

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SANTA
CATARINA – CÂMPUS FLORIANÓPOLIS
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

PAULA ZENNI LODETTI

**ANÁLISE DE ALGORITMO PREDITIVO PARA GERAÇÃO
FOTOVOLTAICA BASEADO EM TÉCNICAS DE *MACHINE
LEARNING***

FLORIANÓPOLIS, 2022.

**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SANTA
CATARINA – CÂMPUS FLORIANÓPOLIS
DEPARTAMENTO ACADÊMICO DE ELETROTÉCNICA
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

PAULA ZENNI LODETTI

**ANÁLISE DE ALGORITMO PREDITIVO PARA GERAÇÃO
FOTOVOLTAICA BASEADO EM TÉCNICAS DE *MACHINE
LEARNING***

Trabalho de Conclusão de Curso
submetido ao Instituto Federal de
Educação, Ciência e Tecnologia de Santa
Catarina como parte dos requisitos para
obtenção do título de Engenheiro(a)
Eletricista.

Orientador: Prof. Edison Antonio Cardoso
Aranha Neto, D. Eng.

Coorientador: Gabriel Henrique dos Santos
Costa, Eng.

FLORIANÓPOLIS, 2022.

Ficha de identificação da obra elaborada pelo autor.

Zenni Lodetti, Paula

Análise de algoritmo preditivo para geração fotovoltaica baseado em técnicas de machine learning / Paula Zenni Lodetti; orientação de Edison Antonio Cardoso Aranha Neto; coorientação de Gabriel Henrique Santos Costa. - Florianópolis, SC, 2022.

52 p.

Trabalho de Conclusão de Curso (TCC) - Instituto Federal de Santa Catarina, Câmpus Florianópolis. Bacharelado em Engenharia Elétrica. Departamento Acadêmico de Eletrotécnica.

Inclui Referências.

1. Microrredes. 2. Previsão de geração. 3. Geração fotovoltaica. 4. Machine learning. I. Cardoso Aranha Neto, Edison Antonio. II. Santos Costa, Gabriel Henrique. III. Instituto Federal de Santa Catarina. IV. Análise de algoritmo preditivo para geração fotovoltaica baseado em técnicas de machine learning.

ANÁLISE DE ALGORITMO PREDITIVO PARA GERAÇÃO FOTOVOLTAICA BASEADO EM TÉCNICAS DE MACHINE LEARNING

PAULA ZENNI LODETTI

Este trabalho foi julgado adequado para obtenção do título de Engenheira Eletricista e aprovado na sua forma final pela banca examinadora do Curso de Graduação em Engenharia Elétrica do Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina.

Florianópolis, 25 de fevereiro de 2022.

Banca Examinadora:



Documento assinado digitalmente
EDISON ANTONIO CARDOSO ARANHA NETO
Data: 14/03/2022 09:22:22-0300
CPF: 005.835.469-78
Verifique as assinaturas em <https://v.ifsc.edu.br>

Prof. Edison A. C. Aranha Neto, D. Eng.

Orientador - IFSC



Documento assinado digitalmente
GABRIEL HENRIQUE DOS SANTOS COSTA
Data: 13/03/2022 11:22:45-0300
Verifique em <https://verificador.iti.br>

Gabriel Henrique dos Santos Costa, Eng.

Coorientador - REIVAX



Documento assinado digitalmente
EVERTHON TAGHORI SICA
Data: 09/03/2022 12:51:40-0300
CPF: 022.992.149-33
Verifique as assinaturas em <https://v.ifsc.edu.br>

Prof. Everthon Taghori Sica, D. Eng.

IFSC

MARCOS AURELIO
IZUMIDA
MARTINS:6605949
6920

Assinado de forma digital por
MARCOS AURELIO IZUMIDA
MARTINS:66059496920
Dados: 2022.03.08 17:16:55 -03'00'

Marcos Aurelio Izumida Martins, D. Eng.

Fundação CERTI

AGRADECIMENTOS

Minha gratidão a toda minha família, por me apoiar incondicionalmente durante toda a graduação, em especial nestes últimos anos. Agradeço por todo o amor e suporte sempre, e por me incentivarem a seguir estudando e buscando conhecimento. Obrigada pela parceria a todos os amigos que adquiri ao longo da graduação, que ajudaram com que todo o processo fosse menos difícil, compartilhando da intensidade de todas as disciplinas e do curso como um todo. Aos amigos do colégio, obrigada por sempre estarem ao meu lado, independente da distância e dos diferentes rumos tomados.

Agradeço ao Instituto Federal de Santa Catarina, e todos os que o compõe, por sua excelência em proporcionar, além de ensino público, gratuito e de qualidade, oportunidades únicas que contribuíram imensamente para minha formação, como estágio e intercâmbio. Obrigada a todos os professores com quem tive contato ao longo da graduação, que compartilharam seu conhecimento e me fizeram adquirir tamanho interesse pela área. Obrigada ao professor Edison Aranha pela orientação, apoio e contribuições, que aperfeiçoaram o processo e o resultado do trabalho.

À Fundação CERTI, agradeço pela oportunidade de estágio e posteriormente trabalho, por proporcionar conhecimento e experiência na área da pesquisa, essenciais para o desenvolvimento deste trabalho e minha formação acadêmica e profissional. Obrigada aos colegas do CES por garantirem um ambiente de trabalho leve e prazeroso durante este processo, especialmente para meu colega, coorientador e amigo Gabriel. Agradeço por todas as reuniões, revisões com prazos apertados, e principalmente por ter me guiado e feito com que eu desenvolvesse este trabalho de forma tranquila e produtiva

RESUMO

Com as redes de energia cada vez mais descentralizadas, o tamanho do mercado mundial de microrredes apresenta um crescimento bastante significativo. O gerenciamento desse tipo de instalação requer um processo de controle e acompanhamento eficiente. Além disso, é usual a presença de fontes de energia renováveis, como solar e eólica, cuja geração tende a ser instável e de difícil previsão. Esse cenário dificulta o processo de gerenciamento otimizado, o que pode afetar negativamente a saúde financeira do proprietário. Em vista disso, é interessante que uma microrrede tenha um bom sistema de otimização, que seja capaz de fornecer uma previsão adequada da geração de energia, facilitando o gerenciamento ótimo do sistema. Assim, o presente trabalho avalia um algoritmo preditivo desenvolvido com a finalidade de ser utilizado em um projeto que usa um otimizador para coordenar a operação de microrredes. A avaliação foi feita através da aplicação do preditor em um sistema fotovoltaico localizado no campus Florianópolis do Instituto Federal de Santa Catarina por um período de 5 meses.

Palavras-chave: Microrredes. Previsão de geração. Geração fotovoltaica. *Machine learning*.

ABSTRACT

With energy networks increasingly decentralized, the size of the global microgrid market has presented a significantly growth. However, the management of this type of installation requires a very intense control and monitoring process. Furthermore, in these cases, the presence of renewable energy generation sources is common, such as solar and wind, which provide an unstable and difficult to predict generation. This scenario hinders the optimized management process, which can negatively affect the owner's financial health. Therefore, it is interesting that a microgrid has an efficient optimization system, capable of providing an adequate energy generation forecast, facilitating the optimal management of the system. Thus, the present work evaluates a predictive algorithm developed with the purpose of being used in a project that uses an optimizer to coordinate the operation of microgrids. The evaluation is held through the application of the predictor in a photovoltaic system located in the Florianópolis campus of the Federal Institute of Santa Catarina for 5 months.

Keywords: Microgrids. Generation forecast. Photovoltaic generation. Machine learning.

LISTA DE FIGURAS

Figura 1 – Esquemático do projeto MR-II	24
Figura 2 – Plataforma <i>Frontend</i>	25
Figura 3 – Esquemático do fluxo de dados do preditor	28
Figura 4 – Exemplo de <i>outliers</i> nas curvas de geração e previsão	34
Figura 5 – <i>Website</i> com dados de geração.....	35
Figura 6 – Evolução mensal do erro.....	40
Figura 7 – Evolução semanal do erro.....	40
Figura 8 – Evolução diária do erro	41
Figura 9 – Curvas de geração e previsão: Dias com MAE mais altos.....	43
Figura 10 – Curvas de geração e previsão: Dias com RMSE mais altos	44
Figura 11 – Curvas de geração e previsão: Dias com MAE e RMSE mais baixos....	45
Figura 12 – Curvas de geração e previsão: Dias ensolarados	47

LISTA DE TABELAS

Tabela 1 – Métodos de <i>machine learning</i>	19
Tabela 2 – Coeficientes de Pearson de correlação de métricas climáticas.....	23
Tabela 3 – MAE e RMSE para todo o período	37
Tabela 4 – MAE e RMSE para cada mês.....	37
Tabela 5 – MAE e RMSE para cada semana.....	37
Tabela 6 – MAE e RMSE para cada dia.....	38
Tabela 7 – Comparação dos resultados.....	41
Tabela 8 – Valores de dias com MAE mais baixos e mais altos	42
Tabela 9 – Valores de dias com RMSE mais baixos e mais altos.....	42
Tabela 10 – Valores de MAE e RMSE de dias ensolarados	46

LISTA DE ABREVIATURAS E SIGLAS

ANEEL – Agência Nacional de Energia Elétrica

API – *Application Programming Interface*

CA – Corrente Alternada

CAGR – *Compound Annual Growth Rate*

CC – Corrente Contínua

CERTI – Centros de Referência em Tecnologias Inovadoras

EPE – Empresa de Pesquisa Energética

FV – Fotovoltaico

GD – Geração Distribuída

GESE – Grupo de Estudos em Sistemas de Energia

IA – Inteligência Artificial

IFSC – Instituto Federal de Santa Catarina

LAPR – Laboratório de Planejamento Integrado de Recursos Energéticos

MAE – *Mean Absolut Error*

MAPE – *Mean Absolute Percentage Error*

ML – *Machine Learning*

MR – Microrredes

MR-II – Microrredes II

MSE – *Mean Squared Error*

P&D – Pesquisa e Desenvolvimento

PCA – Análise de Componentes Principais

PCC – *Point of Common Coupling*

PLD – Preço de Liquidação das Diferenças

pu – por unidade

RMSE – *Root Mean Squared Error*

RNA – Redes Neurais Artificiais

SIN – Sistema Interligado Nacional

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Justificativa	13
1.2	Definição do Problema	14
1.3	Objetivo Geral.....	15
1.4	Objetivos Específicos	15
1.5	Estrutura do trabalho.....	16
2	PREVISÃO DA GERAÇÃO	17
2.1	Métodos de <i>Machine Learning</i>.....	17
2.2	Relação entre Dados Climáticos e Potência Gerada	21
3	PROJETO MICRORREDES II	24
3.1	Algoritmo Preditivo.....	26
3.1.1	Funcionamento	27
4	MÉTODOS DE CÁLCULO DO ERRO	29
4.1	Métodos	29
4.1.1	Método MSE	29
4.1.2	Método RMSE.....	30
4.1.3	Método MAE	31
4.1.4	Método MAPE	31
4.2	Comparação	32
4.3	<i>Outliers</i>.....	33
4.4	Conclusão e Escolha do Método	34
5	ANÁLISE DA PREVISÃO.....	35
5.1	Metodologia	35
5.2	Coleta de Dados	36
5.3	Resultados.....	37
5.4	Análise dos Erros.....	41
6	CONCLUSÃO	48
6.1	Trabalhos Futuros.....	48
	REFERÊNCIAS	50

1 INTRODUÇÃO

Nas últimas décadas, o Brasil enfrentou diversas crises hídricas que somadas a períodos de secas, afetaram os níveis dos reservatórios das usinas hidrelétricas. Visto que a geração hídrica equivale a 65,2% da matriz energética do país (EPE, 2021), essas crises contribuem com a preocupação quanto à diversificação desta matriz. Diante disso, o Brasil promove a adoção de medidas que visem contribuir com o aumento da diversidade de fontes energéticas e reduzir a dependência hidrelétrica do país, aumentando a confiabilidade do Sistema Interligado Nacional (SIN), sempre com a preocupação de manter o mais renovável possível a matriz energética.

Uma alternativa que tem ganhado destaque é a Geração Distribuída (GD), ou seja, geração elétrica eficiente e descentralizada. No ano de 2020, foi observado um crescimento de 137% de GD no Brasil (EPE, 2021). De acordo com INEE (2022) GD designa a geração elétrica realizada junto ou próxima do consumidor independente da potência, tecnologia e fonte de energia.

Além disso, apesar de ainda muito incipiente no país, o conceito de microrredes (MR) também passa a ser mais presente no cenário brasileiro. A Agência Nacional de Energia Elétrica (ANEEL) define uma microrrede como sendo uma rede de distribuição de energia elétrica que pode operar isoladamente do sistema de distribuição, atendida diretamente por uma unidade de geração distribuída (ANEEL, 2022). Essa rede isolada é conectada à rede principal apenas por meio de um ponto chamado PCC, do inglês *Point of Common Coupling*, ou ponto de acoplamento comum. É, também, nesse ponto onde ocorre o controle e gerenciamento da MR.

Com o aumento da descentralização das redes de energia, o tamanho do mercado mundial de microrredes cresce expressivamente. Em 2020, excedeu US\$6 bilhões e deve registrar mais de 27% da taxa de crescimento anual composta (do inglês, CAGR) entre 2021 e 2027, de acordo com relatório da *Global Market Insights Inc* (GMI, 2020).

Os componentes de uma microrrede variam de acordo com cada caso. De maneira geral, elas são compostas por cargas e por pontos de geração. A geração em uma MR pode ser renovável ou não. No caso das fontes renováveis, como solar e eólica, por se tratarem de gerações intermitentes na maioria dos casos, requerem ainda inversores CC/CA e sistemas de armazenamento. Em relação às cargas, elas podem ser controláveis ou não, sendo que no primeiro caso, podem ser alocadas de maneira a otimizar a demanda da rede (OHI, 2018). A operação da MR pode ocorrer de maneira ilhada, ou então conectada à rede de distribuição. Quando ela opera de forma ilhada pode proporcionar benefícios como aumento de confiabilidade e economia (HUAYLLAS, 2015).

