



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ANDRE WAGNER DE BARROS SILVA

PREVISÃO DE GERAÇÃO DE USINA SOLAR FOTOVOLTAICA UTILIZANDO
REDE NEURAL ARTIFICIAL E ALGORITMO PSO

FORTALEZA

2023

ANDRE WAGNER DE BARROS SILVA

PREVISÃO DE GERAÇÃO DE USINA SOLAR FOTOVOLTAICA UTILIZANDO REDE
NEURAL ARTIFICIAL E ALGORITMO PSO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica do Centro de Ciências e Tecnologia da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de concentração: Energias Renováveis e Sistemas de Energia Elétrica

Orientador: Profª. Ph.D. Ruth Pastôra Saraiva Leão.

Coorientador: Prof. Dr. Erick Costa Bezerra

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)

S578p Silva, Andre Wagner de Barros.
Previsão de Geração de Usina Solar Fotovoltaica Utilizando Rede Neural Artificial e Algoritmo PSO /
Andre Wagner de Barros Silva. – 2023.
82 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Tecnologia, Programa de Pós-
Graduação em Engenharia Elétrica, Fortaleza, 2023.

Orientação: Profa. Dra. Ruth Pastôra Saraiva Leão.

Coorientação: Prof. Dr. Erick Costa Bezerra.

1. Previsão. 2. Rede Neural. 3. Algoritmos de Enxame. 4. Geração Fotovoltaica. 5. Aprendizado de
Máquina. I. Título.

CDD 621.3

ANDRE WAGNER DE BARROS SILVA

PREVISÃO DE GERAÇÃO DE USINA SOLAR FOTOVOLTAICA UTILIZANDO REDE
NEURAL ARTIFICIAL E ALGORITMO PSO

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Engenharia Elétrica. Área de concentração: Energias Renováveis e Sistemas de Energia Elétrica.

Aprovada em:

BANCA EXAMINADORA

Prof^a. Ph.D. Ruth Pastôra Saraiva Leão (Orientador)
Universidade Federal do Ceará (UFC)

Dr. Erick Costa Bezerra
Sidia Instituto de Ciência e Tecnologia (Coorientador)

Prof. Dr. Arthur Plínio de Souza Braga
Universidade Federal do Ceará (UFC)

Dr. Felipe Rocha de Araújo
Sidia Instituto de Ciência e Tecnologia

Prof^a. Dra Anna Diva Plasencia Lotufo
UNESP – Campus Ilha Solteira

A Deus.

Minha família, pois tudo o que conquistei na vida devo a eles.

AGRADECIMENTOS

À minha orientadora, Prof^ª. Ph.D. Ruth Pastôra Saraiva Leão e ao meu co-orientador, Eng. Dr. Erick Costa Bezerra por todo o suporte e excelente orientação dada durante as atividades desenvolvidas no mestrado.

Aos gestores da usina solar fotovoltaica de 164 MW pela disponibilização dos dados para o estudo.

À minha companheira, Wiglla Nobre, por ser a parceira de sempre me ajudando em todos os desafios que já se fizeram presentes na minha vida.

Aos meus pais, Altamir Barros e Vania Maria, à minha irmã, Ana Valéria, ao meu sobrinho, Antonio Victor e ao meu cunhado Maikel Rodrigues, por estarem sempre me acompanhando de perto e me auxiliando em tudo o que necessito.

Aos colegas de trabalho da Diretoria de Infraestrutura da UFCA, por todo o apoio dado ao longo da minha pesquisa. Em especial à arquiteta Igliane Teles, que sempre me estimulou e se dispôs a ajudar no que fosse necessário.

À banca examinadora, pelas contribuições dadas a este trabalho, bem como pela disponibilidade em avaliá-lo.

“Em seu coração, a engenharia é sobre usar a ciência para encontrar soluções criativas e práticas. É uma profissão nobre”.

(Rainha Elizabeth II)

RESUMO

Considerando a natureza intermitente e a crescente inserção da geração solar fotovoltaica na matriz energética mundial nos últimos anos, é imperioso o desenvolvimento de modelos de previsão de geração cada vez mais precisos, de modo a permitir um melhor planejamento da operação de usinas fotovoltaicas e do sistema elétrico como um todo. As redes neurais artificiais têm se tornado muito populares por apresentarem resultados promissores devido a assertividade na previsão da geração fotovoltaica e desempenho robusto do modelo. A principal contribuição deste trabalho está na implementação e comparação de modelos de previsão horária de geração fotovoltaica de uma usina de 164 MWp, utilizando redes do tipo *Focused Time Delay Neural Network* (FTDNN). *Backpropagation*, Adam, otimização por enxame de partículas (do inglês, PSO), PSO Caótico (do inglês, CPSO) e PSO com fator de envelhecimento e enfraquecimento (do inglês, PSO-AWF) foram testados no treinamento da rede, enquanto o último algoritmo também foi usado na otimização dos parâmetros da arquitetura da rede FTDNN. Para fins de comparação de desempenho, foram usados os modelos de referência regressão por perceptron multicamadas, regressão linear, regressão por árvore de decisão e persistência. Com base em diferentes métricas estatísticas de desempenho, o modelo FTDNN com técnica de treinamento PSO-AWF obteve o melhor resultado dentre os algoritmos com ajuste manual de parâmetros, com raiz do erro quadrático médio (do inglês, RMSE) 18,354 MW, erro médio absoluto (do inglês, MAE) 13,784 MW, coeficiente de correlação de Pearson (R) 80,042%, raiz do erro quadrático médio normalizado (do inglês, NRMSE) 14,155% e erro médio absoluto normalizado (do inglês, NMAE) 10,631%. Dentre os modelos com ajuste automático de parâmetros e previsão horária (para 1h à frente), a rede FTDNN que utiliza PSO-AWF para estruturação e Adam para treinamento obteve a melhor performance, com RMSE 18,542 MW, MAE 13,565 MW, R 79,631%, NRMSE 14,300% e NMAE 10,462%. Além da previsão para 1h à frente, implementou-se também o modelo com ajuste automático de parâmetros para previsão 3h e 6h à frente (com e sem resolução horária), além da análise do efeito da mudança da quantidade de dados de entrada e da técnica de validação cruzada sob um dos modelos automáticos aplicados. Observou-se uma pequena melhoria no resultado na previsão para 3h e 6h à frente com resolução horária, não sendo identificado ganho para os demais modelos testados com ajuste automático de parâmetros.

Palavras-chave: Previsão; Regressão; Rede Neural; Algoritmos de Enxame; Geração Fotovoltaica; Aprendizado de Máquina.

ABSTRACT

Having the intermittent character and the increasing insertion of solar photovoltaic generation (PV) in the global power plant in recent years, it is imperative to develop even more accurate forecasting models for generation, allowing better planning of the PV plant operation and the entire electrical system. Artificial neural networks have become very popular for presenting promising assertive results in predicting photovoltaic generation and robust model performance. The main contribution of this work is the implementation and comparison of photovoltaic generation hourly forecasting models for a 164 MWp power plant, using types of Focused Time-Delay Neural Networks (FTDNN). Backpropagation, Adam Optimization, Particle Swarm Optimization (PSO), Chaotic PSO (CPSO), and PSO with Aging and Weakening Factor (PSO-AWF) were tested during the network training, although PSO-AWF was also used in the optimization of the architecture parameters for the FTDNN network. For performance comparison purposes, the following reference models were used: multilayer perceptron regression, linear regression, decision tree regression and persistence. Based on different statistical performance metrics, the FTDNN model with PSO-AWF training technique obtained the best result between the algorithms with manual parameter adjustment, with Root Mean Square Error (RMSE) 18.354 MW, Mean Absolute Error (MAE) 13.784 MW, Pearson Correlation Coefficient (R) 80.042 %, Normalized Root Mean Square Error (NRMSE) 14.155%, and Normalized Mean Absolute Error (NMAE) 10.631%. Among the models with automatic adjustment of parameters and forecast for 1h ahead, the FTDNN network that applies PSO-AWF for structuring and Adam for training performed better, with RMSE 18,542 MW, MAE 13,565 MW, R 79,631%, NRMSE 14,300% e NMAE 10,462%. In addition to the forecast for 1h ahead, models for forecasting 3h and 6h ahead (with and without hourly resolution) were also implemented, besides the analysis of the effect of changing the amount of input data and the cross-validation technique on a given automatic model. A small improvement in the result was observed, for the models that provide the forecast for 3h and 6h ahead with hourly resolution, with no improvement being identified for the other models tested with automatic adjustment of parameters.

Keywords: Prediction; Regression; Neural Networks; Swarm Algorithms; Photovoltaic Generation; Machine Learning.

LISTA DE FIGURAS

Figura 1 – Representação genérica dos componentes de um sistema solar fotovoltaico.....	17
Figura 2 – Taxa de crescimento da geração de eletricidade de 2021 a 2022, por tecnologia.....	19
Figura 3 – Previsão da capacidade instalada de energia solar FV no Brasil.....	20
Figura 4 – Curva para diferentes valores de radiação solar.....	22
Figura 5 – Curva para diferentes valores de temperatura ambiente.....	22
Figura 6 – Esquemático da vista superior do complexo fotovoltaico.....	33
Figura 7 – Gráfico de Caixa para temperatura ambiente e velocidade do vento.....	35
Figura 8 – Arquitetura genérica de uma rede FTDNN com uma camada oculta e uma saída.....	36
Figura 9 – Deslocamento de partículas em busca de um ponto mínimo global.....	39
Figura 10 – Arquitetura genérica de uma rede MLP com duas camadas ocultas.....	43
Figura 11 – Gráfico de dispersão e reta de regressão linear.....	44
Figura 12 – Representação de uma árvore de decisão.....	44
Figura 13 – Validação cruzada por amostragem aleatória.....	46
Figura 14 – Validação cruzada por k partições.....	47
Figura 15 – Validação cruzada por unidade.....	47
Figura 16 – Validação cruzada por janela expansiva.....	48
Figura 17 – Gráficos das funções de ativação comumente utilizadas para RNA.....	51
Figura 18 – NMAE por época de treinamento.....	54
Figura 19 – Gráfico de barras para as métricas NRMSE e NMAE.....	56
Figura 20 – Gráfico de barras para a métrica R.....	56
Figura 21 – Dados reais e previstos para o modelo FTDNN&ADAM.....	57
Figura 22 – RMSE por número de gerações.....	58
Figura 23 – Gráfico de barras para as métricas NRMSE e NMAE.....	60
Figura 24 – Gráfico de barras para a métrica R.....	60

Figura 25 – Dados reais e previstos para o modelo FTDNN&PSO-AWF (primeira iteração).....	61
Figura 26 – Dados reais e previstos para o modelo FTDNN&PSO-AWF (última iteração).....	61
Figura 27 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N).....	64
Figura 28 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão para 3h à frente.....	66
Figura 29 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão para 6h à frente.....	67
Figura 30 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 3h à frente (1h).....	69
Figura 31 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 3h à frente (3h).....	69
Figura 32 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 6h à frente (1h).....	70
Figura 33 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 6h à frente (6h).....	70
Figura 34 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão para 1h à frente na abordagem univariada.....	72
Figura 35 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa.....	74
Figura 36 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada.....	75

LISTA DE TABELAS

Tabela 1 – Dados meteorológicos medidos pelas WS.....	32
Tabela 2 – Distribuição das WS dentro do complexo fotovoltaico.....	32
Tabela 3 – Outliers (Z-score e Desvio Absoluto Mediano).....	34
Tabela 4 – Principais informações do conjunto de dados da abordagem multivariada.....	35
Tabela 5 – Funções de ativação comumente utilizadas para RNA.....	50
Tabela 6 – Melhores parâmetros para a FTDNN&BP e FTDNN&ADAM.....	53
Tabela 7 – Parâmetros da busca em grade para os modelos de referência.....	54
Tabela 8 – Resultados obtidos para a FTDNN&BP, FTDNN&ADAM e modelos de referência.....	55
Tabela 9 – Tempo de execução para a FTDNN&BP, FTDNN&ADAM e modelos de referência.....	55
Tabela 10 – Melhores parâmetros para a FTDNN&PSO, FTDNN&PSO-AWF e FTDNN&CPSO.....	57
Tabela 11 – Parâmetros para a rede FTDNN&PSO-AWF.....	58
Tabela 12 – Resultados obtidos para a FTDNN&PSO, FTDNN&PSO-AWF, FTDNN&CPSO e modelos de referência.....	59
Tabela 13 – Tempo de execução para a FTDNN&PSO, FTDNN&PSO-AWF, FTDNN&CPSO e modelos de referência.....	59
Tabela 14 – Parâmetros iniciais e melhores parâmetros definidos automaticamente.....	62
Tabela 15 – Parâmetros relacionados à técnica PSO-AWF nos modelos automatizados.....	63
Tabela 16 – Resultados obtidos para os modelos automatizados.....	63
Tabela 17 – Tempo de execução para os modelos automatizados.....	63
Tabela 18 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente.....	65
Tabela 19 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente.....	65
Tabela 20 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente.....	66

Tabela 21 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente.....	67
Tabela 22 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente.....	68
Tabela 23 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente.....	68
Tabela 24 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada.....	71
Tabela 25 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada.....	71
Tabela 26 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada.....	71
Tabela 27 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa.....	73
Tabela 28 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa.....	73
Tabela 29 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa.....	73
Tabela 30 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada.....	74
Tabela 31 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada.....	75
Tabela 32 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada.....	75
Tabela 33 – Resultados obtidos para todos os modelos implementados.....	76

LISTA DE ABREVIATURAS E SIGLAS

OCDE	Organização para Cooperação e Desenvolvimento Econômico
ANEEL	Agência Nacional de Energia Elétrica
RNA	Rede Neural Artificial
PSO	Otimização por Enxame de Partículas (<i>Particle Swarm Optimization</i>)
PSO-AWF	PSO com Fator de Envelhecimento e Enfraquecimento (<i>Particle Swarm Optimization with Aging and Weakening Factor</i>)
CPSO	PSO Caótico (<i>Chaotic Particle Swarm Optimization</i>)
MLP	Perceptron Multicamadas (<i>Multilayer Perceptron</i>)
RBF	Funções de Base Radial (<i>Radial Base Functions</i>)
RMSE	Raiz do Erro Quadrático Médio (<i>Root Mean Square Error</i>)
MAE	Erro Médio Absoluto (<i>Mean Absolute Error</i>)
NRMSE	Raiz do Erro Quadrático Médio Normalizado (<i>Normalized Root Mean Square Error</i>)
NMAE	Erro Médio Absoluto Normalizado (<i>Normalized Mean Absolute Error</i>)
R	Coefficiente de Correlação de Pearson
FTDNN	Rede Neural <i>Focused Time Delay Neural Network</i>
MLPR	Regressão por Perceptron Multicamadas (<i>Multilayer Perceptron Regressor</i>)
LR	Regressão Linear (<i>Linear Regressor</i>)
DTR	Regressão por Árvore de Decisão (<i>Decision Tree Regressor</i>)

