.

PRIMADIANTO, A.; LU, C.-N. A review on distribution system state estimation. *IEEE Transactions on Power Systems*, IEEE, v. 32, n. 5, p. 3875–3883, 2016.

PÉREZ-CHACÓN, R. et al. Big data analytics for discovering electricity consumption patterns in smart cities. *Energies*, v. 11, n. 3, 2018. ISSN 1996-1073. Disponível em: <<https://www.mdpi.com/1996-1073/11/3/683>>.

RABIE, A. H. et al. A new outlier rejection methodology for supporting load forecasting in smart grids based on big data. *Cluster Computing*, Springer, v. 23, p. 509–535, 2020.

RAMOS, S. et al. Typical load profiles in the smart grid context—a clustering methods comparison. In: IEEE. *2012 IEEE Power and Energy Society General Meeting*. [S.l.], 2012. p. 1–8.

RAMOS, S. et al. A data-mining-based methodology to support mv electricity customers' characterization. *Energy and Buildings*, v. 91, p. 16–25, 2015. ISSN 0378-7788. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378778815000420>>.

RAMOS, S. et al. A data-mining-based methodology to support mv electricity customers' characterization. *Energy and Buildings*, Elsevier, v. 91, p. 16–25, 2015.

RAMOS, S.; VALE, Z. Data mining techniques application in power distribution utilities. In: IEEE. *2008 IEEE/PES Transmission and Distribution Conference and Exposition*. [S.l.], 2008. p. 1–8.

RAMOS, S. et al. Data mining contributions to characterize mv consumers and to improve the suppliers-consumers settlements. In: IEEE. *2007 IEEE Power Engineering Society General Meeting*. [S.l.], 2007. p. 1–8.

RATHOR, S. K.; SAXENA, D. Energy management system for smart grid: An overview and key issues. *International Journal of Energy Research*, Wiley Online Library, v. 44, n. 6, p. 4067–4109, 2020.

RAZA, M. Q.; KHOSRAVI, A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews*, v. 50, p. 1352–1372, 2015. ISSN 1364-0321. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1364032115003354>>.

ROS, F.; RIAD, R.; GUILLAUME, S. Pdbi: A partitioning davies-bouldin index for clustering evaluation. *Neurocomputing*, v. 528, p. 178–199, 2023. ISSN 0925-2312. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0925231223000528>>.

ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, p. 53–65, 1987. ISSN 0377-0427. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0377042787901257>>.

RUTHOTTO, L.; HABER, E. An introduction to deep generative modeling. *GAMM-Mitteilungen*, Wiley Online Library, v. 44, n. 2, p. e202100008, 2021.

SALAZAR, M. et al. Data driven framework for load profile generation in medium voltage networks via transfer learning. In: *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*. [S.l.: s.n.], 2020. p. 909–913.

- SALEH, A. I.; RABIE, A. H.; ABO-AL-EZ, K. M. A data mining based load forecasting strategy for smart electrical grids. *Advanced Engineering Informatics*, v. 30, n. 3, p. 422–448, 2016. ISSN 1474-0346. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1474034616301331>>.
- Scikit-Learn Community. *scikit-learn: Machine Learning in Python*. 2023. <<https://scikit-learn.org/stable/>>.
- SESTINO, A. et al. Internet of things and big data as enablers for business digitalization strategies. *Technovation*, v. 98, p. 102173, 2020. ISSN 0166-4972. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0166497220300456>>.
- SHARMA, S.; SOOD, Y. R. Microgrids: A review of status, technologies, software tools, and issues in indian power market. *IETE Technical Review*, Taylor Francis, v. 39, n. 2, p. 411–432, 2022.
- SHUFIAN, A.; MOHAMMAD, N. Modeling and analysis of cost-effective energy management for integrated microgrids. *Cleaner Engineering and Technology*, v. 8, p. 100508, 2022. ISSN 2666-7908. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2666790822001136>>.
- SICSÚ, A. L.; SAMARTINI, A.; BARTH, N. L. *Técnicas de Machine Learning*. 1ª. ed. São Paulo: Blucher, 2023. 394 p. ISBN 978-65-5506-396-7.
- SILVA, e. a. Luiz Carlos Pereira da. *O Campus Sustentável: um modelo de inovação em gestão energética para a América Latina e o Caribe*. Campinas, São Paulo, Brasil: Editora da Unicamp, 2021. ISBN 978-65-86214-61-1.
- SILVA, W. N. et al. Generating realistic load profiles in smart grids: An approach based on nonlinear independent component estimation (nice) and convolutional layers. *Applied Energy*, v. 351, p. 121902, 2023. ISSN 0306-2619. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0306261923012667>>.
- SILVA, W. N. et al. Energy management system: a case study for a campus microgrid. In: *2022 IEEE PES Generation, Transmission and Distribution Conference and Exposition – Latin America (IEEE PES GTD Latin America)*. [S.l.: s.n.], 2022. p. 1–6.
- SINGH, R. et al. Energy system 4.0: Digitalization of the energy sector with inclination towards sustainability. *Sensors*, v. 22, n. 17, 2022. ISSN 1424-8220. Disponível em: <<https://www.mdpi.com/1424-8220/22/17/6619>>.
- SONG, L.; LI, Y.; LU, N. Profilesr-gan: A gan based super-resolution method for generating high-resolution load profiles. *IEEE Transactions on Smart Grid*, v. 13, n. 4, p. 3278–3289, 2022.