Outra situação vantajosa na qual a MR pode estar inserida é em mercados livres de energia elétrica, permitindo a negociação de valores de compra e venda do excedente de energia gerada e seus respectivos preços com os demais agentes do mercado na Câmara de Comercialização de Energia Elétrica (CCEE). Nesse caso, quando agregado a um modelo de automação de contratos de compra e venda, pode ainda maximizar a eficiência operacional dos geradores e das cargas conectados à MR (FALCÃO, 2019).

1.1 Justificativa

De forma a aprimorar o gerenciamento e o funcionamento das MR, existe, hoje, uma série de métodos e técnicas de otimização a fim de aproveitar de forma ideal todos os recursos presentes, bem como as demandas de carga, da MR. A otimização proporciona uma maior integridade de cargas e gera economia para a rede, no entanto, para que gere bons resultados, ela requer um sistema de previsão da geração e das cargas o mais assertivo possível.

Por mais dinâmica que possa ser a demanda de uma microrrede, a previsão das cargas é fundamental na contribuição para um gerenciamento ótimo da rede. Os fatores que influenciam nessa previsão são bastante variados e incluem aspectos sociais, econômicos, ambientais e temporais (LEKSHMI; SUBRAMANYA, 2019).

Assim como a previsão de carga, a previsão da geração também desempenha um papel crítico no gerenciamento de energia dos sistemas de

potência. No entanto, quando tratamos de fontes renováveis como painéis fotovoltaicos, sua geração não é facilmente previsível com antecedência, e varia com base nas condições climáticas e específicas do local. Essa previsão se torna um pouco mais complexa tendo em vista a sua natureza intermitente e imprevisível. De acordo com Voyant *et al.* (2017), tais características não controláveis da geração fotovoltaica (FV) trazem uma série de outros problemas, como flutuações de tensão, deterioração da qualidade de energia local e questões de estabilidade. Dessa forma, a confiabilidade do sistema passa a depender diretamente de sua capacidade de acomodar mudanças esperadas e inesperadas (na produção e no consumo) e perturbações, mantendo a qualidade e continuidade do serviço aos consumidores.

Diante disso, através de bons modelos de previsão, o gerenciamento otimizado da MR visa manter os valores de geração de energia sempre próximos aos valores da demanda. Modelos de inteligência artificial (IA) mostram-se uma boa solução para prever geração e consumo de energia, podendo diminuir o erro em comparação com outras abordagens (SOBRI; KOOHI-KAMALI; RAHIM, 2018).

1.2 Definição do Problema

Tendo em vista o cenário atual das MR e sua relevância para a diversificação da matriz energética, bem como sua importância na participação no mercado livre de energia, percebe-se a necessidade de garantir que ela opere da melhor forma possível, considerando a possibilidade de compra e venda de energia elétrica como uma oportunidade de operação mais econômica, além de confiável e segura, através de algoritmos para o monitoramento e controle da microrrede.

A melhor maneira para garantir isso é através de um gerenciamento eficiente do conjunto de componentes da rede, utilizando técnicas de otimização, automação e comunicação que aprimorem sua operação. Sabendo que para alcançar esta operação ótima é necessária uma boa metodologia de previsão da geração, são abundantes os estudos sendo desenvolvidos apresentando modelos de previsão, com horizontes de tempo, dados de entrada e estratégias diferentes. Entretanto, é importante que os métodos sejam testados e comprovados de forma a verificar sua eficácia.

Dentre os estudos e métodos de previsão encontrados hoje na literatura, encontram-se exemplos envolvendo técnicas estatísticas, de inteligência artificial e *machine learning* (ML), dividindo-se entre aprendizado supervisionado e não supervisionado. A complexidade de cada método pode variar de acordo com o caso estudado. Assim, encontrar uma maneira simples e efetiva de prever a geração fotovoltaica para uma microrrede pode não ser tão simples, mas é essencial para possibilitar a tomada de ações sob estas informações.

1.3 Objetivo Geral

Realizar análise estatística do desempenho de um algoritmo preditivo de sistema de gerenciamento de microrredes, por meio de dados monitorados em tempo real, referentes à geração de células FV localizadas no Instituto Federal de Santa Catarina (IFSC) em Florianópolis.

1.4 Objetivos Específicos

Considerando o cenário de microrredes e a importância da previsão de potência gerada por sistemas fotovoltaicos para a uma operação otimizada, o presente trabalho tem como objetivos específicos:

- a) elencar os principais métodos de previsão de geração fotovoltaica presentes na literatura;
- b) contextualizar o cenário do projeto no qual o algoritmo preditivo será aplicado;
- c) analisar o comportamento do algoritmo preditivo, fazendo o uso de dados do sistema fotovoltaico do IFSC;
- d) realizar análise dos dados e apontar as principais causas de erros.

1.5 Estrutura do trabalho

O presente documento está estruturado em seis capítulos, sendo o primeiro a introdução, contendo a justificativa, definição do problema, objetivo geral e objetivos específicos.

Na sequência, o capítulo 2 apresenta uma revisão da bibliografia no que diz respeito ao tema previsão da geração, bem como os métodos de *machine learning* envolvidos neste tópico e a relação entre dados climáticos e potência gerada.

O capítulo 3 fornece uma explicação geral relativa ao projeto de pesquisa e desenvolvimento (P&D) Microrredes II, abordando o funcionamento e as características do algoritmo preditivo em análise.

No capítulo 4, são avaliados métodos de cálculo do erro para avaliar algoritmos preditivos, suas características e principais aplicações, para que, na sequência, seja escolhido o mais adequado ao estudo do presente trabalho.

O capítulo 5 apresenta a metodologia empregada e os resultados dos cálculos com suas respectivas análises.

Por fim, a conclusão é realizada no capítulo 6, junto com as sugestões de trabalhos futuro.

2 PREVISÃO DA GERAÇÃO

Por se tratar de um tema de grande relevância para o setor elétrico na atualidade, a quantidade de estudos e trabalhos que referenciam o tema microrredes (MR) na literatura é grande, apesar de ser uma tecnologia ainda pouco aplicada. Ademais, trata-se de um objeto de estudo que engloba as mais variadas áreas e tecnologias do setor. Por isso, este capítulo está centrado nos métodos de previsão de geração fotovoltaica e as técnicas envolvidas nesta ferramenta.

Conforme explicado na introdução deste trabalho, um elemento de grande destaque na operação de microrredes é o algoritmo que “prevê” a potência gerada. Para tanto, são estabelecidos dados de entrada, que normalmente equivalem a informações climáticas, data e geração. Essas informações apresentam forte correlação, e com um período suficiente de amostra, permitem uma previsão apropriada e condizente com os valores reais. De posse dessas informações, o algoritmo utiliza ferramentas que envolvem inteligência artificial para que, com base no histórico de dados e os valores medidos em tempo real, prospecte a potência que será gerada nos próximos instantes. Sabendo desse valor futuro de geração, a gestão da MR pode ocorrer de maneira mais otimizada e adequada para os valores do mercado.

2.1 Métodos de *Machine Learning*

Machine learning (ML), traduzido como aprendizado de máquina, é um ramo da inteligência artificial (IA), ou inteligência computacional, que pode ser usado em vários domínios. Possui como principal vantagem, a possibilidade de resolver problemas que são impossíveis de serem representados por algoritmos explícitos. Em outras palavras, esses modelos encontram relações entre variáveis de entrada e saída, mesmo que a representação matemática seja muito complexa ou até desconhecida. Dessa maneira, permitem a construção e o estudo de sistemas que podem aprender a partir de conjuntos de dados, proporcionando aos computadores a capacidade de aprender sem serem programados explicitamente (VOYANT *et al.*, 2017).

No contexto de previsão de geração de potência, de acordo com Sobri (2018), modelos de IA, especialmente modelos que envolvem ML, podem apresentar uma diminuição de erro, quando comparados a outras abordagens estatísticas.

O algoritmo analisado neste trabalho se baseia em técnicas de ML para realizar a previsão da potência gerada por módulos fotovoltaicos. Em geral, algoritmos de previsão fazem uso de diferentes métodos de ML para atingir seus objetivos e fornecer previsões assertivas para o sistema em que opera. Os métodos se dividem em três principais categorias que possuem vantagens e desvantagens dependendo do propósito com que são utilizados. O trabalho apresentado por Voyant *et al.* (2017) confere uma visão geral dos principais métodos de previsão de irradiância solar usando abordagens de aprendizado de máquina. As principais categorias são:

- a) classificação e preparação de dados;
- b) aprendizagem supervisionada;
- c) aprendizagem não supervisionada;
- d) modelos de aprendizagem de conjunto.

Cada categoria engloba uma série de métodos que devem ser escolhidos de acordo com as entradas disponíveis, tempo de resposta desejável, assertividade e outros fatores. Como os algoritmos de ML aprendem com dados, é fundamental a escolha correta e preparação adequada desses dados para que seja possível a resolução eficaz do problema. Dessa forma Voyant *et al.* (2017) fazem uma introdução aos métodos de classificação e preparação de dados. Na aprendizagem supervisionada, são fornecidos ao computador exemplos de entradas e suas saídas desejadas, e o objetivo é aprender uma regra geral que mapeie entradas em saídas, portanto são métodos que requerem informações rotuladas, e por isso necessitam de intervenção humana.

Ao contrário do modelo de aprendizagem supervisionado, um modelo de aprendizagem não supervisionado não precisa de nenhuma intervenção, de forma que é capaz de encontrar uma estrutura oculta em suas entradas, sem conhecimento das saídas. Além disso, existem também modelos que unem diferentes métodos, por exemplo, a aprendizagem por conjunto. Seu conceito básico

é treinar vários modelos de aprendizagem de base como membros de um conjunto e combinar suas previsões em uma única saída, que deve ter melhor desempenho em média do que qualquer outro membro do conjunto com erro não correlacionado nos conjuntos de dados alvo.

A Tabela 1 apresenta um resumo dos principais modelos de aprendizagem de cada categoria, e também algumas de suas características mais relevantes.

Tabela 1 – Métodos de *machine learning*

Categoria	Método	Características
Classificação e preparação de dados	Análise discriminante e análise de componentes principais (PCA)	Método estatístico. Usa uma transformação ortogonal para transformar um conjunto de observações de variáveis provavelmente correlacionadas, em um conjunto de valores de variáveis linearmente não correlacionadas (componentes principais). O PCA é sensível ao escalonamento relativo das variáveis originais.
	Classificação Naïve Bayes e redes Bayesianas	Classificadores probabilísticos baseados no teorema de Bayes com fortes suposições de independência entre os dados. Altamente escalável.
	Abordagem de mineração de dados – <i>Data Mining</i>	Consiste em extrair as informações mais importantes de um conjunto de dados muito grande. Os dados coletados são analisados para destacar as principais informações.
Aprendizagem Supervisionada	Regressão linear	A ideia de estocasticidade em séries temporais assume que toda série temporal pode ser considerada como a realização de um processo estocástico. Vários métodos de séries temporais foram desenvolvidos desde então. Formula o conceito de modelos autorregressivos e de média móvel.
	Modelos lineares generalizados	Uma generalização flexível de regressão linear ordinária.

Categoria	Método	Características
Aprendizagem Supervisionada	Modelos lineares generalizados	Permite variáveis de resposta que têm modelos de distribuição de erro diferentes de uma distribuição normal.
	Regressão não linear	Redes Neurais Artificiais (RNA) estão sendo cada vez mais usadas para regressão não linear. Particularmente predominante no domínio da previsão de séries temporais com métodos não lineares.
	Máquina de vetores de suporte/ suporte a regressão de vetor – <i>Support vector machines</i>	Com base na aplicação de máquinas de vetores de suporte para problemas de regressão. Aplicado com sucesso a tarefas de previsão de séries temporais.
	Aprendizagem da árvore de decisão – <i>Decision tree learning</i>	Uma resposta ou classe Y das entradas X1, X2, ..., Xp deve ser prevista. Isso é feito crescendo uma árvore binária. Em cada nó da árvore, um teste para uma das entradas é aplicado.
	Vizinho mais próximo – <i>Nearest Neighbor</i>	Tipo de aprendizagem baseada em instância, onde uma função é aproximada apenas localmente e todos os cálculos são atrasados até a classificação. Um dos algoritmos de aprendizado de máquina mais simples.
	Cadeia de Markov – <i>Markov chain</i>	Processo estocástico com a propriedade de Markov. A distribuição de probabilidade do próximo estado depende apenas do anterior.
Aprendizagem Não Supervisionada	<i>Clustering k-means e k-methods</i>	Método de quantização vetorial, originalmente derivado do processamento de sinal. Visa particionar n observações em k <i>clusters</i> em que cada observação pertence ao <i>cluster</i> com a média mais próxima. Concentra-se em extrair informações úteis dos dados com o propósito de modelar o comportamento da série temporal e encontrar padrões do espaço de entrada agrupando os dados.
	Agrupamento hierárquico	Um método de análise de cluster que visa construir uma hierarquia de <i>clusters</i> .