LISTA DE SÍMBOLOS

$v_i(t)$	Velocidade Inicial da Partícula
$v_i(t + 1)$	Velocidade Nova da Partícula
$x_i(t)$	Posição Inicial da Partícula
$x_i(t + 1)$	Posição Nova da Partícula
$Pbest_i(t)$	Melhor Posição Individual da Partícula
$Gbest(t)$	Melhor Posição Global Dentre Todas as Partículas

SUMÁRIO

1 INTRODUÇÃO.....	17
1.1 Geração Fotovoltaica no Brasil e no Mundo.....	18
1.2 Influência de Fatores Meteorológicos na Geração Fotovoltaica.....	21
1.3 Motivação.....	23
1.4 Metodologia.....	24
1.5 Objetivos.....	24
1.6 Organização do trabalho.....	25
2 REVISÃO BIBLIOGRÁFICA.....	26
3 COMPLEXO FOTOVOLTAICO E BASE DE DADOS UTILIZADA.....	31
3.1 Características do Complexo Fotovoltaico.....	31
3.2 Análise Exploratória dos Dados.....	33
4 MODELOS DE PREVISÃO E MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO.....	36
4.1 Rede Neural Artificial tipo FTDNN.....	36
4.2 Adam e <i>Backpropagation</i>	37
4.3 PSO e suas Variações PSO-AWF e CPSO.....	38
4.4 Modelos de Referência Adotados.....	42
4.5 Validação Cruzada.....	45
4.6 Métricas de Avaliação.....	48
4.7 Funções de Ativação para RNA.....	49
5 RESULTADOS E ANÁLISE DE DESEMPENHO.....	52
5.1 Estudo de Caso 1: FTDNN com Técnica de Treinamento BP e Adam.....	53
5.2 Estudo de Caso 2: FTDNN com Técnica de Treinamento PSO, PSO-AWF e CPSO.....	57
5.3 Estudo de Caso 3: FTDNN com Ajuste Automático de Parâmetros.....	62
5.4 Estudo de Caso 4: FTDNN com Ajuste Automático de Parâmetros e Previsão de Potência para 3h e 6h à frente.....	64
5.5 Estudo de Caso 5: FTDNN com Ajuste Automático de Parâmetros e Abordagem Univariada, Multivariada Completa e com Validação Cruzada.....	70
6 CONCLUSÃO.....	78
REFERÊNCIAS.....	79

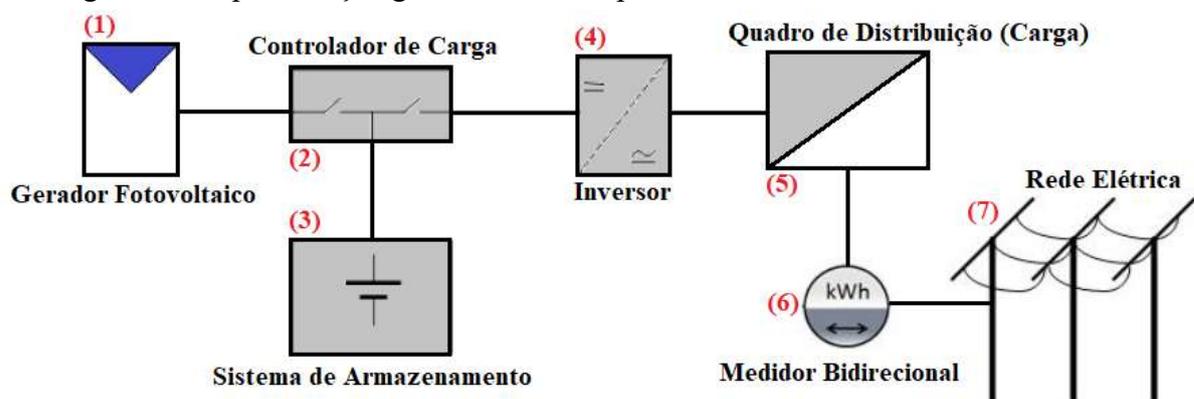
1 INTRODUÇÃO

O aproveitamento da energia gerada pelo sol, a qual é livre e renovável, tanto como fonte de calor quanto de luz, é hoje uma das alternativas mais promissoras para prover eletricidade. A necessidade de preservação do meio ambiente e de recursos energéticos menos poluentes tem alavancado significativamente a utilização de tecnologias de conversão de energia sem emissão de gases de efeito estufa, dentre as quais encontra-se a energia solar fotovoltaica.

Um sistema solar fotovoltaico (FV) é um sistema que absorve a luz do sol e converte energia solar em eletricidade. Basicamente, existem dois tipos de sistemas FV: *on-grid* e *off-grid*. Os sistemas *on-grid* operam conectados à rede elétrica e possuem a vantagem da não dependência exclusiva da conversão fotovoltaica para alimentar as cargas, uma vez que a rede elétrica também é responsável pelo suprimento das cargas. Nesse tipo de sistema existe a possibilidade de fluxo bidirecional de energia, ou seja, caso a energia convertida pelo sistema FV seja maior que a necessária para suprir as cargas, a quantidade excedente pode ser transmitida para a rede elétrica (Moreira, 2019).

Em sua configuração básica, um sistema conectado à rede (*on-grid*) é composto por gerador fotovoltaico, sistema de inversão de energia, quadro de distribuição, medidor bidirecional, carga elétrica e rede elétrica de distribuição. Quando a operação é isolada da rede, o sistema FV é do tipo isolado (*off-grid*), sendo o único responsável pela alimentação das cargas. Assim sendo, neste tipo de sistema faz-se necessário a utilização também de um banco de baterias e de um controlador de cargas. A Figura 01 apresenta, de maneira genérica, os componentes de um sistema FV. Sistemas isolados possuem os itens (1), (2), (3), (4) e (5), já sistemas conectados à rede possuem os itens (1), (4), (5), (6) e (7).

Figura 1 – Representação genérica dos componentes de um sistema solar fotovoltaico



Fonte: adaptado de CEPTEL-CRESESB (2014).

1.1 Geração Fotovoltaica no Brasil e no Mundo

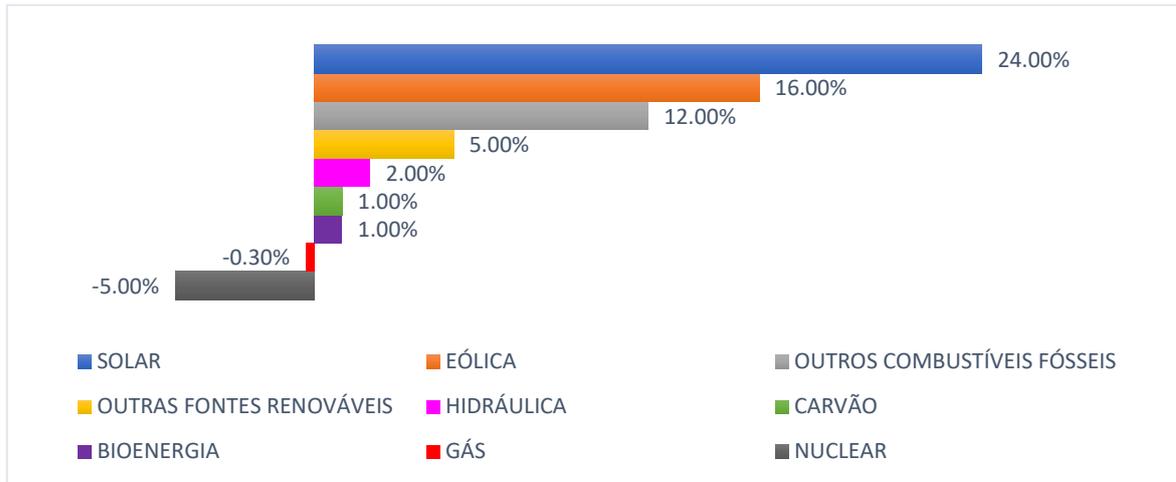
Em termos de capacidade instalada, a geração solar FV mundial apresentou crescimento considerável de 390,2 GW para 843,1 GW entre os anos de 2017 e 2021 (BP, 2022). Dentre os países com maior inserção da geração FV mundial estão a China, a Índia e países membros da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), dentre os quais merecem destaque Estados Unidos, Japão e Alemanha (IEA, 2021). A China, o maior mercado de energia solar FV do mundo, realizou a implantação de quase 100 GW de potência apenas em 2022, uma taxa de crescimento de 72% com relação aos 54,9 GW que haviam sido instalados em 2021 e já representavam um recorde histórico (Solar Power Europe, 2023).

Em 2022, a China sozinha era responsável por 34% da capacidade solar FV mundial, com 403 GW de potência instalada. Nesse mesmo ano, os Estados Unidos ocuparam confortavelmente a segunda posição com 141 GW, seguidos pelo Japão com 84,1 GW. Em adição a isso, o segundo ano forte consecutivo da Índia ajudou o país a passar a Alemanha e garantir o quarto lugar na posição mundial até o final de 2022. A capacidade instalada total da Índia atingiu 77,6 GW e sua participação no mercado global aumentou de 6% para 7%. A Alemanha caiu uma posição em 2022, com um inalterado 6% de participação e 68 GW de capacidade instalada (Solar Power Europe, 2023).

Os cinco países líderes no mercado de geração solar FV mundial ocupam essa posição desde uma longa data e é bastante improvável que essa situação mude em um curto intervalo de tempo. A Austrália, país que ocupa a sexta posição na classificação mundial, possui menos da metade da potência instalada da Alemanha. Levando-se em consideração os dez principais países, uma mudança notável entre os anos de 2021 e 2022 inclui a Espanha, a qual subiu duas posições, passando a ocupar o sétimo lugar com 27,4 GW de potência instalada e ultrapassando a Itália e a Coreia do Sul (Solar Power Europe, 2023).

O imenso potencial do mercado de energia solar mundial indica que o mesmo tende a alcançar uma maior taxa de crescimento anual do que qualquer outra forma de geração de energia na atualidade. Pela Figura 2, percebe-se que a geração solar despontou em termos de crescimento de fontes renováveis de energia, passando de 1040 TWh de geração em 2021 para 1289 TWh em 2022, um aumento de 24%. O vento é a segunda tecnologia que mais cresce quando se trata de geração global de energia, com 16% de taxa de crescimento. Enquanto isso, a eletricidade gerada pelo carvão, a fonte mais suja de geração de energia, cresceu apenas 1% (Solar Power Europe, 2023).

Figura 2 – Taxa de crescimento da geração de eletricidade de 2021 a 2022, por tecnologia



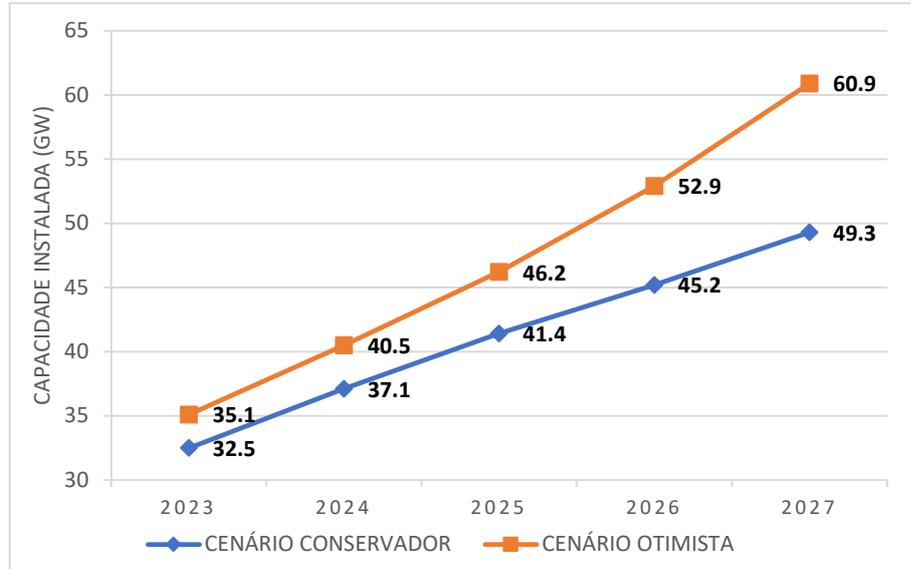
Fonte: adaptado de Solar Power Europe (2023).

O sucesso excepcional da energia solar pode ser atribuído a diversos fatores, mas o fator chave é a significativa redução de custo observado na última década, que impulsionou este tipo de geração a se tornar líder global em termos de competitividade. Apesar do aumento de custo pontual que ocorreu nos últimos dois anos na tecnologia solar, causado principalmente pela pandemia de Covid-19 e pela guerra da Ucrânia, mesmo nesse período esse tipo de energia permaneceu significativamente mais barata do que qualquer combustível fóssil e nuclear (Solar Power Europe, 2023). Associado a este fator, tem-se que os países que mais desenvolveram energia solar contaram com políticas de incentivo a esta tecnologia, incentivos estes que facilitaram a fabricação e a importação dos equipamentos do setor (Fischer, 2021).

No Brasil, observou-se um crescimento da potência total instalada de sistemas solares FV acima de 1000% no período de 2017 a 2021 (BP, 2022), sendo esse crescimento em grande parte devido à integração da micro e minigeração distribuída à rede elétrica. Entretanto, por causa de fatores como extensa área territorial e sistema elétrico interligado, o país também apresenta um grande potencial para aplicação de usinas de geração centralizada de médio e grande porte, de modo que em 2022 foram instalados 10.809 MW de potência de usina solar FV, a qual inclui 8.290 MW em geração distribuída e 2.519 MW em geração centralizada. No total, o país atingiu a marca de 24 GW de capacidade instalada nesse mesmo ano (Solar Power Europe, 2023).

A Associação Brasileira de Energia Solar Fotovoltaica (ABSOLAR) prevê um desempenho robusto para o mercado de energia solar nos próximos anos. De acordo com a Figura 3 e com base na ABSOLAR e nas projeções da Empresa de Pesquisas Energéticas (EPE), a energia solar FV no Brasil pode atingir entre 49.268 MW (cenário conservador) e 60.910 MW (cenário otimista) de capacidade instalada acumulada até 2027.

Figura 3 – Previsão da capacidade instalada de energia solar FV no Brasil



Fonte: adaptado de Solar Power Europe (2023).

Levando-se em conta fatores climáticos e geográficos, o Brasil apresenta altos níveis de irradiação e sua proximidade com a linha do Equador proporciona pouca variabilidade do recurso solar (Gobbo *et al.*, 2018). Assim sendo, ao longo dos últimos anos, a legislação brasileira aposta na diversificação da matriz de energia elétrica, com maior presença das fontes de geração renováveis, em detrimento das fontes fósseis e seus derivados. Nessa perspectiva, merecem destaque os leilões de energia organizados pela Agência Nacional de Energia Elétrica (ANEEL) e a criação da Resolução Normativa nº 482/2012, que estabeleceu as condições gerais para o acesso de sistemas de micro e minigeração ao sistema de distribuição de energia, bem como instituiu o sistema de compensação de energia elétrica (De Barros Silva *et al.*, 2021).

