



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA

RÔMULO GOMES DOS SANTOS

**PREVISÃO DA EFICIÊNCIA DOS MÓDULOS FOTOVOLTAICOS UTILIZANDO
TÉCNICAS DE APRENDIZADO DE MÁQUINA**

FORTALEZA

2023

RÔMULO GOMES DOS SANTOS

PREVISÃO DA EFICIÊNCIA DOS MÓDULOS FOTOVOLTAICOS UTILIZANDO
TÉCNICAS DE APRENDIZADO DE MÁQUINA

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Paulo Cesar Marques Carvalho.

Coorientadora: Ma. Leticia de Oliveira Santos.

FORTALEZA

2023

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

- S238p Santos, Rômulo Gomes dos.
Previsão da eficiência dos módulos fotovoltaicos utilizando técnicas de aprendizado de máquina /
Rômulo Gomes dos Santos. – 2023.
65 f. : il. color.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Tecnologia,
Curso de Engenharia Elétrica, Fortaleza, 2023.
Orientação: Prof. Dr. Paulo Cesar Marques Carvalho.
Coorientação: Profa. Ma. Leticia de Oliveira Santos.
1. Aprendizado de máquina. 2. Previsão da eficiência. 3. Energia solar. 4. Eficiência. I. Título.
CDD 621.3
-

RÔMULO GOMES DOS SANTOS

PREVISÃO DA EFICIÊNCIA DOS MÓDULOS FOTOVOLTAICOS UTILIZANDO
TÉCNICAS DE APRENDIZADO DE MÁQUINA

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia Elétrica do Centro de Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia Elétrica.

Aprovada em: 06/07/2023.

BANCA EXAMINADORA

Prof. Dr. Paulo Cesar Marques
Carvalho (Orientador)
Universidade Federal do Ceará (UFC)

Ma. Leticia de Oliveira Santos (Coorientadora)
Universidade Federal do Ceará (UFC)

Prof. Dr. Raimundo Furtado Sampaio
Universidade Federal do Ceará (UFC)

Dedico este trabalho à minha família, à todos os seus leitores e, principalmente, à minha mãe, que, quanto ao conhecimento das adversidades, apreendido através de uma longa experiência da vida, ofereceu-me, sem presunção, os saberes que não recebi de nenhum crítico de prestígio.

AGRADECIMENTOS

À minha mãe Ana Paula, pelo apoio e confiança, considerando que de sua parte não pôde ser feito maior oferecimento senão dar-me a faculdade de poder, em tempo assaz, compreender tudo aquilo que eu, em poucos anos e com tantos incômodos e perigos, vim a conhecer. À minha irmã Vitória, por perceber os erros que eu não percebo e me fazer refletir seriamente sobre eles. À minha avó Tereza, pelo carinho e amor compartilhado ao longo dos anos. Aos meus tios e tias, pela orientação daquilo que entendem como o verdadeiro bem, e à toda a minha família pela imensa felicidade.

À Universidade Federal do Ceará (UFC), instituição que possibilitou minha formação como engenheiro e à todos os trabalhadores das coordenações, departamentos, laboratórios e restaurante universitário que compõem a Universidade citada, que deram suporte às minhas atividades acadêmicas. Sem o trabalho desses trabalhadores minha formação seria inviável. À empresa que forneceu os dados necessários para a viabilização deste trabalho.

Aos professores que compõem a banca examinadora, quanto ao conhecimento transmitido e contribuições para este trabalho. Especificamente, agradeço imensamente ao Prof. Dr. Bruno de Athayde Prata, que auxiliou-me a enxergar um pouco mais a realidade. Ao Téc. Eliezer Ribeiro da Silva Brandão. Ao Prof. Me. Gustavo Castelo Branco. Ao Prof. Dr. Fernando Luiz Marcelo Antunes. Ao Prof. Dr. Raimundo Furtado Sampaio e a mais recente colega, Ma. Letícia de Oliveira Santos, por estabelecer o formato e esclarecer minhas dúvidas na elaboração deste trabalho, sem a qual não teria concluído.

Aos colegas da turma de graduação, pelas contribuições constantes de suas reflexões teóricas, pela insistência na crítica daquilo que foi estudado e na preocupação em tornarem-se profissionais. É certo que esse obstáculo foi ultrapassado, na prática, pelas análises contidas também nas ideias desses colegas.

”Embora não seja suficiente falar, é preciso tentar compreender, conhecer e explicar. Para fazê-lo, não se deve hesitar em tomar, sem rodeios, os problemas pela raiz.”

(Nicos Poulantzas, 1980, p. 13.)

RESUMO

O avanço da energia solar como uma fonte renovável de eletricidade tem despertado o interesse em entender e prever o desempenho dos módulos fotovoltaicos (FV) em diferentes condições ambientais. Este trabalho tem como objetivo principal a previsão da eficiência dos módulos FV por meio do uso de técnicas de Aprendizado de Máquina (*Machine Learning* - ML). A previsão da eficiência dos módulos é essencial para o planejamento e o monitoramento eficaz de sistemas de energia solar, além de ser um componente-chave na tomada de decisões relacionadas à geração FV. Para tal, de início, são apresentados os princípios e as técnicas de ML utilizadas na previsão da eficiência dos módulos. Dentre os algoritmos de ML explorados estão Redes Neurais Artificiais (RNA), *Support Vector Machine* (SVM) e Regressão Linear (RL). A seleção do algoritmo mais adequado, dentre os explorados, é baseada em critérios como desempenho, interpretabilidade e escalabilidade. A construção dos modelos de previsão é realizada utilizando-se os conjuntos de dados previamente preparados, fundamental para garantir a qualidade e a integridade dos conjuntos de dados utilizados no treinamento e na validação dos modelos. Os modelos são treinados com base em um conjunto de dados históricos que contém informações sobre as variáveis ambientais, como radiação solar, temperatura, velocidade do vento, entre outras, e a eficiência dos módulos correspondente. A avaliação, crucial para determinar sua eficácia e capacidade de generalização, foi baseada em métricas como o Desvio Quadrático Médio (*Root Mean Squared Error* - RMSE) igual a 1,8148; 2,6412; e 3,2831; o Erro Médio Absoluto (*Mean Absolute Error* - MAE) de 1,0972; 2,2471; e 3,1025; o Erro Quadrático Médio (*Mean Squared Error* - MSE) de 6,9841; 8,5032; e 13,4529; o Coeficiente de Determinação (R^2) com resultados iguais a 0,8912; 0,8601; e 0,8107. Dos modelos de ML desenvolvidos, aquele que demonstrou o melhor desempenho foi o modelo RNA, com desempenho, também, acima daqueles apresentados pelos modelos tradicionais da literatura para a previsão da eficiência dos módulos FV.

Palavras-chave: Aprendizado de Máquina. Eficiência. Energia Solar. Previsão da Eficiência.

ABSTRACT

The advancement of solar energy as a renewable source of electricity has sparked interest in understanding and predicting the performance of photovoltaic (PV) modules under different environmental conditions. This work aims to predict the efficiency of PV modules using Machine Learning (ML) techniques. Predicting the efficiency of modules is essential for effective planning and monitoring of solar energy systems, as well as a key component in decision-making related to PV generation. Initially, the principles and ML techniques used in module efficiency prediction are presented. Among the explored ML algorithms are Artificial Neural Networks (ANN), Support Vector Machine (SVM), and Linear Regression (LR). The selection of the most suitable algorithm among the explored ones is based on criteria such as performance, interpretability, and scalability. The construction of prediction models is carried out using pre-prepared datasets, which are crucial to ensure the quality and integrity of the data used in model training and validation. The models are trained based on historical data that includes information about environmental variables such as solar radiation, temperature, wind speed, among others, and the corresponding module efficiency. The evaluation, crucial for determining their effectiveness and generalization capability, is based on metrics such as Root Mean Squared Error (RMSE) equal to 1.8148, 2.6412, and 3.2831; Mean Absolute Error (MAE) of 1.0972, 2.2471, and 3.1025; Mean Squared Error (MSE) of 6.9841, 8.5032, and 13.4529; and Coefficient of Determination (R²) with results of 0.8912, 0.8601, and 0.8107. Among the developed ML models, the Artificial Neural Network (ANN) model performs the best, surpassing the performance of traditional models in the literature for predicting PV module efficiency.

Keywords: Machine Learning. Efficiency. Solar Energy. Efficiency Optimization.

LISTA DE FIGURAS

Figura 1 – Participação de cada fonte de energia na capacidade instalada	17
Figura 2 – Expansão da oferta de geração de energia elétrica em 2022 em (GW) .	17
Figura 3 – Estrutura de modelagem	19
Figura 4 – Evolução da fonte solar FV no Brasil	21
Figura 5 – Evolução dos investimentos de PDeD em eficiência energética no Brasil	22
Figura 6 – Fluxograma de implementação de modelo	24
Figura 7 – Algoritmos ML empregados	33
Figura 8 – Fases da metodologia	34
Figura 9 – Disposição dos componentes do arranjo FV	35
Figura 10 – Disposição dos demais componentes	36
Figura 11 – Inclinação dos módulos	36
Figura 12 – Atividades envolvidas na construção de um modelo preditivo	38
Figura 13 – Técnicas de pré-processamento de dados	39
Figura 14 – Fluxograma para implementação do treinamento	41
Figura 15 – A segmentação de ML	42
Figura 16 – Comportamento dos atributos meteorológicos: a) V_w ; b) G ; c) η ; d) T_c e e) T_{amb}	45
Figura 17 – Distribuição média dos parâmetros	46
Figura 18 – Velocidade média do vento para os períodos chuvosos e secos de Fortaleza	47
Figura 19 – Gráfico da distribuição anual da velocidade do vento do banco de dados do GB em Fortaleza	48
Figura 20 – Variação anual da radiação solar na cidade de Fortaleza	49
Figura 21 – Gráfico da distribuição anual da radiação solar do banco de dados do GB em Fortaleza	50
Figura 22 – Gráfico da distribuição média anual da temperatura ambiente do banco de dados do GB em Fortaleza	51
Figura 23 – Registros termográficos dos módulos FV do dia 18/02/2021 as 11:00 . .	53
Figura 24 – Distribuição média anual da eficiência do banco de dados do GB	55
Figura 25 – Gráficos de dispersão de: RNA, RL e SVM	57
Figura 26 – Gráficos dos valores observados para dias típicos de março e setembro .	58
Figura 27 – Gráficos dos valores estimados pelos modelos ao longo do ano de 2022 .	59

Figura 28 – Gráfico dos valores dos modelos comparativo e estimado ao longo do
ano de 2022 61

LISTA DE TABELAS

Tabela 1 – Capacidade instalada das regiões do mundo, em GW	16
Tabela 2 – Resumo das principais características de revisão da literatura	32
Tabela 3 – Resultados dos parâmetros estatísticos de treinamento	56
Tabela 4 – Resultados dos parâmetros estatísticos de treinamento	60

LISTA DE ABREVIATURAS E SIGLAS

R^2	R-squared
ABSOLAR	Associação Brasileira de Energia Solar Fotovoltaica
ANEEL	Agência Nacional de Energia Elétrica
CA	Corrente Alternada
CC	Corrente Contínua
DT	Decision Trees
EIA	Estudo de Impacto Ambiental
EPE	Empresa de Pesquisa Energética
FV	Fotovoltaico
GB	Gradient Boosting
GD	Geração Distribuída
IA	Inteligência Artificial
MAE	Mean Absolute Error
ML	Machine Learning
MMGD	Micro e Minigeração Distribuída
MPPT	Rastreamento do Ponto de Potência Máxima
MSE	Mean Squared Error
NOCT	Nominal Operating Cell Temperature
ONU	Organização das Nações Unidas
PDeD	Pesquisa, Desenvolvimento e Demonstração
RL	Regressão Linear
RMSE	Root Mean Squared Error
RN	Redes Neurais
RNA	Redes Neurais Artificiais
RT	Random Trees
SVM	Support Vector Machine

LISTA DE SÍMBOLOS

T_m	Temperatura dos módulos
T_{amb}	Temperatura ambiente
η	Eficiência elétrica
G	Irradiância solar
V_w	Velocidade do vento
V	Tensão
I	Corrente elétrica
P	Potência
R	Resistência
t	Tempo

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Objetivos	20
1.2	Estrutura do trabalho	20
2	REVISÃO DA LITERATURA	21
2.1	Relevância da energia solar e o panorama de crescimento da geração FV	21
2.2	Aprendizagem de máquina na geração FV	23
3	METODOLOGIA	34
3.1	Etapa 1: Coleta e pré-processamento dos dados	35
3.2	Etapa 2: Escolha do algoritmo mais adequado	39
3.3	Etapa 3: Construção do modelo e treinamento	40
3.4	Etapa 4: Avaliação e validação do modelo	42
4	RESULTADOS	44
4.1	Configuração dos atributos meteorológicos e representação dos dados	44
4.1.1	<i>Velocidade do vento</i>	47
4.1.1.1	<i>Regimes do vento em Fortaleza</i>	48
4.1.2	<i>Radiação solar</i>	49
4.1.2.1	<i>Regimes de radiação solar em Fortaleza</i>	50
4.1.2.2	<i>Relação da radiação solar com outros parâmetros meteorológicos</i>	51
4.1.3	<i>Temperatura ambiente</i>	51
4.1.3.1	<i>Regimes da temperatura ambiente de Fortaleza</i>	52
4.1.3.2	<i>Relação da temperatura ambiente com outros parâmetros meteorológicos</i>	52
4.1.4	<i>Temperatura dos módulos</i>	52
4.1.4.1	<i>Inspeção termográfica</i>	52
4.1.4.2	<i>Relação da temperatura dos módulos com parâmetros meteorológicos</i>	54
4.1.5	<i>Eficiência</i>	54
4.2	Modelo de previsão da eficiência dos módulos	55
4.2.1	<i>Treinamento e performance do modelo</i>	55
4.2.2	<i>Resultados do modelo</i>	58
4.2.3	<i>Análise comparativa</i>	60
5	CONCLUSÕES E TRABALHOS FUTUROS	62

REFERÊNCIAS 64

1 INTRODUÇÃO

O percentual de fontes renováveis de energia no plano internacional e nacional é alto (ONU, 2018), sobretudo nesse último, que possui grande potencial de participação de fontes de energia renovável na matriz energética, comparado ao restante do mundo. Contudo, ganhou destaque nos últimos anos a geração (FV), destacadamente os resultados alcançados pela região Nordeste do Brasil, sendo essa região posta em condição de relevo no cenário nacional. A título de comparação com o panorama mundial, a capacidade instalada só de energia solar no Brasil, nos anos de 2017 a 2022, teve um avanço de cerca de 652% (EPE, 2022), enquanto o mundo apresentou, para o mesmo período, o percentual de 98%. Isso corresponde a uma capacidade instalada, em GW, igual a 9,1, no panorama nacional, e 773,6 (EIA, 2019) no panorama mundial. Com esses valores nacionais, temos que a capacidade instalada sofreu expansão de 3,9% (EPE, 2022) com relação aos anos de 2020 e 2021, conforme a Tabela 1.

Tabela 1 – Capacidade instalada das regiões do mundo, em GW

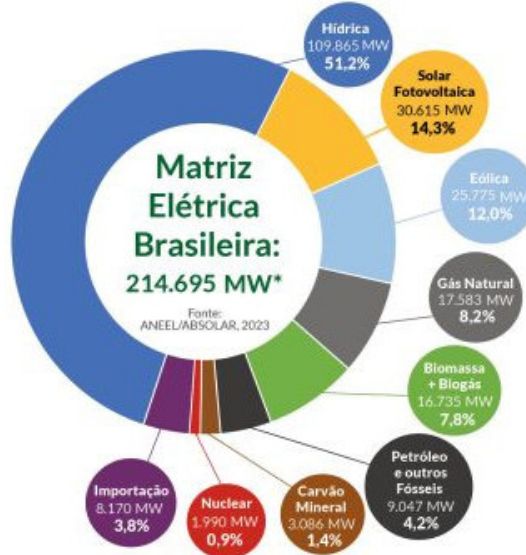
Região	2018	2019	2020	2021	%(2021/2020)
África	231,0	239,0	243,0	245,0	0,7%
América do Norte	1.352,0	1.365,0	1.387,0	1.425,0	2,7%
América do Sul e Central	371,0	384,0	394,0	405,0	2,8%
Ásia & Oceania	3.265,0	3.430,0	3.680,0	3.853,0	4,0%
Eurásia	401,0	410,0	419,0	429,0	2,4%
Europa	1.240,0	1.258,0	1.284,0	1.322,1	3,0%
Oriente Médio	298,0	304,0	307,0	310,0	0,9%

Fonte: (EIA, 2022).

O crescimento da energia solar é uma condição imprescindível para a garantia, internamente, de uma matriz energética sustentável, visto que, com a apresentação de estudos recentes acerca das usinas hidrelétricas, que compõem cerca de 53,4% da oferta de energia elétrica do país, elas tendem a gerar danos ambientais, com impactos na fauna e flora locais (MORAIS, 2018). O fortalecimento do desenvolvimento da energia solar no Brasil se dará desde que haja eficiência energética nos sistemas FV para isso e viabilidade econômica, com custos cada vez menores, o que já é um grande desafio para o país (NEJAT, 2015). Além disso, com a abertura do mercado livre, o consumidor terá a oportunidade de outros fornecedores de eletricidade, com a possibilidade de contratação de planos, provável

economia e garantia do contínuo crescimento acelerado do setor de energia solar. A Figura 1 mostra as principais fontes de energia do Brasil.

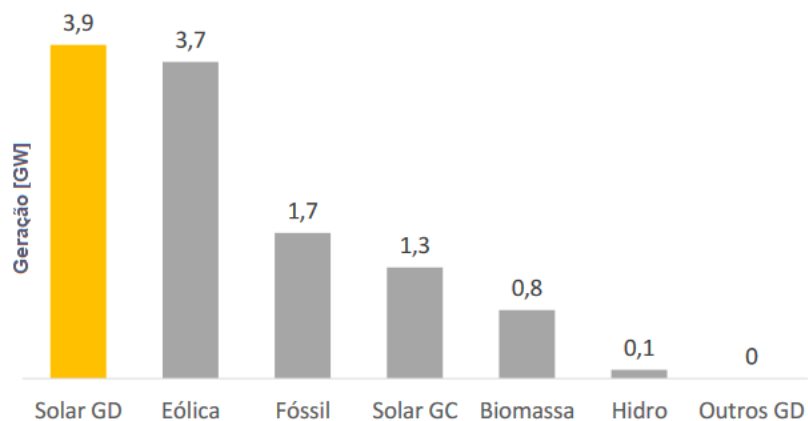
Figura 1 – Participação de cada fonte de energia na capacidade instalada



Fonte: ABSOLAR, 2023.

Incentivado por rigorosas ações governamentais (ABDEL-BASSET et al., 2021), a micro e minigeração distribuída (MMGD) de energia elétrica apresentou um alto crescimento no decorrer dos anos no Brasil, com valores, em 2022 alcançando 9,8 GWh e potência instalada igual a 9,0 MW, destacando-se nesse período a fonte solar FV, com 9,0 GWh e 8,8 MW de geração e potência instalada - respectivamente. A Figura 2 simboliza o protagonismo da MMGD na capacidade instalada do Brasil em 2022. Dessa figura, destaca-se a fonte solar distribuída, que superou a expansão de todas as fontes centralizadas.