Categoria	Método	Características
Aprendizagem Não Supervisionada	Agrupamento hierárquico	Algoritmos para agrupamento hierárquico são geralmente aglomerativos ou divisivos.
	Modelos de mistura gaussiana	Desenvolvimento relativamente recente em modelagem não linear. Uma generalização de uma distribuição gaussiana multivariada para muitas variáveis.
	Avaliação de <i>cluster</i>	A avaliação mais direta, mas demorada, especialmente se é necessário um grande número de estudos.
Modelos de aprendizagem de conjunto	Impulsioneamento – <i>Boosting</i>	Surgiu como uma forma de combinar muitos classificadores fracos para produzir um conjunto poderoso. É um processo iterativo que dá cada vez mais importância à má classificação.
	Ensacamento – <i>Bagging</i>	Um meta-algoritmo de conjunto de aprendizado de máquina projetado para melhorar a estabilidade e a precisão dos algoritmos de aprendizado de máquina. Também reduz a variância e ajuda a evitar <i>overfitting</i> .
	Subespaço aleatório	Melhorias substanciais na precisão da classificação obtidas cultivando um conjunto de árvores e permitindo que votassem na classe mais popular.
	Conjunto de preditores	A prática atual sugere que as previsões devem ser compostas. Seja por meio de uma série de previsões simples ou produzindo uma previsão simples a partir de outras previsões simples (não apenas previsões pontuais, mas também probabilísticas). Isso leva a ganhos de desempenho, em relação às previsões de contribuição.

Fonte: Elaboração própria (2021).

2.2 Relação entre Dados Climáticos e Potência Gerada

Os dados de entrada para o sistema de previsão equivalem, em sua maioria, a dados climáticos, os quais possuem diferentes graus de influência na geração de energia e consequentemente na previsão da geração. Para avaliar esse

grau de influência, Sharma *et al.* (2011) apresentam um estudo que, através de dados de previsão do tempo coletados do site do Serviço Nacional de Meteorologia dos Estados Unidos por um período de 10 meses, e dados de irradiância coletados de uma estação meteorológica, elaboram gráficos que contribuem no entendimento sobre como a irradiância e a geração de energia do painel FV dependem de uma combinação de várias métricas de clima, e como os dados se conectam entre si.

A primeira análise de relação avaliada refere-se ao dia do ano, que como era de se esperar, influencia na irradiância, já que a tendência é que durante o inverno a irradiância seja menor, e no verão, maior. No entanto, esse padrão não é observado em todos os dias avaliados, havendo dias em janeiro (inverno no hemisfério Norte) com maior intensidade que dias de verão. Isso porque, nesses casos, há grande interferência dos demais fatores na irradiância medida. Os demais fatores cuja influência foi avaliada são: temperatura, ponto de orvalho, velocidade do vento, cobertura do céu, probabilidade de precipitações e humidade relativa do ar.

Por fim, os resultados encontrados referentes à relação desses dados com a irradiância foram que a velocidade do vento, o ponto de orvalho e a temperatura não estão altamente correlacionados com a intensidade solar, sendo que para a velocidade, quase não há correlação e para a temperatura e ponto de orvalho, há alguma correlação quando estes valores são elevados. Já para cobertura do céu, umidade relativa e a chance de precipitação a correlação apresentada é altamente negativa, uma vez que a irradiância tende a diminuir com o aumento destes valores.

De maneira a resumir as correlações descritas, através do coeficiente de correlação produto-momento de Pearson, a Tabela 2 mostra os coeficientes de correlação para cada métrica climática, sendo 1 o valor máximo de correlação e 0 o valor mínimo. O método do coeficiente de Pearson divide a covariância das duas variáveis pelo produto de seus desvios-padrão. Vale ressaltar que, apesar de ser o método mais utilizado, há ainda dois coeficientes de relação também utilizados para este tipo de análise, o método de Kendall e o de Spearman.

É comum que os três métodos sejam usados para análise de correlação de sistemas de ML, no entanto, os coeficientes de Kendall e Spearman são resolvidos principalmente pela classificação da matriz de dados, enquanto o de

Pearson é usado principalmente para medir a correlação linear entre os dados (CHEN *et al.*, 2021).

Tabela 2 – Coeficientes de Pearson de correlação de métricas climáticas

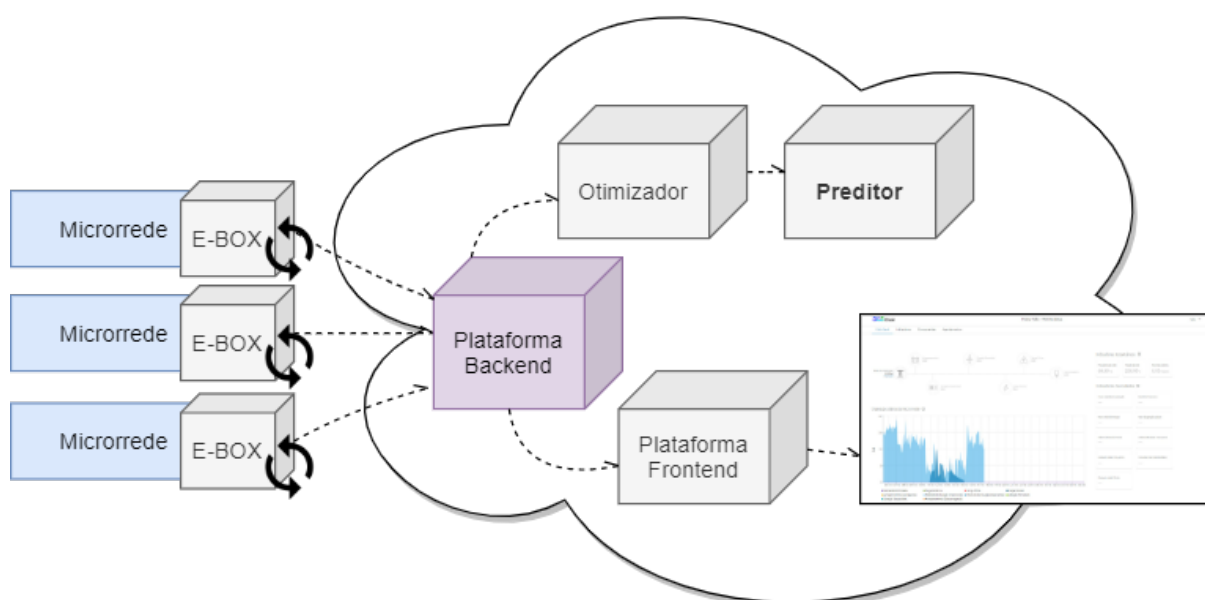
	Dia	Temp.	Orvalho	Vento	Cobert.	Precip.	Humid.
Dia	1,0000	0,7022	0,7007	-0,0711	-0,1034	-0,0571	0,0645
Temp.	0,7022	1,0000	0,9212	-0,1994	-0,2582	-0,1279	-0,0791
Orvalho	0,7007	0,9212	1,0000	-0,2251	0,0455	0,1491	0,3081
Vento	-0,0711	-0,1994	-0,2251	1,0000	-0,0192	0,0340	-0,1025
Cobert.	-0,1034	-0,2582	0,0455	-0,0192	1,0000	0,7067	0,7525
Precip.	-0,0571	-0,1279	0,1491	0,0340	0,7067	1,0000	0,7475
Humid.	0,0645	-0,0791	0,3081	-0,1025	0,7525	0,7475	1,0000

Fonte: Adaptado de Sharma *et al.* (2011).

3 PROJETO MICRORREDES II

De maneira a contextualizar o presente estudo e o algoritmo analisado por este trabalho, apresenta-se o projeto denominado Microrredes II (MR-II), um projeto de Pesquisa e Desenvolvimento – P&D regulado pela Agência Nacional de Energia Elétrica – ANEEL desenvolvido pelo Centro de Energia Sustentável da Fundação Centros de Referência em Tecnologias Inovadoras – CERTI para a empresa AES Brasil. O projeto tem como principal objetivo viabilizar para o cliente um sistema de gestão inteligente de microrredes. A Figura 1 apresenta um esquemático que permite uma visão geral das etapas componentes do projeto.

Figura 1 – Esquemático do projeto MR-II

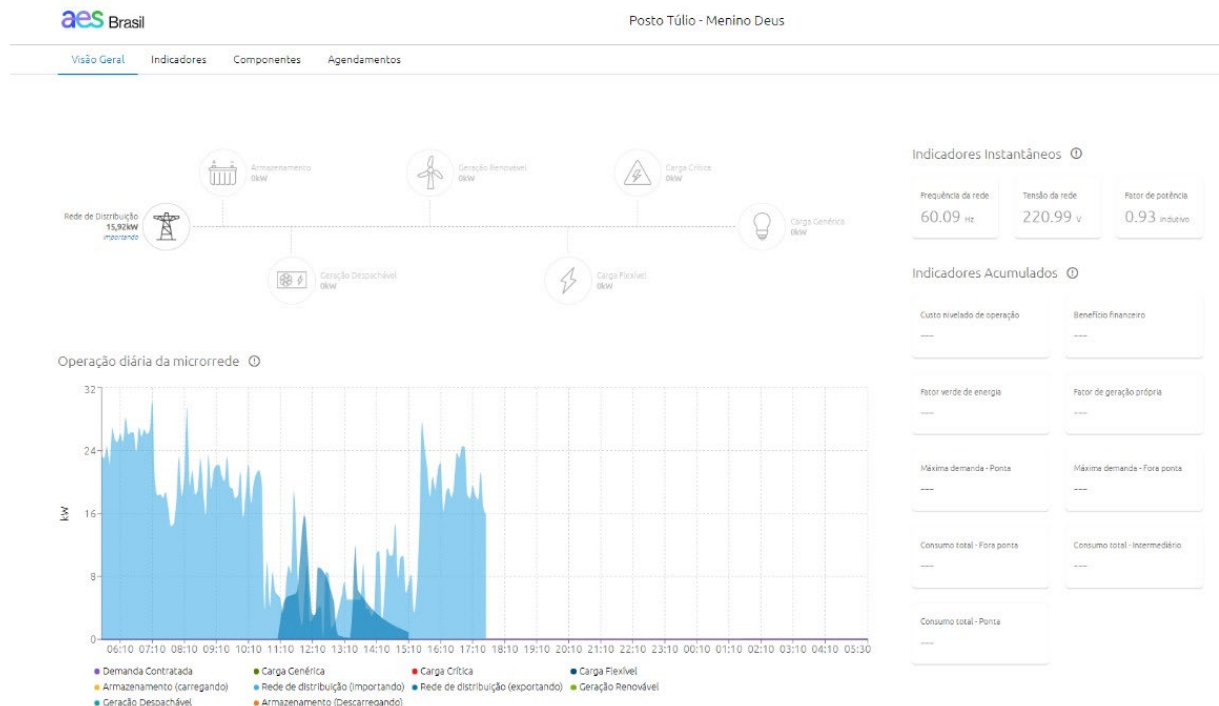


Fonte: Elaboração própria (2021).

O projeto contou com a instalação de quadros de automação em determinadas microrredes. Cada um dos quadros faz toda a supervisão da instalação através de medidores alocados em pontos estratégicos da instalação elétrica. Além dos medidores, o controlador chamado E+BOX é responsável pelo gerenciamento local e envio de dados para processamento na nuvem. As E+BOX coletam dados de medição da microrrede e enviam ao *Backend* da plataforma. A plataforma *Backend* envia então os dados à plataforma *Frontend* (Figura 2) que equivale a uma página na *web* onde o cliente, proprietário da microrrede, pode

verificar os dados da mesma em tempo real e tomar as ações de gestão, como despachar geração ou manobrar cargas.

Figura 2 – Plataforma *Frontend*



Fonte: Elaboração própria (2021).

Os dados coletados pela E+BOX são também fornecidos ao otimizador do processo, onde é realizada uma otimização para a operação futura. Por isso, o otimizador tem como *input*, além do custo de operação, as previsões geradas pelo preditor, objeto de estudo do presente trabalho. Conhecendo a previsão de demanda e geração, o otimizador visa à minimização do custo de operação da microrrede, atuando em cima das cargas flexíveis e na manipulação da geração despachável, como banco de baterias e geradores a combustível.

No processo de otimização são analisados a prioridade de cargas e o custo de operação, visto que, por vezes, há mais vantagens no despacho do gerador do que na compra de energia da distribuidora, especialmente no horário de ponta. A avaliação do custo de operação pelo otimizador leva em conta o Preço de Liquidação das Diferenças (PLD) coletado da CCEE, os preços dentro e fora do horário de ponta, o preço do combustível e por fim, o preço de despacho da bateria. Com esses valores otimizados, os comandos são enviados à E+BOX e executados na MR.

O preditor que fornece as previsões ao otimizador corresponde ao algoritmo preditivo analisado por este trabalho. Este algoritmo foi desenvolvido com inteligência capaz de correlacionar os dados de potência com os dados meteorológicos para cada horário; criar grupos (*clusters*) automaticamente com o passar do tempo; e fazer reagrupamentos periódicos quando a quantidade máxima de *clusters* for atingida.

O elemento que faz o repasse de dados entre a E+BOX e o preditor é o *Backend* ou API – *Application Programming Interface*. Os *inputs* do modelo são dados meteorológicos, como a previsão de temperatura, previsão da quantidade de nuvens sobre a região, entre outros. Há também um banco de dados que armazena os dados de medição da MR, os dados do clima e os modelos de previsão (*clusters*). Mais detalhes do preditor são explicados na sequência do presente trabalho.

Sendo assim, faz-se necessário que o preditor opere corretamente e forneça os dados de previsão de potência suficientemente próximos do valor real para o otimizador, de forma que a microrrede seja gerida da maneira mais apropriada possível. Dessa maneira, este trabalho tem como finalidade a avaliação do preditor, ou algoritmo preditivo desenvolvido para a aplicação no projeto descrito, sendo que para este caso, será avaliada apenas a utilização do algoritmo para a previsão da geração.

3.1 Algoritmo Preditivo

O algoritmo preditivo, desenvolvido para o projeto MR-II, faz uso de um método de *machine learning* de aprendizagem não supervisionada, conhecido como *k-means*. Esta abordagem foi utilizada para desempenhar previsões de geração fotovoltaica a serem utilizadas na otimização de microrredes

Conforme apresentado na Tabela 1, que aborda os principais métodos de ML, uma característica deste método é possibilitar a extração de informações úteis dos dados com a finalidade de modelar o comportamento de séries temporais e encontrar padrões do espaço de entrada agrupando os dados.