Além disso, algumas políticas e práticas foram criadas ao longo do tempo para incentivar a inserção de sistemas fotovoltaicos no Brasil (Fischer, 2021). Segundo Paixão (2019), dentre os programas criados, tem-se:

- Venda direta a consumidores;
- Descontos na Tarifa de Uso do Sistema de Distribuição (TUSD) e Tarifa de Uso do Sistema de Transmissão (TUST);
- Programa de apoio ao desenvolvimento tecnológico da indústria de semicondutores;
- Redução do imposto de renda;
- Condições diferenciadas de financiamentos pela Caixa Econômica Federal e pelo Banco Nacional de Desenvolvimento Econômico e Social (BNDES);
- Fundo clima;
- Inova Energia;

- Programas de fomento a projetos P&D;
- Criação do sistema de compensação de energia.

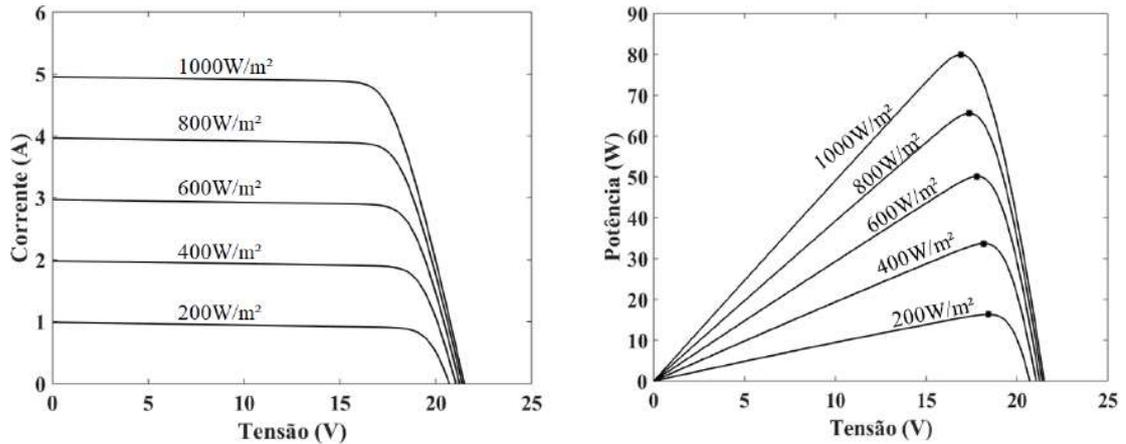
De todas as regiões brasileiras, a Nordeste é a que apresenta o maior potencial solar, tendo os valores médios diários totais da irradiação global horizontal e da componente direta normal na ordem de, respectivamente, 5,49 kWh/m² e 5,05 kWh/m² (Adece; Fiec; Sebrae, 2019). No contexto nordestino, 93% do território do Ceará estão no semiárido, que se caracteriza pela escassez e irregularidade pluviométrica, alta taxa de evaporação e elevada irradiância solar. Tais características são vantajosas para o uso da energia eólica e solar no estado (De Barros Silva *et al.*, 2021).

1.2 Influência de Fatores Meteorológicos na Geração Fotovoltaica

Diversos fatores influenciam na eficiência dos módulos FV e, conseqüentemente, na geração da usina solar que é composta pelos mesmos. De acordo com Francisco *et. al* (2019), a radiação solar possui relação direta com a geração de energia FV, sendo a variável que mais teve influência dentre todas as analisadas pelo autor. Para Ruviaro *et al.* (2018) e Silva (2018) outra grandeza meteorológica que interfere diretamente na geração solar fotovoltaica é a temperatura ambiente. Além disso, Glenn (2018) afirma que, além das variáveis já mencionadas, a umidade do ar e a velocidade do vento também possuem certa influência na geração de energia FV.

Levando-se em consideração as grandezas meteorológicas que interferem na geração de energia elétrica de sistemas solares FV, a radiação solar é o fator que mais influencia neste quesito, exercendo efeito direto tanto na tensão de circuito aberto quanto na corrente de curto circuito de um módulo FV (Francisco *et al.*, 2019). A Figura 4 mostra cinco curvas IV e PV simuladas para diferentes condições de radiação incidente sobre um módulo FV genérico. A partir da curva IV, percebe-se que o aumento da radiação aumenta significativamente a corrente de curto circuito e tem menor influência na tensão de circuito aberto, resultando em uma elevação da potência máxima. A curva PV evidencia o aumento da potência máxima a partir do aumento da radiação solar.

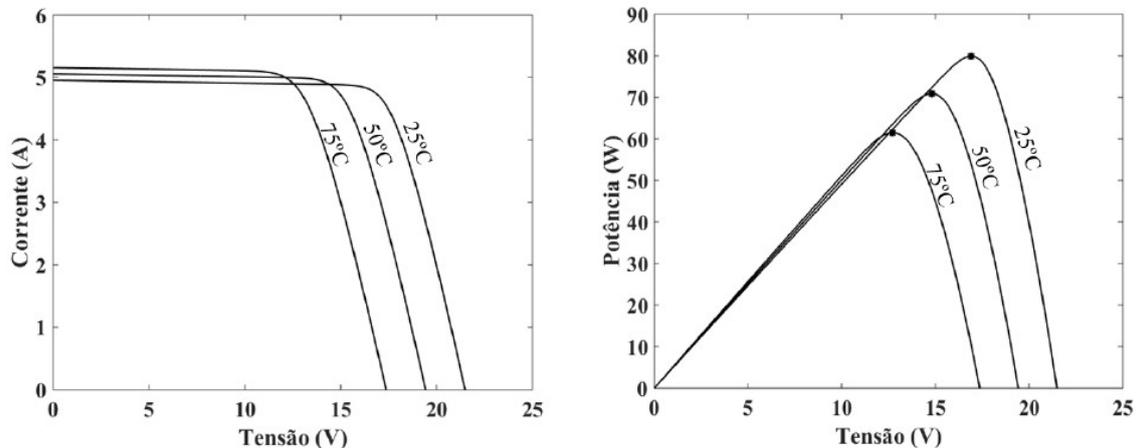
Figura 4 – Curva para diferentes valores de radiação solar



Fonte: (Silva, 2018).

A temperatura ambiente é outra grandeza que exerce influência na geração de energia de uma usina solar FV. De acordo com Ruviano *et al* (2018), todos os painéis solares perdem eficiência quando aquecem. Na Figura 5 são apresentadas três curvas características simuladas para a temperatura de 25°C, 50°C e 75°C, destacando a influência da temperatura nas curvas IV e PV. Pela curva IV, é possível se observar que a variação da temperatura modifica significativamente a tensão de circuito aberto, enquanto resulta em pequenas alterações na corrente de curto-circuito. Embora a corrente de curto-circuito aumente com a temperatura, esse aumento não é suficiente para compensar a queda na tensão de circuito aberto do módulo fotovoltaico e, conseqüentemente, a potência gerada diminui com o aumento da temperatura (curva PV).

Figura 5 – Curva para diferentes valores de temperatura ambiente



Fonte: (Silva, 2018).

Diante do exposto, tanto a radiação solar quanto a temperatura ambiente possuem grande relação com a geração de energia fotovoltaica. O aumento da temperatura interna do

módulo FV gera uma redução na saída do sistema de geração, necessitando, dessa maneira, de uma radiação maior para suprir a perda de eletricidade devido ao aumento da temperatura (Silva, 2018).

Outro fator meteorológico que influencia, ainda que indiretamente, no desempenho dos sistemas FV é a velocidade do vento. O vento apresenta uma relação positiva na geração de energia, ou seja, quanto maior a velocidade do vento, maior será a energia FV gerada (Francisco *et al.*, 2019). Esta eficiência na produção de energia pode ser explicada devido às altas velocidades do vento resultar no resfriamento dos módulos e, por consequência, no aumento da potência gerada. De acordo com Glenn (2019), o efeito de resfriamento causado pela velocidade do vento sobre módulos fotovoltaicos faz diferença na geração, de modo que painéis resfriados em 1°C são 0,05% mais eficientes, e essa porcentagem aumenta com o tempo.

Por fim, a umidade relativa do ar também gera certo impacto sobre a geração de energia de uma usina solar FV. Quando, em dias ensolarados e com alta umidade do ar, a luz solar atinge as partículas de água podem ocorrer três situações: refração, reflexão ou difração. Em qualquer uma das situações citadas, a quantidade de radiação que atinge os módulos será inferior à quantidade que atingiria sem as partículas de água, efeito causado pelo fato da umidade alterar a irradiação de forma não linear, gerando variações tanto na tensão de circuito aberto quanto na corrente de curto-circuito dos módulos fotovoltaicos e resultando, assim, na diminuição da potência gerada (Fischer, 2021). Além do efeito já citado, segundo Glenn (2019) a umidade do ar associada a climas quentes pode degradar os módulos fotovoltaicos ao longo de sua vida útil. Isso se aplica tanto para células de silício cristalino quanto para módulos de filme fino.

1.3 Motivação

Dado que a geração FV depende em grande parte da irradiância solar que é vulnerável a mudanças meteorológicas, como temperatura ambiente, velocidade do vento, precipitação, umidade e cobertura de nuvens, ela é por natureza estocástica. A alta volatilidade na produção FV cria desafios para a integração de sistemas de geração FV à rede elétrica, desta maneira, uma previsão segura de geração FV para dado horizonte futuro no tempo tende a diminuir incertezas.

O Operador do Sistema de Transmissão (*Transmission System Operator* - TSO) e o Operador do Sistema de Distribuição (*Distribution System Operator* - DSO) são agentes responsáveis por gerenciar a segurança do sistema de potência em tempo real e coordenar o fornecimento e a demanda de eletricidade, de forma a evitar flutuações na frequência ou interrupções no fornecimento. Eles são os responsáveis finais por toda a rede. O desbalanço

entre oferta e demanda de energia causa violações de restrições técnicas e penalidades econômicas aos atores da rede elétrica.

Existem muitos modelos e soluções que fornecem previsões de consumo, de preço, de manutenção, sendo a previsão de geração de significativa relevância para a segurança da rede e economicidade no uso de recursos energéticos. Neste trabalho serão investigados modelos de previsão de geração solar fotovoltaica de curto prazo usando rede neural artificial.

1.4 Metodologia

A metodologia aplicada neste trabalho propõe o uso de uma rede neural artificial do tipo *Focused Time Delay Neural Network* (FTDNN) para previsão de geração de energia de curto prazo de um complexo solar FV com diferentes técnicas de treinamento, sendo os parâmetros da rede ajustados tanto de maneira manual quanto automática. De maneira sucinta, as etapas executadas compreendem:

- Implementar, verificar o desempenho e selecionar uma FTDNN a partir do ajuste manual dos parâmetros da mesma e utilizando-se diferentes técnicas de treinamento: *backpropagation*, Adam, PSO, PSO Caótico e PSO-AWF;
- Aplicar diferentes algoritmos de aprendizado de máquina já consolidados na literatura como referência para comparar desempenho com as FTDNN testadas;
- Utilizar o algoritmo PSO-AWF para ajustar automaticamente os parâmetros (número de neurônios da camada oculta e atraso) da FTDNN e a técnica Adam para treinar a rede;
- Utilizar o algoritmo PSO-AWF tanto para ajustar automaticamente os parâmetros (número de neurônios da camada oculta e atraso) da FTDNN quanto para treinar a rede;
- Alterar o horizonte de previsão, a quantidade de dados de entrada e aplicar a técnica validação cruzada nos modelos com ajuste automático de parâmetros para fins de comparação de resultados.

1.5 Objetivos

Este trabalho tem como objetivo geral desenvolver, implementar e avaliar o desempenho de dezoito modelos distintos para a previsão de geração de energia elétrica de um sistema solar FV para o horizonte de curto prazo.

Como objetivos específicos, tem-se:

- Analisar o comportamento de uma rede neural FTDNN, a qual utiliza como entrada uma série temporal com atrasos, com diferentes técnicas de treinamento (*backpropagation*, Adam, PSO, PSO Caótico e PSO-AWF);
- Realizar o ajuste dos parâmetros da rede FTDNN de maneira manual e automática, comparando-se os resultados;
- Comparar os resultados dos modelos que utilizam FTDNN com algoritmos de aprendizado de máquina já consolidados na literatura;
- Medir o desempenho dos modelos de previsão testados utilizando diferentes métricas estatísticas: raiz do erro quadrático médio (*Root Mean Square Error* – RMSE), erro médio absoluto (*Mean Absolute Error* - MAE), coeficiente de correlação de Pearson (R), raiz do erro quadrático médio normalizado (*Normalized Root Mean Square Error* –NRMSE) e erro médio absoluto normalizado (*Normalized Mean Absolute Error* - NMAE);
- Verificar o efeito da mudança do horizonte de previsão, da quantidade de dados de entrada e da técnica de validação cruzada nos resultados obtidos;

1.6 Organização do trabalho

O restante do texto está organizado em cinco capítulos adicionais. No capítulo 2 é feita uma revisão bibliográfica sobre o assunto abordado. O capítulo 3 detalha o complexo FV e o conjunto de dados utilizado e o capítulo 4 apresenta os modelos de previsão e as métricas de avaliação implementadas. No capítulo 5 os resultados são apresentados, analisados e discutidos e finalmente o capítulo 6 traz a conclusão do trabalho.

2 REVISÃO BIBLIOGRÁFICA

Os modelos de previsão de geração FV são classificados geralmente em três categorias: modelos físicos, estatísticos e híbridos. Os modelos físicos ou numéricos utilizam mecanismos matemático e físico que descrevem o comportamento da atmosfera para prever a geração de energia elétrica, já os modelos estatísticos incluem todos os métodos baseados em dados históricos de saída, possivelmente combinados com variáveis meteorológicas. Os modelos estatísticos abrangem tanto a modelagem estatística clássica quanto os modelos no campo da inteligência computacional. Modelos clássicos (p.ex. SARIMA, ARIMA, ARMA etc.), persistência e aprendizado de máquinas (p.ex. rede neural artificial, árvore de decisão, regressão linear, floresta aleatória etc.) são todos modelos estatísticos. Já o método híbrido é uma combinação de dois métodos diferentes, um físico e um estatístico, ou dois ou mais métodos estatísticos (Mayer; Gróf, 2021). Nesse contexto, as abordagens de inteligência computacional têm ganhado notoriedade nos últimos anos, pelo fato de possuírem potencial de produzir previsões mais precisas e de maneira rápida.