Figura 2 – Expansão da oferta de geração de energia elétrica em 2022 em (GW)



Fonte: EPE, com dados da (ANEEL)

Recentemente, em 2022, a fonte solar FV (BEN, 2022; HEIN, 2022) já superou os valores impostos pelos anos de 2020 e 2021, alcançando o patamar de capacidade instalada de 13,2 GW, com destaque para o setor da indústria (FINA, 2019), que tem optado por fontes de energias renováveis, visto que dependem da eficiência energética no dia a dia dos seus processos produtivos e para não sofrerem com as consequências negativas da força da lei (como impostos).

Apesar disso, os dados oriundos da plataforma INOVA-E' apontam para investimentos em projetos de eficiência energética, entre 2012 e 2018, de 1 bilhão de reais em pesquisa, desenvolvimento e demonstração (PDeD) majoritariamente públicos. Também, nas residências, verificam-se várias medidas de eficiência energética implementadas por meio de políticas de padrões e de etiquetagens obrigatórias ou até mesmo voluntárias quando a fonte de energia é FV. Essas iniciativas são executadas desde a mais de 10 anos no país com a implementação dos marcos da energia solar e outras ações de taxação da fonte solar FV. Isto à parte, temos o dever de reconhecer, através dos avanços tecnológicos e computacionais, as técnicas de melhorias e previsão de eficiência energética, considerando para isso as variáveis que a energia solar está suscetível, como fatores meteorológicos, níveis de irradiância solar, temperatura, chuvas, vento, sujeidade dos painéis solares, dentre outros, que podem influenciar a melhoria ou a precisão da eficiência energética (EPE, 2022).

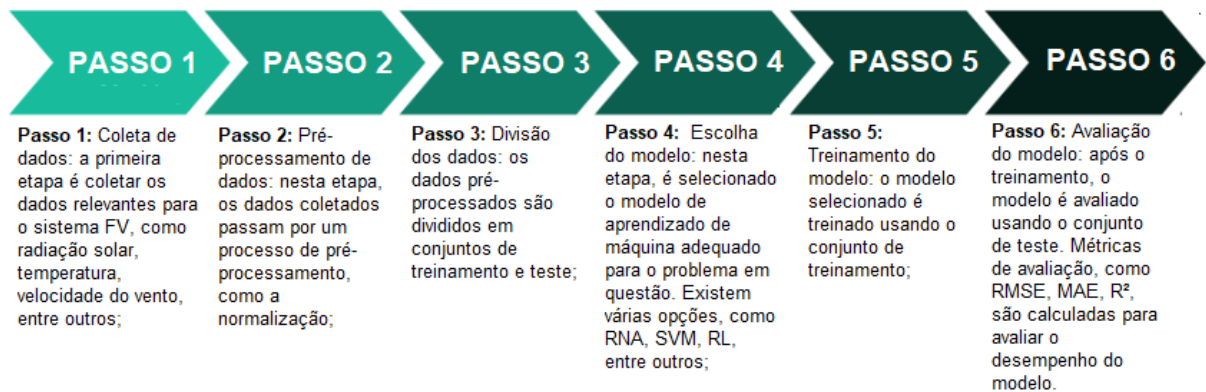
Existem várias abordagens no campo da tecnologia que se baseiam em modelos computacionais para estimar a eficiência do módulos FV. O primeiro grupo de métodos comumente utilizado para melhorar ou prever a eficiência é o método físico, seguido pelo computacional e híbrido. O método físico envolve a aplicação de equações ou correlações simples para analisar a performance do sistema FV. O método computacional utiliza técnicas de ML e análise estatística para prever a eficiência dos sistemas FV. E o método híbrido combina elementos do método estatístico com o método físico. Nesse contexto da previsão de eficiência para sistemas FV e para horizontes temporais de um dia à frente ou mais longos, os modelos físicos são os mais comuns e amplamente adotados pelos softwares. Essa abordagem consiste em algumas etapas, com modelos para cada uma delas. Basicamente, as etapas de modelagem são:

- 1) Coleta de dados: a primeira etapa é coletar os dados relevantes para o sistema FV, como radiação solar, temperatura, velocidade do vento, entre outros;
- 2) Pré-processamento de dados: nesta etapa, os dados coletados passam por um processo

- de pré-processamento, como a normalização;
- 3) Divisão dos dados: os dados pré-processados são divididos em conjuntos de treinamento e teste;
 - 4) Escolha do modelo: nesta etapa, é selecionado o modelo de aprendizado de máquina adequado para o problema em questão. Existem várias opções, como RNA, SVM, RL, entre outros;
 - 5) Treinamento do modelo: o modelo selecionado é treinado usando o conjunto de treinamento;
 - 6) Avaliação do modelo: após o treinamento, o modelo é avaliado usando o conjunto de teste. Métricas de avaliação, como RMSE, MAE, R^2 , são calculadas para avaliar o desempenho do modelo.

Por ser uma abordagem específica que se concentra em utilizar algoritmos e modelos de ML para estimar a eficiência dos módulos com base em dados históricos e outras variáveis relevantes, a previsão da eficiência usando ML difere dos métodos tradicionais de previsão de eficiência dos módulos FV. Além disso, a previsão da eficiência de módulos FV usando ML oferece algumas vantagens significativas em relação a métodos tradicionais de previsão, como a precisão, otimização de desempenho e adaptabilidade à medida que o modelo é treinado com novos dados coletados. A Figura 3 ilustra as etapas de modelagem.

Figura 3 – Estrutura de modelagem



Fonte: Elaborada pelo autor

1.1 Objetivos

O principal objetivo é desenvolver os modelos (RNA, SVM e RL) de ML capazes de prever a eficiência dos módulos FV com precisão, contribuindo para a melhoria do desempenho e eficiência dos módulos FV,

Para tanto, têm-se como objetivos específicos:

- Realizar uma revisão bibliográfica a nível internacional dos trabalhos e melhores praticas de estudo e previsão da eficiência dos módulos FV;
- Analisar variáveis que impactam a eficiência dos módulos FV, como os fatores meteorológicos (irradiância solar, velocidade do vento, temperatura ambiente) e temperatura dos módulos;
- Comparar os modelos de ML implementados com base no desempenho de cada algoritmo aplicado, utilizando métricas estatísticas, como RMSE, MAE, MSE e R^2 , além de realizar uma análise comparativa dos resultados obtidos pelo modelo RNA e do modelo tradicional NOCT, a fim de identificar qual método apresenta o melhor desempenho.

1.2 Estrutura do trabalho

Este trabalho é constituído de cinco capítulos, os quais apresentam de forma estruturada e detalhada as etapas e resultados alcançados na previsão da eficiência dos módulos FV utilizando técnicas de ML. No capítulo 1, é realizada uma introdução ao tema do trabalho, abordando o crescimento da energia solar, a importância da previsão da eficiência dos módulos FV, seus desafios e a relevância da aplicação de técnicas de ML. Adicionalmente, são apresentados os objetivos do trabalho e a estrutura do restante do documento. No capítulo 2, é apresentada uma revisão da literatura sobre os conceitos fundamentais relacionados à eficiência dos módulos FV, técnicas de ML e os métodos utilizados para previsão. No capítulo 3, é detalhada a metodologia adotada para o desenvolvimento do trabalho. No capítulo 4, são apresentados os resultados obtidos a partir da aplicação dos modelos de ML na previsão da eficiência dos módulos FV. No capítulo 6, são apresentadas as conclusões do trabalho, destacando os principais resultados e contribuições obtidas. Além disso, são discutidas as limitações do estudo e são apresentadas sugestões para trabalhos futuros, visando aperfeiçoar e expandir o conhecimento na área.

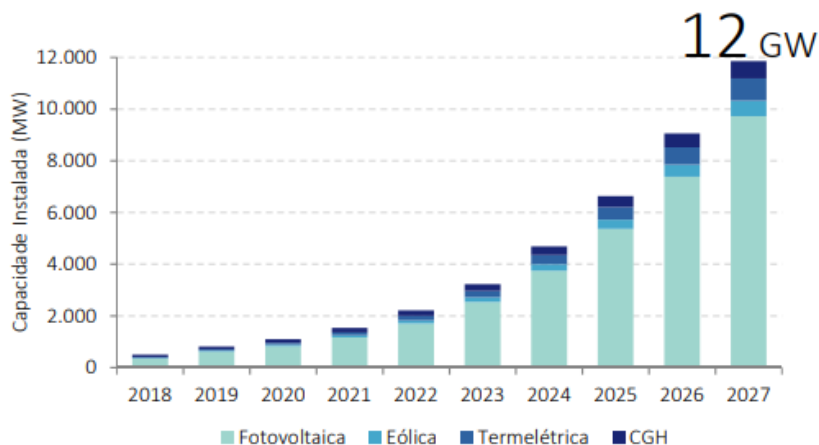
2 REVISÃO DA LITERATURA

No presente capítulo, a revisão de literatura é desenvolvida por meio da apresentação do estado da arte da geração FV e da aprendizagem de máquina no contexto, em que, antes de tudo, são expostas as principais contribuições desde os tipos de metodologias, as adequações considerando os efeitos das variáveis da geração FV, os dispositivos e as vantagens das metodologias nos tempos atuais; em seguida, são apresentadas a aplicação de geração FV no contexto da aprendizagem de máquina, a melhoria da eficiência do sistema por meio da modelagem e simulação.

2.1 Relevância da energia solar e o panorama de crescimento da geração FV

A energia solar é uma fonte de energia renovável e limpa que está ganhando cada vez mais importância no Brasil e no mundo como uma alternativa aos combustíveis fósseis. Essa fonte de energia é gerada a partir da energia do sol, que é capturada por meio de módulos FV e convertida em eletricidade. Nos últimos anos, a geração FV tem experimentado um crescimento sem precedentes em todo o mundo, impulsionado pelo aumento da conscientização sobre a necessidade de fontes de energia limpas, renováveis e também pela fonte solar ser uma forma de geração distribuída (GD), podendo ser produzida em locais próximos aos pontos de consumo. Esse crescimento se deve em grande parte à redução dos custos dos painéis solares e ao aumento da eficiência na conversão da energia solar em eletricidade. Conforme ilustrado na Figura 4, fica claro o protagonismo da geração FV no cenário nacional.

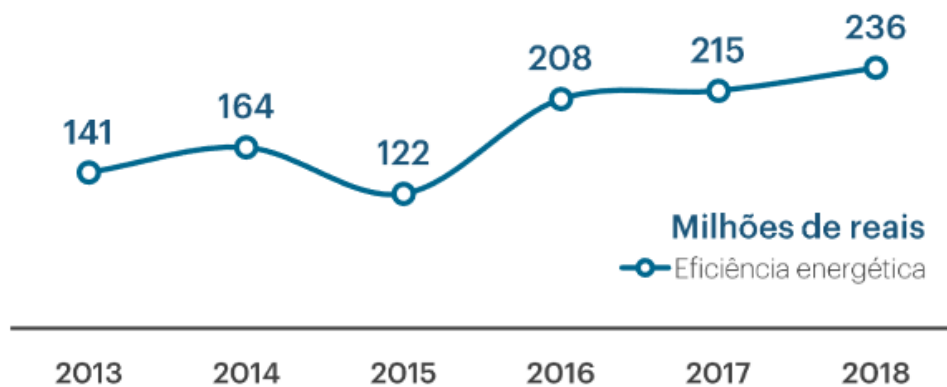
Figura 4 – Evolução da fonte solar FV no Brasil



Fonte: EPE, 2023

No entanto, apesar do crescimento rápido da geração FV, ainda há desafios a serem enfrentados, como a necessidade de aumentar a eficiência e reduzir ainda mais os custos de produção. A pesquisa e o desenvolvimento contínuos em tecnologias de energia solar, bem como o apoio político e financeiro adequado, são fundamentais para impulsionar a adoção da energia solar em todo o mundo e para uma transição para um futuro mais sustentável e com baixo carbono. Vale salientar, ainda, que a eficiência da conversão de energia solar em eletricidade tem aumentado significativamente ao longo dos anos, principalmente com o aprimoramento das tecnologias de células FV, com o desenvolvimento de novos materiais para absorção de luz e com o investimento em eficiência energética. Na Figura 5, é possível identificar os investimentos feitos ao longo do ano em eficiência energética.

Figura 5 – Evolução dos investimentos de PDeD em eficiência energética no Brasil



Fonte: EPE, 2022

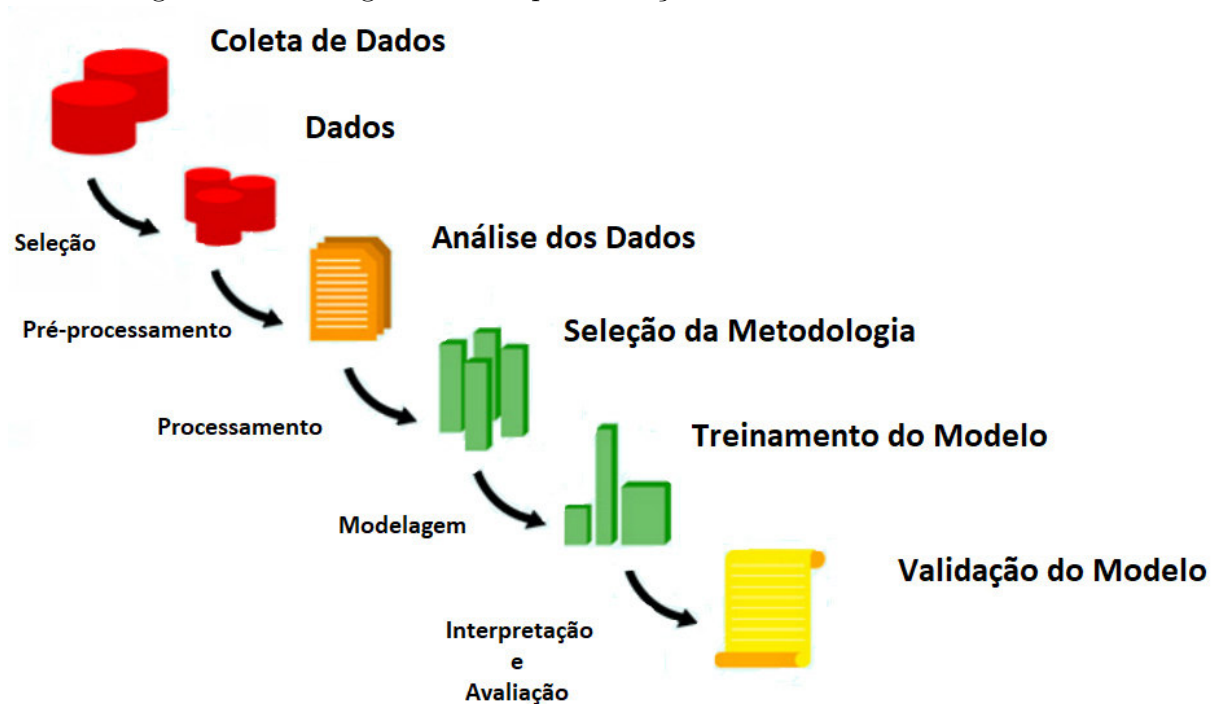
Visando maior eficiência, o mercado dispõe de três tipos principais de tecnologias de células FV em uso atualmente: células de silício monocristalino, células de silício policristalino e filme fino. As células de silício monocristalino são consideradas as mais eficientes, mas também as mais caras de produzir. As células de silício policristalino são mais econômicas, porém com menor eficiência. Os filmes finos são os mais versáteis e apresentam menor custo de produção, mas com eficiência ainda mais reduzida. Cada tecnologia tem suas próprias vantagens e desvantagens, e sua escolha depende das necessidades e restrições de cada aplicação específica. Com a crescente distribuição de dados, é importante o uso adequado de sistemas de gerenciamento ou tratamento desses dados de forma ordenada e objetiva. Com isso, aplica-se muito o ML para analisar a eficiência energética de sistemas FV.

2.2 Aprendizagem de máquina na geração FV

Tecnologias FV encontram uso em diversas aplicações (AKHTER et al. 2019). Essas tecnologias têm sido utilizadas como sistemas autônomos de exploração do sol nos domicílios e indústria. Os métodos mais comuns para melhorar ou prever a eficiência dos sistemas FV podem ser divididos em três grupos. O primeiro deles, é o método físico (ou convencional), que se refere à aplicação de equações ou correlações simples para analisar a performance do sistema. O segundo método (SIVAGAMI; JOTHISWAROOPAN, 2021) é o computacional (ou estatístico) e, como especificamente trata o presente trabalho, o de ML. Há ainda um terceiro método, descrito por (MAYER, 2021) como híbrido, que intuitivamente é a relação do estatístico com o físico. Existem vários métodos de ML que podem ser usados para melhorar ou prever a eficiência de sistemas FV, incluindo: 1. o aprendizado supervisionado, que usa dados históricos de desempenho de sistemas FV para treinar um modelo que pode ser usado para prever o desempenho futuro; 2. aprendizado não supervisionado, através do uso de dados de desempenho de sistemas FV para identificar padrões e tendências sem um alvo predefinido; 3. aprendizado profundo, que usa redes neurais para modelar complexas relações entre os dados de entrada e a saída desejada (eficiência de sistemas FV); 4. aprendizado por reforço, sendo esse modelo treinado por meio de experimentos para maximizar uma recompensa (eficiência de sistemas FV) e 5. aprendizado por otimização, no qual o modelo é treinado para encontrar os melhores parâmetros de configuração para maximizar a eficiência de sistemas FV (HAN et al., 2019). Cada método tem suas vantagens e desvantagens. Escolher o método certo dependerá das características específicas dos dados e dos objetivos do projeto. No esquemático introduzido na Figura 6, é mostrado um fluxograma para a implementação de um modelo que emprega ML.

Os métodos descritos anteriormente dispõem, em sua aplicação, de técnicas ou modelos (HANAFY, 2019) que podem ser aplicados para melhorar e prever a eficiência do sistema. Essas técnicas incluem (PRIYADHARSIN, 2021): (RL), que é um modelo simples que pode ser usado para prever a eficiência de sistemas FV com base em variáveis preditivas, como a irradiação solar e a temperatura; Árvores de Decisão (*Decision Trees* - DT): uma técnica que pode ser usada para identificar fatores que afetam a eficiência de sistemas FV e tomar decisões baseadas nesses fatores; Redes Neurais (RN): técnica avançada que pode ser usada para modelar relações complexas entre variáveis preditivas e a eficiência de

Figura 6 – Fluxograma de implementação de modelo



Fonte: Elaborada pelo autor

sistemas FV; Árvore Aleatória (*Random Trees - RT*): um modelo baseado em RT que usa múltiplas árvores para melhorar a precisão das previsões; Aumento do Gradiente (*Gradient Boosting - GB*): é uma técnica que consiste em adicionar modelos sucessivos para melhorar a precisão das previsões; SVM, que é uma técnica de ML que pode ser usada para prever a eficiência de sistemas FV a partir de dados multivariados; Algoritmos genéticos, que é uma técnica inspirada na evolução que busca otimizar parâmetros para maximizar a eficiência dos sistemas FV. Esses são apenas alguns modelos que encontram implementação, existindo outros tantos a serem empregados de acordo com as características e objetivos do projeto.

Recorrendo a um conjunto de técnicas de ML, recentemente (KHILAR, 2022) propôs a melhoria da eficiência de painéis FV após a análise de dados de medição de desempenho dos painéis, como tensão e corrente. Nesse contexto, por intermédio do treinamento das técnicas RN, DT e RL foi possível adaptar os dados coletados a uma simulação computacional como ferramenta de otimização na geração de energia solar. Os estudos apontaram para uma acurácia, medida comparando as previsões dos modelos treinados com os valores reais e descrita em valor de erro RMS. O menor valor de erro RMS obtido foi 1,53 e corresponde a uma eficiência de 99%, um aumento de 31% em comparação com os resultados anteriores, que foi de 68%. As técnicas usadas alcançaram

resultados positivos, uma vez que o modelo proposto atingiu uma alta taxa de acurácia.