De acordo com Martins (2018), agrupamento, ou *clustering* representa uma técnica de mineração de dados, cujo propósito é a organização dos dados de

um conjunto em grupos (*clusters*), conforme similaridade existente entre os dados de um mesmo grupo, com base em algum critério predeterminado. Como não conta com classes rotuladas, é uma forma de aprendizado não supervisionado. Dentre suas vantagens incluem-se:

- e) permitir a descrição das características particulares de cada grupo identificado de forma mais eficiente;
- f) apresentar um panorama detalhado acerca dos dados;
- g) permitir o desenvolvimento de novos esquemas de classificação;
- h) auxiliar na descoberta de correlações entre os atributos dos dados que, até então, eram desconhecidas.

Além disso, o método *k-means* é um dos algoritmos de aprendizagem não supervisionada mais utilizado em problemas de agrupamento (MARTINS, 2018). Os algoritmos *k-means* para *clustering* são considerados simples e de fácil implementação, mas eficientes e de sucesso empírico, possuindo várias extensões desenvolvidas em várias formas (CASSIANO, 2014). Ademais, é um método eficiente que costuma convergir para uma configuração estável em poucas iterações (MARTINS, 2018).

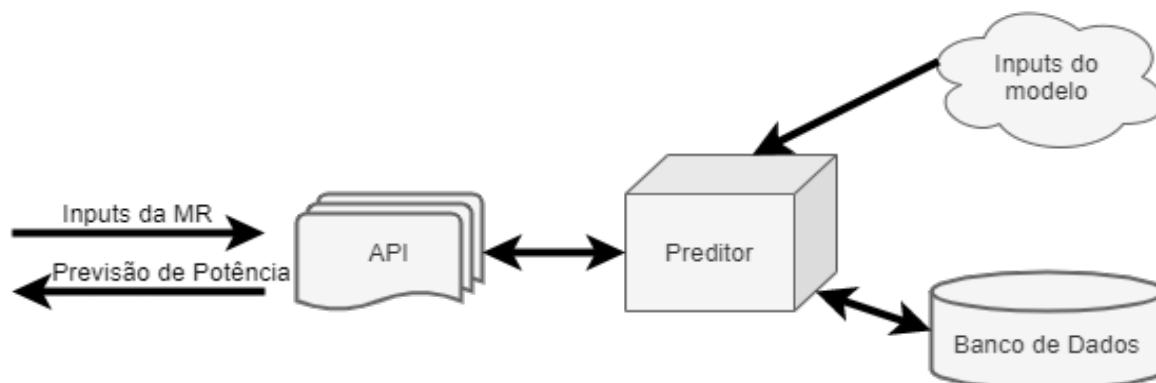
Dadas essas características e o contexto no qual o algoritmo foi desenvolvido, o método se mostrou aplicável e vantajoso para o cenário, principalmente devido à simplicidade de implementação. Além disso, buscava-se um método que fosse escalável com o tempo e capaz de aprender sozinho. Por conta da característica de método não supervisionado do *k-means*, foi possível que o algoritmo coletasse os dados da operação e rodasse o agrupamento (*clustering*) na busca dos melhores e recém-encontrados modelos para desempenhar as previsões. Dito isso, na sequência deste capítulo, é dada uma explicação básica a respeito do funcionamento do preditor.

3.1.1 Funcionamento

O funcionamento do algoritmo, baseia-se no recebimento e envio de dados a uma API, que é responsável pela validação dos dados. Além disso, é ela

quem faz a solicitação da previsão ao algoritmo e repassa esta previsão à MR. Um esquemático do processo é apresentado na Figura 3.

Figura 3 – Esquemático do fluxo de dados do preditor



Fonte: Elaboração própria (2021).

Os *inputs* recebidos pela API, provenientes da MR, equivalem aos dados de medição do sistema, assim como o horário das medições e os dados de localização. Além disso, o horizonte e a taxa de amostragem da previsão também são dados de entrada da API. A requisição desses dados é realizada com uma periodicidade de 15 minutos. A previsão de potência que a API envia ao sistema corresponde a um vetor de previsão e o horizonte e taxa de amostragem requisitada, com valores a partir do último horário de medição.

Com relação ao preditor em si, inicialmente, ao receber a solicitação de previsão da geração, o algoritmo busca dados climáticos para a localização do requisitor. Como a requisição da previsão é feita a cada 15 minutos, durante este período, o preditor permanece “cego”, ou seja, sem saber se a previsão está correta. Por isso, ele a compara ao último valor passado pelo requisitor, com a última previsão encontrada. Em seguida, ele realiza uma normalização dos dados, quando a diferença entre os dois for superior a 80%. Por último, o algoritmo preditivo converte o resultado para os parâmetros de horizonte e amostragem do requisitor e passa a resposta para o API, que encaminha ao requisitor.

Algumas características do algoritmo preditivo são que ele permite e lida com a mudança de dia, sendo que o maior horizonte permitido para a previsão é de 24 horas e o mínimo de uma hora. Além disso, ele desempenha os cálculos com os dados amostrados a cada 5 minutos e na sequência, faz o ajuste para dados de horizonte e amostragem requeridos.

4 MÉTODOS DE CÁLCULO DO ERRO

As maneiras de avaliar a assertividade de um sistema baseado em *machine learning* são diversas e suas vantagens e desvantagens variam de acordo com o tipo e a finalidade do código. No cenário de algoritmos de ML não supervisionados, esta avaliação dos resultados possui menos avanços, quando comparada ao aprendizado supervisionado (PALACIO-NIÑO; BERZAL, 2019). Além disso, Palacio-Niño e Berzal (2019) ilustram que esta validação pode ser estatística, ou não. E quando for, pode levar em consideração critérios externos ou internos ao sistema. Para o caso estudado neste trabalho, a análise do erro será estatística, levando em conta os valores medidos da geração fotovoltaica e os valores previstos pelo algoritmo preditivo.

4.1 Métodos

De maneira a avaliar a previsão fornecida pelo algoritmo preditivo em questão, buscou-se na literatura os métodos de análise de erros mais indicados para este tipo de sistema. Os métodos estatísticos mais utilizados para isso, encontrados no decorrer da pesquisa foram os seguintes: *mean squared error* (MSE), *root mean squared error* (RMSE), *mean absolute error* (MAE) e *mean absolute percentage error* (MAPE).

Neste capítulo são abordados esses métodos de cálculo de erro, ressaltando suas principais características, vantagens e desvantagens, para que sejam testados e avaliados. Após a avaliação dos métodos, o mais adequado e apropriado para este trabalho será escolhido e aplicado aos dados de previsão do preditor e geração real do sistema. Assim, será possível analisar propriamente o sistema de previsão de geração.

4.1.1 Método MSE

O método *mean squared error* (MSE), traduzido como erro quadrático médio, é amplamente utilizado para avaliar métodos de previsão (JJ, 2016). Nele, a

diferença entre o valor previsto ($\hat{y}_i$) e o valor medido (y_i) é elevada ao quadrado, conforme observado na Equação 1.

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

Sendo assim, o MSE é um método bastante sensível a erros grandes, uma vez que ele eleva ao quadrado a diferença entre valor testado e o valor medido. Em outras palavras, se a diferença for grande, o operador quadrático torna o valor final muito maior, sendo adequada para erros mais reduzidos.

4.1.2 Método RMSE

O método *root mean squared error* (RMSE) equivale à raiz quadrada do MSE, ou seja, calcula "a raiz quadrática média" dos erros entre os valores observados e predições (PRAYUDANI *et al.*, 2019). A Equação 2 representa a fórmula deste método.

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

Uma característica do MSE e RMSE é que a diferença dos valores é elevada ao quadrado antes de ter a média calculada, e por isso o valor final tende a ser bem mais elevado quando há a presença de grandes diferenças. Em casos de *outliers*, por exemplo, seu peso será maior para o cálculo destes métodos e, por conseguinte, prejudicará sua métrica deixando-a maior (JJ, 2016).

Muitos algoritmos usam MSE, pois é mais rápido de computar e mais fácil de manipular em comparação com o RMSE. Porém, não é dimensionado para o erro original, já que a diferença dos valores é elevada ao quadrado e resulta em um índice que não pode ser relacionado à escala original (VANDEPUT, 2021).

Esta característica do MSE dificulta a relação direta com os valores das amostras. Ou seja, se a amostra está em cm, o MSE vai estar em cm², por exemplo, o que não acontece no caso do RMSE, uma vez que a raiz quadrada é extraída no final do cálculo.

4.1.3 Método MAE

O método do erro médio absoluto, do inglês *mean absolut error* (MAE) é o método que calcula a magnitude média dos erros em um conjunto de previsões (JJ, 2016). Em outras palavras ele equivale ao valor médio das diferenças absolutas entre o valor previsto e o medido. A Equação 3 representa o cálculo deste método.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (3)$$

O MAE fornece uma visão de quão próximas das medições as previsões estão, em escala absoluta (LI *et al.*, 2016). Vantagem que possui em comum com o RMSE, sendo que os dois métodos fornecem valores em escala de medição.

4.1.4 Método MAPE

Calcula-se o método *mean absolute percentage error* (MAPE), ou erro percentual absoluto médio, através dos valores de erros absolutos em cada período dividido pelos valores reais de observação para aquele período (PRAYUDANI *et al.*, 2019), conforme Equação 4.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \times 100\% \quad (4)$$

De acordo com Prayudani *et al.* (2019), essa abordagem é útil quando o tamanho das variáveis previstas é importante para avaliar a precisão das previsões.

Brolese (2019) estabelece que o método MAPE é usualmente utilizado nas previsões tanto de demanda quanto de geração por sua simplicidade de interpretação da dimensão do erro, visto que este é representado de forma proporcional a um valor real. Por outro lado, Wang *et al.* (2018), que propõem uma nova abordagem para a previsão FV de curto prazo com base em uma rede de unidades recorrentes com portas, ressaltam que para este caso, o MAPE não é adequado para avaliar os resultados da previsão. Os autores justificam que, como a potência fotovoltaica de saída por vezes é igual à zero, o MAPE se tornará infinito. Outra desvantagem do método é que erros altos durante os períodos de baixa demanda afetarão significativamente o valor final (VANDEPUT, 2021).

4.2 Comparação

Por se tratar de um tema de grande relevância no cenário de aprendizado de máquina, há várias discussões a respeito do melhor método para o cálculo do erro. Li *et al.* (2016) definem a principal diferença entre os métodos sendo que o MAE fornece uma visão de quão próximas as previsões estão das medições em escala absoluta. O RMSE amplifica e pune severamente grandes erros, por conta de usar a forma quadrada.

Chugh (2020) faz uma comparação entre MAE, MSE, RMSE, e outros métodos não abordados neste trabalho (coeficiente de determinação e *Adjusted R Squared*), e eleger o RMSE como uma melhor opção por ser simples de calcular e diferenciável. No entanto, se o conjunto de dados tiver *outliers*, sugere a utilização de MAE em vez de RMSE.

Ao fazer uma comparação dos métodos RMSE, MAE e MAPE, Vandeput (2021) traz ainda o conceito de *bias*, que pode ser traduzido como viés. Esse valor representa o erro médio histórico e sua formulação está representada pela Equação 5.

$$bias = \frac{1}{n} \sum_{i=1}^n \hat{y}_i - y_i \quad (5)$$

A conclusão obtida pelo autor confirma que o MAE protege os *outliers*, enquanto o RMSE garante a obtenção de uma previsão imparcial. Caso o método MAE for escolhido e resultar em um *bias* elevado, o método RMSE pode ser mais adequado. Por outro lado, se o conjunto de dados contiver muitos outliers, resultando em uma previsão distorcida, convém utilizar o MAE (VANDEPUT, 2021).

No intuito de otimizar o modelo, ainda é possível que mais de um método seja aplicado para relatar o erro da previsão. Além disso, outra forma de aprimorar a análise é agregar os valores medidos a diferentes horizontes de tempo (mensal, semanal, trimestral, etc.). Dessa forma, é possível a utilização do MAE, amenizando os picos de demanda (VANDEPUT, 2021).

4.3 Outliers

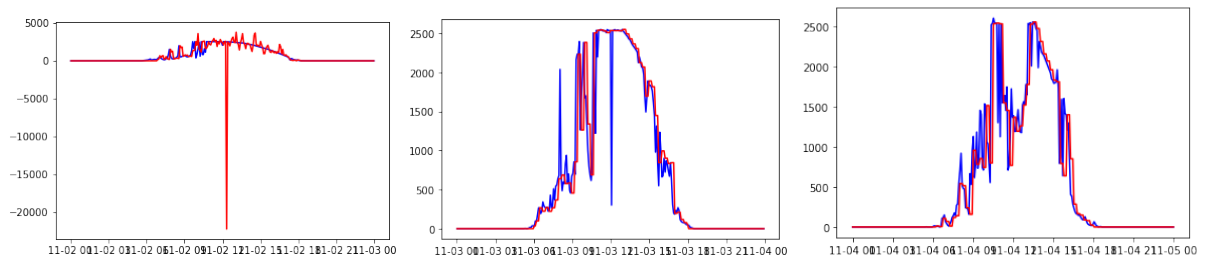
Conforme abordado em diversas das referências supracitadas, a presença de *outliers* no banco de dados analisado é um critério crucial para a escolha do índice de erro aplicado. Ben-Gal (2005) reúne uma série de definições para *outliers*. A primeira define esse evento como uma observação que se desvia tanto de outras a ponto de levantar suspeitas de que foi gerada por um mecanismo diferente. Outra definição é: um desvio acentuado de outros membros da amostra em que ocorre. Ou ainda, uma observação em um conjunto de dados que parece ser inconsistente com o restante do conjunto.