Diversos trabalhos na literatura têm utilizado redes neurais artificiais (RNA) e otimização por enxame de partículas (*Particle Swarm Optimization* – PSO) para previsão de geração de energia em diferentes horizontes de tempo. Em De Barros Silva *et al.* (2021) foi realizada a previsão horária de geração FV de duas plantas de microgeração (de 6,03 kWp e 7,37 kWp) instaladas na cidade de Fortaleza-CE. A metodologia aplicada fez uso de duas RNA: perceptron multicamadas (*Multilayer Perceptron* - MLP) e funções de base radial (*Radial Base Functions* - RBF). Os modelos foram treinados com dados horários de geração de energia coletados durante o ano de 2020 e a performance dos mesmos foi avaliada a partir do coeficiente de correlação (R), coeficiente de eficiência de Nash-Sutcliffe (NSE) e viés relativo (VR). Ambas RNA mostraram bom potencial para aprender o comportamento das plantas, com R acima de 80%, VR próximo a zero e NSE acima de 0,50.

Em Su *et al.* (2019) foram implementados dez modelos diferentes de aprendizado de máquina para previsão de geração de energia em uma planta solar FV de 460 MWp localizada no Reino Unido. Esses modelos incluem seis redes neurais recorrentes e quatro outros métodos de aprendizado de máquinas, os quais tiveram performances comparadas através da métrica de avaliação NRMSE. Além disso, implementou-se ainda um modelo híbrido a partir dos modelos que obtiveram os melhores resultados, o qual apresentou a melhor performance dentre todas as avaliadas com NRMSE anual de 6,74%.

Uma análise comparativa utilizando rede neural recorrente do tipo memória de longo e curto prazo (*Long Short Term Memory* - LSTM) é apresentada em Da Cunha (2021) para previsão horária de geração FV, com diferentes variáveis de entrada e diferentes atrasos. Os dados meteorológicos e de geração de energia utilizados foram coletados no período de um ano em uma central fotovoltaica de 54 kWp localizada em Santa Catarina. O modelo proposto, que prevê geração FV a partir apenas de dados de energia, obtém um resultado bem próximo do melhor resultado obtido com quatro variáveis meteorológicas de entrada. As métricas de erro consideradas foram RMSE e NRMSE, com valores 0,398 e 0,110 para o modelo com quatro variáveis e 0,419 e 0,116 para o modelo proposto, respectivamente.

Redes neurais FTDNN foram treinadas com os algoritmos *backpropagation* (BP), evolução diferencial (ED), PSO e PSO com fator de envelhecimento e enfraquecimento para previsão da velocidade do vento nos horizontes de 30min, 1h, 3h e 6h à frente em Bezerra *et al.* (2017). O PSO com fator de envelhecimento e enfraquecimento (*Particle Swarm Optimization with Aging and Weakening Factor* - PSO-AWF), usa o conceito de expectativa de vida modelada por estes dois fatores. O PSO-AWF foi também usado para parametrizar o número de neurônios ocultos e o número de atrasos da rede FTDNN. Os resultados com a rede FTDNN treinada com PSO-AWF e parâmetros da rede ajustados por tentativa e erro e de forma automática superaram aqueles obtidos com FTDNN treinada por BP, ED e PSO em todos os horizontes de previsão, e seu uso reduziu consideravelmente o tempo de treinamento.

Levando-se em consideração que a previsão de geração FV a curto prazo é de extrema importância para o controle e a operação dos sistemas elétricos de potência, em Guo *et al.* (2019) os autores propõem um método de aprendizado de máquina para previsão FV de curto prazo. A abordagem é construída a partir dos modelos XGBoost, LSTM e LightGBM, sendo a melhor performance obtida através da combinação das vantagens destes três modelos. A simulação é realizada utilizando o conjunto de dados de uma usina solar fotovoltaica real localizada na China, o qual inclui dados meteorológicos locais e de potência gerada em junho de 2018. A resolução dos dados é de 15 minutos, e o dia 30 de junho de 2018 foi selecionado aleatoriamente como a data para se prever a geração no horizonte de 0:00 a.m a 12:00 p.m. A partir dos resultados obtidos, o modelo criado pela combinação dos algoritmos XGBoost, LSTM e LightGBM apresentou melhor performance que os modelos tradicionais Máquina de Vetores de Suporte (*Support Vector Machine* - SVM) e ARIMA, com RMSE 0,1736; MAE 0,1034 e Erro Percentual Absoluto Médio (*Mean Absolute Percentage Error* – MAPE) 0,0341.

Em Pelisson (2021) o autor utiliza duas abordagens com o intuito de realizar a previsão de geração de energia de uma planta PV a curto prazo. Na primeira estratégia, a qual é baseada

em métodos individuais, realizou-se a comparação de cinco modelos de aprendizado de máquina, quais sejam MLP, regressão por vetores suporte (*Support Vector Regression* – SVR), k-vizinhos mais próximos (*K-nearest neighbors* – KNN), árvore de decisão e regressão linear. Os modelos foram treinados e validados por meio de diferentes configurações dos seus parâmetros e diferentes subconjuntos de atributos de entrada. A segunda estratégia, baseada em modelos de comitês de métodos, envolve a utilização em separado de modelos individuais e de comitês, e para os modelos de comitês foram usados os métodos *Stacking Ensemble* de Intervalos Rígidos e *Stacking Ensemble* de Intervalos Fuzzy Ponderados.

Para cada uma das estratégias foi utilizada uma base de dados distinta, sendo ambas coletadas de uma estação solarimétrica localizada no laboratório Fotovoltaica da Universidade Federal de Santa Catarina (UFSC), situado na cidade de Florianópolis. A análise dos resultados apontou a MLP e o SVR como os métodos que obtiveram os melhores desempenhos na estratégia baseada em métodos individuais, tendo como melhores parâmetros de avaliação RMSE 45,92 e 52,20, respectivamente, e R 0,99 para ambos os modelos. Já para a estratégia baseada em modelos de comitês de métodos, os melhores resultados foram obtidos por meio dos modelos *Stacking Ensemble* de Intervalos Rígidos e *Stacking Ensemble* de Intervalos Fuzzy Ponderados, com RMSE final 50,26 e 46,97, respectivamente.

Em Dantas (2020), o autor apresenta uma proposta de otimização para redes MLP utilizando duas categorias de algoritmos de otimização: algoritmos genéticos (AG) e PSO. O AG foi aplicado na otimização da arquitetura da MLP, determinando a quantidade de camadas ocultas da rede neural bem como a quantidade de neurônios nestas camadas. Já o PSO foi utilizado com a finalidade de otimizar os pesos sinápticos da rede MLP, tudo de forma automática. Para avaliar estes dois experimentos, utilizou-se os conjuntos de dados de problemas de classificação *Iris*, *Wine*, *Glass*, *Heart*, *Ecoli* e *Diabetes*, todos obtidos através do repositório de bases de aprendizado de máquina *UC Irvine* (UCI).

Os resultados demonstraram que as redes treinadas com PSO, quando comparadas com as redes treinadas com a técnica BP, apesar de não terem obtido as melhores acurácias de classificação para todas as bases mostraram-se uma boa alternativa para otimizar os pesos sinápticos da MLP, visto que nas bases *Glass* e *Ecoli* o PSO superou o BP em termos de desempenho, com desvio padrão de 3,9 e 4,1 para o cenário de 30% das instancias utilizadas no treinamento, respectivamente. A aplicação de AG na otimização da arquitetura das redes mostrou-se também eficiente, uma vez que conseguiu selecionar redes com menores complexidades do que as obtidas no primeiro experimento, sem perda de performance.

Um método de desenvolvimento de redes neurais artificiais para previsão de geração fotovoltaica foi implementado em Fischer (2021). O modelo aplicado utiliza rede neural do tipo MLP e quatro variáveis meteorológicas são utilizadas como entradas: radiação (W/m^2), temperatura ambiente ($^{\circ}C$), velocidade do vento (m/s) e umidade (%). A definição do número de neurônios da MLP foi realizada empiricamente, através da avaliação estatística dos resultados obtidos com o número de neurônios escolhido dentro de uma faixa estipulada (de 5 a 300). Como função de ativação utilizou-se a tangente hiperbólica e para treinamento dos dados foi aplicado o algoritmo de Levenberg-Marquadt, o qual realizou a atualização dos pesos sinápticos das redes. O conjunto de dados foi dividido em 70% para treinamento, 15% para validação e 15% para teste, sendo composto por dois grupos distintos de informações. O primeiro grupo de dados utilizados refere-se aos dados meteorológicos da estação do Instituto Nacional de Meteorologia (INMET). Já o segundo conjunto de informações corresponde aos dados de geração FV da usina de geração distribuída da Universidade Federal de Santa Maria. Todos os resultados obtidos foram satisfatórios e capazes de comprovar a capacidade de predição das redes neurais artificiais desenvolvidas, sendo obtido coeficiente de correlação acima de 0,9 para todos os casos em estudo.

Em Mukilan *et al.* (2022) é realizada a previsão do potencial solar FV de sistemas instalados em cobertas a partir de aprendizado de máquina. No trabalho, o autor utiliza um modelo de aprendizado profundo (*deep learning*) chamado *Restricted Boltzmann Machine* (RBM), sendo os resultados da simulação obtidos através do pacote R, ferramenta que possui várias bibliotecas para previsão do potencial solar de sistemas FV. Os dados de alimentação do modelo são divididos em treinamento e teste, sendo aplicado aos mesmos o método k-partições de validação cruzada (k foi feito igual a 5). Os resultados demonstraram que o método proposto atingiu uma alta taxa de acurácia (acima de 90%), sendo superior a outros modelos utilizados como referência no trabalho.

Andrade *et al* (2021) faz uso de rede neural do tipo LSTM e da ferramenta Facebook Prophet para previsão da produção de energia elétrica de uma usina solar fotovoltaica localizada na Índia. O conjunto de dados utilizado, intitulado “*Solar Power Generation Data*”, foi encontrado no site Kaggle, sendo composto por informações meteorológicas e de geração. Diversos hiperparâmetros foram ajustados para a rede LSTM, dentre os quais estão número de camadas escondidas: 3; função de ativação da camada de saída: Sigmoid; função de otimização: Adam e número de épocas de treinamento: 100. Para a Prophet, os parâmetros escolhidos foram sazonalidade anual: falso; sazonalidade diária: verdadeiro e intervalo de previsão: 15 minutos. Os resultados obtidos alcançaram precisão superiores a 90%, onde foi observado que o modelo

com LSTM obteve melhor performance com menor quantidade de dados. Por outro lado, mesmo com uma performance inferior, o Prophet teve uma precisão aceitável considerando o menor tempo de processamento.

Modelos de inteligência artificial (IA) são utilizados para previsão da geração de energia de um sistema solar FV na universidade de Yarmouk, Jordânia, em Alhmoud *et al* (2022). Os modelos implementados possuem dois tipos de atributos: baseados em energia e baseados em tempo. Os atributos baseados em energia são: irradiação, temperatura ambiente e temperatura dos módulos, já os atributos baseados em tempo incluem dias, semanas, anos e horas do dia. Toda essa informação é agregada em um conjunto de dados e utilizada para treinar diferentes modelos de aprendizado de máquina. Aplicou-se ainda uma técnica de validação cruzada com dez partições, na qual em um procedimento sequencial escolhe-se 10% do conjunto de dados para teste e 90% para treinamento (as outras nove partições). Desta maneira, garante-se que todo o conjunto de dados estará envolvido nas etapas de treinamento e teste, reduzindo o risco de ocorrência de vieses nos resultados. Dentre todos os modelos implementados, as abordagens baseadas em árvores superaram as demais, tendo a floresta aleatória os melhores resultados com RMSE 172,07 MWh e MAE 68,70 MWh.

O diferencial deste trabalho está na utilização de uma rede neural FTDNN, a qual é de fácil implementação, com séries temporais e diferentes técnicas de treinamento para previsão da potência gerada em um complexo FV no horizonte de curto prazo. O ajuste dos parâmetros das redes implementadas foi feito tanto de maneira manual quanto automática, e os resultados desmontaram que, a depender do modelo, diferentes valores foram obtidos com relação a métricas de avaliação e ao tempo de execução. Além disso, para fins de comparação de resultados, implementou-se ainda quatro modelos de referência, quais sejam: regressão por perceptron multicamadas, regressão linear, regressão por árvore de decisão e persistência. Escolheu-se como referência quatro modelos amplamente utilizados na literatura e com filosofias distintas, de modo a se avaliar diferentes cenários.

3 COMPLEXO FOTOVOLTAICO E BASE DE DADOS UTILIZADA

A pesquisa realizada neste trabalho pode ser caracterizada como quantitativa com obtenção de dados coletados por instrumentos de medição e análise de dados por métodos estatísticos, e de natureza aplicada, sendo dirigida à solução de problemas de previsão de geração de energia elétrica. A seguir será descrito o complexo fotovoltaico de onde foram obtidos os dados de entrada para previsão da geração de energia elétrica.

3.1 Características do Complexo Fotovoltaico

O complexo fotovoltaico em estudo está localizado no interior do Ceará e é subdividido em quatro usinas fotovoltaicas, com nove estações de transformação (*Transformer Stations - TS*) em cada usina, numeradas de acordo com o número da usina e TS (como exemplo 1.1, TS 1 da usina 1), perfazendo um total de 36 TS (TS 1.1 a 4.9). Todas as estações de transformação possuem quatro inversores, com exceção das TS 1.5, 2.5, 3.5 e 4.5 que possuem apenas dois. Cada inversor é conectado a cinco estruturas denominadas *string combine box* (SCB) a qual é alimentada por doze circuitos duplos. Um circuito duplo é composto por dois circuitos simples conectados em paralelo, sendo cada circuito simples constituído por 31 módulos fotovoltaicos de 325 W. Desta maneira, todo o complexo é composto por 505.920 módulos de 325 W, totalizando 164,424 MW de potência instalada.

Ao longo das quatro usinas solares que compõem o complexo existem treze estações meteorológicas (*Weather Stations - WS*) as quais medem e registram diferentes grandezas. Os dados meteorológicos são registrados com resolução horária, assim como acontece com os dados de potência registrados nos inversores das TS. A Tabela 1 indica as grandezas meteorológicas registradas em cada WS da planta fotovoltaica e, para fins de estudo neste trabalho, foram utilizados como dados de entrada a potência elétrica gerada pela usina bem como quatro variáveis meteorológicas medidas pela WS01: irradiância horizontal, temperatura ambiente, umidade e velocidade do vento. A esta situação convencionou-se chamar neste trabalho de abordagem multivariada. Escolheu-se utilizar a WS01 devido a mesma ser a mais completa dentre as estações, ou seja, a que realiza a medição e registro da maior variedade de grandezas meteorológicas.