TIAN, Y.; ZHANG, Y. A comprehensive survey on regularization strategies in machine learning. *Information Fusion*, Elsevier, v. 80, p. 146–166, 2022.

TIGHTIZ, L.; YOO, J. A review on a data-driven microgrid management system integrating an active distribution network: Challenges, issues, and new trends. *Energies*, v. 15, n. 22, 2022. ISSN 1996-1073.

TOKTAROVA, A. et al. Long term load projection in high resolution for all countries globally. *International Journal of Electrical Power & Energy Systems*, v. 111, p. 160–181, 2019. ISSN 0142-0615. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0142061518336196>>.

TOKUDA, E. K.; COMIN, C. H.; COSTA, L. da F. Revisiting agglomerative clustering. *Physica A: Statistical Mechanics and its Applications*, v. 585, p. 126433, 2022. ISSN 0378-4371. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378437121007068>>.

UNICAMP, D. D. E. *Detalhamento dos serviços – ENERGIA*. 2023. <<https://www.prefeitura.unicamp.br/servicos/divisao-de-agua-e-energia/diretoria-de-energia/detalhamento-dos-servicos-energia/>>. Prefeitura Universitária, Universidade Estadual de Campinas, Acesso em 11 de junho de 2023.

WANG, C.; TINDEMANS, S. H.; PALENSKY, P. Generating contextual load profiles using a conditional variational autoencoder. In: *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. [S.l.]: IEEE, 2022.

WANG, W.-T. et al. Adaptive density-based spatial clustering of applications with noise (dbscan) according to data. In: *2015 International Conference on Machine Learning and Cybernetics (ICMLC)*. [S.l.: s.n.], 2015. v. 1, p. 445–451.

WANG, Z.; HONG, T. Generating realistic building electrical load profiles through the generative adversarial network (gan). *Energy and Buildings*, v. 224, p. 110299, 2020. ISSN 0378-7788.

WEN, X. et al. Economic mining of thermal power plant based on improved hadoop-based framework and spark-based algorithms. *Journal of Supercomputing*, 2023. Disponível em: <<https://doi-org.ez25.periodicos.capes.gov.br/10.1007/s11227-023-05443-5>>.

YAMASHITA, R. et al. Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, Springer, v. 9, p. 611–629, 2018.

YANG, L.; ZHENG, Z.; ZHANG, Z. An improved mixture density network via wasserstein distance based adversarial learning for probabilistic wind speed predictions. *IEEE Transactions on Sustainable Energy*, v. 13, n. 2, p. 755–766, 2022.

ZHANG, F. et al. Advantages and challenges of dc microgrid for commercial building a case study from xiamen university dc microgrid. In: *2015 IEEE First International Conference on DC Microgrids (ICDCM)*. [S.l.: s.n.], 2015. p. 355–358.

ZHANG, L.; ZHANG, B. Scenario forecasting of residential load profiles. *IEEE Journal on Selected Areas in Communications*, v. 38, n. 1, p. 84–95, 2020.

ZHANG, L.; ZHANG, B. Scenario forecasting of residential load profiles. *IEEE Journal on Selected Areas in Communications*, v. 38, n. 1, p. 84–95, 2020.

ZHANG, W.; HE, Y.; YANG, S. Day-ahead load probability density forecasting using monotone composite quantile regression neural network and kernel density estimation. *Electric Power Systems Research*, v. 201, p. 107551, 2021. ISSN 0378-7796. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378779621005320>>.

ZHANG, Y. et al. Generation of sub-item load profiles for public buildings based on the conditional generative adversarial network and moving average method. *Energy and Buildings*, v. 268, p. 112185, 2022. ISSN 0378-7788.