No caso apresentado neste trabalho, o sistema está propenso à presença de *outliers*, que podem ocorrer por uma série de motivos, e com intensidades variadas. Uma possível causa são erros na medição ou então, acontecimentos climáticos inesperados, como uma única nuvem encobrindo o sistema FV em um dia ensolarado. O mesmo pode ocorrer com os valores medidos, em caso de distúrbios na medição ou na plataforma onde os dados são coletados.

A Figura 4 mostra as curvas de previsão e geração relativas a três dias consecutivos, com diferentes níveis de *outliers*. O primeiro gráfico, referente ao dia 02/11/2021, apresenta um *outlier* bastante acentuado na curva de previsão (vermelha) por volta das 12h20min, representando uma possível falha momentânea na medição. Já no segundo gráfico, do dia 03/11/2021, notam-se dois *outliers* na curva de medição (azul), um às 8h e outro às 12h. Neste caso, pode ser que tenha ocorrido algum fenômeno climático inesperado, como mencionado anteriormente. Por fim, no terceiro gráfico, que mostra os dados do dia 04/11/2021, não se identificam *outliers* em nenhuma das curvas.

Neste sentido, é esperado que o método de cálculo do erro priorize a amenização destes outliers, como ocorre com o MAE. Outra solução plausível seria o desenvolvimento de um filtro que suavizasse os efeitos destes distúrbios, e então utilizar outro método, como o RMSE.

Figura 4 – Exemplo de *outliers* nas curvas de geração e previsão



Fonte: Elaboração própria (2021).

4.4 Conclusão e Escolha do Método

Com a revisão dos métodos realizada neste capítulo, e com todos os fatores avaliados e considerados, optou-se por aplicar dois métodos de cálculo do erro da previsão. O primeiro a ser aplicado será o método do erro médio absoluto (MAE), por conta da presença de *outliers* nos dados utilizados, e por ser um método amplamente utilizado como métrica da assertividade de algoritmos preditivos. Além disso, também será aplicado um filtro nos dados medidos e previstos para atenuar o efeito dos outliers mais acentuados. Possibilitando, assim, o cálculo também da raiz quadrática média (RMSE) do sistema, sem que este valor seja afetado pelos distúrbios.

5 ANÁLISE DA PREVISÃO

Neste capítulo é realizada a análise do algoritmo preditivo. Inicialmente os dados de previsão e medição do período estudado são coletados. Na sequência estes dados são filtrados eliminando os *outliers*. Por fim, o erro é calculado e os valores, analisados.

5.1 Metodologia

Neste trabalho é realizada a avaliação do algoritmo preditor desenvolvido pela fundação CERTI para aplicação em microrredes para projeto de P&D. O sistema FV utilizado para os testes do algoritmo corresponde a um conjunto de módulos localizados no IFSC e gerenciados pelo Grupo de Estudos em Sistemas de Energia (GESE), mais especificamente pelo Laboratório de Planejamento Integrado de Recursos Energéticos (LAPR). O laboratório conta com um *website* (<https://gese.florianopolis.ifsc.edu.br/lapr/>) que fornece em tempo real os dados referentes à geração FV. Esses dados de geração foram extraídos e utilizados como dados de entrada pelo preditor, além dos dados meteorológicos da localidade (IFSC). A Figura 5 apresenta uma captura da tela feita do *website*.

Figura 5 – Website com dados de geração



Fonte: GESE (2021).

O monitoramento das previsões do algoritmo foi realizado por um período de 5 meses, sendo que a data a qual deu-se o início ao funcionamento do preditor foi 11 de agosto de 2021, às 11:20. A última coleta de dados ocorreu no dia 11 de janeiro de 2022, totalizando 5 meses, 22 semanas e 154 dias. Passado o período de monitoramento previsto, é possível uma avaliação da assertividade dos valores previstos pelo algoritmo. Para avaliar o sistema e garantir que a análise dos valores previstos seja a melhor possível, realizou-se uma pesquisa referente aos métodos de cálculo de erro usualmente aplicados para avaliação de algoritmos de previsão.

Depois de selecionado o melhor método para o caso estudado, o cálculo é aplicado em bases diária, semanal e mensal, para que a análise possa ser feita para cada período de maneira isolada. Assim, extraídos os dados e a partir do método escolhido, é calculado o erro médio de cada período. Dessa forma, é possível observar e avaliar o comportamento do erro com o passar do tempo, verificando a sua variação no decorrer do período de monitoramento. Na sequência, são investigados os motivos para este comportamento, com a seleção dos piores casos de desvio.

5.2 Coleta de Dados

Conforme explicitado no subcapítulo 5.1 (Metodologia), o início do monitoramento ocorreu no dia 11 de agosto de 2021. A partir deste dia, o algoritmo preditivo rodou coletando os dados referentes ao sistema do LAPR. Esse laboratório foi instituído como um "Laboratório Piloto" por meio do projeto de pesquisa aplicada e desenvolvimento tecnológico intitulado "Gestão Energética pelo Lado da Demanda em unidades consumidoras de Baixa Tensão por meio de Microgeração fotovoltaica" aprovado na Chamada MEC/SETEC/CNPq n. 94/2013 (GESE, 2022). O projeto contou com, além de outras, a instalação de um sistema *on-grid* de 3kWp (12 módulos solares), utilizado neste trabalho para fins de análise. Dessa maneira, o monitoramento deste sistema gerou um total de 44.037 amostras de dados, contendo a data e o horário da medição, bem como o valor medido e o valor previsto da geração, sendo este último fornecido pelo algoritmo preditor.

Com os dados extraídos, é realizada uma filtragem da amostra. Foram removidos da base de dados os valores em que a medição e/ou a previsão da

geração fossem maiores que 3.000W e menores que zero. Do total de 44.037 amostras, 487 foram eliminadas, ou seja, apenas 1,11%. Essas amostras foram removidas por representarem dados inconsistentes, especialmente erros na medição. Por exemplo, quando o valor de medição ou previsão é igual a -9.999, ele indica que houve um erro na medição, que pode ocorrer por diversos motivos, por isso, não é interessante para a análise levar estes valores em consideração.

5.3 Resultados

Nesta seção do trabalho estão expostos da Tabela 3 à Tabela 6 os resultados obtidos. Na Tabela 3 estão os valores de MAE e RMSE total do período e na Tabela 4, os valores relativos a cada mês. A Tabela 5 e a Tabela 6 apresentam os valores semanais e diários, respectivamente. Ademais, a Figura 6, a Figura 7 e a Figura 8 representam os gráficos da evolução dos erros de cada período (mensal, semanal e diário). Os valores encontrados e apresentados nas tabelas supracitadas foram calculados de acordo com o Capítulo 4 e as equações nele presentes.

Tabela 3 – MAE e RMSE para todo o período

	Início	Fim	MAE	RMSE
total	11/08/2021	11/01/2022	0,0347	0,0930

Fonte: Elaboração própria (2022).

Tabela 4 – MAE e RMSE para cada mês

	Início	Fim	MAE	RMSE
mês 1	11/08/2021	10/09/2021	0,0280	0,0760
mês 2	11/09/2021	10/10/2021	0,0305	0,0820
mês 3	11/10/2021	10/11/2021	0,0354	0,0925
mês 4	11/11/2021	10/12/2021	0,0402	0,1078
mês 5	11/12/2021	11/01/2022	0,0393	0,1027

Fonte: Elaboração própria (2022).

Tabela 5 – MAE e RMSE para cada semana

	Início	Fim	MAE	RMSE
semana 1	11/08/2021	17/08/2021	0,0251	0,0726
semana 2	18/08/2021	24/08/2021	0,0262	0,0668
semana 3	25/08/2021	31/08/2021	0,0272	0,0734
semana 4	01/09/2021	07/09/2021	0,0320	0,0865
semana 5	08/09/2021	14/09/2021	0,0300	0,0796
semana 6	15/09/2021	21/09/2021	0,0297	0,0803
semana 7	22/09/2021	28/09/2021	0,0473	0,1185

	Início	Fim	MAE	RMSE
semana 8	29/09/2021	05/10/2021	0,0226	0,0556
semana 9	06/10/2021	12/10/2021	0,0202	0,0516
semana 10	13/10/2021	19/10/2021	0,0382	0,0973
semana 11	20/10/2021	26/10/2021	0,0357	0,0937
semana 12	27/10/2021	02/11/2021	0,0406	0,1039
semana 13	03/11/2021	09/11/2021	0,0319	0,0867
semana 14	10/11/2021	16/11/2021	0,0352	0,0924
semana 15	17/11/2021	23/11/2021	0,0290	0,0814
semana 16	24/11/2021	30/11/2021	0,0412	0,1096
semana 17	01/12/2021	07/12/2021	0,0461	0,1203
semana 18	08/12/2021	14/12/2021	0,0472	0,1168
semana 19	15/12/2021	21/12/2021	0,0484	0,1203
semana 20	22/12/2021	28/12/2021	0,0336	0,1007
semana 21	29/12/2021	04/01/2022	0,0269	0,0730
semana 22	05/01/2022	11/01/2022	0,0478	0,1143

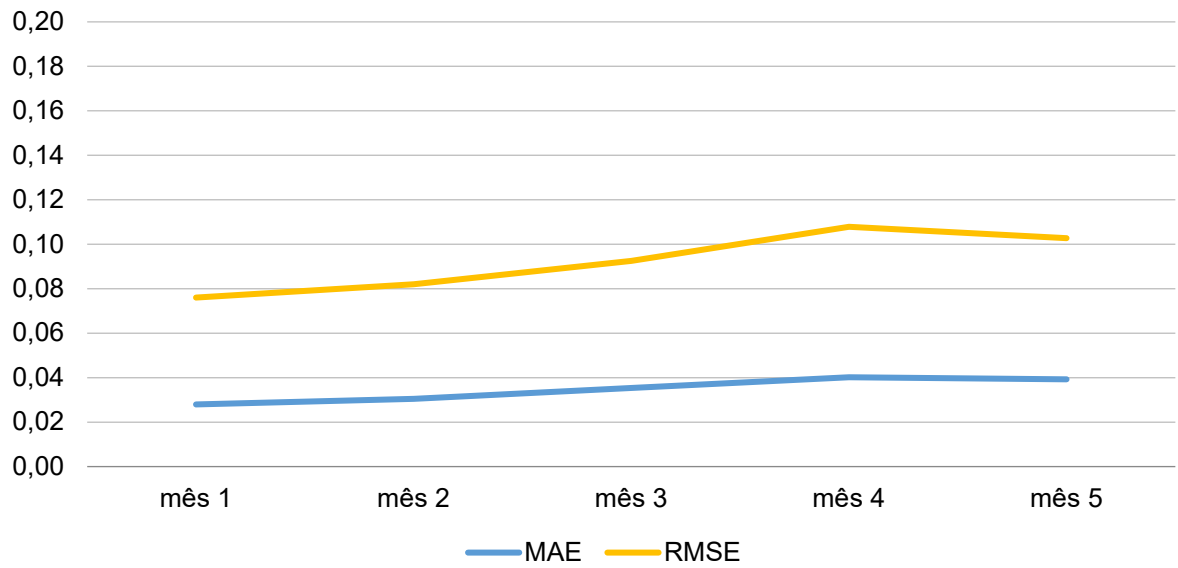
Fonte: Elaboração própria (2022).