Entretanto, para fins de comparação, no capítulo 5 é apresentado também um estudo de previsão de geração utilizando-se apenas dados de geração (abordagem univariada) bem como dados de geração da planta e meteorológicos de todas as WS do complexo (abordagem multivariada completa). As quatro variáveis meteorológicas utilizadas na abordagem

multivariada deste trabalho foram escolhidas tendo em vista a influência das mesmas sobre a geração de energia de uma usina FV, conforme mencionado no capítulo 1. Além disso, analisou-se também a correlação entre os atributos de entrada e a variável alvo (potência gerada), sendo a grandeza irradiância horizontal a que apresentou correlação mais forte, de 82,76%, seguida pela temperatura ambiente com 51,42%. A velocidade do vento apresentou correlação, também positiva, de 37,03% e, por fim, a umidade do ar teve correlação negativa de 50,84%. Estes valores de correlação confirmam a influência das grandezas relacionadas na variável alvo do estudo, tendo em vista que evidenciam que a correlação entre as mesmas não é fraca.

Tabela 1 – Dados meteorológicos medidos pelas WS

Estação Meteorológica	Grandezas Medidas
WS01	Irradiância difusa, irradiância direta, irradiância global, temperatura ambiente, pressão barométrica, umidade, precipitação, irradiância horizontal, irradiância inclinada, direção do vento, velocidade do vento.
WS02 a WS13	Albedo, temperatura ambiente, temperatura dos módulos, irradiância horizontal, irradiância inclinada, direção do vento, velocidade do vento.

Fonte: o próprio autor.

A Tabela 2 mostra a distribuição das WS dentro do complexo FV, ou seja, em qual TS está localizada cada WS.

Tabela 2 – Distribuição das WS dentro do complexo fotovoltaico

Estação Meteorológica	Localização
WS 01	Próximo a TS 1.5
WS 02	TS 1.2
WS 03	TS 1.5
WS 04	TS 1.8
WS 05	TS 2.1
WS 06	TS 2.3
WS 07	TS 2.8
WS 08	TS 3.1
WS 09	TS 3.3
WS 10	TS 3.9
WS 11	TS 4.4
WS 12	TS 4.6
WS 13	TS 4.9

Fonte: o próprio autor.

Outra análise realizada foi a existência de anomalias (*outliers*), que são dados que desviam extremamente do comportamento bem definido dos demais presentes no conjunto de dados, podendo ou não ser útil removê-los. A identificação e o tratamento de *outliers* é bastante importante no contexto da preparação dos dados, tendo em vista que estes, se presentes nos dados de treinamento dos modelos, podem introduzir vieses indesejados nos algoritmos computacionais, distorcendo os resultados finais (Smiti, 2020).

Foram aplicadas as técnicas Z-score e Desvio Absoluto Mediano (Peres, 2021) ao conjunto de dados da abordagem multivariada. O primeiro método identificou quatro pontos para o atributo temperatura ambiente e dois pontos para a velocidade do vento, já o Desvio Absoluto Mediano identificou seis pontos para a temperatura ambiente e dois pontos para a velocidade do vento. A Tabela 3 resume datas, horários e valores identificados como *outliers* pelos métodos previamente citados.

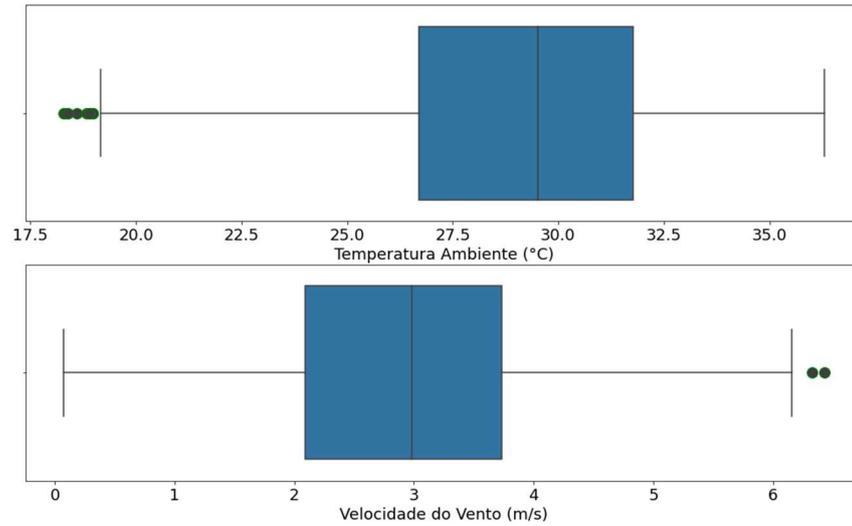
Tabela 3 – *Outliers* (Z-score e Desvio Absoluto Mediano)

Método	<i>Outliers</i>			
	Temp. Ambiente (Data e hora)	Temp. Ambiente (Valor)	Velocidade do Vento (Data e hora)	Velocidade do Vento (Valor)
Desvio Absoluto Mediano	13/07/2020 às 05:00h;	18,60 °C		
	14/07/2020 às 05:00h;	18,82 °C		
Z-score	22/07/2020 às 05:00h;	18,29 °C	03/08/2020 às 08:00h	6,33 m/s
	28/07/2020 às 05:00h;	18,35 °C	31/08/2020 às 11:00h	6,42 m/s
Desvio Absoluto Mediano	29/07/2020 às 05:00h;	18,30 °C		
	24/08/2020 às 05:00h	18,39 °C		

Fonte: o próprio autor.

Observa-se que todos os pontos identificados como *outliers* pelo método Z-score também o foram pelo método Desvio Absoluto Mediano, o que indica que estas duas metodologias estão apontando para resultados semelhantes. A Figura 7 apresenta o gráfico de caixa (*boxplot*) (Peres, 2021) para as grandezas temperatura ambiente e velocidade do vento. Percebe-se que a quantidade de pontos identificados por esta técnica também está em acordo com os resultados obtidos pelo método Desvio Absoluto Mediano.

Figura 7 – Gráfico de Caixa para temperatura ambiente e velocidade do vento



Fonte: o próprio autor.

A Tabela 4 resume as principais informações sobre os atributos do conjunto de dados da abordagem multivariada, quais sejam: média, desvio padrão, valor mínimo, valor máximo, quartis e número de amostras.

Tabela 4 – Principais informações do conjunto de dados da abordagem multivariada

Grandeza	Temp. Amb. (°C)	Umidade (%)	Irrad. Horiz. (W/m ²)	Vel. do vento (m/s)	Potência (MW)
Média	29,07	64,02	456,89	2,91	63,53
Desvio Padrão	3,55	18,15	325,33	1,11	40,43
Min	18,29	21,10	0,0447	0,0750	0
25%	26,69	49,63	143,31	2,09	24,56
50%	29,51	62,63	439,79	2,98	71,17
75%	31,77	77,60	748,20	3,73	99,61
Max	36,30	100	1191,27	6,42	129,66
Amostras	8866				

Fonte: o próprio autor.

A partir do exposto, tendo em vista a pequena quantidade de *outliers* identificados em comparação com a quantidade total de amostras e por seus valores próximos da média, optou-se por não os desconsiderar.

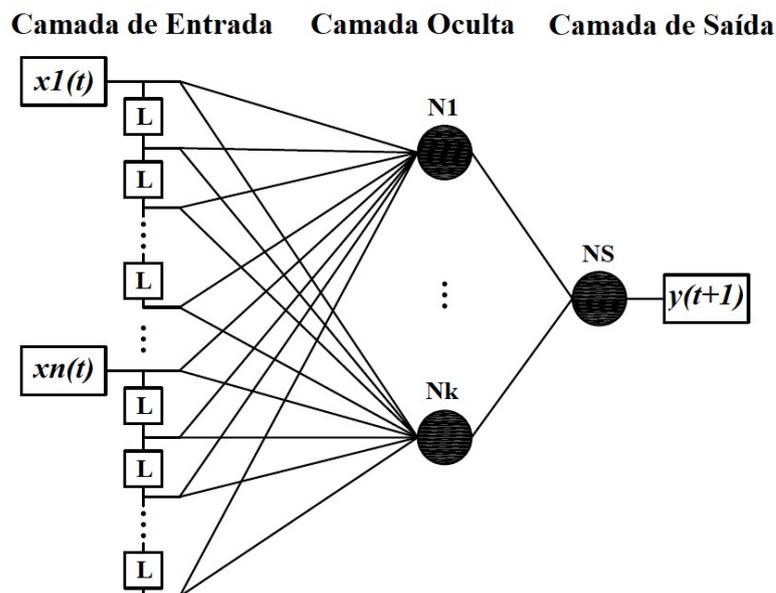
4 MODELOS DE PREVISÃO E MÉTRICAS DE AVALIAÇÃO DE DESEMPENHO

Os modelos implementados neste trabalho usam RNA tipo FTDNN com diferentes técnicas de treinamento, além de algoritmos de referência para fins de comparação de resultados. Os dezoito modelos, cujos resultados encontram-se sumarizados na Tabela 33, fazem uso de dados meteorológicos e de potência para previsão da potência gerada pela usina FV em determinado instante de tempo. O computador utilizado trata-se de um Dell modelo Precision 5820, 32 GB de memória RAM, disco rígido de 1TB e processador Intel Xeon W-2133.

4.1 Rede Neural Artificial tipo FTDNN

A rede neural artificial utilizada é uma MLP com atrasos colocados na entrada a fim de torná-la dinâmica, a qual é denominada FTDNN (Príncipe; Euliano; Lefebvre, 1999). Em redes dinâmicas, os dados de saída não dependem apenas dos dados de entrada atuais, mas também dos dados de entrada anteriores. A FTDNN é uma rede multicamadas, sem realimentação, na qual os pesos sinápticos podem ser ajustados por diferentes técnicas. A topologia FTDNN tem sido aplicada com sucesso em identificação de sistemas não lineares, previsão de séries temporais e reconhecimento de padrões (Príncipe; Euliano; Lefebvre, 1999). A Figura 8 apresenta a arquitetura de uma rede neural FTDNN, sendo $x1(t) \dots xn(t)$ os elementos de entrada no instante t , L o atraso considerado, $N1 \dots Nk$ o número de neurônios da camada oculta, NS o neurônio da camada de saída e $y(t+1)$ o dado de saída da rede neural no instante $t+1$.

Figura 8 – Arquitetura genérica de uma rede FTDNN com uma camada oculta e uma saída



Fonte: o próprio autor.

Como uma típica RNA do tipo *feed-forward*, o sinal passa pela FTDNN em apenas um sentido e direção, da entrada para a saída, e os neurônios de uma mesma camada não são conectados entre si (Bezerra, 2022). A saída de uma rede FTDNN é dada pela Equação 1 (Engelbrecht, 2007).

$$y_j = f_{y_j} \left(\sum_{k=1}^{K+1} w_{jk} f_{h_k} \left(\sum_{n=1}^N \sum_{t=0}^L v_{k,n(t)} x_n(t) + z_{N+1} v_{k,N+1} \right) \right) \quad (1)$$

em que L é o atraso, K é o número de neurônios da camada oculta, N é o número de entradas, y_j é a saída da rede, f_{y_j} e f_{h_k} são as funções de ativação para a camada de saída e a camada oculta, respectivamente, w_{jk} é o valor do peso entre a camada de saída e a camada oculta, $v_{k,n(t)}$ é o valor do peso entre a camada oculta e a camada de entrada e $x_n(t)$ é a entrada da RNA. O termo $z_{N+1} v_{k,N+1}$ leva em consideração o *bias* da rede neural.

4.2 Adam e *Backpropagation*

Conforme mencionado no capítulo 1, diferentes técnicas de treinamento para a FTDNN foram implementadas neste trabalho, sendo os dois primeiros algoritmos aplicados o Adam e o BP. O algoritmo Adam trata-se de um método de otimização estocástica baseado em descida de gradiente, o qual foi apresentado por Kingma e Ba (2014) e baseia-se nos métodos *AdaGrad* e *RMSProp* (Bouzon, 2021). Esta técnica mostra-se efetiva em base de dados de alta dimensionalidade e que possuem distribuições não-estacionárias, ou seja, que possuem média e variância que mudam ao longo do tempo (Bouzon, 2021).

Já o BP é um algoritmo utilizado para treinar RNA através da retro-propagação dos erros da camada de saída para as camadas intermediárias. Os pesos da rede são otimizados de modo que ela aprenda como mapear corretamente as entradas para as saídas. O método foi apresentado por Rumelhart *et al.* (1986) e possui as seguintes etapas (Bouzon, 2021):

- *Feed-Forward*: É feita a leitura das entradas, que se propagam pela rede até a camada de saída;
- *Feed-Backward*: Calcula-se o erro de saída da rede, que é propagado da camada de saída até a primeira camada escondida, atualizando os pesos das conexões entre os neurônios.

4.3 PSO e suas Variações PSO-AWF e CPSO

As demais técnicas de treinamento aplicadas neste trabalho foram o algoritmo PSO e suas variações PSO-AWF e CPSO. O PSO é um método de otimização inspirado no comportamento social cooperativo e competitivo de animais que se organizam em bandos e cardumes (Engelbrecht, 2007). Esta técnica foi criada por Kennedy e Eberhart (1995), e mantém uma nuvem de partículas que fluem através de um espaço multidimensional pré-definido, de modo que cada partícula é ajustada de acordo com sua própria experiência e a de seus vizinhos.

Em um espaço de busca multidimensional, inicialmente cada partícula é posicionada de maneira aleatória, e suas respectivas posições iniciais $x_i(t)$ são definidas como seus *Pbests*, que podem ser descritos como as suas melhores posições individuais até o momento. A melhor solução dentre todas as partículas é o *Gbest* da nuvem. Usualmente a velocidade inicial $v_i(t)$ de cada partícula é definida como zero e a nova velocidade $v_i(t + 1)$ que orienta todo o processo de otimização é descrita conforme Equação 2.

$$v_i(t + 1) = v_i(t) + C_1 R_1 (Pbest_i(t) - x_i(t)) + C_2 R_2 (Gbest(t) - x_i(t)) \quad (2)$$

em que C_1 e C_2 são respectivamente as constantes sociais e cognitivas, enquanto R_1 e R_2 são valores aleatórios entre $[0,1]$. A nova posição $x_i(t + 1)$ é então computada pela Equação 3.