Os estudos de (PHOTANA, 2021) abordam, também, o uso de técnicas de aprendizagem de máquina para melhorar o desempenho de painéis FV. O autor destaca que a aprendizagem de máquina pode ser útil para prever a produção de energia de um painel FV de acordo com a orientação e inclinação dos painéis, visando maximizar a captação de luz solar e aponta que como fatores meteorológicos (radiação solar, velocidade do vento, temperatura ambiente) e a umidade afetam o desempenho dos painéis solares, mesmo sendo possível o ajuste manual, implicando em um baixo rendimento na eficiência energética. Dessa forma, visando contribuir para uma abordagem mais dinâmica, com a matriz energética e o desenvolvimento sustentável pela aplicação de técnicas de RL, SVM, DT e RT. Esse trabalho apresentou um erro RMS de 5,98 ao fazer o uso das características da RL, merecendo destaque nos programas e periódicos de energia renovável.

São cada vez maiores os estudos de simulação computacional como ferramenta de otimização na geração FV. Como exemplo disso, (TAKURI, 2020) desenvolveu um método eficaz para melhorar a eficiência da produção de eletricidade a partir de painéis FV com a modelização das características FV utilizando SVM em modo de regressão. A SVM é implementada em duas fases: 1. a fase de treino que modela a relação de temperatura e irradiância com a tensão de referência; 2. a fase de estimação da tensão de referência para novos níveis de temperatura e irradiância. A técnica proposta resultou em uma estimativa precisa dos valores de tensão de referência que correspondem à potência máxima gerada por painéis FV em determinada temperatura e em dados valores de irradiação. Isso assegurou uma ótima produção de energia a partir dos painéis FV e, com a técnica implementada, foi possível prever a saída dos elementos de saída, como por exemplo tensão e corrente de saída, com base nos dados usados. Com isso, a ferramenta encontra uso para o planejamento eficiente de recursos de energia renováveis em locais urbanos, merecendo destaque, ainda, por não ultrapassar um erro RMS de 5,0, o que é muito positivo e apresenta um grau de acurácia ótimo para a melhoria da eficiência de painéis FV. Apesar disso, o modelo proposto não considerou variáveis importantes que podem gerar resultados mais precisos e resultar em uma melhoria de desempenho maior, como a detecção de ângulo ou umidade.

(FARAYOLA & HASAN 2018) observaram que o sistema de Rastreamento do Ponto de Potência Máxima (do inglês, MPPT) somado ao emprego de painéis FV podem ser utilizados para melhorar a eficiência. Para isso, foram implementados algoritmos

matemáticos que contribuem para interligar as ferramentas de modelagem, incluindo Máquina de Vetores de Suporte Gaussiano (CGSVM, na sigla em inglês) e RNA, permitindo prever com precisão o ponto de potência máxima e garantindo a adaptação as mudanças nas condições de operação. Desta forma, combinando as técnicas supracitadas, os autores melhoraram a eficiência do sistema de geração FV e sua adaptabilidade as diferentes condições climáticas. A acurácia e precisão da implementação desses autores foram atestadas com os testes RMS e o erro médio, ambos com valores mínimos iguais a 0. Essa abordagem promissora, no entanto, conta com um conjunto de dados de teste insuficientes para implementar a técnica CGSVM e os erros de ajuste contidos no PSIM para implementação dos controladores afetam a técnica CGSVM.

Outros autores, como (ANIS, 2022), investigaram a eficiência energética de sistemas de baixa tensão em corrente contínua (CC) com aplicação residencial. Esse trabalho apresenta, inclusive, um panorama de utilização de sistemas CC em edifícios residenciais e descreve os principais benefícios desses sistemas, incluindo a redução da perda de energia durante a conversão de CA para CC e a facilidade de conexão de cargas CC como equipamentos eletrônicos. Além disso, ao fazerem uso das características dos conversores CC-CC, controladores e inversores CC-AC afim de desenvolverem dispositivos mais eficazes, e, por consequência, contribuindo para a a otimização de processos e ampliação da geração FV. Apesar da pesquisa não desenvolver novos materiais e dispositivos, a abordagem se diferencia das demais por utilizar as características da CC e alcançarem uma eficiência de 20% mais elevada do que aquela que utilizaria CA. A acurácia do modelo foi medida em valores RMS e demonstrou um resultado igual a 6.

De acordo com (KLOTTER, 2018), é possível, usando SVM, prever a eficiência de produção em larga escala de uma planta FV, uma vez que com a coleta do conjunto de dados será permitido identificar eventuais problemas de geração de energia e eficiência. O arranjo selecionado de energia solar e técnica SVM apresentam bons resultados, contribuindo para uma previsão próxima daquilo realmente obtido para a eficiência. Contudo, outras técnicas podem apresentar melhores resultados em função da estrutura e conjunto de dados tratados pelo autor. Uma análise comparativa do erro foi apresentada pelo autor, junto com uma resposta dinâmica e a eficiência realizada com o arranjo implementado, o que colaborou para (KLOTTER, 2018) descrever a acurácia do modelo como muito alta.

Em (LI, 2018), apresentam-se resultados que provam a melhoria da eficiência

dos sistemas para praticamente todos os tipos de falhas e condições, chegando a uma taxa de sucesso superior a 90%. Entre os tipos de técnicas utilizadas, se destaca com a taxa de sucesso a RNA Perceptron com Múltiplas Camadas (*Multiplayer Perceptron* - MLP) e a Convolutacional (*Convolutional Neural Network* - CNN), As limitações, conforme o autor aponta, podem ser verificadas já de início, com a baixa taxa de dados coletados observada em alguns casos. Contudo, isso pode ser corrigido com uma nova coleta de dados em campo para refletir as condições operacionais reais do sistema FV. Como descrito no anteriormente, a taxa de acurácia, nesse caso, apresentada em valores percentuais, pode chegar a 90%.

Em (HUSSAIN, 2020) é proposto um algoritmo matemático de identificação de falhas em sistemas FV baseado na Função de Base Radial (*Radial Basis Function* - RBF) para a melhoria da eficiência, onde é desenvolvido, implantado e validado um sistema de detecção de falhas, necessitando apenas de dois parâmetros como entrada para a RNA: irradiância e potência FV. O modelo é testado em outro sistema FV, permitindo avaliar o comportamento da RNA com e sem sombreamento. O processo de validação fornece uma precisão de detecção de falhas acima de 97%. A acurácia do modelo foi descrita como alta, acima de 90%, podendo alcançar valores maiores.

Combinando *Grey Wolf Optimizer* (GWO) com o algoritmo MLP para melhoria da eficiência, (COLAK, 2019) utilizou um conjunto de dados que contempla radiação solar horizontal difusa (*Diffuse Horizontal Irradiance* - DHI), umidade relativa e temperatura do ar. Na análise, a melhoria da eficiência é utilizada como variável de saída, enquanto as demais grandezas são utilizadas como variáveis de entrada. O algoritmo é testado para três funções de ativação: sigmoide, seno e tangente hiperbólica; a melhor performance é obtida para a função sigmoide, com entradas DHI, umidade relativa e temperatura do ar. Para esta função é obtido um erro percentual médio absoluto (*Mean Absolute Percentage Error* - MAPE) de 3,03%, um MAE de 0,022 e um coeficiente de determinação (CD) de 0,98. Ainda, são apontados diferentes critérios estatísticos utilizados para avaliar o desempenho de uma RNA.

Sob diferentes condições climáticas e influência da sujidade dos painéis FV, (GHOLAMI, 2018) mostra que os resultados de eficiência podem diferir. O autor afirma que quando a distribuição de sujidade não é homogênea nas superfícies do módulo FV, para além das perdas relacionadas com a reduzida absorção de irradiação, há também perdas

devido à descoordenação elétrica entre as células limpas e sujas do módulo. Isto pode comprometer o funcionamento do módulo como um todo devido ao sombreamento parcial, causando a acumulação de sujeira numa região do módulo, o que, por sua vez, pode causar perdas de energia desproporcionais à sua área. O estudo feito por (GHOLAMI, 2018) aponta que a densidade de deposição de sujeira permaneceu praticamente constante no início da proposta, possivelmente devido ao período chuvoso durante os primeiros dias da experiência, o que limpou naturalmente as superfícies. A deposição de sujeira aumentou abruptamente entre os 130^o e 210^o dias da experiência, quando ocorreu a menor precipitação acumulada (0,4 mm). A maior densidade de deposição de sujeira foi de 1280 mg m⁻², no 204^o dia da experiência, 48 dias após uma pluviosidade de apenas 0,3 mm. No entanto, os resultados mostram estimativas fiáveis da tendência de deposição de sujeira em módulos FV para áreas na região semi-árida do mundo. Após a estimativa da densidade de deposição de sujeira e da reduzida potência de saída, foi possível relacionar as duas quantidades e obter um resultado para a eficiência, mostrando que a mesma alcançou valores superiores a 80%, segundo o responsável pela experiência. Não são apresentados dados de acurácia da aplicação.

Contribuindo para a melhoria da eficiência dos painéis FV, (YOANN, 2020) propõe um método baseado na aprendizagem de máquina para realizar a otimização simultânea de multiprocessos. Utilizando a variação natural de uma linha de produção, treinou-se modelos de aprendizagem de máquina para investigar a relação entre parâmetros de processo e eficiência celular, empregando algoritmos para identificar novos parâmetros de processo, a fim de maximizar a eficiência celular. Usando RN é construído um modelo preciso para prever as eficiências celulares a partir de parâmetros de processo de entrada com erros de < 0,03% de eficiência absoluta. Em cinco iterações, a eficiência média das células aumenta de 18,07% para 19,45%, evidenciando a aplicação promissora feita no trabalho. Fornecendo uma forte monitorização do processo e um rastreamento preciso, o método proposto é diretamente aplicável a conjuntos de dados do tipo produção, permitindo à indústria FV construir fábricas inteligentes e juntar-se à quarta revolução industrial.

Em (SOBRI, 2018) são propostos uma revisão e um comparativo entre RNA e modelos estatísticos de série temporal, modelos físicos e modelos de conjunto para previsão da geração FV. São analisadas RNA e SVM devido à sua capacidade de resolver modelos de previsão complexos e não lineares. A avaliação mostra que os modelos de Inteligência

Artificial (IA) podem diminuir o erro, em comparação com outras abordagens estatísticas, o que reforça a efetividade da IA nos estudos de previsão FV. Comparando os modelos RNA e SVM é verificado que RNA apresenta um melhor resultado em previsão da irradiação; porém, para entradas diferentes, a SVM tem um desempenho melhor. MLP é aplicado com horas de sol, temperatura mínima e máxima, média relativa de umidade, velocidade mínima, média e máxima do vento, irradiação em céu nublado, ensolarado e a quantidade de material particulado como parâmetros de entrada e, para saída, a irradiação solar global. Com isso, são obtidos valores medidos e previstos próximos de irradiação solar diária global, tendo o segundo modelo o melhor desempenho, com acurácia de 98,80%, erro médio percentual (*Mean Percentage Error* - MPE) de 1,7501 e MSE de 0,0131. Portanto, dos resultados apresentados, é possível verificar que a aplicação é muito positiva para melhorar a eficiência dos sistemas FV e contribuir para o avanço tecnológico da área de energia solar.

Sob condições de sombreamento variável (SUMEDHA, 2020), os módulos FV numa matriz solar experimentam condições variáveis de irradiação solar. Nesse sentido, a pesquisa em questão tem como foco a melhorar a eficiência máxima de rastreamento do ponto de máxima potência em sombreamento parcial através de ML. Usando conceitos de matemática, (SUMEDHA, 2020) define que um ponto máximo global e múltiplos pontos máximos locais de irradiação estão relacionados com as curvas de tensão e potência do sistema sob tais condições variáveis de irradiação solar. A aprendizagem de máquina é empregada na experiência para solucionar dois problemas principais: treinar um algoritmo para prever o verdadeiro valor do ponto de máxima potência e a limitação do conjunto de dados de amostragem coletados. A técnica empregada com maior sucesso foi a RT, e as amostras agrupadas de acordo com as diferentes zonas de sombreamento e com base na região na qual os painéis estão. A precisão do experimento realizado foi de 96% utilizando RT considerando todos os dados coletados. A acurácia, obtida em pontos percentuais alternados em 33%, 50%, 66% e 100%, foram de 55,55% - 96,29% para zonas com vizinhança e 100% para zonas com vizinhança.

Vários métodos de rastreamento do ponto de máxima potência, o conhecido acrônimo em inglês MPPT foram estudados e implementados em arranjos FV, segundo o (XIAO, 2010). Os métodos MPPT podem ser classificados como: técnicas off-line, e técnicas on-line. As técnicas off-line requerem um modelo de painel FV e a medição da

temperatura e irradiação solar. As técnicas on-line não requerem a medição da temperatura e irradiância solar. De acordo com (XIAO, 2010), entre os recursos mais desejáveis nas técnicas de MPPT estão os seguintes: 1. Estabilidade; 2. Resposta dinâmica rápida; 3. Erro pequeno no estado estacionário; 4. Robustez a distúrbios; e 5. Eficiência em uma grande faixa de potência. Os métodos convencionais on-line podem ser classificados como: Tensão Constante, Perturbação e Observação (P&O) e Condutância Incremental. Algumas variações desses métodos também são apresentadas em outras literaturas. O método de Perturbação e Observação (P&O), frequentemente usado em arranjos FV, opera perturbando o valor de referência para mover o ponto de operação em direção ao MPP, abreviação de Maximum Power Point. Essa técnica apresenta resposta dinâmica lenta e erro no estado estacionário. A escolha de altos valores de perturbação fornece um rastreamento rápido para a tensão no MPP, mas possui grandes oscilações. Se a perturbação tiver um valor baixo, o MPPT será mais lento, mas terá pequenas oscilações ao redor do MPP.

(WANG, 2017) não se limita a reproduzir as técnicas implementadas na extensiva documentação que trata da abordagem de ML no campo da energia solar, mas acrescenta, por meio de todos os dados que dispunha, a proposta de utilização de modelos de Previsão em Sequência de Padrões (*Pattern Sequence-based Forecast – PSF*). O intuito de tal abordagem para a previsão de eficiência de usinas solares, é apresentar uma metodologia híbrida, que combina as técnicas PSF e RNA. Contudo, mesmo com essa promissora combinação, o trabalho considera a irradiância, a sujidade dos painéis FV, inclinação dos módulos, velocidade do vento e fatores meteorológicos, condições comuns encontradas na literatura para gerar os dados tratados em modelo computacional. A acurácia da técnica empregada foi superior aos trabalhos mais comuns e o resultado obtido foi de 5,22, medida em erro RMS.

Em (PARK, 2019), foram utilizados métodos com baixo emprego de simulação e análises para investigar a eficiência energética. Com a aplicação mínima de algoritmos hierárquicos de RL, o autor visa perceber a relação e o impacto dos agrupamentos de temperatura, irradiância, diferentes condições meteorológicas e condições das células FV. Como resultado do estudo, foi possível classificar os agrupamentos em diferentes níveis de impacto no sistema FV, destacando que determinados agrupamentos podem alterar mais do que outros, influenciando os resultados dos sistemas FV e gerando menores porcentagens

de eficiência. O autor conclui, então, que os sistemas FV exigem, com o progresso das novas tecnologias, um gerenciamento eficaz e eficiente da qualidade de geração e distribuição da energia gerada, de modo a serem alcançados novos progressos. A acurácia da metodologia aplicada foi baixa, alcançando apenas 42% de eficiência com o estudo e um erro RMS muito alto, de 9,76.

Em (ALI, 2018), foi apresentada uma revisão de técnicas de ML utilizadas para prever a geração FV. Para isso, foi utilizado um conjunto de dados contendo informações sobre temperatura, radiação solar, umidade e tensão de circuito aberto. As técnicas de ML empregadas na pesquisa foram regressão linear múltipla (MLR, da sigla em inglês) e RNA. A técnica escolhida foi o SVR e os dados foram coletados de uma usina solar no México. Os resultados mostraram que a RNA apresentou melhor desempenho do que a MLR, com uma acurácia de 95,5% na previsão da eficiência do sistema FV. A pesquisa também mostrou que a radiação solar e a temperatura foram os fatores mais significativos na determinação da eficiência do sistema FV. Essa pesquisa é importante, pois demonstra que as técnicas de ML podem ser utilizadas para melhorar a estimativa da eficiência de sistemas FV, o que pode contribuir para um melhor dimensionamento e operação desses sistemas, bem como para a expansão da energia solar como fonte renovável de energia.

Um grupo de pesquisadores da Arabia Saudita, (RIAZ & HASAN, 2017), desenvolveram um modelo preciso para prever a saída de energia de um sistema FV baseado em condições ambientais, usando técnicas de ML. Os dados utilizados foram coletados de um sistema FV de 5 kWp instalado no telhado de um edifício no campus da Universidade Rei Saud. Os autores usaram três algoritmos de ML diferentes: RL, DT e RNA. Treinaram cada algoritmo usando os dados coletados e mediram a precisão do modelo usando a raiz do erro quadrático médio (RMSE) e a correlação de Pearson (r). Os resultados mostraram que a RNA teve melhor desempenho, com uma RMSE de 0,021 e uma correlação de Pearson de 1,00. Isso significa que o modelo foi capaz de prever a saída de energia com alta precisão com base nas condições ambientais, maximizando a eficiência e o desempenho do sistema FV.

(KUMAR, 2017), propôs o uso de uma Rede Neural Profunda (RNP) para prever a produção de energia elétrica de um sistema FV com base nas condições meteorológicas locais. A rede neural foi treinada e testada com dados de um sistema FV localizado em um instituto de engenharia na Índia. Foram coletados dados meteorológicos, dados de

produção de energia elétrica do sistema FV e dados de previsão meteorológica de um site local de previsão do tempo. Os dados foram divididos em conjuntos de treinamento e teste e usaram uma rede neural com várias camadas ocultas para fazer as previsões. A acurácia da previsão da produção de energia elétrica do sistema FV foi avaliada com o coeficiente de determinação (R^2) e a raiz do erro médio quadrático (RMSE). A rede neural profunda apresentou um R^2 de 0,98 e um RMSE de 0,04.

Na Tabela 2 é ilustrado um resumo das publicações e os algoritmos empregados pelos autores de cada publicação.