Tabela 6 – MAE e RMSE para cada dia

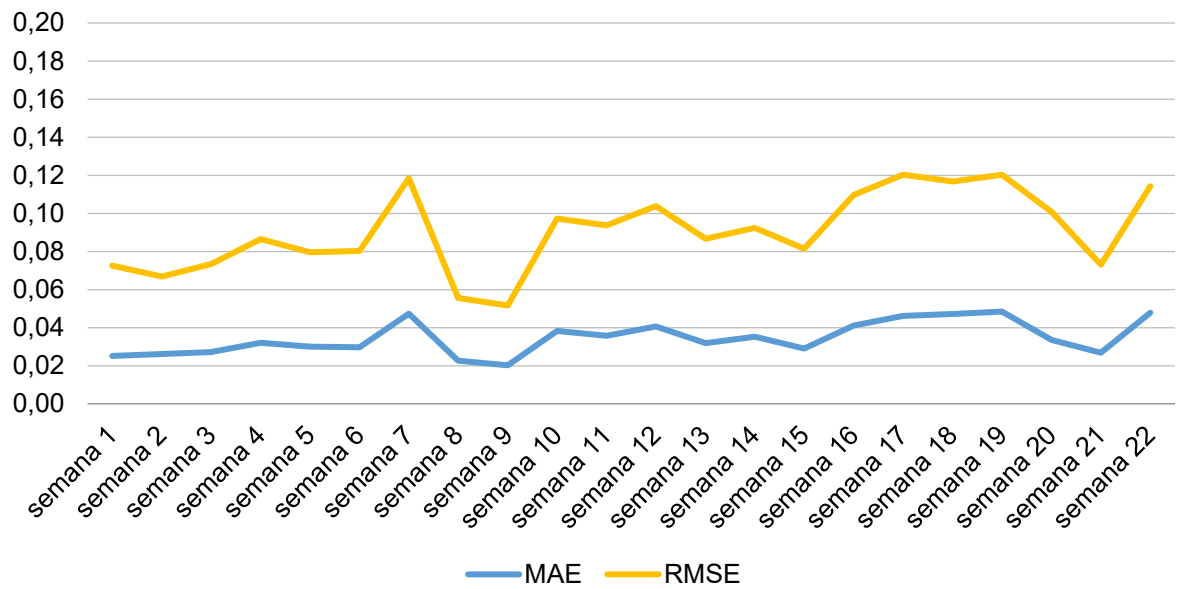
Data		MAE	RMSE	Data		MAE	RMSE
dia 1	11/08/2021	0,0171	0,0309	dia 78	27/10/2021	0,0645	0,1260
dia 2	12/08/2021	0,0123	0,0260	dia 79	28/10/2021	0,0780	0,1673
dia 3	13/08/2021	0,0144	0,0563	dia 80	29/10/2021	0,0116	0,0272
dia 4	14/08/2021	0,0370	0,0971	dia 81	30/10/2021	0,0112	0,0269
dia 5	15/08/2021	0,0476	0,1189	dia 82	31/10/2021	0,0238	0,0630
dia 6	16/08/2021	0,0145	0,0321	dia 83	01/11/2021	0,0372	0,0942
dia 7	17/08/2021	0,0297	0,0740	dia 84	02/11/2021	0,0596	0,1340
dia 8	18/08/2021	0,0297	0,0773	dia 85	03/11/2021	0,0357	0,0992
dia 9	19/08/2021	0,0255	0,0621	dia 86	04/11/2021	0,0401	0,0980
dia 10	20/08/2021	0,0270	0,0763	dia 87	05/11/2021	0,0393	0,0993
dia 11	21/08/2021	0,0318	0,0809	dia 88	06/11/2021	0,0134	0,0429
dia 12	22/08/2021	0,0237	0,0560	dia 89	07/11/2021	0,0500	0,1131
dia 13	23/08/2021	0,0186	0,0470	dia 90	08/11/2021	0,0110	0,0239
dia 14	24/08/2021	0,0271	0,0608	dia 91	09/11/2021	0,0337	0,0903
dia 15	25/08/2021	0,0330	0,0821	dia 92	10/11/2021	0,0291	0,0656
dia 16	26/08/2021	0,0091	0,0241	dia 93	11/11/2021	0,0209	0,0415
dia 17	27/08/2021	0,0175	0,0387	dia 94	12/11/2021	0,0234	0,0770
dia 18	28/08/2021	0,0193	0,0507	dia 95	13/11/2021	0,0286	0,0711
dia 19	29/08/2021	0,0340	0,0867	dia 96	14/11/2021	0,0420	0,0913
dia 20	30/08/2021	0,0226	0,0549	dia 97	15/11/2021	0,0470	0,1222
dia 21	31/08/2021	0,0553	0,1259	dia 98	16/11/2021	0,0564	0,1344
dia 22	01/09/2021	0,0464	0,1075	dia 99	17/11/2021	0,0317	0,0711
dia 23	02/09/2021	0,0532	0,1146	dia 100	18/11/2021	0,0054	0,0109
dia 24	03/09/2021	0,0452	0,1089	dia 101	19/11/2021	0,0363	0,0881
dia 25	04/09/2021	0,0086	0,0186	dia 102	20/11/2021	0,0482	0,1228
dia 26	05/09/2021	0,0355	0,1015	dia 103	21/11/2021	0,0137	0,0300
dia 27	06/09/2021	0,0282	0,0725	dia 104	22/11/2021	0,0198	0,0698
dia 28	07/09/2021	0,0077	0,0186	dia 105	23/11/2021	0,0485	0,1132
dia 29	08/09/2021	0,0037	0,0086	dia 106	24/11/2021	0,0208	0,0462
dia 30	09/09/2021	0,0468	0,1069	dia 107	25/11/2021	0,0397	0,0988

Data		MAE	RMSE	Data		MAE	RMSE
dia 31	10/09/2021	0,0419	0,0974	dia 108	26/11/2021	0,0489	0,1101
dia 32	11/09/2021	0,0477	0,1099	dia 109	27/11/2021	0,0624	0,1502
dia 33	12/09/2021	0,0271	0,0630	dia 110	28/11/2021	0,0623	0,1519
dia 34	13/09/2021	0,0310	0,0809	dia 111	29/11/2021	0,0403	0,1150
dia 35	14/09/2021	0,0126	0,0309	dia 112	30/11/2021	0,0140	0,0363
dia 36	15/09/2021	0,0319	0,0803	dia 113	01/12/2021	0,0701	0,1811
dia 37	16/09/2021	0,0169	0,0420	dia 114	02/12/2021	0,0196	0,0401
dia 38	17/09/2021	0,0062	0,0149	dia 115	03/12/2021	0,0242	0,0747
dia 39	18/09/2021	0,0666	0,1426	dia 116	04/12/2021	0,0347	0,0834
dia 40	19/09/2021	0,0216	0,0520	dia 117	05/12/2021	0,0434	0,0992
dia 41	20/09/2021	0,0296	0,0766	dia 118	06/12/2021	0,0557	0,1303
dia 42	21/09/2021	0,0360	0,0899	dia 119	07/12/2021	0,0750	0,1662
dia 43	22/09/2021	0,0714	0,1741	dia 120	08/12/2021	0,0853	0,1843
dia 44	23/09/2021	0,0704	0,1620	dia 121	09/12/2021	0,0414	0,0944
dia 45	24/09/2021	0,0352	0,0984	dia 122	10/12/2021	0,0425	0,1263
dia 46	25/09/2021	0,0471	0,1139	dia 123	11/12/2021	0,0607	0,1270
dia 47	26/09/2021	0,0269	0,0537	dia 124	12/12/2021	0,0424	0,1098
dia 48	27/09/2021	0,0285	0,0565	dia 125	13/12/2021	0,0307	0,0655
dia 49	28/09/2021	0,0524	0,1156	dia 126	14/12/2021	0,0274	0,0644
dia 50	29/09/2021	0,0351	0,0787	dia 127	15/12/2021	0,0538	0,1275
dia 51	30/09/2021	0,0248	0,0562	dia 128	16/12/2021	0,0510	0,1267
dia 52	01/10/2021	0,0185	0,0407	dia 129	17/12/2021	0,0688	0,1502
dia 53	02/10/2021	0,0290	0,0687	dia 130	18/12/2021	0,0289	0,0585
dia 54	03/10/2021	0,0180	0,0447	dia 131	19/12/2021	0,0861	0,1863
dia 55	04/10/2021	0,0160	0,0465	dia 132	20/12/2021	0,0205	0,0515
dia 56	05/10/2021	0,0170	0,0416	dia 133	21/12/2021	0,0300	0,0755
dia 57	06/10/2021	0,0312	0,0789	dia 134	22/12/2021	0,0351	0,0920
dia 58	07/10/2021	0,0232	0,0569	dia 135	23/12/2021	0,0481	0,1383
dia 59	08/10/2021	0,0132	0,0276	dia 136	24/12/2021	0,0540	0,1482
dia 60	09/10/2021	0,0136	0,0425	dia 137	25/12/2021	0,0125	0,0321
dia 61	10/10/2021	0,0173	0,0375	dia 138	26/12/2021	0,0176	0,0660
dia 62	11/10/2021	0,0215	0,0483	dia 139	27/12/2021	0,0413	0,1117
dia 63	12/10/2021	0,0215	0,0534	dia 140	28/12/2021	0,0267	0,0603
dia 64	13/10/2021	0,0650	0,1392	dia 141	29/12/2021	0,0548	0,1210
dia 65	14/10/2021	0,0331	0,0827	dia 142	30/12/2021	0,0556	0,1159
dia 66	15/10/2021	0,0122	0,0259	dia 143	31/12/2021	0,0177	0,0368
dia 67	16/10/2021	0,0648	0,1443	dia 144	01/01/2022	0,000	0,0000
dia 68	17/10/2021	0,0159	0,0331	dia 145	02/01/2022	0,0152	0,0379
dia 69	18/10/2021	0,0199	0,0384	dia 146	03/01/2022	0,0254	0,0686
dia 70	19/10/2021	0,0577	0,1284	dia 147	04/01/2022	0,0194	0,0410
dia 71	20/10/2021	0,0382	0,1058	dia 148	05/01/2022	0,0558	0,1392
dia 72	21/10/2021	0,0357	0,0929	dia 149	06/01/2022	0,0373	0,0920
dia 73	22/10/2021	0,0201	0,0550	dia 150	07/01/2022	0,0342	0,0801
dia 74	23/10/2021	0,0397	0,1069	dia 151	08/01/2022	0,0399	0,0929
dia 75	24/10/2021	0,0540	0,1311	dia 152	09/01/2022	0,0474	0,1163
dia 76	25/10/2021	0,0229	0,0645	dia 153	10/01/2022	0,0386	0,0887
dia 77	26/10/2021	0,0395	0,0769	dia 154	11/01/2022	0,0882	0,1734

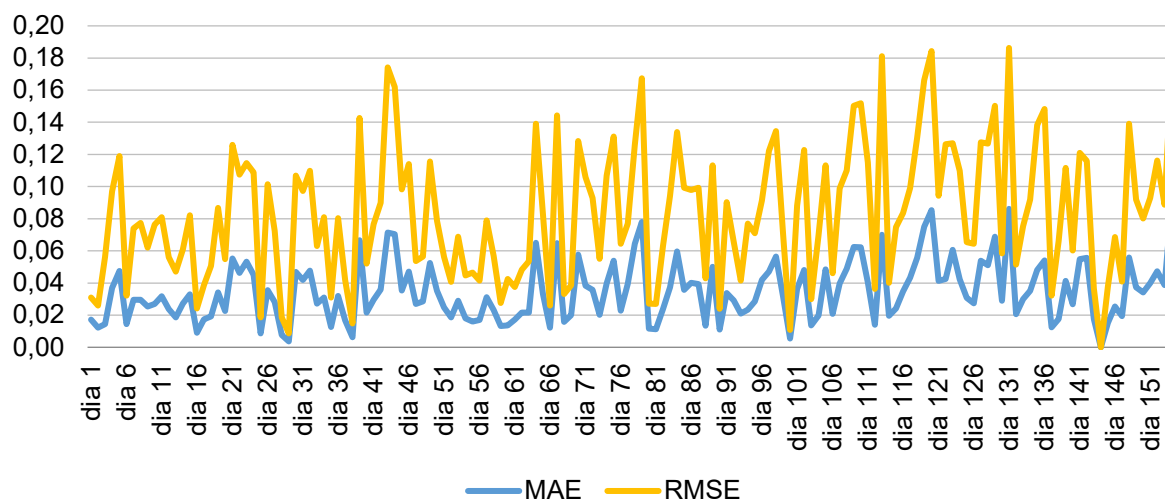
Fonte: Elaboração própria (2022).

Figura 6 – Evolução mensal do erro

Fonte: Elaboração própria (2022).

Figura 7 – Evolução semanal do erro

Fonte: Elaboração própria (2022).

Figura 8 – Evolução diária do erro

Fonte: Elaboração própria (2022).

5.4 Análise dos Erros

O resultado encontrado para os valores em de MAE e RMSE, em pu (por unidade) foram, respectivamente, 0,0347 e 0,0930, considerando os dados filtrados de todo o período analisado. Para fins de comparação, foram pesquisados na literatura trabalhos semelhantes a este, que tivessem calculado o erro dessa maneira. A Tabela 7 mostra essas referências e os valores resultantes de cada uma.

Tabela 7 – Comparação dos resultados

Referência	MAE	RMSE
Karasu <i>et al.</i> (2017)	0,1500	0,1600
	0,0300	0,0500
Wang, Liao e Chang (2018)	0,0399	0,0708
Elmousalami (2020)	-	0,0109

Fonte: Elaboração própria (2022).

Tendo como base os valores da Tabela 7, os valores encontrados indicam uma boa previsão fornecida pelo algoritmo. Isto é, a assertividade do sistema está acima da média encontrada para este tipo de trabalho.

No entanto, o comportamento indicado nos gráficos da Figura 6, da Figura 7 e da Figura 8 foram diferentes do esperado. A expectativa era de que, com o passar do tempo, o erro fosse diminuir, já que o algoritmo tende a ficar mais preciso conforme o tempo de execução. Todavia, é evidente que para ambos MAE e RMSE,

a curva apresenta uma leve subida. Uma razão para este evento pode ser a troca de estações. O início do monitoramento ocorreu em agosto, inverno no hemisfério sul, e seguiu passando pela primavera e encerrando em janeiro, no auge do verão. Esse período é marcado por inúmeras mudanças climáticas, tanto na temperatura, quanto no quesito nebulosidade e chuvas. Esses fatores podem ter ocasionado o leve aumento nos valores de erro. Outro ponto relevante para esta análise é o período de 15 minutos que o preditor permanece “cego”, que pode estar ligado aos erros de previsão.

Contudo, mesmo com valores de MAE e RMSE levemente acima do esperado, ao extrair gráficos com as curvas de medição e previsão do sistema de diferentes dias de monitoramento, nota-se que os valores estão significativamente próximos. Para ilustrar esta ocorrência, foram selecionados os três dias com valores de MAE mais elevados e mais inferiores. O mesmo foi feito com os valores de RMSE. A Tabela 8 e a Tabela 9 indicam respectivamente as datas selecionadas e os valores de MAE e RMSE de cada data.

Tabela 8 – Valores de dias com MAE mais baixos e mais altos

		Data	MAE	RMSE
<i>Valores mais baixos</i>	dia 29	08/09/2021	0,0037	0,0086
	dia 100	18/11/2021	0,0054	0,0109
	dia 38	17/09/2021	0,0062	0,0149
<i>Valores mais altos</i>	dia 120	08/12/2021	0,0853	0,1843
	dia 131	19/12/2021	0,0861	0,1863
	dia 154	11/01/2022	0,0882	0,1734

Fonte: Elaboração própria (2022).