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \quad (3)$$

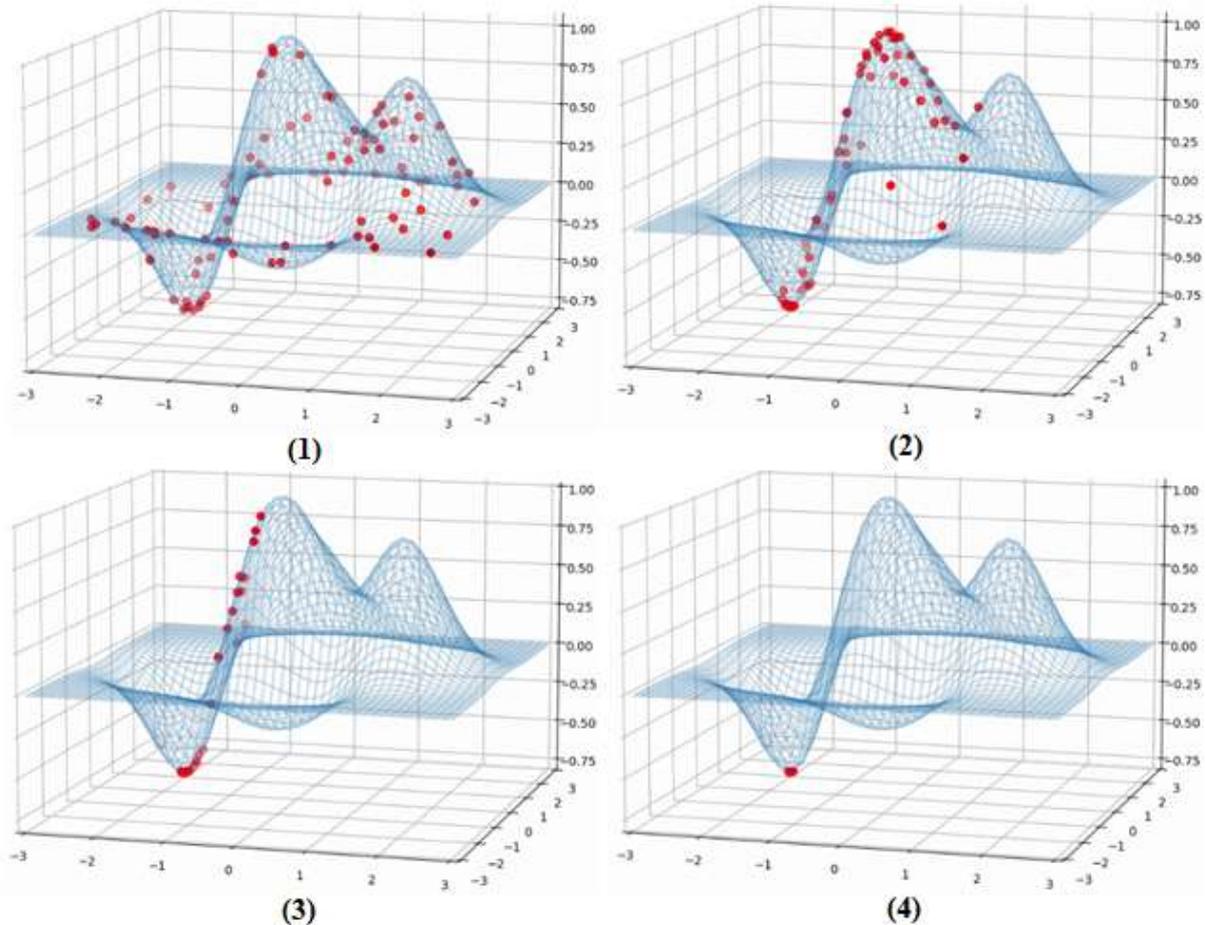
Todas as partículas da nuvem tendem a se deslocar para o ponto ótimo. Durante a busca no espaço multidimensional, caso o *Gbest* não seja a solução ótima, uma das partículas pode encontrar uma solução melhor e tomar o lugar do *Gbest*. Esse ciclo continua até que a solução ótima global ou uma restrição de parada do algoritmo seja encontrada (Engelbrecht, 2007). O algoritmo do PSO é descrito nos passos que seguem.

- Inicialização: A população (quantidade de partículas) e o número máximo de gerações (iterações) são definidos como números inteiros pelo usuário. As posições iniciais ($x_i(t)$) de todas as partículas são geradas randomicamente dentro de um espaço de busca pré-definido, enquanto as velocidades iniciais ($v_i(t)$) são inicializadas em zero. A melhor partícula do enxame é selecionada como líder (*Gbest*);
- Atualização da velocidade e posição: Todas as partículas seguem a regra de atualização de velocidade da Equação 2 e de posição da Equação 3 para ajustar suas velocidades e posições;

- Atualização do $Pbest$ e $Gbest$: Para todas as partículas se a posição recém gerada ($x_i(t)$) é melhor que o seu $Pbest_i$, então ($x_i(t)$) assume o lugar do $Pbest_i$. Adicionalmente, se a melhor posição encontrada nesta iteração é melhor que o $Gbest$, então o $Gbest$ é atualizado para esta nova posição;
- Condição de parada: Se o número de iterações é maior que o máximo pré-definido, o algoritmo para, caso contrário, ele volta para o passo 2 para uma nova iteração.

A Figura 9 apresenta quatro etapas do deslocamento de partículas através da técnica PSO em um ambiente tridimensional, as quais se deslocam buscando um ponto de mínimo global. Em termos computacionais, esta técnica requer apenas operadores matemáticos primitivos, apresentando baixo custo computacional tanto em termos de memória quanto de velocidade (Kennedy; Eberhart, 1995).

Figura 9 – Deslocamento de partículas em busca de um ponto mínimo global



Fonte: (GBHAT, 2021).

O algoritmo PSO-AWF trata-se de uma variação do PSO que, além da adequação dos indivíduos ser determinada pelo valor da função objetivo (*Fitness*), é atribuído um valor de

idade para as partículas da nuvem, o que o faz ser um algoritmo evolutivo. Assim, caso a idade da partícula (*Particle Age* - PA) chegue ao fim, ou seja, atinja o valor de expectativa de vida (*Life Expectancy* - LE) previamente definido, novas partículas emergem para desafiar e tomar a liderança da nuvem. Desta maneira, o mecanismo de envelhecimento e enfraquecimento prevê oportunidades para outras partículas liderarem o enxame e assim trazer diversidade, minorando o risco de um ótimo local (Bezerra *et al.*, 2017).

Duas variáveis compõem a idade de uma partícula no PSO-AWF, o fator de envelhecimento (*Aging Factor* – AF) e o fator de enfraquecimento (*Weakening Factor* – WF). O WF é utilizado toda vez que uma partícula não obtém um novo resultado melhor que o anterior, assim sendo, mesmo partículas com AF iguais possuirão diferentes vidas úteis e deixarão de existir em diferentes momentos. O algoritmo do PSO-AWF é descrito nos passos que seguem.

- Inicialização: Os fatores LE, AF e WF, bem como a população e o número máximo de gerações (iterações) são definidos como números inteiros pelo usuário. As posições iniciais ($x_i(t)$) de todas as partículas são geradas randomicamente dentro de um espaço de busca pré-definido, enquanto as velocidades iniciais ($v_i(t)$) e as idades das partículas (PA) inicializadas em zero. A melhor partícula do enxame é selecionada como líder (*Gbest*);
- Atualização da velocidade e posição: Todas as partículas seguem a regra de atualização de velocidade da Equação 2 e de posição da Equação 3 para ajustar suas velocidades e posições;
- Atualização do *Pbest* e *Gbest*: De maneira semelhante ao PSO convencional, para todas as partículas se a posição recém gerada ($x_i(t)$) é melhor que o seu *Pbest_i*, então ($x_i(t)$) assume o lugar do *Pbest_i*. Adicionalmente, se a melhor posição encontrada nesta iteração é melhor que o *Gbest*, então o *Gbest* é atualizado para esta nova posição;
- Controle da vida útil: Neste passo, a idade (PA) de cada uma das partículas é atualizada. O *PA_i* do *Gbest* é definido como zero, enquanto o restante do enxame obedece à seguinte regra:

Se *Fitness_i* é pior que o *Pbest_i*, **então**

$$PA_i(t) = PA_i(t - 1) + AF + WF \text{ senão}$$

$$PA_i(t) = PA_i(t - 1) + AF$$

Fim

Se a idade de uma partícula é exaurida, ou seja $PA_i \geq LE$, então a partícula é eliminada.

- Gerando desafios: Quando o número de gerações atinge o valor máximo pré-definido pelo usuário, um novo exame surgirá conforme descrito no passo 1, o qual será utilizado para desafiar o *Gbest*;
- Condição de parada: Se o número de iterações é maior que o máximo pré-definido, o algoritmo para, caso contrário, ele volta para o passo 2 para uma nova iteração.

No dicionário, a palavra caos significa um estado de completa desordem e confusão. Em ciência, entretanto, não existe uma definição padrão para caos, o qual pode ser considerado como determinístico e capaz de ser representado por regras e equações. No caos, o futuro é completamente determinado pelo passado e, além disso, uma pequena mudança pode levar a um resultado bastante diferente, situação chamada de “efeito borboleta” (Hefny; Azab, 2010). As principais características do caos são:

- Determinístico: os sistemas seguem uma regra determinística;
- Não linear: as regras são não lineares;
- Irregular: o comportamento dos sistemas mostra irregularidade;
- Sensível às condições iniciais: pequenas mudanças no estado inicial de sistemas caóticos podem levar a grandes mudanças no resultado final.

No PSO caótico (*Chaotic Particle Swarm Optimization - CPSO*), além de ser avaliada a função objetivo, realiza-se uma busca caótica de modo a se tentar evitar mínimos locais. Entretanto, essa busca caótica normalmente necessita de um grande número de iterações para atingir um ótimo global e não é eficaz em grandes espaços de busca (Hefny; Azab, 2010). O algoritmo do CPSO é descrito nos passos que seguem.

- Inicialização: A população (quantidade de partículas) e o número máximo de gerações (iteraões) são definidos como números inteiros pelo usuário. As posições iniciais ($x_i(t)$) de todas as partículas são geradas randomicamente dentro de um espaço de busca pré-definido, enquanto as velocidades iniciais ($v_i(t)$) são inicializadas em zero. A melhor partícula do exame é selecionada como líder (*Gbest*);
- Atualização da velocidade e posição: Todas as partículas seguem a regra de atualização de velocidade da Equação 2 e de posição da Equação 3 para ajustar suas velocidades e posições;
- Atualização do *Pbest* e *Gbest*: Para todas as partículas se a posição recém gerada ($x_i(t)$) é melhor que o seu *Pbest_i*, então ($x_i(t)$) assume o lugar do *Pbest_i*. Adicionalmente, se

a melhor posição encontrada nesta iteração é melhor que o *Gbest*, então o *Gbest* é atualizado para esta nova posição;

- Inicialização da busca caótica local, com separação das $n/5$ melhores partículas, sendo n o número total (população);
- Criação de variáveis caóticas, as quais constituem novas partículas criadas a partir das $n/5$ melhores;
- Avaliação da nova solução;
- Caso a nova solução seja melhor que o *Gbest* já encontrado ou o número máximo de iterações seja alcançado, encerra-se a busca caótica local. Caso contrário, realiza-se os quatro passos que seguem;
- Redução do espaço de busca;
- Criação aleatória de $4n/5$ novas partículas dentro do espaço de busca reduzido;
- Criação da nova população, constituída pelas $n/5$ melhores mais as $4n/5$ partículas criadas aleatoriamente;
- Retorno ao passo 2 para uma nova iteração.

Neste trabalho os algoritmos PSO e CPSO foram utilizados para o treinamento da rede FTDNN (otimização dos pesos sinápticos), já o algoritmo PSO-AWF foi utilizado tanto para treinamento quanto para otimização dos parâmetros número de neurônios da camada escondida (NNHL) e atraso (LAG) da rede FTDNN.

4.4 Modelos de Referência Adotados

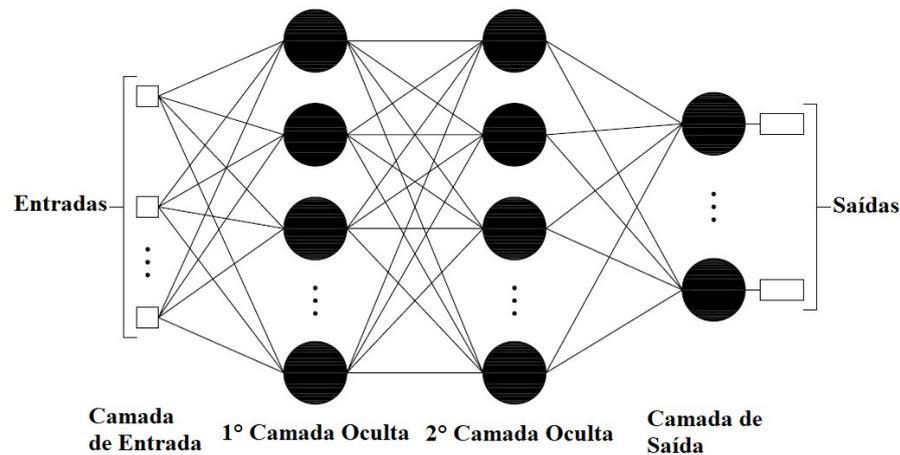
Para fins de referência, três modelos da biblioteca Scikit-learn para análise preditiva de dados desenvolvidas em Python foram utilizados. São eles regressão por perceptron multicamadas (*Multilayer Perceptron Regressor* - MLPR), regressão linear (*Linear Regressor* - LR) e regressão por árvore de decisão (*Decision Tree Regressor* - DTR).

A função MLPR, apesar de guardar certa semelhança com a rede FTDNN, difere da mesma por não utilizar os dados de entrada de maneira sequencial, ou seja, neste modelo, assim como acontece também com os demais modelos de referência implementados, a entrada não é aplicada na forma de uma série temporal com unidades de atraso. As redes MLP foram apresentadas por Rumelhart *et al.* (1986) e são caracterizadas pela existência de pelo menos uma camada intermediária (oculta) de neurônios entre a camada de entrada e a camada de saída. Elas são consideradas umas das arquiteturas mais versáteis quanto à aplicabilidade, possuindo aplicações nas mais variadas áreas de conhecimento, como reconhecimento de padrões,

previsão de séries temporais, otimização de sistemas, dentre outras (Silva; Spatti; Flauzino, 2016).

Assim como a rede FTDNN, a MLP também pertence à arquitetura *feed-forward*, de modo que o fluxo de informações na estrutura da rede se inicia na camada de entrada, percorre em seguida as camadas intermediárias (ocultas), sendo então finalizado na camada de saída. A Figura 10 ilustra a estrutura genérica de uma rede MLP com duas camadas ocultas.

Figura 10 – Arquitetura genérica de uma rede MLP com duas camadas ocultas

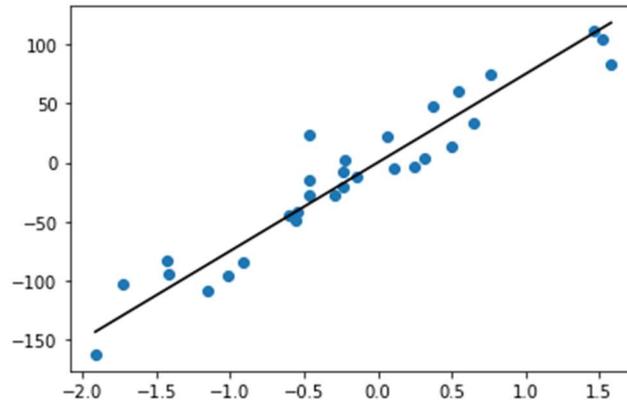


Fonte: adaptado de Silva, Spatti e Flauzino (2016).