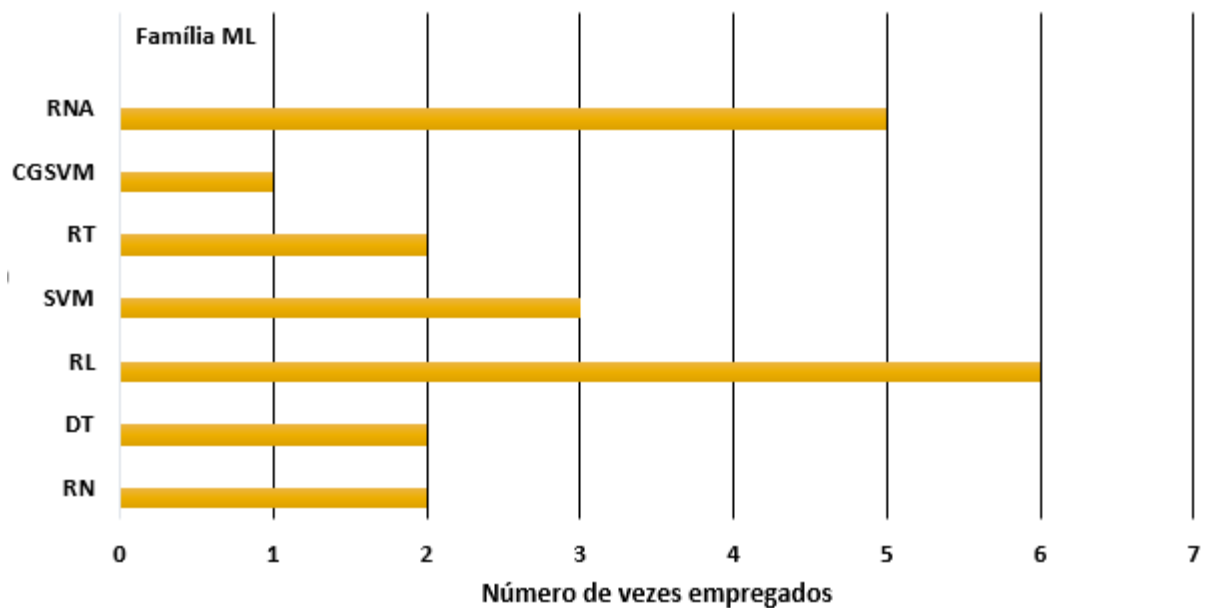
Tabela 2 – Resumo das principais características de revisão da literatura

Autor	País	Ano	Algoritmo	Acurácia
Khilar	Índia	2022	RN, DT, RL	Não se aplica
Photana	Alemanha	2021	RL, SVM, DT, RT	5,98 RMS para RL
Takuri	Japão	2020	SVM	5,0 RMS
Farayola & Hasan	Espanha	2018	CGSVM, RNA	0 RMS
Anis	EUA	2022	RL	6,0 RMS
Kloter	EUA	2018	SVM	Não se aplica
Li	China	2018	RNA	90%
Hussain	EUA	2020	RNA	91%
Colar	EUA	2019	RNA	0,022 em erro absoluto
Gholami	Canadá	2018	RL	Não se aplica
Yoann	EUA	2020	RN	Não se aplica
Sobri	Inglaterra	2018	RNA, SVM	98,80% para RNA
Sumedha	EUA	2020	RT	55,55% - 76,29%
Xiao	China	2010	RL	Não se aplica
Park	Coreia do Sul	2019	RL	9,76 RMS
Ali	México	2019	RLM e RNA	95,5% para RNA
Riaz & Hasan	Arabia Saudita	2017	RL, DT, RNA	0,021 RMS para RNA
Kumar	Índia	2017	RNP	0,98 RMS

Fonte: Elaborada pelo autor

Na Figura 7, são apresentadas as escolhas mais tomadas pelos pesquisadores para desenvolverem os seus estudos baseados em ML. A partir disso, nota-se que o modelo mais comum entre os estudos consultados é o de RL, com 6 trabalhos que empregaram o modelo. Acompanha o modelo RL, com 5 aplicações, o modelo de RNA. Por último, apresentando boas aplicações, está o modelo de SVM, com 3 aplicações.

Figura 7 – Algoritmos ML empregados



Fonte: Elaborada pelo autor

Sendo assim, a aprendizagem de máquina apresenta capacidade para ser uma forte aliada da geração FV na diversificação da matriz energética brasileira, tendo em vista o potencial das metodologias, muito acima da média dos métodos tradicionais utilizados para melhorar a eficiência dos sistemas. No entanto, apesar do potencial de implementação crescente nos sistemas FV, ainda há problemas passíveis de melhorias. As propostas contém diferentes metodologias e abordagens, mas com um propósito único voltado para a melhoria da eficiência dos sistemas FV ou de previsão dessa eficiência. Através dos resultados obtidos por cada autor por meio de diferentes metodologias, percebe-se a ferramenta poderosa que é o ML, com resultados superiores às técnicas que não utilizam uma abordagem computacional, podendo atender as normas e requisitos da indústria ou residências. Com isso, neste trabalho foram aplicados os modelos de RNA, SVM e RL, que conforme os estudos citados possuem alta precisão.

3 METODOLOGIA

O presente capítulo tem como objetivo apresentar a metodologia utilizada para a estimação da eficiência de sistemas FV aplicando ML. A metodologia inclui a coleta e pré-processamento dos dados de irradiação (G), velocidade do vento (V_w), temperatura ambiente (T_{amb}), temperatura dos módulos (T_c) e eficiência (η), a escolha do algoritmo de ML mais adequado para o problema em questão, o treinamento e avaliação do modelo, bem como a validação do modelo com novos dados. Espera-se que essa metodologia possa contribuir para o desenvolvimento de soluções eficientes e sustentáveis para a geração FV.

Esquematizando todas as fases, obtém-se a Figura 8.

Figura 8 – Fases da metodologia



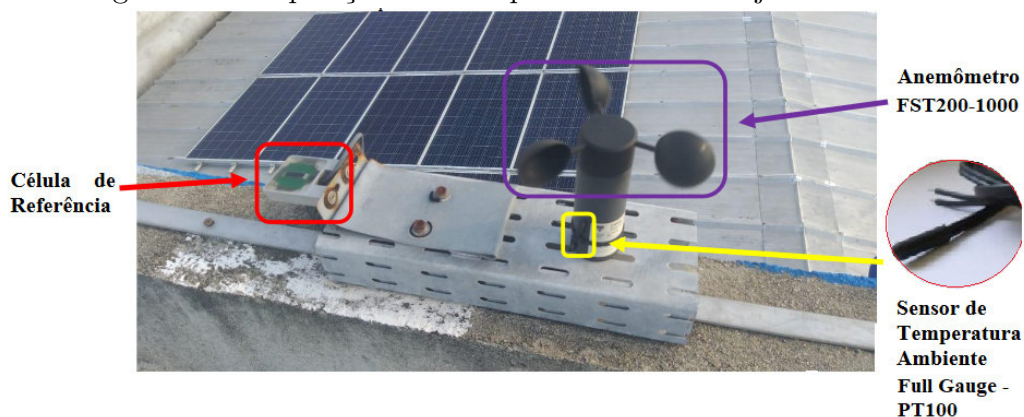
Fonte: Elaborada pelo autor.

3.1 Etapa 1: Coleta e pré-processamento dos dados

Essa etapa tem como objetivo melhorar a qualidade dos dados que serão apresentados ao algoritmo. Os dados utilizados neste trabalho são de geração elétrica diária (kWh), de temperatura ambiente, irradiância, velocidade do vento, potência (kW) e temperatura dos painéis instalados na cobertura de uma unidade do Grupo Boticário (GB) em Fortaleza com capacidade de 22,26 kWp. O sistema FV é conectado à rede elétrica da concessionária, e utiliza o sistema de compensação de energia elétrica estabelecido no capítulo 3 da Resolução Normativa 482/2012 da ANEEL. Com uma coleta de dados adequada, é possível garantir que os dados utilizados na estimação da eficiência do sistema FV sejam confiáveis e representativos do sistema como um todo. Isso é essencial para garantir a precisão e confiabilidade do modelo de ML. Descrita como uma das etapas mais importantes (HUANG, 2017), a coleta de dados é iniciada após a instalação e configuração dos dispositivos necessários para coletar os dados.

Os dados foram coletados continuamente ao longo do tempo de fevereiro de 2020 a julho de 2022, totalizando 29 meses e armazenados em um banco de dados para análise posterior. O sistema FV é composto de um arranjo de 84 módulos, um inversor e quatro strings (conjunto de módulos FV conectados em série). O inversor utilizado é de fabricação ABB e modelo ABB TRIO 20-TL-OUT-D. Os módulos FV instalados são do modelo CS6P-265P do fabricante Canadian Solar e apresentam potência nominal de 265 Wp. A eficiência pode ser calculada dividindo a potência máxima do módulo nas condições STC pela área total do mesmo multiplicada pela irradiância. A Figura 9 apresenta alguns elementos do arranjo FV.

Figura 9 – Disposição dos componentes do arranjo FV



Fonte: Elaborada pelo autor

A Figura 10 mostra os demais componentes do sistema FV utilizados como forma de obtenção dos dados. O quadro “Caixa Meteorológica” recebe dados da célula de referência, do sensor de temperatura Full Gauge PT100 e do anemômetro FST200-1000 instalados.

Figura 10 – Disposição dos demais componentes



Fonte: Elaborada pelo autor

O plano em que se encontra o arranjo FV, conforme é ilustrado na Figura 11 possui um ângulo de inclinação de 7° , medido utilizando equipamento adequado SUN 2 da BENNING. A medição foi feita colocando o equipamento sobre o módulo FV.

Figura 11 – Inclinação dos módulos

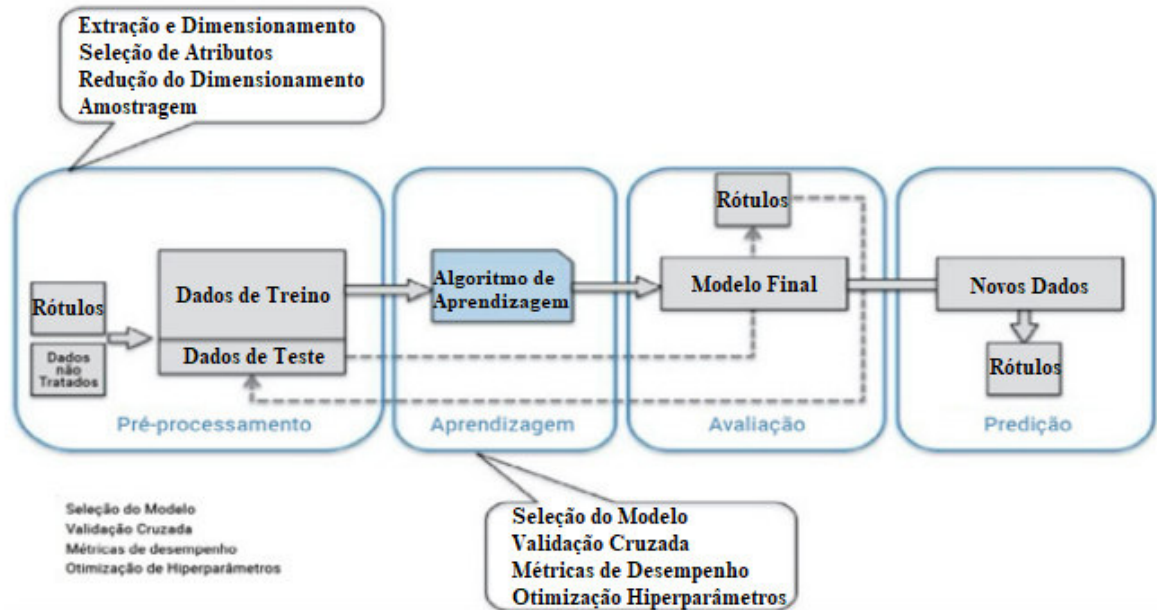


Fonte: Elaborada pelo autor

Ao criar um modelo de ML, nem sempre os dados estão prontos para aplicação, uma vez que esses dados podem conter erros, valores faltantes, dados duplicados e outros problemas. Os dados coletados, ainda, podem estar em escalas diferentes e precisam ser normalizados para que possam ser comparados de forma significativa. A normalização pode ser feita por meio de diferentes técnicas, como normalização min-max, normalização Z-score, entre outras. Neste trabalho, a técnica utilizada para normalizar os dados foi a normalização min-max para remover valores redundantes encontrados no conjunto de dados. Daí, garante-se que os dados estejam em um formato adequado para serem utilizados pelo algoritmo de ML e que o modelo resultante seja preciso e confiável. A limpeza de dados é algo a ser feito quando esses problemas são identificados. Ainda, mesmo não sendo o caso deste trabalho, os valores faltantes podem ser preenchidos com médias ou valores interpolados. Assim, é necessário realizar algumas transformações nos dados antes de apresentá-los ao algoritmo (ENAP, 2020).

Na fase de pré-processamento, os dados são divididos em dados de treino e dados de teste. Os dados de treino são aqueles apresentados ao algoritmo para que ele aprendam o relacionamento entre as variáveis e seja criado o modelo. Já os dados de teste são utilizados para avaliar o quanto o algoritmo aprendeu. Ao apresentar os dados de teste ao modelo, as previsões são realizadas tomando-se como base o que foi aprendido na fase de treinamento. Essas previsões são então comparadas com as respostas esperadas para calcular o desempenho do modelo. Uma vez criado e validado, o modelo pode ser utilizado para que sejam realizadas novas previsões quando for apresentado a novos dados. A Figura 12, adaptada de ENAP (2021), apresenta um esquema destacando as atividades envolvidas na construção de um modelo preditivo. Na etapa de pré-processamento, os dados brutos são preparados para serem utilizados pelo modelo de aprendizado de máquina. Na aprendizagem, o modelo de aprendizado de máquina é treinado usando os dados de treinamento preparados na etapa anterior. Após o treinamento, é necessário avaliar a performance do modelo. Isso é feito utilizando os dados de teste, que não foram usados no treinamento. Após a avaliação do modelo, ele pode ser usado para fazer previsões ou classificações em novos dados

Figura 12 – Atividades envolvidas na construção de um modelo preditivo

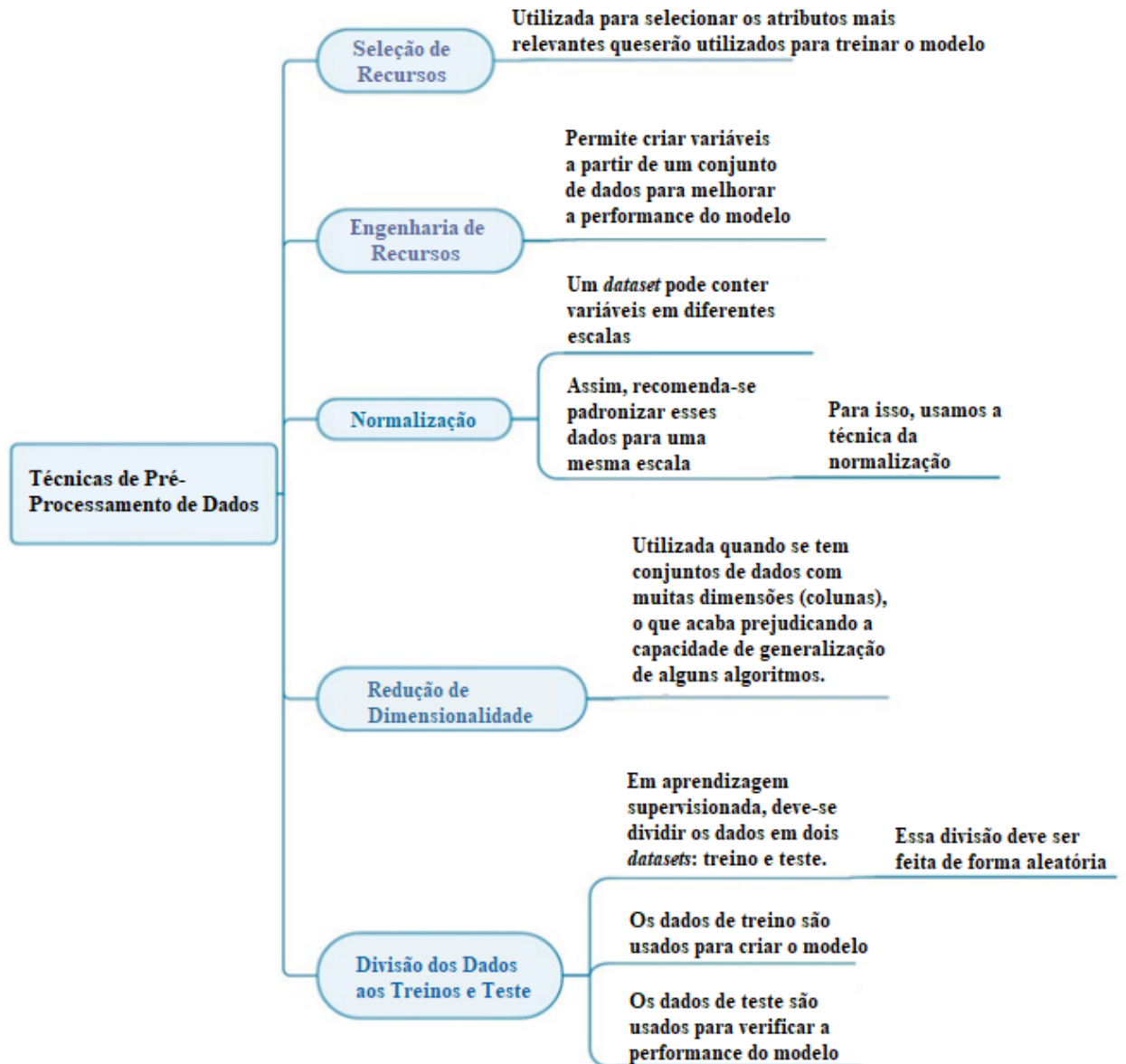


Fonte: Adaptado de (ENAP, 2021)

Após o pré-processamento, realiza-se uma análise exploratória dos dados para compreender melhor as relações entre as variáveis. A seleção de variáveis é feita para escolher as mais relevantes para o modelo. Os dados são divididos em conjunto de treinamento e teste. Em seguida, é construído o modelo, que é treinado utilizando o conjunto de treinamento. A avaliação do modelo é feita utilizando o conjunto de teste, e o modelo é ajustado e otimizado conforme necessário. Após a validação do modelo, ele pode ser implantado em um ambiente de produção. Por fim, é importante monitorar e realizar a manutenção do modelo ao longo do tempo para garantir sua eficácia.

Assim, e com o conhecimento prévio do supervisor sobre o resultado (rótulos/classes), pode-se guiar o aprendizado mapeando as entradas em saídas por meio do ajuste de parâmetros em um modelo capaz de prever rótulos desconhecidos. O ML, portanto, como subconjunto da inteligência artificial (IA), segmento da ciência da computação que se concentra na criação de computadores que "pensam" quase da maneira que os humanos pensam, aprende com seus erros e acertos e é capaz de fazer previsões e tomar decisões baseadas em sua experiência. Algumas técnicas utilizadas nesta etapa são destacadas e explicadas na Figura 13.

Figura 13 – Técnicas de pré-processamento de dados



Fonte: Elaborada pelo autor

3.2 Etapa 2: Escolha do algoritmo mais adequado

Etapa importante para este trabalho, a escolha do algoritmo mais adequado foi feita com base em diversos fatores, como por exemplo os dados disponíveis, o objetivo deste trabalho e as limitações de recurso ou computacionais. Por meio da característica do problema tratado neste trabalho, a técnicas de RNA foram a melhor escolha, pois são capazes de modelar relações complexas. Isso é particularmente útil em sistemas FV, onde a eficiência pode ser afetada por vários fatores, como a temperatura dos módulos, a temperatura ambiente, a radiação solar, entre outros. Além disso, a técnica adotada possui uma boa performance em grandes conjuntos de dados, uma vez que as RNA são

capazes de lidar com grandes conjuntos de dados, o que pode ser útil em sistemas FV, onde muitas medições são coletadas. Isso permite que o modelo capture a variabilidade dos sistemas FV e produza resultados mais precisos. Ainda, há muitas ferramentas de desenvolvimento disponíveis para redes neurais, o que torna a sua implementação mais fácil para cientistas de dados e engenheiros. Isso também significa que há muita documentação e suporte disponíveis para a utilização desses algoritmos. Contudo, é importante ressaltar que outras técnicas de ML, como RL, DT e SVM, também podem ser eficazes para estimar a eficiência de sistemas FV, dependendo do problema em questão. É importante avaliar cuidadosamente as opções disponíveis e selecionar o algoritmo mais adequado para o problema em questão.