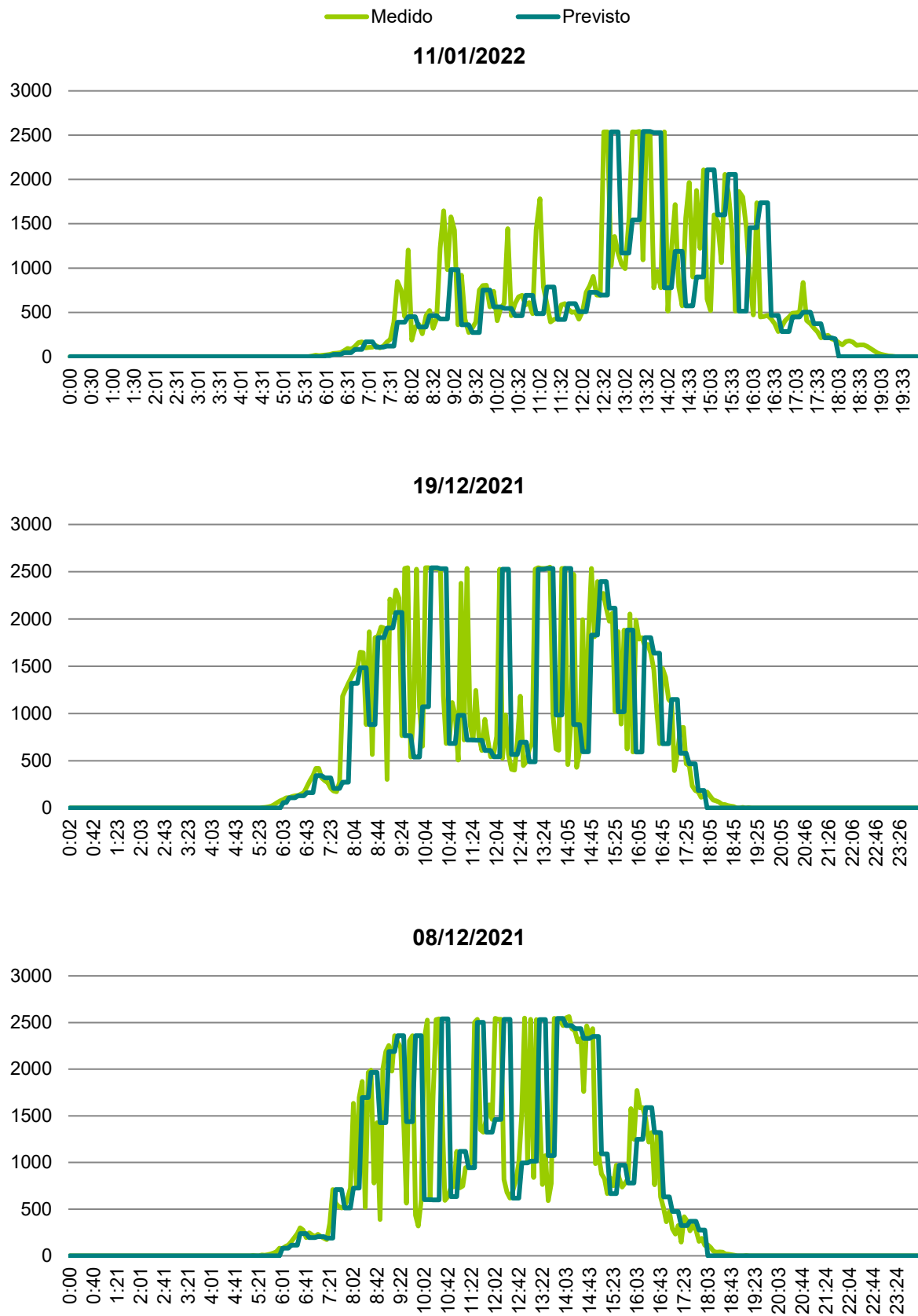
Tabela 9 – Valores de dias com RMSE mais baixos e mais altos

		Data	MAE	RMSE
<i>Valores mais baixos</i>	dia 29	08/09/2021	0,0037	0,0086
	dia 100	18/11/2021	0,0054	0,0109
	dia 38	17/09/2021	0,0062	0,0149
<i>Valores mais altos</i>	dia 113	01/12/2021	0,0701	0,1811
	dia 120	08/12/2021	0,0853	0,1843
	dia 131	19/12/2021	0,0861	0,1863

Fonte: Elaboração própria (2022).

A Figura 9 mostra as curvas dos dias com maior valor de MAE. As curvas dos dias com valores mais altos de RMSE estão na Figura 10. A Figura 11 mostra as curvas dos dias com menores MAE e RMSE.

Figura 9 – Curvas de geração e previsão: Dias com MAE mais altos



Fonte: Elaboração própria (2022).

Figura 10 – Curvas de geração e previsão: Dias com RMSE mais altos

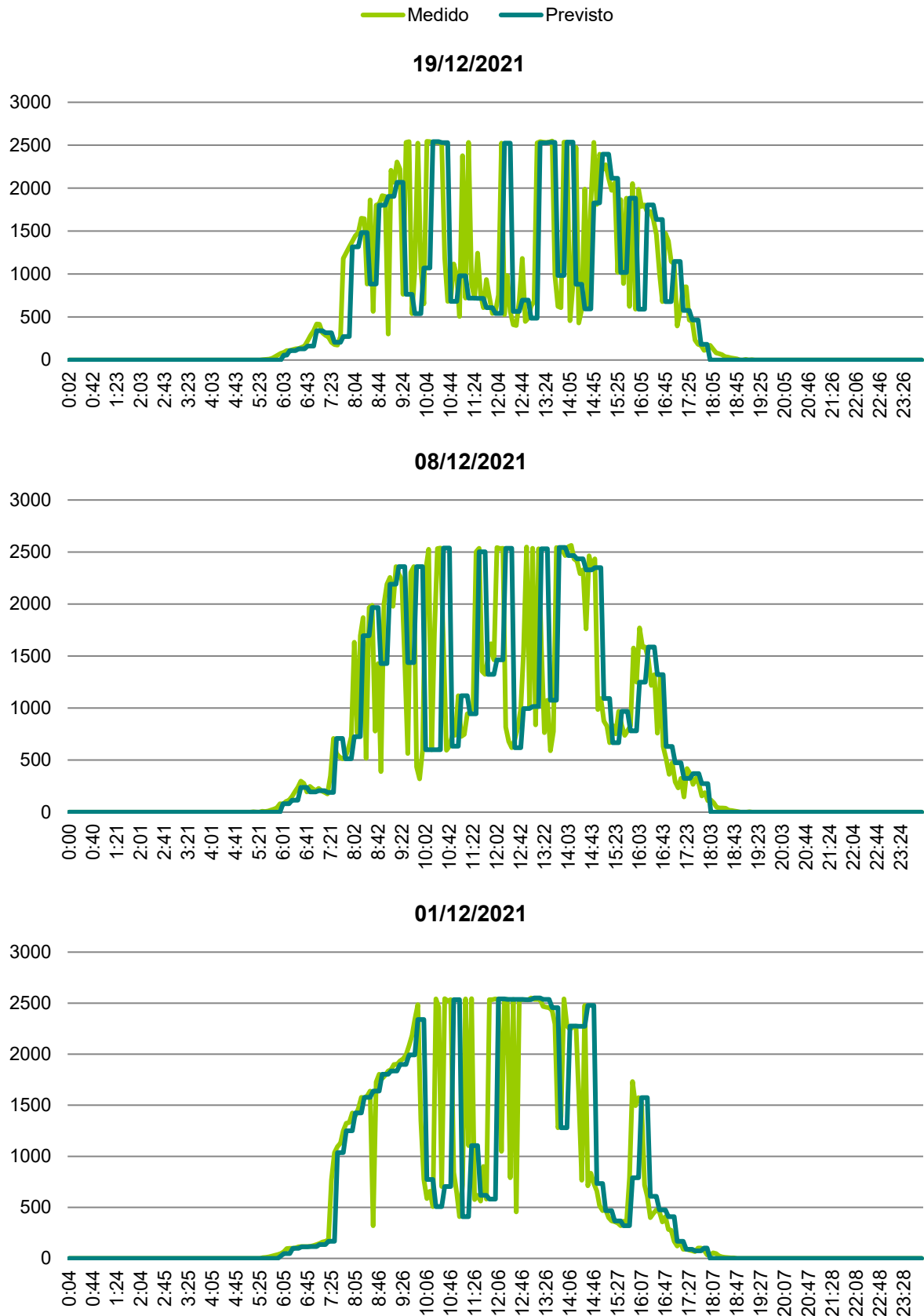
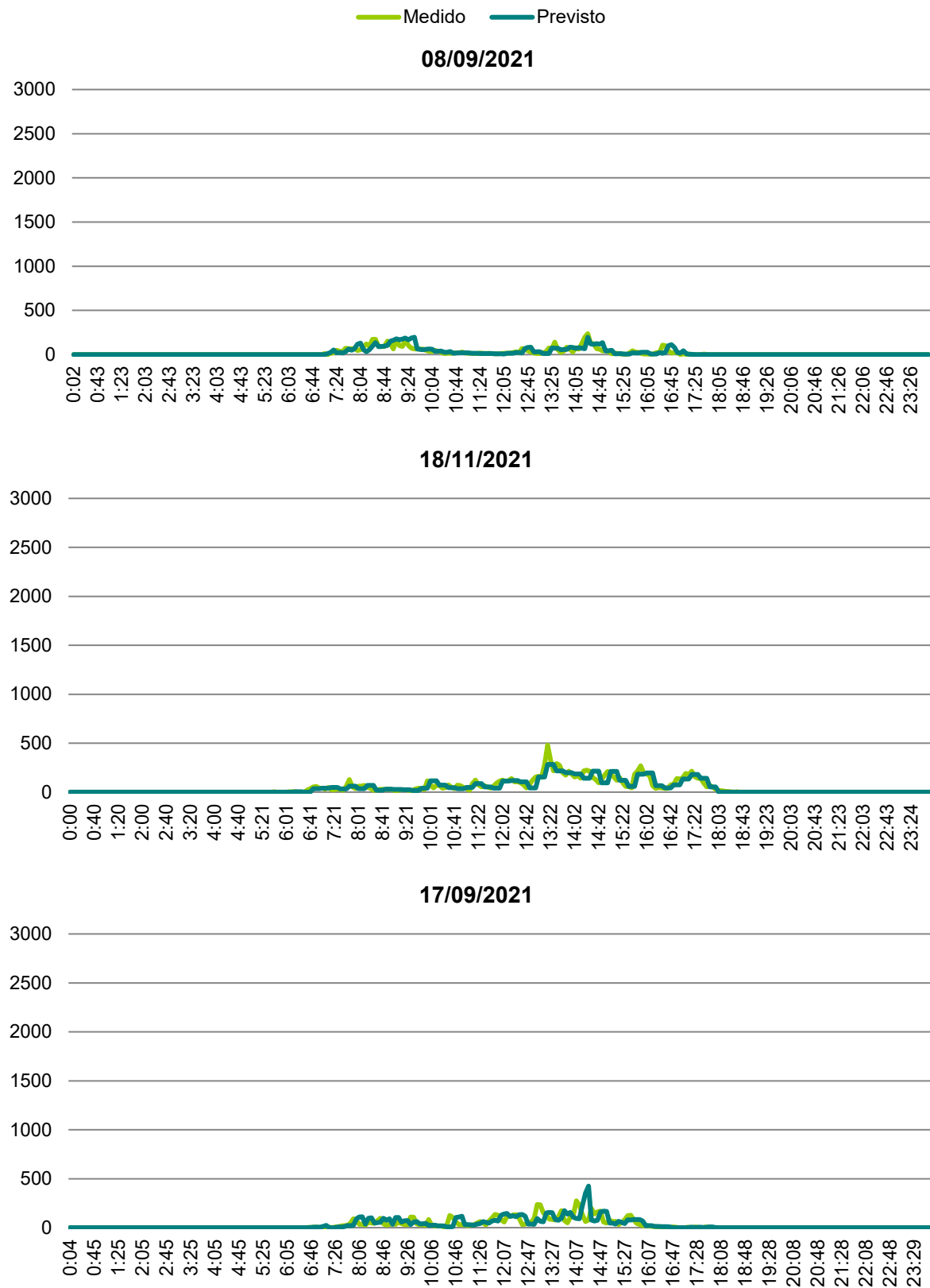


Figura 11 – Curvas de geração e previsão: Dias com MAE e RMSE mais baixos



Fonte: Elaboração própria (2022).

É perceptível que os dias com menores valores de erro, tanto MAE quanto RMSE, foram dias em que a geração foi baixa, com valores de pico inferiores a 500W, por terem sido dias majoritariamente encobertos. Como tais valores de geração foram baixos, quando comparados com outros dias, é possível deduzir que os valores baixos de MAE e RMSE se deram mais em decorrência disso do que da assertividade do preditor propriamente dita. Vale ressaltar também que nos dias com valores altos de erro, as curvas de medição e previsão se mostram suficientemente próximas uma da outra, indicando uma boa previsão.

Ainda a título de exemplo, foram extraídos gráficos de outros dias com comportamento diferentes. Na Figura 12 estão as curvas de medição e previsão dos dias 103, 137 e 145, tendo sido estes dias ensolarados e com ausência de nuvens. É esperado que em dias assim a previsão seja muito próxima da geração real, como é evidente nos gráficos. Por outro lado, ao verificar os valores de MAE e RMSE destes dias (Tabela 10), nota-se que são valores relativamente altos.