O termo regressão surgiu em 1885 com o antropólogo, matemático e estatístico Francis Galton, num estudo onde ele demonstrou que a altura dos filhos não tende a refletir a altura dos pais (Damaceno, 2020). O modelo LR é bastante conhecido e amplamente aplicado no campo do aprendizado de máquina, tratando-se de um algoritmo que assume uma relação linear entre uma ou mais variáveis de entrada e uma única variável de saída. Em outras palavras, o valor de saída pode ser calculado a partir de uma combinação linear dos valores de entrada. Quando existe uma única variável de entrada o método é chamado *Simple Linear Regressor*, porém, quando há múltiplas variáveis de entrada é chamado *Multiple Linear Regressor* (Brownlee, 2020).

A variável que se deseja prever é chamada de variável dependente, enquanto a variável que é usada para prever o valor de outra é chamada de variável independente. O algoritmo aprende o melhor ajuste de uma linha reta entre as entradas (variáveis independentes) e as saídas (variáveis dependentes), reduzindo o erro estimado a partir dos valores reais e previstos da variável de saída (Alhmoud *et al.*, 2022). A Figura 11 demonstra um exemplo genérico de gráfico de dispersão com a reta de regressão linear dos pontos.

Figura 11 – Gráfico de dispersão e reta de regressão linear

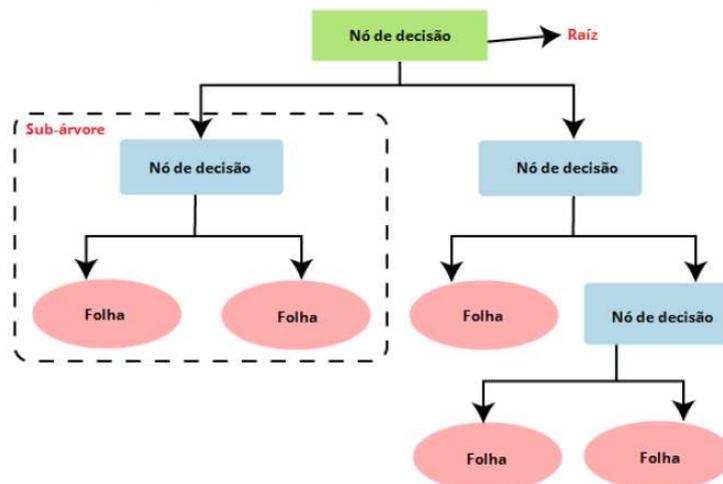


Fonte: o próprio autor

O primeiro algoritmo para gerar árvores de decisão foi apresentado por Quinlan (1986). A DTR é um modelo estatístico do campo do aprendizado de máquinas, no qual os dados são divididos obedecendo regras definidas. De acordo com Gama *et al.* (2004), os modelos DTR são funções de decisão que dividem os dados de acordo com suas características (atributos) e com determinados parâmetros (regras), de modo que um problema complexo seja dividido em subproblemas mais simples e, recursivamente, a mesma estratégia seja aplicada aos subproblemas até a solução do problema como um todo.

Em uma árvore de decisão (*Decision Tree - DT*), a raiz ou nó raiz é o começo da estrutura hierárquica da árvore, onde a primeira decisão será tomada e a árvore será dividida em subconjuntos (sub-árvores). O nó é a parte onde as variáveis são divididas obedecendo uma determinada regra e a folha é o último estágio, no qual é alcançada a resposta prevista. A Figura 12 demonstra uma representação simplificada de uma árvore de decisão, com uma raiz, três nós de decisão e cinco folhas.

Figura 12 – Representação de uma árvore de decisão



Fonte: adaptado de Arain (2021).

Além dos modelos já citados, a persistência também foi utilizada. Os modelos de previsão baseados em métodos persistentes são simples e, na maioria das vezes, utilizados como referência para outros modelos mais precisos e com custo computacional mais alto (Pelisson, 2021). A persistência assume que a geração futura no instante $(t+1)$ será igual à geração passada (no instante t), método de fácil implementação e bastante aplicado na literatura (Chu *et al.*, 2021).

Para as funções de referência da biblioteca Scikit-learn foi realizada a técnica da busca em grade (*grid Search*) na busca dos melhores parâmetros. Este método consiste numa busca exaustiva pela combinação ótima de hiperparâmetros, dentro de um espaço de combinações possíveis. Assim, avalia-se cada configuração possível do modelo, de acordo com os valores pré-definidos para os hiperparâmetros, selecionando-se ao final a combinação que apresentar o menor erro (Malik, 2020).

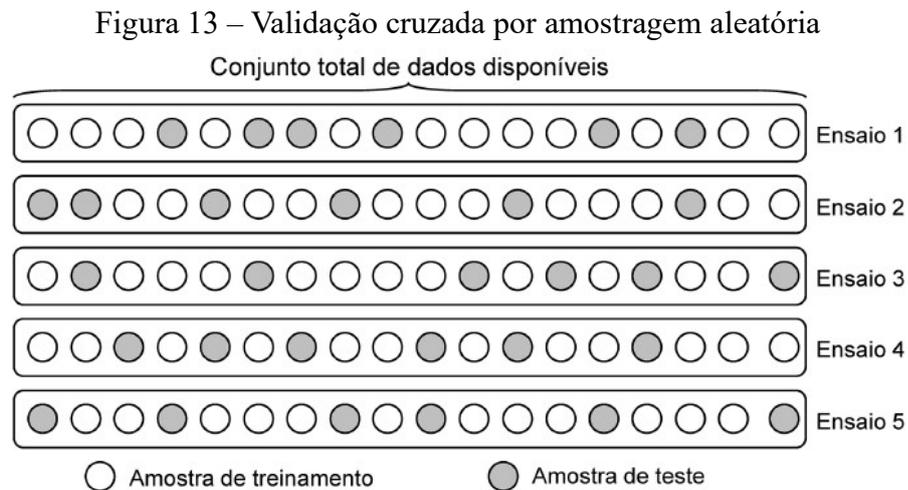
4.5 Validação Cruzada

Inúmeros métodos de séries temporais e técnicas de aprendizado de máquina exigem otimização de hiperparâmetros. Uma prática comum é utilizar validação cruzada (*Cross-Validation - CV*) para determinar a melhor escolha destes parâmetros. Geralmente, três métodos de validação cruzada são empregados no processo de seleção de hiperparâmetros de redes MLP: validação cruzada por amostragem aleatória (*random subsampling cross-validation*), validação cruzada por k-partições (*k-fold cross-validation*) e validação cruzada por unidade (*leave-one-out cross validation*) (Silva; Spatti; Flauzino, 2016).

Na validação cruzada por amostragem aleatória o conjunto total de dados é aleatoriamente dividido em duas partes, quais sejam subconjunto de treinamento e subconjunto de validação. O subconjunto de treinamento será utilizado para treinar todas as topologias de RNA candidatas a melhor rede, enquanto o subconjunto de validação é somente aplicado para selecionar a RNA com melhor performance. Uma vez que as amostras do subconjunto de validação não participam do processo de aprendizagem, torna-se então possível avaliar o desempenho da generalização proporcionada em cada uma das topologias candidatas. Normalmente, nesta metodologia a partir do conjunto de dados disponível cerca de 60 a 90% são aleatoriamente escolhidos para o subconjunto de treinamento, ficando o restante alocado no conjunto de validação. Este procedimento de partição deve ser repetido várias vezes durante o processo de aprendizagem das RNA candidatas, permitindo-se a possibilidade de contemplação de amostras diferentes tanto no subconjunto de treinamento quanto no de validação (Silva;

Spatti; Flauzino, 2016). A performance global de cada RNA candidata será então a média dos desempenhos individuais em cada experimento.

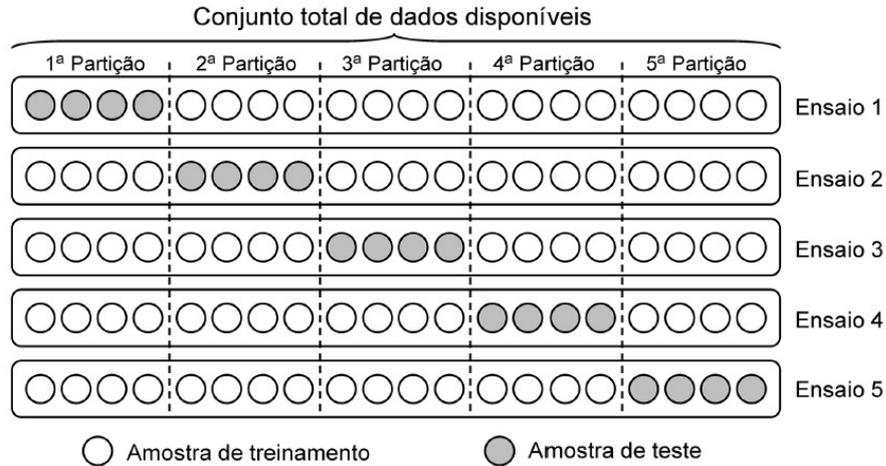
A Figura 13 ilustra a validação cruzada por amostragem aleatória. Para cada um dos cinco ensaios, escolhem-se aleatoriamente seis que pertencerão ao subconjunto de validação, ficando as demais (doze) alocadas no subconjunto de treinamento. Em cada ensaio realizado, o sorteio das amostras para o subconjunto de validação será sempre realizado considerando todas as amostras disponíveis.



Fonte: (Silva; Spatti; Flauzino, 2016).

Na validação cruzada por k -partições (k -fold) realiza-se a divisão do conjunto total de amostras em k partições, sendo $(k-1)$ usadas para compor o subconjunto de treinamento e a partição restante constituirá o subconjunto de validação. Assim sendo, o processo de aprendizado se repete k vezes até que todas as partições tenham sido utilizadas como subconjunto de validação. O valor de k está relacionado com a quantidade total de amostras disponíveis, sendo usualmente utilizado um número compreendido entre 5 e 10. A performance global de cada RNA candidata será também a média dos desempenhos individuais na aplicação das k partições. A Figura 14 sistematiza o mecanismo desta técnica para um total de 20 amostras, onde k foi feito igual a 5.

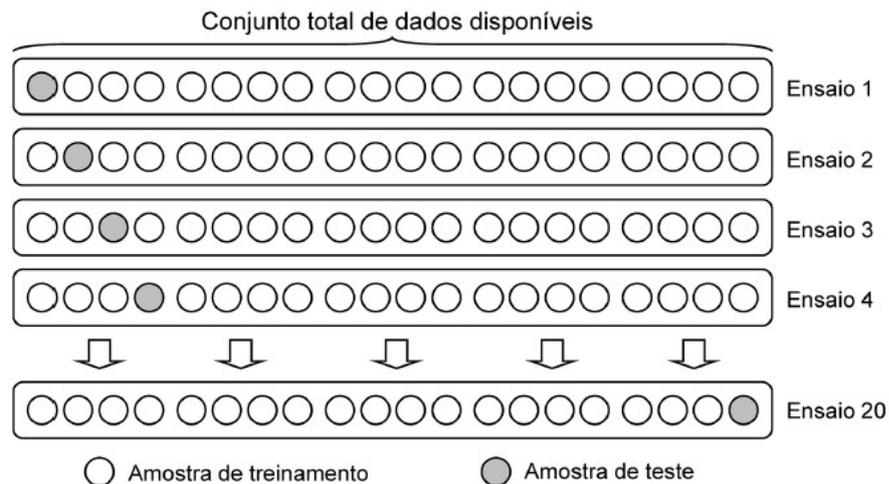
Figura 14 – Validação cruzada por k partições



Fonte: (Silva; Spatti; Flauzino, 2016).

Já a validação cruzada por unidade (*leave-one-out*) consiste em uma única amostra para o subconjunto de validação, sendo as demais alocadas para o subconjunto de treinamento. O processo de treinamento é então repetido até que todas as amostras sejam utilizadas como subconjunto de validação (Silva; Spatti; Flauzino, 2016). Esta técnica é na realidade um caso particular do método de k-partições, envolvendo, entretanto, um esforço computacional bem maior. A Figura 15 ilustra a validação cruzada por unidade para um total de 20 amostras disponíveis.

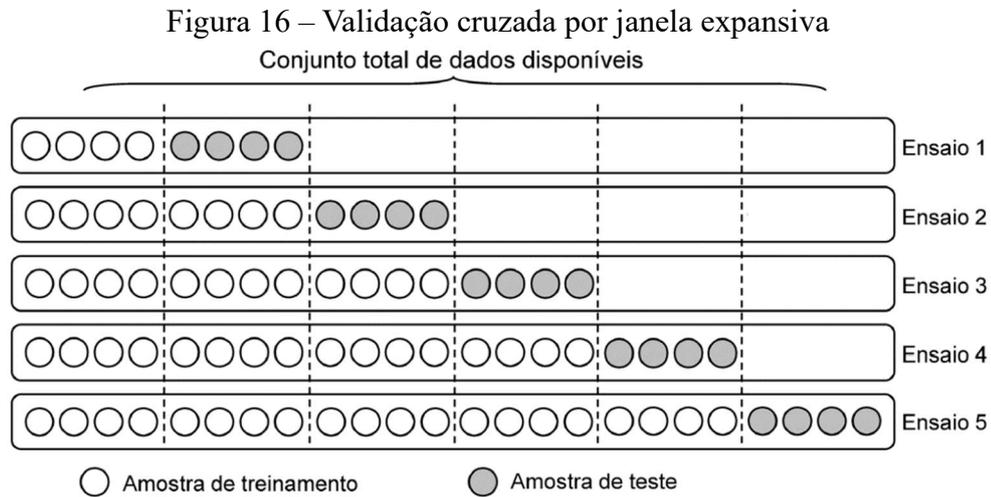
Figura 15 – Validação cruzada por unidade



Fonte: (Silva; Spatti; Flauzino, 2016).

Neste trabalho, utilizou-se uma técnica de validação cruzada para séries temporais chamada validação por janela expansiva. Esta técnica é semelhante à validação por k partições, entretanto, as partições são aplicadas de modo a expandir a janela de treino com o tempo e manter a de validação fixa, sempre respeitando a ordem cronológica dos dados, ou seja,

treinando no passado e validando no futuro (Mario Filho, 2023). A Figura 16 demonstra a validação cruzada por janela expansiva utilizando-se k igual a 5, valor adotado também neste trabalho tendo em vista que comumente o mesmo é escolhido entre 5 e 10 (Isaksson; Conde, 2018).



Fonte: adaptado de Silva, Spatti e Flauzino (2016).

4.6 Métricas de Avaliação

Para avaliação dos resultados, utilizou-se as métricas de desempenho RMSE, MAE, R, NRMSE e NMAE, tendo em vista que são indicadores bastante utilizados para estudar e compreender a precisão de dados previstos por meio de modelos computacionais (Botchkarev, 2019).