3.3 Etapa 3: Construção do modelo e treinamento

Nesta etapa, o modelo é construído a partir dos dados que são apresentados ao algoritmo (ENAP, 2021). Algumas técnicas utilizadas são:

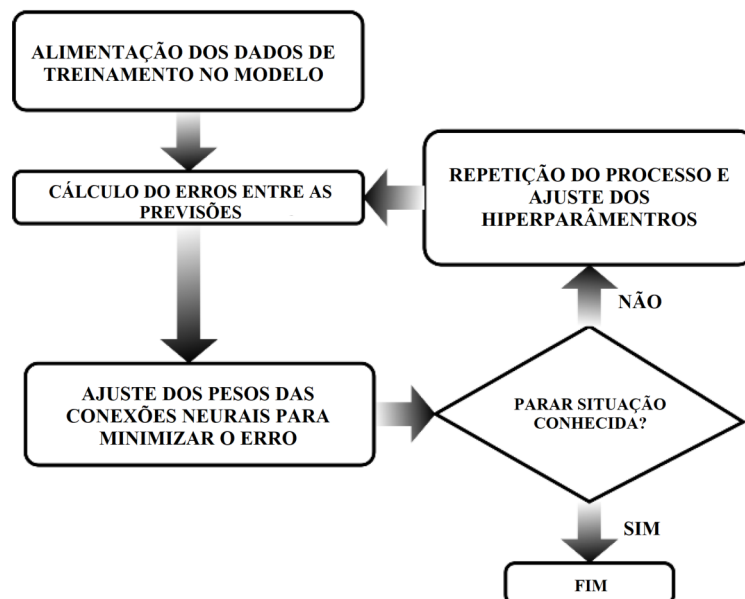
- **Cross-validation**, utilizada para treinar e validar um modelo com o mesmo conjunto de dados, dividindo-os em partições;
- **Métricas de desempenho**, com uso de métricas para medir o desempenho de um modelo. Como exemplo, é possível medir a acurácia (o percentual de previsões corretas em problemas de classificação);
- **Otimização de parâmetros**, com cada algoritmo possuindo um conjunto de hiperparâmetros que podem ser alterados. Essa técnica busca encontrar a combinação certa de valores com o objetivo de melhorar a performance do modelo. Assim, a cada iteração, o algoritmo troca os dados de treino e teste com o objetivo de obter um melhor desempenho.

Outra fase de extrema importância para o sucesso do método implementado neste trabalho é a fase de treinamento. É nessa etapa que o modelo de ML é ajustado para que possa realizar a previsão da eficiência do sistema FV com maior precisão. O processo de treinamento de um modelo de ML envolve a utilização de um conjunto de

dados de treinamento, que consiste em exemplos de entrada e saída conhecidos. Cada exemplo de entrada é uma combinação de valores de entrada, como por exemplo dados meteorológicos, inclinação dos painéis solares, temperatura ambiente, entre outros. A saída de cada exemplo corresponde à eficiência real do sistema FV em questão. O objetivo do treinamento neste trabalho foi de ajustar os parâmetros do modelo de aprendizado de máquina para que ele possa mapear corretamente os valores de entrada para os valores de saída desejados. O processo de treinamento envolve a execução do modelo com os dados de treinamento e a avaliação do desempenho do modelo. Com base na avaliação do desempenho, os parâmetros do modelo são ajustados de forma iterativa até que o desempenho seja satisfatório.

Durante o treinamento, é importante realizar validação cruzada para avaliar a capacidade do modelo de generalização para dados que não foram utilizados no treinamento. A validação cruzada envolve a divisão dos dados de treinamento em vários conjuntos menores, com alguns conjuntos sendo utilizados para o treinamento e outros para a validação. Esse processo é repetido várias vezes para avaliar o desempenho médio do modelo. A Figura 14 apresenta, em forma de fluxograma, a implementação tomada neste trabalho para a etapa de treinamento. A alimentação dos dados de treinamento do modelo consiste em fornecer os dados corretos e representativos para que o modelo possa aprender. Após isso, por meio das métricas estatísticas, pode-se calcular os erros para minimizar os erros de predição do modelo e continuar o aprendizado do método.

Figura 14 – Fluxograma para implementação do treinamento



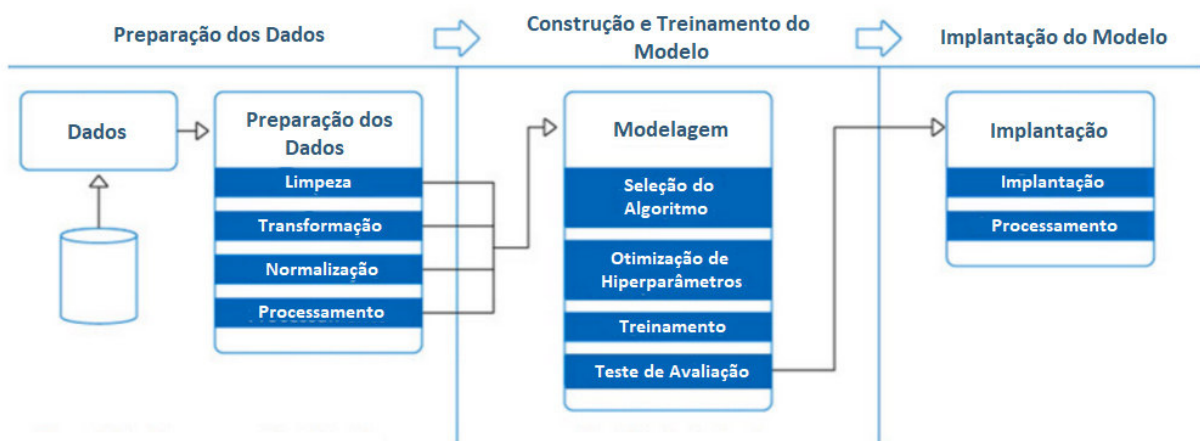
Fonte: Elaborada pelo autor

3.4 Etapa 4: Avaliação e validação do modelo

Nesta etapa, os dados de teste são apresentados ao modelo e, com isso, são geradas previsões. Essas previsões são comparadas com os resultados desejados para avaliar o desempenho do modelo, com base na primeira estratégia deste trabalho, que dividia o conjunto de dados em dados de treinamento e de teste. Isso é feito a partir do método k-fold, que trata-se de um método de validação de modelos preditivos usado para avaliar o desempenho de um modelo de ML, por meio da divisão do conjunto de dados feita anteriormente. Esse método irá permitir fazer uma estimativa imparcial do desempenho do modelo. Dessa forma, a divisão foi feita seguinte o critério de 80/20, com 80% dos dados sendo para teste e apenas 20% para treino.

O conjunto de passos entre a definição de um problema e a escolha de um modelo preditivo para solucionar o problema, pode ser definido seguindo-se a linha ilustrada na Figura 15. A preparação dos dados é a fase em que os dados brutos são coletados e preparados para serem utilizados pelo modelo. Na construção e treinamento do modelo, é selecionado o tipo de modelo de ML a ser utilizado, como redes neurais, árvores de decisão, SVM, entre outros. A implantação do modelo envolve disponibilizá-lo em um ambiente de produção, onde ele pode receber dados de entrada, realizar previsões e retornar os resultados.

Figura 15 – A segmentação de ML



Fonte: Elaborada pelo autor

O critério para avaliar com rigor o modelo proposto foi baseado nos parâmetros estatísticos de RMSE, MSE, R^2 e MAE (RUTH LEAO, 2018). Essas métricas foram aplicadas aos valores experimentais e simulados para medir a exatidão dos modelos preditivos e podem ser vistas através das equações 3.1, 3.2, 3.3, 3.4. O RMSE é uma medida que indica o desvio quadrático médio entre as previsões do modelo e os valores reais. Quanto menor o valor do RMSE, mais precisas são as previsões do modelo. O MAE é uma medida que indica a média dos erros absolutos entre as previsões do modelo e os valores reais. Quanto menor o valor do MAE, mais precisas são as previsões do modelo. O R^2 é uma medida que indica a proporção da variabilidade dos dados que pode ser explicada pelo modelo. Valores de R^2 mais próximos de 1 indicam que o modelo é capaz de explicar uma grande parte da variabilidade dos dados. O MSE é uma medida que indica a média dos erros ao quadrado entre as previsões do modelo e os valores reais. Quanto menor o valor do MSE, mais precisas são as previsões do modelo.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (3.1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (3.2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (3.3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \quad (3.4)$$

Neste capítulo, foram descritas as etapas fundamentais para a implementação dos modelos de ML utilizados na previsão da eficiência dos módulos FV. Inicialmente, foram coletados os dados da usina FV e as informações ambientais relevantes, passando por uma etapa de pré-processamento para garantir a qualidade e a integridade dos dados. A etapa de treinamento dos modelos envolveu o uso de algoritmos como RNA, SVM e RL, otimizando seus parâmetros para obter a melhor performance na previsão da eficiência dos módulos. A avaliação dos modelos foi realizada utilizando métricas estatísticas como RMSE, MAE, MSE e R^2 , que permitiram uma análise precisa do desempenho dos modelos. Os resultados obtidos e a análise detalhada serão apresentados no próximo capítulo.

4 RESULTADOS

Este capítulo se constitui em caracterizar os resultados da análise dos dados produzidos, visando contribuir com diversidade metodológica para o aprofundamento da previsão da eficiência dos módulos em sistemas FV. Destaca-se, também, os resultados dos modelos empregados e a comparação dos resultados de um modelo tradicional com aquele de melhor desempenho dentre os propostos. São consideradas as especificidades meteorológicas da região de Fortaleza como um dos elementos naturais de primeira ordem e a interação desses elementos meteorológicos para prever a eficiência de sistemas FV.

4.1 Configuração dos atributos meteorológicos e representação dos dados

A velocidade do vento (V_w), mostrada na Figura 16a), pode variar ao longo do dia, dependendo das condições meteorológicas locais. Em Fortaleza, pode haver uma variação diurna na velocidade do vento, com uma tendência a ser mais baixa durante a noite e aumentar durante o dia. No entanto, a velocidade do vento diurna pode ser quase constante, com pouca variação dos valores registrados.

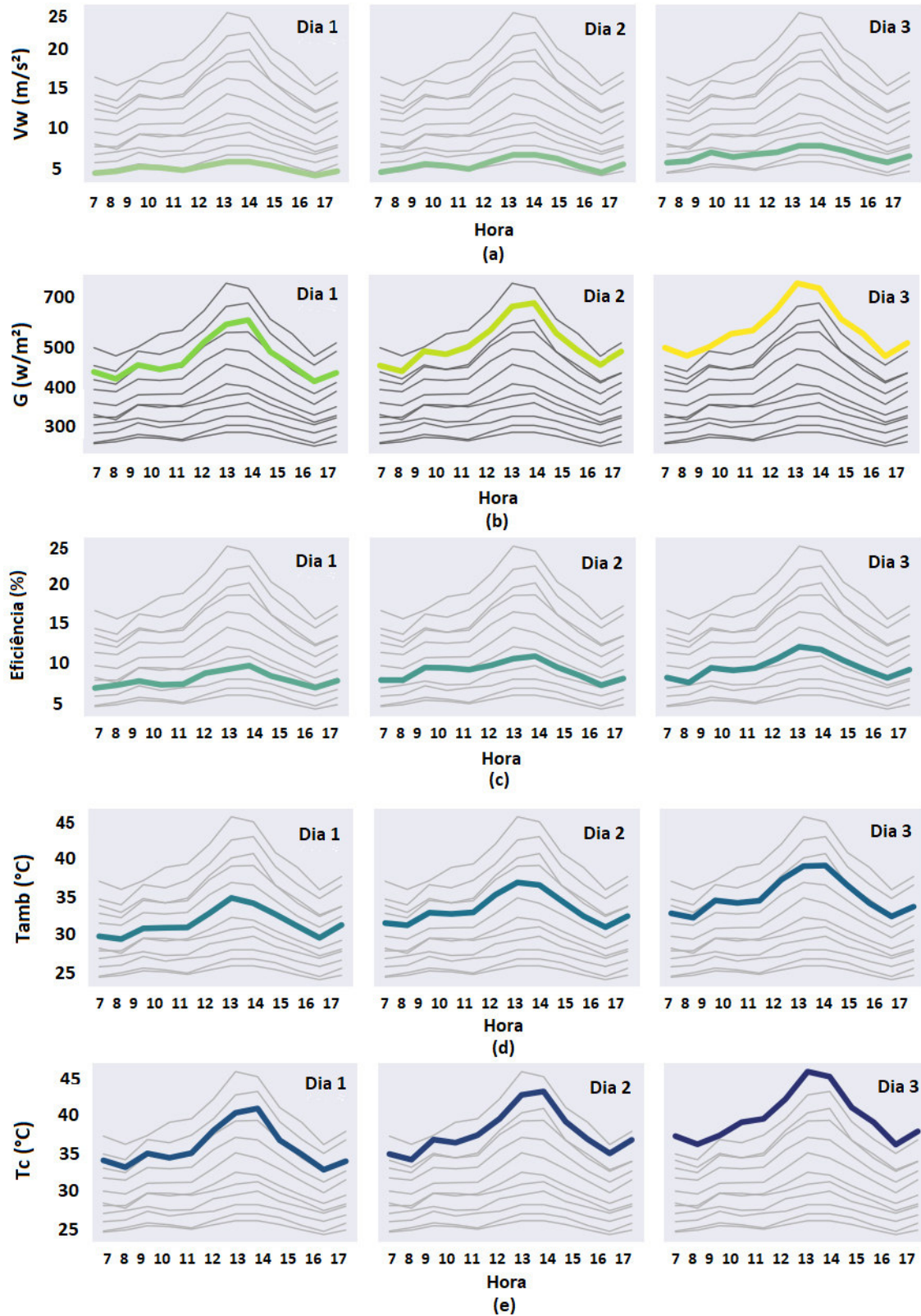
Observa-se, através da Figura 16b), que a distribuição diária da radiação solar (G) em um sistema FV geralmente segue um padrão semelhante ao ciclo diário de luz solar. A radiação solar começa a aumentar gradualmente ao amanhecer, atinge seu pico durante o meio-dia ou às 13:00 horas e depois diminui à medida que o sol se põe. Essa distribuição é influenciada pela posição geográfica, períodos secos e chuvosos.

A eficiência dos módulos FV (η), Figura 16c), pode variar ao longo do dia devido à interação entre a radiação solar e a temperatura dos módulos. Geralmente, espera-se que a eficiência atinja seu valor máximo durante o período em que a radiação solar é mais intensa, ou seja, durante o meio-dia ou início da tarde. Conforme a radiação solar diminui e a temperatura dos módulos também, a eficiência pode diminuir ligeiramente.

A temperatura ambiente (T_{amb}), na Figura 16d), segue um padrão diário semelhante. No início da manhã e durante a noite, a temperatura ambiente tende a ser mais baixa. À medida que o sol se eleva, a temperatura ambiente começa a aumentar, iniciando em torno de 27°C e atingindo seu pico durante o meio-dia ou início da tarde, com valor em torno de 30°C . Após o pico, a temperatura ambiente começa a diminuir. A temperatura dos módulos (T_c) em um sistema FV também segue uma distribuição diária.

Durante a manhã e noite estão com temperatura reduzida, entre 25°C 35°C. Ao meio dia, no período de seca, a temperatura dos módulos alcança valores iguais a 45°C.

Figura 16 – Comportamento dos atributos meteorológicos: a) V_w ; b) G ; c) η ; d) T_c e e) T_{amb}

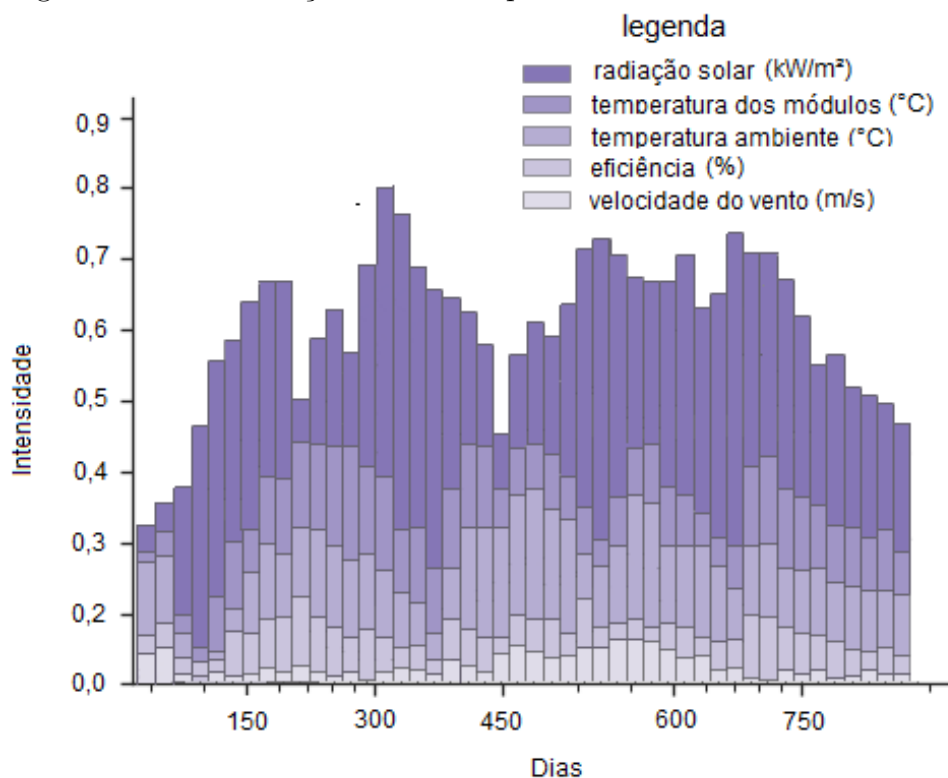


Fonte: Elaborada pelo autor

Outrossim, na Figura 16, os gráficos individuais são organizados entre três dias, permitindo visualizar várias séries temporais lado a lado. Isso facilita a comparação e a identificação de padrões ou tendências comuns ou distintas entre as séries. Nessa figura, as curvas auxiliares, em cinza, constituem os registros dos outros parâmetros (V_w ; G ; η ; T_c e T_{amb}), facilitando analisar a correlação e como o comportamento de um parâmetro influencia o outro. As curvas coloridas, por outro lado, constituem os registros propriamente dito para aquele parâmetro específico.