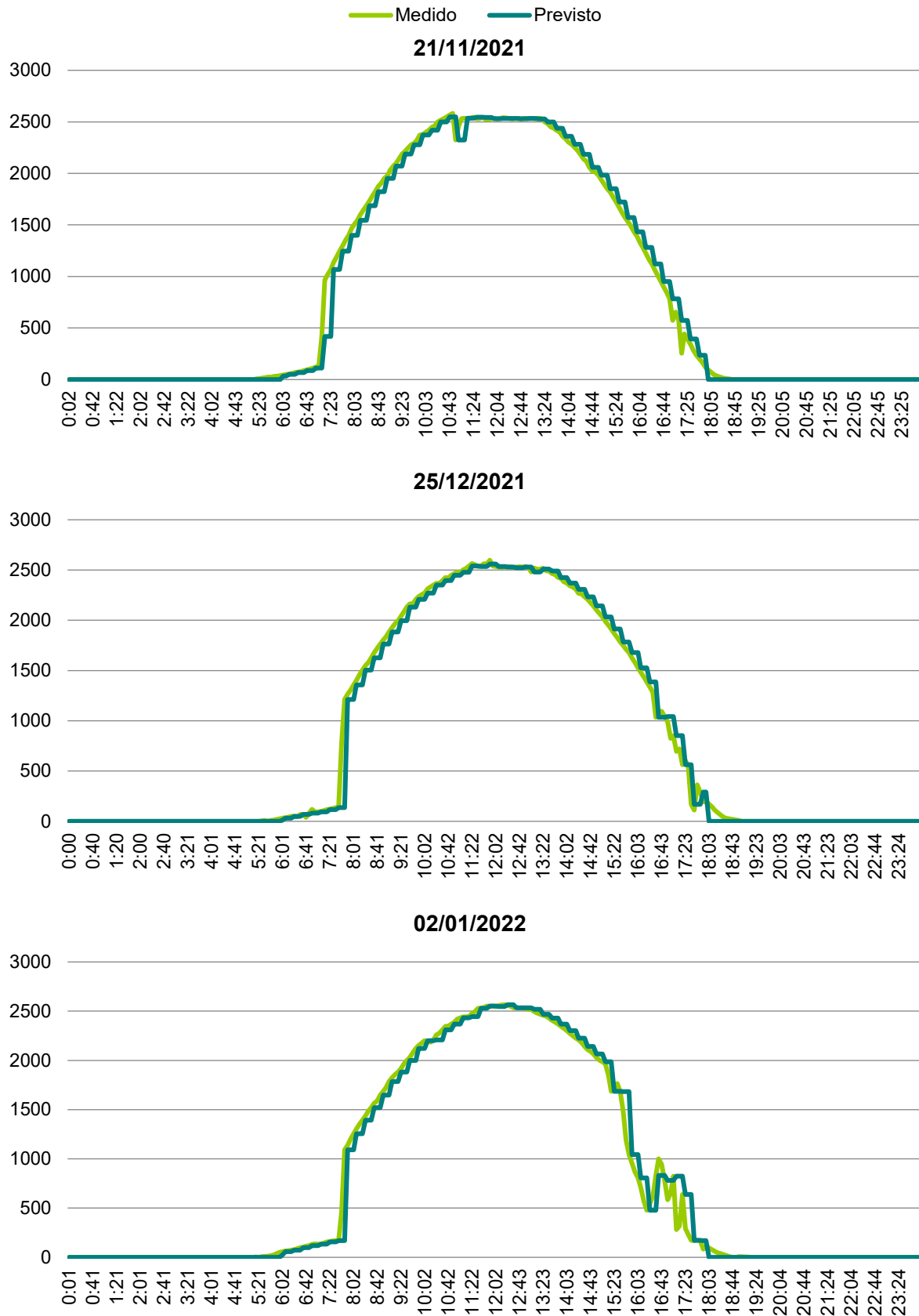
Tabela 10 – Valores de MAE e RMSE de dias ensolarados

	Data	MAE	RMSE
dia 103	21/11/2021	0,0137	0,0300
dia 137	25/12/2021	0,0125	0,0321
dia 145	02/01/2022	0,0152	0,0379

Fonte: Elaboração própria (2022).

Com isso, pode-se considerar a possibilidade de o método MAE e RMSE não representar tão efetivamente a assertividade da previsão fornecida pelo algoritmo preditivo. Uma vez que os valores por vezes representam um erro elevado, mas na prática as curvas de previsão se mostram próximas o suficiente da geração real. Isto é, mesmo que, no geral, os valores obtidos tenham sido satisfatórios, ainda não condizem com o que se observa nas curvas de medição e previsão.

Figura 12 – Curvas de geração e previsão: Dias ensolarados



Fonte: Elaboração própria (2022).

6 CONCLUSÃO

Levando em consideração o atual cenário brasileiro das microrredes e sua relevância para a diversificação da matriz energética, e também a importância de garantir que ela opere de forma otimizada, este trabalho realizou a análise de um algoritmo desenvolvido para prever a geração fotovoltaica de microrredes baseado em aprendizado de máquina não supervisionado.

Inicialmente foi apresentada uma revisão bibliográfica sobre o tópico previsão da geração e uma contextualização sobre o projeto de P&D Microrredes II desenvolvido pela fundação CERTI. Em seguida, foram abordados os principais métodos para o cálculo do erro de algoritmos preditivos e escolhidos os dois mais adequados. O algoritmo preditivo do projeto Microrredes II foi testado em um sistema de geração fotovoltaica localizado no IFSC Florianópolis por um período de 5 meses. Os dados reais de geração e a previsão fornecida pelo preditor foram extraídos, filtrados e analisados. A análise se deu através do cálculo do MAE e RMSE para o período total, bem como para cada mês, semana e dia monitorado.

Os valores totais encontrados foram de 0,0347 para o MAE e 0,0930 para o RMSE. A conclusão alcançada foi que comparado a outros trabalhos semelhantes, o resultado foi satisfatório, considerando o curto período de execução do algoritmo. Além disso, ao analisar os gráficos comparativos da previsão em relação à geração real, foi possível notar que, na maioria dos casos, as curvas estavam suficientemente próximas umas das outras, mesmo nos casos em que os valores de MAE e RMSE diários estavam elevados.

6.1 Trabalhos Futuros

Para trabalhos futuros, seria interessante monitorar o preditor por um período maior de tempo, para que seu aprendizado pudesse incluir também situações como outras estações não englobadas no período estudado e diferentes períodos climáticos, uma vez que a localização geográfica do sistema apresenta estações do ano bem definidas. E desta forma, verificar se os valores de MAE e RMSE apresentariam uma redução em relação ao presente resultado.

Outra possibilidade seria avaliar outros métodos de cálculo do erro da previsão, de maneira que os valores encontrados sejam mais consistentes e adequados para a análise do algoritmo e os gráficos comparativos de geração e previsão.

REFERÊNCIAS

- ANEEL. Agência Nacional de Energia Elétrica. **Micro-rede**. Disponível em: https://www.aneel.gov.br/home?p_p_id=101&p_p_lifecycle=0&p_p_state=maximized&p_p_mode=view&_101_struts_action=%2Fasset_publisher%2Fview_content&_101_returnToFullPageURL=%2F&_101_assetEntryId=15050867&_101_type=content&_101_groupId=656835&_101_urlTitle=micro-rede&inheritRedirect=true. Acesso em: 10 jan. 2022.
- BEN-GAL, Irad. Outlier detection. *In: Data mining and knowledge discovery handbook*. Springer, Boston, MA, 2005. p. 131-146.
- BROLESE, Roan Roberto. **Previsão de Geração de Energia Fotovoltaica Utilizando Método de Aprendizado de Máquina**. 2019. 84 f. TCC (Graduação) - Curso de Engenharia Elétrica, Universidade de Caxias do Sul, Caxias do Sul, 2019.
- CASSIANO, Keila Mara. **Análise de Séries Temporais Usando Análise Espectral Singular (SSA) e Clusterização de Suas Componentes Baseada em Densidade**. 2014. 155 f. Tese (Doutorado) - Curso de Engenharia Elétrica, Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2014.
- CHEN, Hao; ZHANG, Chao; JIA, Ninghong; DUNCAN, Ian; YANG, Shenglai; YANG, YongZhi. A machine learning model for predicting the minimum miscibility pressure of CO₂ and crude oil system based on a support vector machine algorithm approach. **Fuel**, v. 290, p. 120048, 2021.
- CHUGH, Akshita. **MAE, MSE, RMSE, Coefficient of Determination, Adjusted R Squared — Which Metric is Better?** 2020. Disponível em: <https://medium.com/analytics-vidhya/mae-mse-rmse-coefficient-of-determination-adjusted-r-squared-which-metric-is-better-cd0326a5697e>. Acesso em: 06 nov. 2021.
- ELMOUSALAMI, Haytham H. Comparison of Artificial Intelligence Techniques for Project Conceptual Cost Prediction: A Case Study and Comparative Analysis. **IEEE Transactions on Engineering Management**, v. 68, p. 183-196, 2021.
- EPE. **Balanço Energético Nacional 2021: Ano base 2020**. Rio de Janeiro, 2021.
- FALCÃO, Djalma M. Smart Grids e Microrredes: o futuro já é presente. **Simpósio de Automação de Sistemas Elétricos**. Rio de Janeiro, Rio de Janeiro, 2009.
- GESE (Florianópolis SC). Planejamento Integrado de Recursos Energéticos. **Potência ativa em tempo real**. 2021. Disponível em: <https://gese.florianopolis.ifsc.edu.br/lapr/potencia.php>. Acesso em: 22 ago. 2021.
- GESE (Florianópolis SC). Planejamento Integrado de Recursos Energéticos. **Quem somos? – O Laboratório (LAPR)**. 2022. Disponível em: <https://gese.florianopolis.ifsc.edu.br/lapr/about.php>. Acesso em: 10 jan. 2022.

GMI. Global Market Insights. **Microgrid Market**. Disponível em: <https://www.gminsights.com/industry-analysis/microgrid-market>. Acesso em: 10 jan. 2022.

HUAYLLAS, Tesoro Elenadel Carpio. **Micro-Redes Elétricas: Estado da Arte e Contribuição para o Dimensionamento, Aplicação e Comercialização da Energia**. Produzida. 2015. 162 f. Tese (Doutorado) - Curso de Sistemas de Potência, Departamento de Engenharia de Energia e Automação Elétricas, Escola Politécnica da Universidade de São Paulo, São Paulo, 2015.

INEE. Instituto Nacional de Eficiência Energética. **O que é geração distribuída**. Disponível em: [http://www.inee.org.br/forum_ger_distrib.asp#:~:text=Gera%C3%A7%C3%A3o%20Distribu%C3%ADa%20\(GD\)%20%C3%A9%20uma,incluir%20pot%C3%Aancias%20cada%20vez%20menores](http://www.inee.org.br/forum_ger_distrib.asp#:~:text=Gera%C3%A7%C3%A3o%20Distribu%C3%ADa%20(GD)%20%C3%A9%20uma,incluir%20pot%C3%Aancias%20cada%20vez%20menores). Acesso em: 10 jan. 2022

JJ. **MAE and RMSE — Which Metric is Better?**: mean absolute error versus root mean squared error. Mean Absolute Error versus Root Mean Squared Error. 2016. Disponível em: <https://medium.com/human-in-a-machine-world/mae-and-rmse-which-metric-is-better-e60ac3bde13d>. Acesso em: 06 nov. 2021.

KARASU, Seckin; ALTAN, Aytac; SARAC, Zehra; HACIOGLU, Rifat. Prediction of solar radiation based on machine learning methods. **The journal of cognitive systems**, v. 2, p. 16-20, 2018.

LEKSHMI, M.; SUBRAMANYA, KN Adithya. Short-Term Load Forecasting of 400kV Grid Substation Using R-Tool and Study of Influence of Ambient Temperature on the Forecasted Load. In: 2019 **Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)**. IEEE, Gangtok, India, 2019. p. 1-5.

LI, Zhaoxuan; RAHMAN, SM Mahbobur; VEGA, Rolando; DONG, Bing. A Hierarchical Approach Using Machine Learning Methods in Solar Photovoltaic Energy Production Forecasting. **Energies**, San Antonio, EUA, v. 11, n. 8, p. 2163-2174, 19 jan. 2016.

MARTINS, Rodrigo Pizetti. **Previsão Probabilística de Energia Solar Utilizando Redes Neurais com Inferência Bayesiana e Mistura de Gaussianas**. 2018. Dissertação (Mestrado), Universidade Federal Fluminense, Niterói, Rio de Janeiro.

OHI, Daniel Kenji de Alencar. **Solução de Gerenciamento e Controle de Carga Distribuído para Gestão de Energia Elétrica Residencial**. 2018. 90 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Universidade Federal do Ceará, Fortaleza, 2018.

PALACIO-NIÑO, Julio-Omar; BERZAL, Fernando. Evaluation metrics for unsupervised learning algorithms. **arXiv preprint arXiv:1905.05667**, Granada, Espanha. 2019.

PRAYUDANI, S., HIZRIADI, A., LASE, Y. Y., FATMI, Y. E AL-KHOWARIZMI. Analysis Accuracy Of Forecasting Measurement Technique On Random K-Nearest

Neighbor (RKNN) Using MAPE And MSE. In: **Journal of Physics: Conference Series**. IOP Publishing, 2019. p. 012089.

SHARMA, Navin; SHARMA, Pranshu; IRWIN, David; SHENOY, Prashant. Predicting solar generation from weather forecasts using machine learning. **2011 IEEE International Conference on Smart Grid Communications, SmartGridComm 2011**, 528–533.

SOBRI, Sobrina; KOOHI-KAMALI, Sam; RAHIM, Nasrudin Abd. Solar photovoltaic generation forecasting methods: A review. **Energy Conversion and Management**, v. 156, p. 459-497, 2018.

VANDEPUT, Nicolas. **Forecast KPIs: RMSE, MAE, MAPE & Bias**. 2019. Disponível em: <https://towardsdatascience.com/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d>. Acesso em: 06 nov. 2021.

VOYANT, Cyril; NOTTON, Gilles; KALOGIROU, Soteris; NIVET, Marie Laure; PAOLI, Christophe; MOTTE, Fabrice; FOUILLOY, Alexis. Machine learning methods for solar radiation forecasting: A review. **Renewable Energy**, v. 105, p. 569-582, 2017.

WANG, Yusen, LIAO, Wenlong e CHANG, Yuqing. Gated Recurrent Unit Network-Based Short-Term Photovoltaic Forecasting. **Energies**, Basel, Suíça, v. 11, n. 8, p. 1-14, 18 ago. 2018.