Para entender o desempenho e comportamento das variáveis de entrada (V_w ; G ; η ; T_c e T_{amb}), está ilustrado na Figura 17, além da distribuição anual da radiação solar, que mostra um padrão sazonal, com maior incidência durante os meses de verão. A temperatura dos módulos é influenciada pela radiação solar, temperatura ambiente e velocidade do vento, com temperaturas mais altas durante o verão e em climas mais quentes. Para comportar os valores de G , que nesse exemplo está kW/m^2 , foi elaborado um eixo vertical em decimal ao se reduzir a proporção dos valores dividindo-os por 100 e trabalhando-os entre 0,0 a 1,0. A intensidade, como foi denominado o eixo vertical se refere aos valores de todos os parâmetros sem precisar indicar suas unidades, permitindo exibir todos em uma única figura

Figura 17 – Distribuição média dos parâmetros



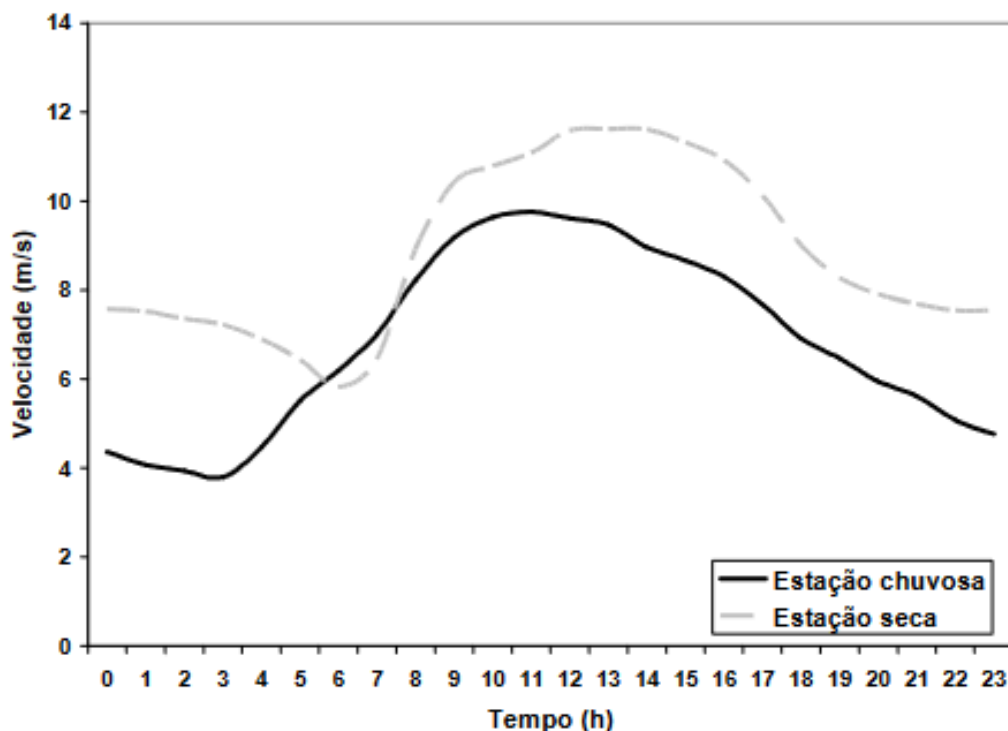
Fonte: Elaborada pelo autor

E, com efeito, na Figura 17, a relação entre tempo (em dias) e todos os parâmetros (V_w ; G ; T_{amb} ; T_c e η) possuem valores coerentes, com variações em determinados períodos, visivelmente no período seco e chuvoso.

4.1.1 Velocidade do vento

Esse parâmetro exibiu ao longo da história da cidade de Fortaleza uma evolução (FUNCEME), com seus valores indo da calmaria para um recorte sazonal e apresentando um gradiente baixo no primeiro semestre. A responsável pela variação, modulação e caracterização dos ventos na cidade é a Zona de Convergência Intertropical, que atua como uma banda de nuvens na região. Fundamentada em dados de reanálise do modelo global (MERRA), a Figura 18 indica as intensidades mais significativas que Fortaleza está suscetível, com intensidades mais significativas da velocidade do vento nas regiões litorâneas. Devido ao vento ser sensivelmente influenciado pela topografia e vegetação em suas proximidades, características distintas podem ser observadas mesmo em regiões próximas aos pontos indicados nas medições da Figura 18, onde efeitos locais podem predominar.

Figura 18 – Velocidade média do vento para os períodos chuvosos e secos de Fortaleza

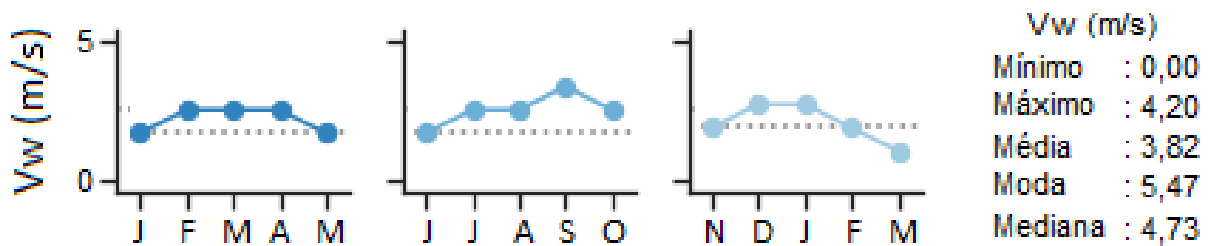


Fonte: Henrique do Nascimento, 2018

4.1.1.1 Regimes do vento em Fortaleza

Os meses iniciais do ano, março e abril, são os que possuem as baixas na velocidade do vento. Em oposição a isso, no segundo semestre, verificam-se valores mais significativos, com altas nos meses finais do ano. Em Fortaleza, a velocidade se tornou a variável mais alterada (FUNCEME) devido às mudanças na urbanização da cidade, com a verticalização contribuindo muito para esse recorte. No caso da cidade de Fortaleza, com respeito a velocidade do vento durante o dia, o principal fator de influência desse atributo reside no fato urbano, em termos de uma ascendente verticalização, devido a especulação imobiliária, notadamente presente nas áreas de orla marítima e demais bairros considerados privilegiados. A Figura 19, exhibe a distribuição anual da velocidade do vento com base nos dados coletados.

Figura 19 – Gráfico da distribuição anual da velocidade do vento do banco de dados do GB em Fortaleza



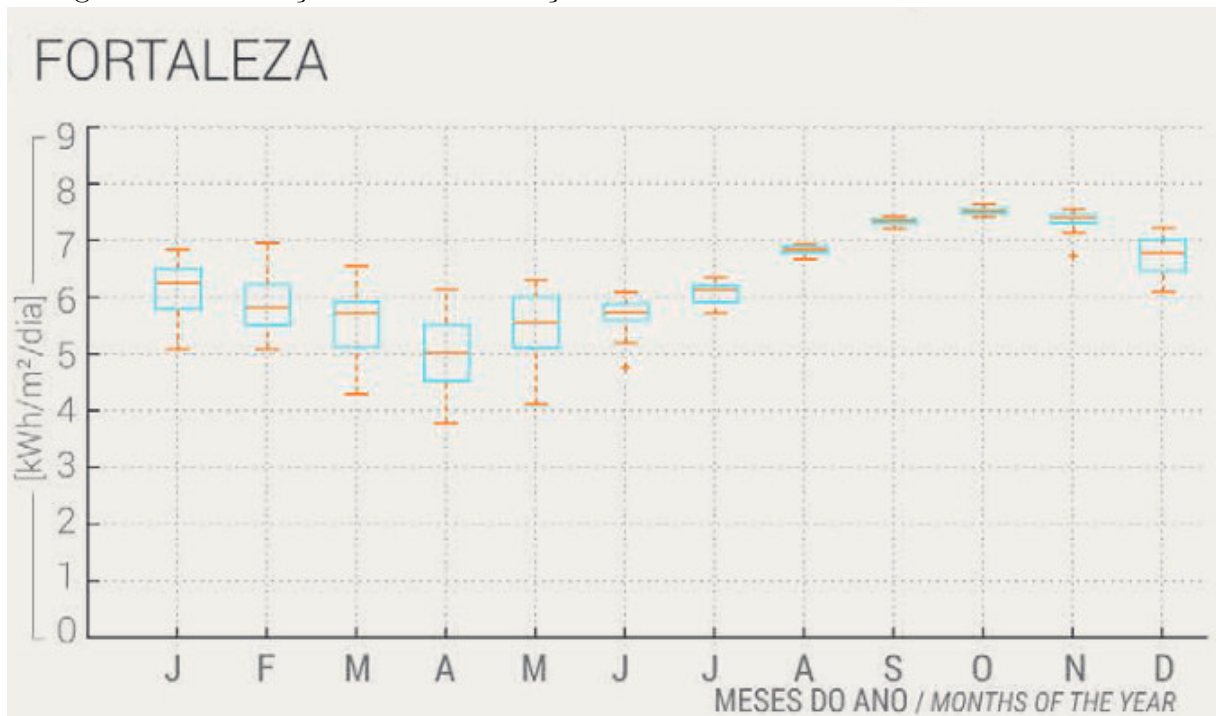
Fonte: Elaborada pelo autor

Na Figura 19, observa-se a modulação anual da velocidade do vento, conforme dados médios mensais para o período de chuvoso e seco, elaborado a partir de medições anemométricas coletadas. De acordo com essa figura, fica claro que os ventos em Fortaleza exibem velocidades médias durante o período seco, principalmente de setembro a dezembro, e menores médias mensais no ápice do período chuvoso, de março a abril. Mensalmente, a velocidade do vento distribui-se como exibido na Figura 19, ao compararmos com as informações da (FUNCEME). O comportamento das curvas da Figura 19 deixa claro como a velocidade do vento tende a alcançar valores maiores no segundo semestre que no primeiro semestre, mantendo-se os valores entre 1 m/s e 5 m/s, para mínimo e máximo. A média, além disso, se demonstrou igualmente coerente, com valor de 3,92 m/s. .

4.1.2 Radiação solar

Em termos de radiação solar, paralelamente ao elemento tratado anteriormente, a cidade de Fortaleza ocupa uma posição de destaque, pois é uma das cidades mais ensolaradas do Brasil. Na cidade, a radiação solar possui um comportamento específico que demonstra por altos níveis de radiação solar ao longo do dia e do ano. Isso ocorre devido à localização da cidade, que está situada próxima à linha do Equador e pela influência exercida por algumas Zonas de Convergência, como é o caso da Zona de Convergência Intertropical (FUNCEME). A Figura 20 mostra como incide a radiação solar em toda a cidade de Fortaleza. Observam-se os maiores índices de radiação no mês de outubro e menor dispersão dos dados na segunda metade do ano. No início do ano, apesar dos índices de radiação serem menores que no final do ano, a dispersão dos dados é maior que a segunda metade do ano.

Figura 20 – Variação anual da radiação solar na cidade de Fortaleza

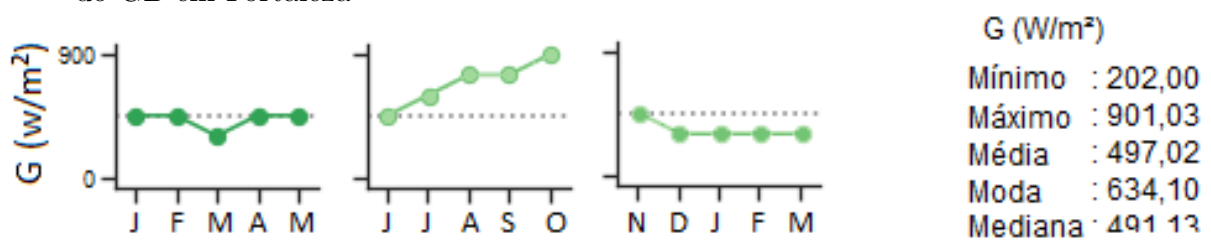


Fonte: ATLAS SOLAR

4.1.2.1 Regimes de radiação solar em Fortaleza

No decorrer do ano, a metrópole experimenta uma variação sazonal na radiação solar, com maior intensidade durante o período seco e menor intensidade durante o período chuvoso. Durante os meses finais do ano, principalmente os de agosto a dezembro, a radiação solar atinge seu pico, uma vez que a quantidade de nuvens e a cobertura do céu podem ser menores. Nesse intervalo, as condições climáticas tendem a ser mais favoráveis para uma maior incidência de radiação solar, com céu mais claro e menor cobertura de nuvens. Entretanto, mesmo durante os meses chuvosos, Fortaleza ainda recebe níveis significativos de radiação solar, devido à sua posição tropical. A presença de poucas nuvens ao longo do ano também contribui para altos índices de radiação solar em Fortaleza. No que diz respeito ao comportamento da radiação solar ao longo do dia em Fortaleza, geralmente segue um padrão ascendente desde o nascer do sol, atingindo seu pico próximo ao meio-dia, e diminuindo gradualmente até o pôr do sol. No período de pico, a radiação solar é mais intensa, proporcionando um maior potencial para a geração FV. A Figura 21 exhibe as características distintas da radiação solar ao longo do ano, com base nos dados colhidos na unidade do GB.

Figura 21 – Gráfico da distribuição anual da radiação solar do banco de dados do GB em Fortaleza



Fonte: Elaborada pelo autor

A Figura 21 mostra a distribuição anual da radiação solar a partir dos maiores valores para o mês registrado. Nessa figura, podemos verificar a sazonalidade que se manifesta em Fortaleza, reforçando, assim, aquilo difundido na literatura. Nota-se como no segundo semestre, de agosto a outubro, os valores de radiação solar são maiores em decorrência da incidência solar e da baixa cobertura de nuvens. No primeiro semestre, de forma diametralmente oposta, os valores se apresentam com menos intensidade.

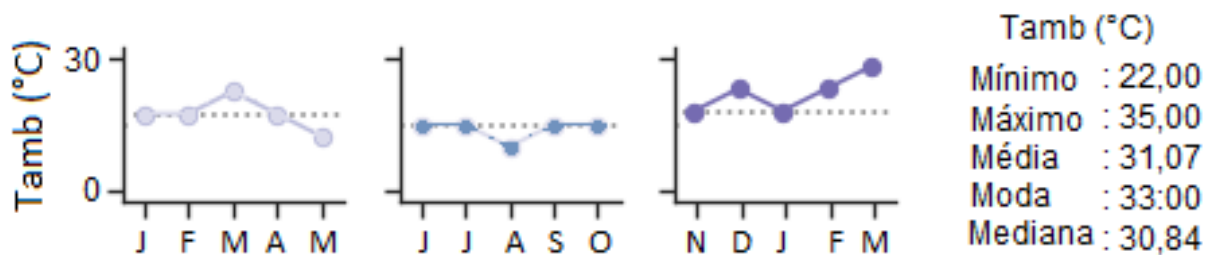
4.1.2.2 Relação da radiação solar com outros parâmetros meteorológicos

A radiação solar, de igual modo, está diretamente relacionada a outros parâmetros meteorológicos, como a temperatura ambiente. A quantidade de radiação solar recebida em Fortaleza é um dos principais fatores que influenciam a temperatura ambiente. Quanto maior a quantidade de radiação solar recebida, maior tende a ser a temperatura ambiente. Decorre que, quando a radiação solar incide sobre a superfície terrestre, fornece energia térmica, fazendo com que a temperatura aumente. Em uma localização como a de Fortaleza, suscetível a receber mais radiação solar, a tendência é ter temperaturas mais altas, enquanto que, regiões com menor incidência de radiação solar apresentam temperaturas mais baixas. Certifica-se, através daquilo apresentado na Figura 15(b), que conforme a radiação solar aumentou, a temperatura ambiente descreveu um movimento ascendente. A partir da Figura 15(b), também, fica claro que a irradiância atingiu valores maiores que 700 W/m^2 e ainda, a irradiância não variou mais que 20% durante aproximadamente 1 hora anterior ao registro seguinte.

4.1.3 Temperatura ambiente

A temperatura ambiente de Fortaleza é influenciada pelo clima tropical úmido da cidade de Fortaleza, contribuindo assim para as altas temperaturas registradas (INMETRO) ao longo do ano. Em média, a capital apresenta temperaturas elevadas durante a maior parte do ano. Os meses iniciais, notadamente os de janeiro a março, são considerados os mais quentes, com médias que podem ultrapassar os $30 \text{ }^\circ\text{C}$. No decurso desse período de verão, é comum ocorrerem dias com temperaturas próximas ou acima de $35 \text{ }^\circ\text{C}$. A Figura 22 caracteriza a distribuição da temperatura ambiente em Fortaleza.

Figura 22 – Gráfico da distribuição média anual da temperatura ambiente do banco de dados do GB em Fortaleza



Fonte: Elaborada pelo autor

4.1.3.1 Regimes da temperatura ambiente de Fortaleza

É importante destacar, pela Figura 22 que Fortaleza também possui variações sazonais nas temperaturas. Durante os meses de junho a agosto, que correspondem ao inverno no hemisfério sul, as temperaturas tendem a ser um pouco mais amenas, com médias em torno de 27°C. Além das variações sazonais, as temperaturas diárias em Fortaleza também podem apresentar diferenças significativas entre o dia e a noite. Durante o dia, com a presença de luz solar intensa, as temperaturas podem subir rapidamente, alcançando valores elevados. Já durante a noite, as temperaturas tendem a cair, mas ainda se mantêm em níveis relativamente altos devido à influência do clima tropical.

4.1.3.2 Relação da temperatura ambiente com outros parâmetros meteorológicos

Importante destacar a íntima relação que a temperatura ambiente guarda com a radiação solar, aquela apresentando elevações quando essa demonstra aumento na intensidade. A Figura 22 nos permite identificar essa relação, sendo possível, ademais, afirmar, com o auxílio da Figura 16(c) que o pico de temperatura ocorre às 13:00 horas. Após às 15:00 horas, a curva da temperatura ambiente descende.

4.1.4 Temperatura dos módulos

A temperatura dos módulos de um sistema FV é um fator importante a ser considerado, pois afeta diretamente a eficiência dos módulos e o desempenho do sistema. A temperatura dos módulos pode variar de acordo com diferentes fatores, como a radiação solar incidente, a temperatura ambiente, a circulação do vento e, em alguns casos até mesmo o projeto do sistema FV pode variar a temperatura dos módulos.

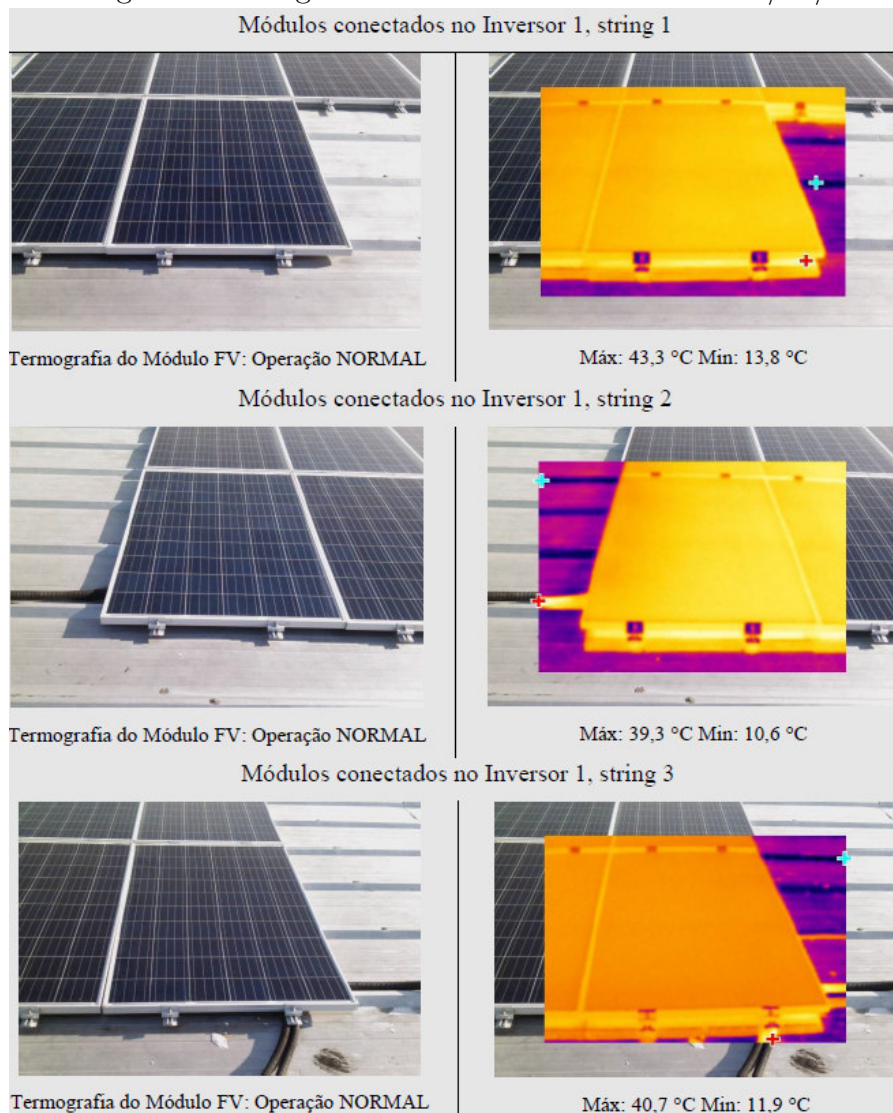
4.1.4.1 Inspeção termográfica

A inspeção termográfica, realizada com sensores de raios infravermelhos, objetiva a procura de temperaturas anormais no sistema FV como um todo. A inspeção foi realizada quando o sistema estava em pleno funcionamento e para efetiva validação dos testes, certificou-se que a irradiância no momento dos testes atingiram valores iguais ou maiores que 800 W/m² e ainda, a irradiância não variou mais do que 20% no decorrer dos 10 minutos anteriores aos testes. Ressalta-se, por meio da Figura 23, que não foi encontrada

célula FV danificada.

Em situações em que a temperatura de uma célula seja inferior à média das células de um mesmo painel FV não necessariamente configura células danificadas, mas apenas indica a presença de sujeira ou sombreamento. Realizou-se a inspeção termográfica nos módulos FV em busca de pontos quentes e variações de temperatura anormais entre diferentes pontos de um mesmo módulo. Detalhes dos registros termográficos são mostrados na Figura 23. A análise termográfica do sistema mostrou que todos os componentes do sistema estavam operando com temperatura normal. Além disso, nenhum módulo FV ou outro componente do sistema estava com temperatura próxima de 100°C. Dessa forma, o teste de termovisão não identificou nenhum equipamento danificado ou necessitando de substituição.

Figura 23 – Registros termográficos dos módulos FV do dia 18/02/2021 as 11:00



Fonte: Relatório de inspeção termográfica

A análise da temperatura dos módulos mostrou que os módulos FV do sistema estavam operando com temperatura normal. A inspeção revelou uma distribuição uniforme de temperatura nos módulos, indicando um funcionamento adequado e eficiente. A comparação das temperaturas entre diferentes módulos, também indicou equilíbrios de desempenho, constatando a alta eficiência que eles alcançaram.

4.1.4.2 Relação da temperatura dos módulos com parâmetros meteorológicos

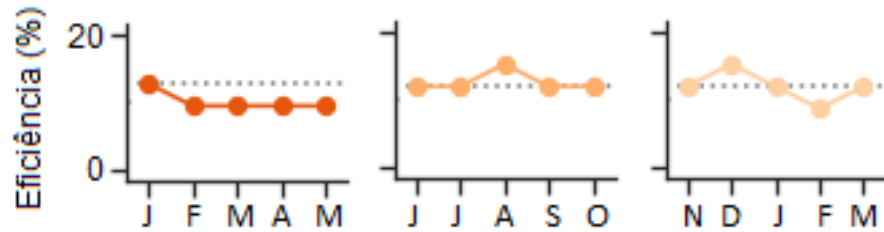
Na operação normal de um sistema FV, os módulos absorvem a radiação solar e convertem parte da energia solar em eletricidade. No entanto, esse processo também gera calor nos módulos, resultando em um aumento da temperatura. À medida, portanto, que a quantidade de radiação solar incide nos módulos, afeta diretamente a sua temperatura. Assim, quanto maior a radiação solar, maior será o aquecimento dos módulos, devido à absorção dessa radiação. Por esse motivo, é comum que a temperatura dos módulos FV geralmente seja mais elevada do que a temperatura ambiente. A temperatura dos módulos eleva-se com o aumento da irradiação incidente e após cerca de 15:00 horas começa a reduzir.

4.1.5 Eficiência

A eficiência é uma das principais métricas utilizadas para avaliar o desempenho dos módulos FV. Quanto maior a eficiência, maior a quantidade de energia elétrica que pode ser gerada a partir da mesma quantidade de radiação solar. Os módulos FV comerciais disponíveis no mercado geralmente têm eficiências que variam de 15% a 25%. Em alguns casos, existem módulos de alta eficiência disponíveis que podem alcançar eficiências de 30%. A eficiência dos módulos depende de vários fatores, incluindo a tecnologia e o material semicondutor utilizados, o projeto e a fabricação dos módulos, além das condições ambientais em que são instalados.

Os valores de eficiência dos módulos FV, conforme informado previamente na seção anterior, foram cedidos pela fabricante e determinados a partir da relação entre potência e área dos módulos FV, levando em consideração as especificações e as condições de operação para determinar a eficiência. Com isso, tomando a Figura 24 como exemplo, nota-se que a eficiência dos módulos aproxima-se de 20%, sobretudo no verão com maiores intensidades de radiação solar.

Figura 24 – Distribuição média anual da eficiência do banco de dados do GB



Fonte: Elaborada pelo autor

Entre os detalhes a serem considerados na observação da eficiência, com base na Figura 24, estão o detalhamento das perdas na estimativa mensal de energia. Os horários em que a eficiência dos módulos FV alcança valores máximos e mínimos podem variar dependendo de vários fatores, incluindo a localização geográfica, a orientação e a inclinação dos módulos, as condições climáticas e a tecnologia dos módulos utilizados. Contudo, pode-se destacar que a eficiência é máxima nos horários em que a radiação solar é mais intensa, notoriamente, ao meio dia, em que o sol está mais elevado no céu. Nesse instante, como visto, a radiação solar é mais intensa e permite que a eficiência seja maior. Em contrapartida, a eficiência dos módulos é mínima naqueles horários em que a radiação solar é mínima, como no início da manhã e no final da tarde.

4.2 Modelo de previsão da eficiência dos módulos

Selecionamos o modelo com melhor desempenho para avaliar as nossas características de previsão propostas. Utilizou-se o modelo de previsão, treinado com dados históricos de V_w , G , T_{amb} , T_c e η , avaliando o conjunto de dados dividido em treinamento e teste as análises de impacto.

4.2.1 Treinamento e performance do modelo

Para prever os valores das etapas de tempo futuro de uma sequência usando RNA - de alimentação direta -, é necessário treinar o modelo usando os parâmetros temporais (dias e horas), G , T_{amb} , T_c e V_w e η . Com isto, para se obter um melhor ajuste e evitar que o treinamento seja divergente, os dados usados no treinamento e validação foram de 80% daqueles contidos no banco de dados. Os dados de teste são avaliados usando os mesmos parâmetros dos dados de treinamento. Em outro passo, as RNA aprendem a prever os valores futuros através do treinamento dos 20% dos dados restantes do conjunto

de dados, ou seja, de teste. Com base em dados de séries temporais, as RNA são adequadas para fazer previsões.

Existem vários métodos para determinar a exatidão ou performance de um modelo. Um deles é o MAE, que é a média de todos os erros absolutos, indicando as diferenças entre os valores reais e os previstos. Além desse método, são empregados outros que nos permite analisar o desempenho da técnica utilizada, como é o caso do RMSE, R^2 e MSE. A Tabela 3 apresenta uma avaliação inicial das características do algoritmo usando os parâmetros estatísticos de RMSE, MAE e MSE, expostos na magnitude da variável de referência, a eficiência (em %), com o MSE sendo a unidade da variável ao quadrado. Já o R^2 varia de 0 a 1, onde 1 indica uma explicação perfeita da variabilidade dos dados.

Tabela 3 – Resultados dos parâmetros estatísticos de treinamento

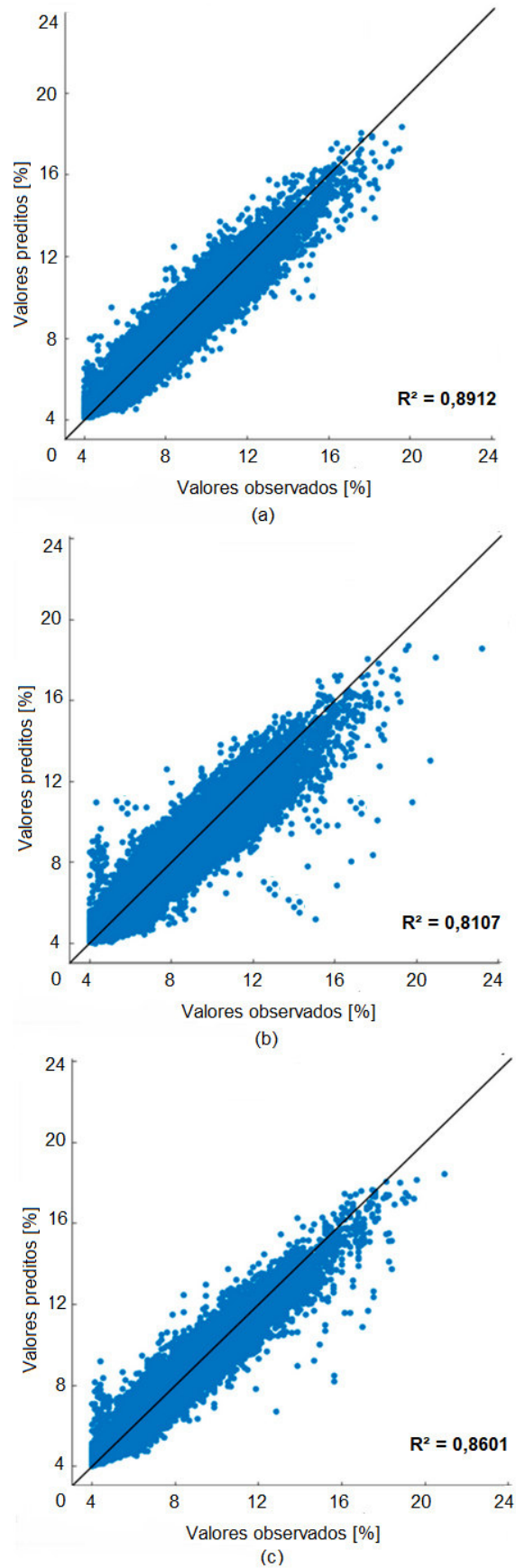
Modelo	Aplicabilidade	R^2	MAE	RMSE	MSE
RNA	Treinamento	0,8897	1,1164	1,9971	7,2213
	Teste	0,8912	1,0972	1,8148	6,9841
SVM	Treinamento	0,8589	2,3843	2,7824	8,8340
	Teste	0,8601	2,2451	2,6412	8,5032
RL	Treinamento	0,8096	3,2317	3,3495	13,9836
	Teste	0,8107	3,1025	3,2831	13,4529

Fonte: Elaborada pelo autor

A Figura 25, que consiste em três gráficos, ilustra a dispersão das previsões dos três modelos em relação aos dados de teste. Na figura, os pontos representam as previsões dos modelos em relação aos valores reais dos dados de teste. Dessa forma, cada ponto no gráfico representa uma amostra de dados e sua posição no eixo y corresponde à previsão do modelo, enquanto sua posição no eixo x corresponde ao valor real (ou observado). Os gráficos permitem avaliar o quão bem as previsões dos modelos se aproximam dos valores reais. Idealmente, os pontos estariam alinhados em uma linha diagonal, o que indicaria uma correspondência perfeita entre as previsões e os valores reais.

De acordo com a visualização dos gráficos da Figura 25, a precisão e a acurácia das previsões dos modelos em relação aos valores reais diferencia-se pouco deste, com a maior dispersão dos pontos em torno da linha diagonal representada pelo gráfico da Figura 25(b) do modelo RL, em seguida pela Figura 25(c) do modelo SVM e a menor dispersão observada pela Figura 25(a) do modelo RNA. Com isto, os gráficos baseiam-se na finalidade expressa sobre o quão bem o modelo está ajustado aos dados e identificar possíveis padrões ou discrepâncias nas previsões.

Figura 25 – Gráficos de dispersão de: RNA, RL e SVM

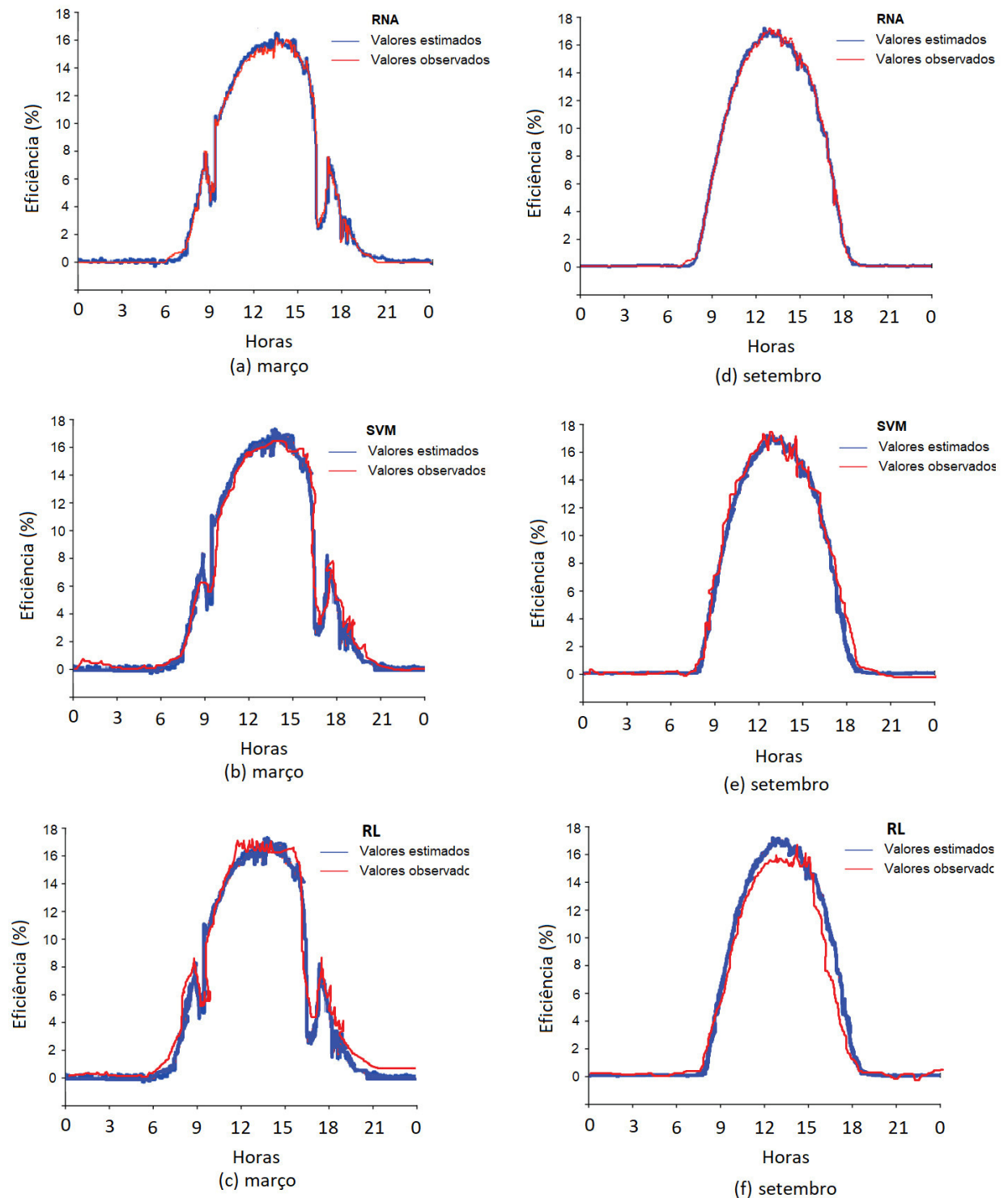


Fonte: Elaborada pelo autor

4.2.2 Resultados do modelo

Os gráficos da Figura 26 são gráficos de estimativa para a eficiência. Esses gráficos fornecem informações sobre as métricas de eficiência estimadas com base nos dados da cidade de Fortaleza.

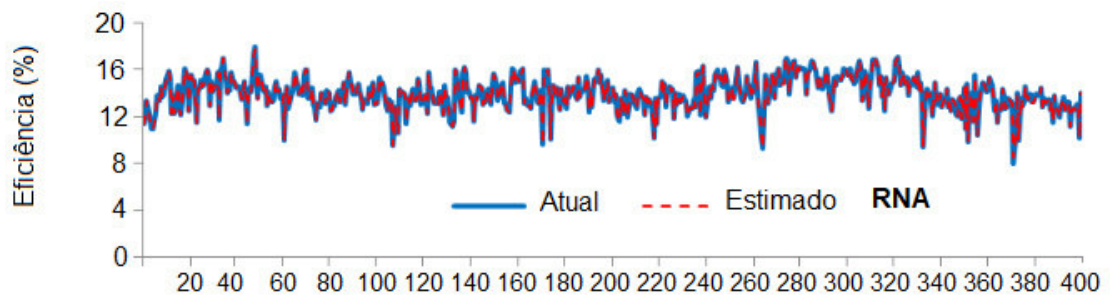
Figura 26 – Gráficos dos valores observados para dias típicos de março e setembro



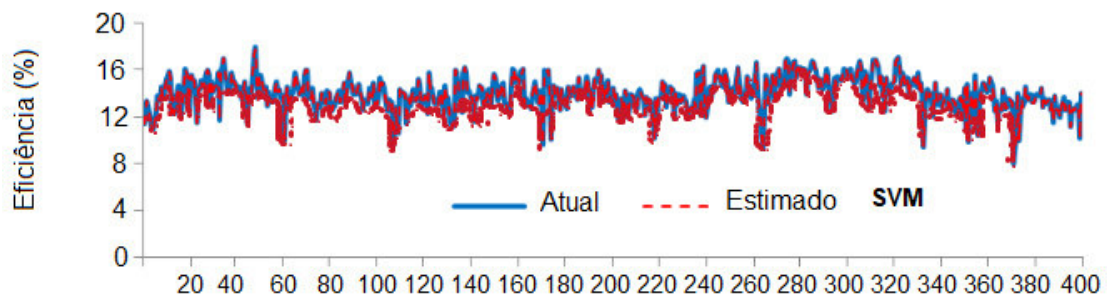
Fonte: Elaborada pelo autor

Esses gráficos, ainda, de estimativa obtidas através do modelo são valiosos para obter uma visão geral das características e tendências da eficiência, fornecendo uma representação visual das informações dos dados observados de eficiência e os preditos. Os gráficos ajudam a identificar padrões, variações e possíveis relações entre diferentes métricas estimadas. Com isto, fica claro que a eficiência, influenciada pela irradiância e temperatura dos módulos, é alta quando estes parâmetros também são altos, notadamente no início da tarde. A partir das métricas estatísticas apresentadas anteriormente, com o apoio dos gráficos da Figura 26, observa-se, novamente, a adequação do modelo implementado, com poucas variações em relação aos valores reais coletados. A Figura 27 expõe os valores estimados com os modelos aplicados ao longo do ano.

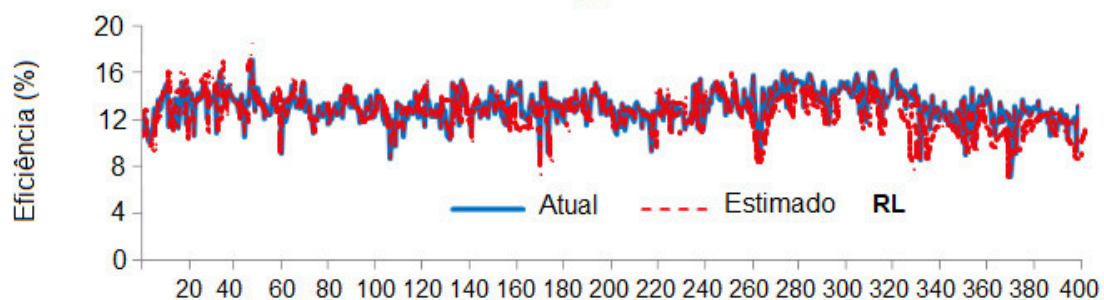
Figura 27 – Gráficos dos valores estimados pelos modelos ao longo do ano de 2022



(a) Dias



(b) Dias



(c) Dias

Fonte: Elaborada pelo autor

4.2.3 Análise comparativa

Os resultados obtidos anteriormente indicam que o modelo RNA é o mais capaz de capturar a previsão da eficiência dos módulos do sistema FV, uma vez que ao comparar as previsões do modelo RNA com os valores reais de eficiência dos módulos, observou-se uma concordância significativa. Apesar disso, os dados de teste do modelo RNA foram, também, defrontados com um modelo comparativo, descrito na literatura como eficiência do módulo no NOCT (Temperatura Nominal de Operação das Células) para analisar ainda mais seu rigor em prever a eficiência dos módulos. A equação 4.1 representa a relação entre a eficiência elétrica (η_{el}) de um módulo FV em um determinado momento e a eficiência nominal de operação em condições normais de temperatura (η_{NOCT}), considerando a variação da temperatura do módulo (t_p) em relação a uma temperatura de referência de 25°C.

$$\eta_{el} = \eta_{NOCT}[1 + C(t_p - 25)] \quad (4.1)$$

Onde η_{NOCT} é a eficiência do módulo FV no NOCT (eficiência fornecida pelo fabricante). A Tabela 4 apresenta os resultados obtidos após a implementação do modelo comparativo.

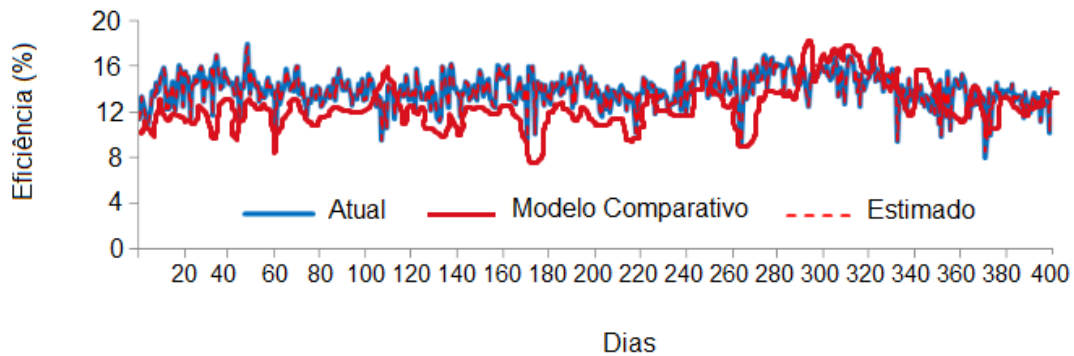
Tabela 4 – Resultados dos parâmetros estatísticos de treinamento

Modelo	R ²	MAE	RMSE	MSE
RNA	0,8912	1,0972	1,8648	6,9841
NOCT	0,5963	5,0176	7,3951	51,2648

Fonte: Elaborada pelo autor

Conforme demonstrado na Tabela 4, em comparação com o modelo comparativo apresentado na literatura, para analisar o rigor do modelo de previsão da eficiência dos módulos proposto, este consegue alcançar uma precisão maior. Para isso, tomamos os resultados da Tabela 4, atentando-se para os valores métricos resultantes da comparação, como os menores (MAE, RMSE e MSE) obtidos para o modelo RNA e o melhor R². O valor de RMSE 1,8148 do modelo RNA no conjunto de teste indica que o erro médio absoluto do modelo em estimar a eficiência dos módulos é de 1,8148%. Na Figura 28, são apresentados os valores previstos pelo modelo RNA, bem como pelos modelos comparativos. É possível notar que o modelo RNA acompanha os padrões de variação da eficiência dos módulos de forma próxima à linha dos valores medidos.

Figura 28 – Gráfico dos valores dos modelos comparativo e estimado ao longo do ano de 2022



Fonte: Elaborada pelo autor

Neste trabalho, o valor baixo de RMSE (1,8648) indica que as previsões do modelo estão muito próximas dos valores reais, o que demonstra a capacidade do modelo em realizar previsões precisas. Outra métrica que apresentou um valor baixo foi a MAE (1,0972), indicando que as previsões do modelo têm uma precisão consistente, estando muito próximas dos valores reais em média. Um valor baixo de MSE (6,9841) indica que as previsões do modelo estão muito próximas dos valores reais, com uma boa capacidade de ajuste aos dados observados. Por último, o R^2 é uma métrica que mede a proporção da variabilidade dos valores reais que é explicada pelo modelo. Um valor alto de R^2 (89%) indica que o modelo é capaz de explicar uma grande parte da variação dos dados, ou seja, as previsões são altamente correlacionadas com os valores reais.

A comparação entre os valores estimados e reais, possível através da análise dos gráficos das Figuras 26, 27 e 28, permite visualizar a semelhança dos valores estimados no modelo RNA com os valores reais e a discrepância dos valores apresentada nos valores estimados nos modelos de SVM e RL, principalmente. Com base na sobreposição dos valores estimados e reais desses gráficos, é possível avaliar a precisão do modelo RNA, com os pontos estimados próximos aos pontos reais, indicando que o modelo RNA está capturando com acurácia os padrões e tendências da eficiência dos módulos.

Esses resultados evidenciam a confiabilidade e a eficácia do modelo de previsão da eficiência dos módulos FV. Com baixos valores de RMSE, MAE e MSE, juntamente com um alto valor de R^2 , o modelo demonstra sua capacidade de realizar previsões precisas e explicar uma parcela significativa da variabilidade dos dados. Isso permite que o modelo seja aplicado com confiança em diversas aplicações práticas, contribuindo para a previsão da eficiência dos módulos FV.

5 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, foi explorada a utilização de técnicas de ML para prever a eficiência dos módulos de usinas FV. O objetivo foi prever a eficiência dos módulos das usinas dessa natureza e característica para beneficiar a integração de tecnologias como a implementada neste estudo e o crescimento da energia solar. Para isso, foram desenvolvidos e avaliados diferentes modelos como RNAs, SVM e RL. Os modelos foram alimentados com dados históricos de eficiência da usina e também com séries históricas de dados meteorológicos. Essa abordagem permitiu a comparação de um método tradicional com o RNA, a análise do desempenho e a capacidade dos modelos em realizar previsões precisas e confiáveis da eficiência dos módulos em usinas FV. Dessa maneira, o emprego de técnicas de ML sobressai como uma abordagem promissora para a obtenção de previsões precisas e confiáveis da eficiência dos módulos. A previsão da eficiência usando ML difere dos métodos tradicionais de previsão de eficiência dos módulos.

Na primeira etapa deste estudo, foi realizada uma revisão da literatura recente de estudos da eficiência dos módulos, que possibilitou demonstrar que os modelos de ML, em especial o de RL, têm sido implementados nos estudos da eficiência dos módulos. Em seguida, foi apresentado a metodologia com as principais tomadas de decisões. Entre as etapas garantidoras do desenvolvimento da metodologia deste trabalho, o primeiro passo desenvolvido foi a coleta dos dados de η , G , V_w , T_{amb} e T_c . Antes de passar para a implementação do modelo propriamente dito, foi realizada a análise dos dados utilizados e se eles apresentavam incoerências nos seus valores. Com isso, foi possível atestar a qualidade dos dados coletados, demonstrando que os mesmos estão dentro dos padrões e seguem o comportamento esperado se comparado a dados meteorológicos regionais.

Por fim, a terceira etapa consistiu na aplicação dos dados coletados nos anos de 2020 a 2022 à três modelos distintos, sendo eles RNA, SVM e RL. Os resultados obtidos de RMSE para o conjunto de dados de treinamento e teste do modelo RNA foi de 1,9976 e 1,8648, respectivamente. Para o mesmo conjunto de dados do modelo SVM, os resultados obtidos de RMSE foram de 2,7824 e 2,6412. E o parâmetro RMSE para o conjunto de dados de treinamento e teste do modelo RL foi de 3,3495 e 3,2831. A análise comparativa do modelo de melhor desempenho (o RNA) com um modelo tradicional (NOCT) contido na literatura, resultou em um valor de RMSE igual a 7,3951 para o modelo NOCT. Isso significa que a magnitude média do erro na eficiência dos módulos é

de 7,3951%, representando cerca de quatro vezes o valor do modelo de RNA. Dessa forma, fica evidenciado um maior poder de previsão do modelo RNA em comparação com os demais, quando aplicado aos dados coletados. A elaboração desse modelo torna possível obter previsões mais precisas e confiáveis sobre a eficiência dos módulos FV, auxiliando na tomada de decisões e no planejamento adequado da geração de energia solar.

As principais indicações feitas para os trabalhos futuros são: a exploração de diferentes modelos de ML, além dos modelos explorados neste trabalho, como redes neurais profundas e comparar com o desempenhos obtidos. Outra indicação pertinente seria o aperfeiçoamento do modelo, com ajuste de hiperparâmetros e validação cruzada para otimizar o seu desempenho. Além dos dados ambientais (radiação solar, temperatura ambiente e velocidade do vento), eficiência dos módulos FV e temperatura dos módulos utilizados neste trabalho, é possível considerar a incorporação de outros dados, como informações sobre a inclinação e orientação dos módulos, sombreamento e sujidade. Essas informações podem melhorar a precisão das previsões e fornecer conclusões adicionais sobre a eficiência dos módulos FV.

REFERÊNCIAS

- ABSOLAR. **Panorama da solar fotovoltaica no Brasil e no mundo ABSOLAR**. 2023. Disponível em: <https://www.absolar.org.br/mercado/infografico/>. Acesso em: 21/05/2023.
- AHMADI, M.; BAGHBAN, A.; SADEGHZADEH, M.; ZAMEN, M.; MOSAVI, A.; SHAMSHIRBAND S KUMAR, R. M.-K. M.; MARAM, E.; UCHIMURA, T. T. Evaluation of electrical efficiency of photovoltaic thermal solar collector. **Mac. Sol.**, ELSEVIER, v. 14, n. 1, p. 4–9, 2020.
- AL-SHAHRI, O.; ISMAIL, F.; HANNAN, M.; MAL-SHETWI, A. L.; BEGUM, R.; AL-MUHCEN, N.; SOUJERI, E. Solar photovoltaic energy optimization methods, challenges and issues: A comprehensive review. **Tecnol. Sol. Energy**, ELSEVIER, v. 284, n. 2, p. 6–13, 2021.
- AMMOUS, A.; ALSAEDI, A.; ALAHMADI, A.; AMMOUS, K. Energy efficiency of low voltage direct current supplies including photovoltaic sources for future homes. **Photov. Sustai. Energy**, ELSEVIER, v. 21, n. 3, p. 5–10, 2022.
- AN, Q.; TANG, R.; SU, H.; ZHANG, J.; LI, X. Robust configuration and intelligent mppt control for building integrated photovoltaic system based on extreme learning machine. **Effici. Photov. Energy**, ELSEVIER, v. 40, n. 6, p. 12283–12300, 2021.
- ANYOHA, R. History of Artificial Intelligence **Science in the News**. 2020. Disponível em: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>. Acesso em: 07/12/2022.
- BONACCORSO, G. **Machine Learning Algorithms**. [S. l.]: Birmingham: Packt Publishing,, 2017.
- BOSCHETTI, A.; MASSARON, L. **Python Data Science Essentials**. [S. l.]: Birmingham: Packt Publishing,, 2016.
- BRANCO, H. **Overfitting e Underfitting em Machine Learning**. ABRACD, 2022. Disponível em: <https://abracd.org/overfitting-e-underfitting-em-machine-learning/>. Acesso em: 25/01/2023.
- BURATTI, Y.; EIJKENS, C.; HAMEIRI, Z. Optimization of solar cell production lines using neural networks and genetic algorithms. **Monit. Photov. Indust**, ELSEVIER, v. 3, n. 11, p. 10317–10322, 2020.
- CHAOUCH, H.; CHARFEDDINE, S.; AOUN, S.; HLEIVA, V. J. Multiscale monitoring using machine learning methods: New methodology and an industrial application to a photovoltaic system. **Optim. Ratio. Predic**, ELSEVIER, v. 10, n. 6, p. 7–16, 2020.
- CHINNAMGARI, S. K. R. **Machine Learning Projects: Implement supervised, unsupervised, and reinforcement learning techniques using R 3.5**. [S. l.]: Birmingham: Packt Publishing,, 2019.
- DASGUPTA, N. **Practical Big Data Analytics**. [S. l.]: Birmingham: Packt Publishing,, 2018.

DATASCIENCEACADEMY. **Linguagem Python Para Data Science**. 2022. Disponível em: <https://www.datascienceacademy.com.br/bundle/formacao-linguagem-python-para-data-science>. Acesso em: 09/11/2022.

DHANALAKSHMI, S.; NANDINI, P.; RAKSHIT, S.; RAWAT, P.; NARAYANAMORTHY, R.; KUMAR, R.; SENTHIL, R. Fiber bragg grating sensor-based temperature monitoring of solar photovoltaic panels using machine learning algorithms. **Sol. Fiber. Energy**, ELSEVIER, v. 69, n. 4, p. 9–14, 2022.

ENAP. **Inteligência Artificial no Contexto do Serviço Público ENAP**. 2020. Disponível em: <https://www.escolavirtual.gov.br/curso/377>. Acesso em: 04/06/2023.

EPE. **Painel 2050 do Plano Nacional de Energia EPE**. 2023. Disponível em: <http://shinyepe.brazilsouth.cloudapp.azure.com/pne/shiny.html>. Acesso em: 03/04/2023.

EPE. **Relatório Online do Anuário Estatístico de Energia Elétrica EPE**. 2023. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 03/04/2023.

FARAYOLA, A.; HASAN, A.; ALI, A. Efficient photovoltaic mppt system using coarse gaussian support vector machine and artificial neural network techniques. **PV. Neural. Efficiency**, ELSEVIER, v. 14, n. 1, p. 323–339, 2018.

FERLITO, S.; VITO, S. D.; FRANCA, G. D. Detect anomalies in photovoltaic systems using isolation forest (preliminary results). **Forest. Photov. Detect**, ELSEVIER, v. 753, n. 13, p. 231–238, 2021.

GHANNAM, R.; KLAINE, P.; IMRAN, M. Artificial intelligence for photovoltaic systems. **Artificial. Systems. Sol**, ELSEVIER, v. 3, n. 2, p. 121–142, 2019.

GOLIATT, L.; YASEEN, Z. Development of a hybrid computational intelligent model for daily global solar radiation prediction. **Sol. Energy. Models**, ELSEVIER, v. 212, n. 7, p. 12–35, 2023.

HABA, C.; G. Monitoring solar panels using machine learning techniques. **Moder. Power. Systems**, ELSEVIER, v. 2, n. 1, p. 11–18, 2019.

IMAI, H.; BÓRIO, L.; SILVA, A. da; BERTO, L.; YAMAGUCHI, N.; HEREK, R.; L ezende. Computational simulation as an optimization tool for photovoltaic solar energy generation | simulação computacional como ferramenta de otimização na geração de energia solar fotovoltaica. **Solar. Energy. Systems**, URBE, v. 12, n. 2, p. 6–10, 2020.

K, K.; RASHID, S.; MANSOOR, M.; KHAN, A.; RAZA, H.; ZAFAR, M.; AKHTAR, N. Data-driven green energy extraction: Machine learning-based mppt control with efficient fault detection method for the hybrid pv-teg system. **Green. Hybrid. Energy**, ELSEVIER, v. 9, n. 15, p. 3604–3623, 2023.

KHILAR, R.; SUBA, G.; KUMAR, T.; SAMSON; ISAAC, J.; SHINDE, S.; RAMYA, S.; PRABHU, V.; ERKO, K. Improving the efficiency of photovoltaic panels using machine learning approach. **Smart. Solar. Energy**, ELSEVIER, v. 20, n. 3, p. 5–12, 2022.

KLÖTER, B. Application of machine learning for production optimization. **Conv. Photov. Energy**, IEEE, v. 112, n. 9, p. 3489–3489, 2018.

LIU, W.; SHEN, Y.; AUNGKULANON, P.; GHALANDARI, M.; LE, B.; ALVIZ-MEZA, A.; CÁRDENAS-ESCROCIA, Y. Machine learning applications for photovoltaic system optimization in zero green energy buildings. **Nano. Sol. Energy**, ENERGY REPORTS, v. 9, n. 7, p. 2787–2793, 2023.

ONU. **Agenda 2030 ONU**. 2019. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 27/05/2023.

POTHANA, S.; GUPTA, N. Optimizing the performance of solar photovoltaic panels using machine learning. **Optimi. Photov. Energy**, IEEE, v. 2, n. 1, p. 2–9, 2021.

PRASAD, K.; ISAAC, J. S.; PONSUDHA, P.; NITHYA, N.; SHINDE, S.; GOPAL, S.; SAROJWAL, A.; KARTHIKUMAR, K.; HADISH, K. A machine learning-based novel energy optimization algorithm in a photovoltaic solar power system. **Novel. Power. Sol**, ELSEVIER, v. 21, n. 6, p. 5–17, 2022.

PRIYADHARSINI, K.; KUMAR, J. D.; BABU, C. G.; SRIKANTH, A.; SOUNDAR, V.; SENTHAMILSELVAN, M. Elegant method to improve the efficiency of remotely located solar panels using iot. **Clean. Energy. Improve**, ELSEVIER, v. 45, n. 10, p. 8094–8104, 2021.

RAO, S.; KATOCH, S.; NARAYANASWAMY, V.; MUNIRAJU, G.; TEPEDELENLIOGLU, C.; ATURAGA, P. S.; AYYANAR, R.; SRINIVASAN, D. Machine learning for solar array monitoring, optimization, and control. **Clean. Energy. Improve**, ELSEVIER, v. 7, n. 15, p. 1–91, 2020.

SILVA, A.; BARCELOS, S.; SANTOS, B. Study cases of consumers migration from captive energy market to free energy market. **Market. Energy. Free**, SBSE, v. 1, n. 1, p. 1–6, 2020.

SOHANI, A.; SAYYAADI, H.; CORNARO, C.; SHAHVERDIAN, M.; PIERRO, M.; MOSER, D.; KARIMI, N.; DORANEHGARD, M.; LI, L. Using machine learning in photovoltaics to create smarter and cleaner energy generation systems: A comprehensive review. **Tecnol. Smart. Cleaner**, ELSEVIER, v. 364, n. 19, p. 2–18, 2022.

SOUZA, T. de; RJÚNIOR, C. L. Dirt on photovoltaic modules and efficient energy generation in the brazilian semiarid | sujidade em módulos fotovoltaicos e eficiência na geração de energia no semiárido brasileiro. **Module. Power. Dirt**, ELSEVIER, v. 26, n. 5, p. 321–326, 2022.

SPANIAS, A. Solar energy management as an internet of things (iot) application. **Sol. IoT. Applic**, ELSEVIER, v. 2, n. 6, p. 1–4, 2018.

TAKRURI, M.; FARHAT, M.; SUNIL, S.; RAMOS-HERNANZ, J.; BARAMBONES, O. Support vector machine for photovoltaic system efficiency improvement. **Improv. Renew. Energy**, ELSEVIER, v. 8, n. 3, p. 441–451, 2020.

WANG, N.; LI, J.; HO, S.; QIU, C. Distributed machine learning for energy trading in electric distribution system of the future. **Tecnol. Power. Comput**, SBSE, v. 8, n. 3, p. 441–451, 2020.

ZHANG, K.; ZOU, G. Photovoltaic output prediction method based on weather forecast and machine learning. **Forecast. Predic. Output**, SBSE, v. 4, n. 19, p. 2–10, 2022.