O MAE fornece informações úteis sobre os erros de previsão, uma vez que ele representa a soma das diferenças entre os valores previstos e reais dividido pelo número de observações. No RMSE os maiores erros são enfatizados, pois a diferença entre o valor real e o medido é elevada ao quadrado. O NRMSE e NMAE constituem a normalização das métricas RMSE e MAE calculadas, respectivamente, o que facilita a comparação dos resultados obtidos entre modelos com conjunto de dados em escalas diferentes (Da Cunha, 2021). As Equações 4, 5, 6 e 7 representam o cálculo matemático para se obter as métricas de erro MAE, RMSE, NRMSE e NMAE.

$$MAE = \frac{1}{N} \sum_{t=1}^N |\hat{y}(t) - y(t)| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{y}(t) - y(t))^2} \quad (5)$$

$$NRMSE = \frac{RMSE}{\max(y) - \min(y)} \quad (6)$$

$$NMAE = \frac{MAE}{\max(y) - \min(y)} \quad (7)$$

em que $y(t)$ é o valor real no instante t , $\hat{y}(t)$ é o valor previsto no instante t , $\max(y)$ e $\min(y)$ constituem o maior e o menor valor existente em toda a série temporal de valores medidos e N é o número total de amostras previstas.

O coeficiente de correlação de Pearson (R) constitui uma forma bastante comum de se medir uma correlação linear. É representado por um número entre -1 e 1, o qual mede a força e a direção entre duas variáveis quaisquer. Valores de R entre 0 e 1 representam uma correlação positiva (quando uma variável muda, a outra muda na mesma direção), valores entre 0 e -1 representam uma correlação negativa (quando uma variável muda, a outra muda na direção oposta) e R igual ou próximo a 0 indica que não existe correlação entre as variáveis. A Equação 8 apresenta a expressão matemática para a métrica R (De Barros Silva, 2021).

$$R = \frac{\sum_{i=1}^N (x_i - x_{med}) \cdot (y_i - y_{med})}{\sqrt{\sum_{i=1}^N (x_i - x_{med})^2 \cdot \sum_{i=1}^N (y_i - y_{med})^2}} \quad (8)$$

em que x_i é o valor real, x_{med} é o valor médio do conjunto de dados reais, y_i é o valor previsto, y_{med} é o valor médio do conjunto de dados previstos e N é o número total de amostras previstas.

4.7 Funções de Ativação para RNA

As funções de ativação são essenciais para dar capacidade representativa às redes neurais artificiais, introduzindo um componente de não linearidade, que faz com que elas possam aprender mais do que relações lineares entre as variáveis dependentes e independentes (Facure, 2017). De acordo com Silva, Spatti e Flauzino (2016), as funções de ativação são comumente divididas em dois tipos:

- Funções de ativação parcialmente diferenciáveis: aquelas que possuem pontos cujas derivadas de primeira ordem são inexistentes;
- Funções de ativação totalmente diferenciáveis: aquelas cujas derivadas de primeira ordem existem e são conhecidas em todos os pontos de seu domínio de definição.

Segundo Silva (2021), dentre as principais funções parcialmente diferenciáveis tem-se: função degrau, a qual produz em seus resultados os valores 0 ou 1, ou seja, o resultado produzido pela aplicação da função assumirá valores unitários positivos quando o valor do potencial de ativação for maior ou igual a zero, caso contrário, o resultado assumirá valores nulos; função degrau bipolar, também conhecida como função sinal, onde o resultado produzido pela aplicação assumirá valores unitários positivos quando o valor do potencial de ativação for maior que zero, valor nulo quando o potencial de ativação for também nulo e valores unitários negativos quando o potencial de ativação for menor que zero e função rampa simétrica, que constitui a mais simples, cujos valores retornados são iguais aos valores do potencial de ativação.

Já as funções de ativação totalmente diferenciáveis comumente utilizadas para RNA são: Função Sigmoid, que produz resultados entre zero e um e tende a ser similar à função degrau, entretanto, diferentemente da função degrau, ela é totalmente diferenciável no seu domínio de definição; função tangente hiperbólica, cujo resultado de saída fica entre -1 e 1 e função Relu, a qual retorna zero quando o valor do potencial de ativação for menor que zero e o próprio valor quando o potencial de ativação for positivo. É uma função que envolve baixo custo computacional, entretanto, não é centrada em zero, o que pode fazer com que a RNA tenha dificuldade para convergir. A Tabela 5 apresenta as principais funções de ativação totalmente diferenciáveis comumente utilizadas para RNA.

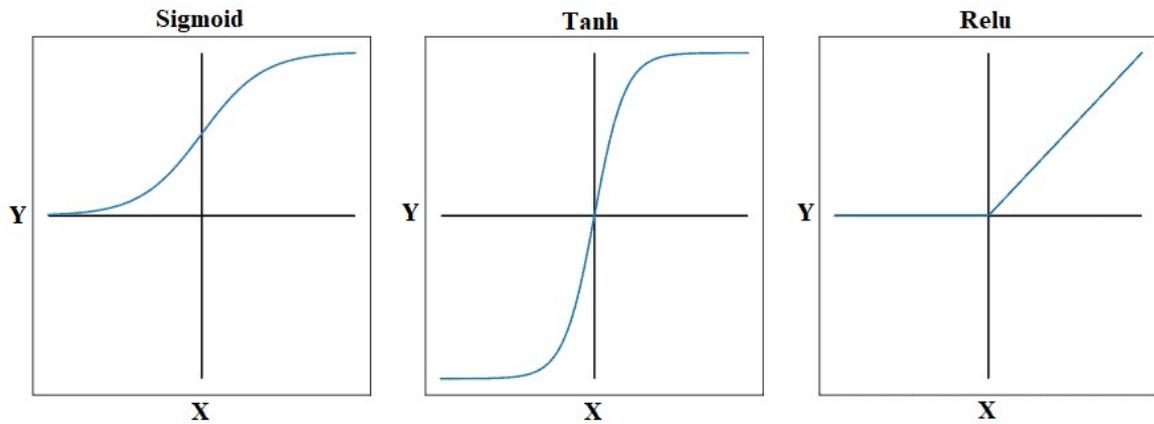
Tabela 5 – Funções de ativação comumente utilizadas para RNA

Função de Ativação	Fórmula
Sigmoid	$f(v) = \frac{1}{1 + e^{-v}}$
Tangente Hiperbólica	$f(v) = \frac{e^v + e^{-v}}{e^v - e^{-v}}$
Relu	$f(v) = \max [v, 0]$

Fonte: adaptado de Isaksson e Conde (2018).

A Figura 17 reúne os gráficos das funções de ativação comumente utilizadas para redes neurais artificiais.

Figura 17 – Gráficos das funções de ativação comumente utilizadas para RNA



Fonte: adaptado de Isaksson e Conde (2018)

Para a previsão de geração solar FV realizada neste trabalho, utilizou-se as funções de ativação Sigmoid, Relu e Tanh, as quais são as mais comumente utilizadas para este tipo de previsão (Isaksson; Conde, 2018).

5 RESULTADOS E ANÁLISE DE DESEMPENHO

Cinco experimentos foram feitos com a série temporal da usina solar de geração centralizada localizada no Ceará, identificados como:

- Estudo de Caso 1: FTDNN com técnica de treinamento BP e Adam e ajuste manual dos parâmetros da rede para previsão 1h à frente. Além disso, neste estudo de caso implementou-se também os modelos de referência MLPR, LR, DTR e persistência para fins de comparação de desempenho;
- Estudo de caso 2: FTDNN com técnica de treinamento PSO, PSO-AWF, e CPSO e ajuste manual dos parâmetros da rede para previsão 1h à frente;
- Estudo de caso 3: FTDNN com ajuste automático dos parâmetros para previsão 1h à frente por dois métodos de otimização:
 - Método 1: PSO-AWF para otimização dos parâmetros NNHL e LAG e Adam para treinamento da rede;
 - Método 2: PSO-AWF tanto para otimização dos parâmetros NNHL e LAG quanto para treinamento da rede.
- Estudo de caso 4: FTDNN com técnica de treinamento PSO-AWF para otimização automática dos parâmetros NNHL e LAG e Adam para treinamento da rede, sendo a previsão realizada para os horizontes de 3h à frente, 6h à frente, 3h à frente com resolução horária (de 1h até 3h) e 6h à frente com resolução horária (de 1h até 6h);
- Estudo de caso 5: FTDNN com ajuste automático dos parâmetros e previsão para 1h à frente em três diferentes aplicações:
 - Aplicação 1: PSO-AWF para otimização dos parâmetros NNHL e LAG e Adam para treinamento da rede, em uma abordagem univariada (utilizando-se apenas dados de potência para prever potência);
 - Aplicação 2: PSO-AWF para otimização dos parâmetros NNHL e LAG e Adam para treinamento da rede, em uma abordagem multivariada completa (utilizando-se os dados meteorológicos de todas as estações existentes na planta FV);
 - Aplicação 3: PSO-AWF para otimização dos parâmetros NNHL e LAG e Adam para treinamento da rede, com utilização de validação cruzada. Para este caso, foi utilizado o mesmo conjunto de dados dos demais estudos de caso, com dados meteorológicos da WS01 e de potência gerada para prever potência gerada (abordagem multivariada).

5.1 Estudo de Caso 1: FTDNN com Técnica de Treinamento BP e Adam

Aplicou-se ao conjunto de dados o modelo FTDNN com técnicas de treinamento BP (FTDNN&BP) e Adam (FTDNN&ADAM). Estas duas técnicas de treinamento foram implementadas a partir da biblioteca Tensorflow do Python. Os parâmetros testados por tentativa e erro foram número de neurônios da camada escondida (NNHL): 2; 5; 10; 25; 50; atraso (LAG): 1; 3; 6; 10; função de ativação (ACT): Sigmoid, Relu, Tanh; e taxa de aprendizagem (LR): 0,1; 0,01; 0,001; 0,0001; 0,5; 0,05; 0,0005; 0,9; 0,09; 0,0009.

A escolha dos valores ajustados manualmente foi feita baseando-se em trabalhos científicos relacionados à área. De acordo com Dantas (2020), caso a topologia de uma RNA possua uma pequena quantidade de neurônios e conexões, a rede pode não aprender de maneira satisfatória. Por outro lado, se a topologia contém muitos neurônios e conexões, a rede pode acabar memorizando os padrões específicos de treinamento (*overfitting*), perdendo a capacidade de generalização. Os valores de atraso praticados neste trabalho foram utilizados também em Da Cunha (2021) e Bezerra (2022), pesquisas que utilizaram redes FTDNN para aplicações em geração solar e eólica e segundo Isaksson (2018), as funções de ativação Sigmoid, Relu e Tanh são as mais comumente utilizadas para previsão solar fotovoltaica. A taxa de aprendizagem depende muito dos valores inerentes a cada problema, não tendo assim um valor determinado (Dantas, 2020). Deste modo, levando-se em consideração que a taxa de aprendizagem padrão na biblioteca Scikit-learn do Python é 0,001, variou-se esse valor dentro de uma determinada faixa e aplicou-se ao modelo.

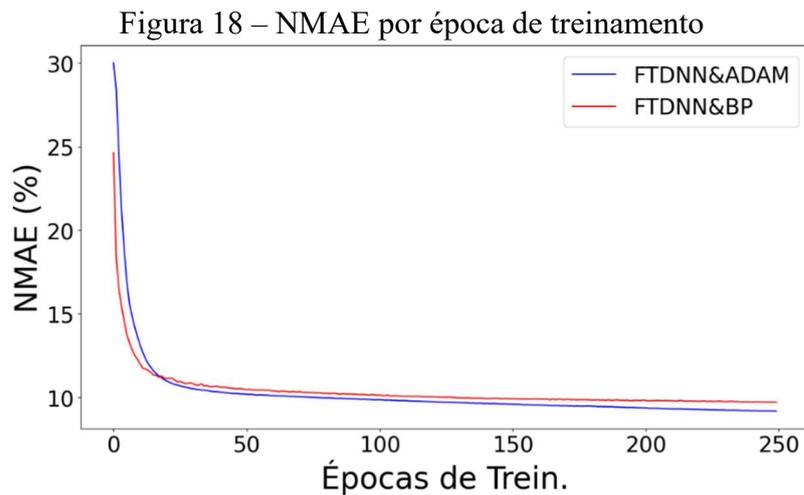
Os melhores parâmetros para a FTDNN e o número total de redes criadas pela combinação dos parâmetros ajustados manualmente (NR1) são listados na Tabela 6. Adotou-se como número de épocas de treinamento 250, uma vez que graficamente em torno desse valor já se percebe a convergência do modelo. O tamanho do lote (*batch size*) foi adotado em 32, tendo em vista que este tamanho normalmente funciona bem em modelos com RNA (Brownlee, 2021).

Tabela 6 – Melhores parâmetros para a FTDNN&BP e FTDNN&ADAM

MODELO	NNHL	LAG	ACT	LR	NR1
FTDNN&BP	5	10	Sigmoid	0,05	600
FTDNN&ADAM	10	10	Sigmoid	0,001	600

Fonte: o próprio autor.

Na busca de se eliminar a ocorrência de vieses, executou-se cada modelo três vezes, sendo considerada a média dos resultados obtidos. A Figura 18 apresenta o gráfico NMAE por época de treinamento para a melhor dentre as três execuções dos modelos FTDNN&BP e FTDNN&ADAM, sendo o último o que apresentou melhor performance (erros menores para a mesma quantidade de épocas).



Fonte: o próprio autor.

Para fins de comparação de resultados, aplicou-se também os modelos de referência MLPR, LR, DTR e persistência. Conforme informado na seção anterior, para as funções de referência foi realizada a técnica da busca em grade (*grid Search*) na busca dos melhores parâmetros, variando-se algumas entradas das funções pré-existentes do Python. A Tabela 7 mostra os parâmetros aplicados na busca em grade.

Tabela 7 – Parâmetros da busca em grade para os modelos de referência

Parâmetros MLPR:	
<i>hidden_layer_sizes:</i>	(2), (5), (10), (25), (50)
<i>activation:</i>	sigmoid, tanh, relu
<i>solver:</i>	SGD, Adam
<i>alpha (regularization term):</i>	0,1; 0,01; 0,001; 0,0001; 0,5; 0,05; 0,0005; 0,9; 0,09; 0,0009
<i>learning rate</i>	<i>constant, adaptive</i>
Parâmetros LR:	
<i>copy_X:</i>	<i>true, false</i>
<i>fit_intercept:</i>	<i>true, false</i>
<i>positive:</i>	<i>true, false</i>

Parâmetros DTR:	
<i>criterion:</i>	<i>squared_error, friedman_mse, absolute_error</i>
<i>max_depth:</i>	4, 5, 6, 7, 8, 9, 10

Fonte: o próprio autor.

A Tabela 8 resume os resultados médios obtidos após três execuções dos modelos FTDNN&BP, FTDNN&ADAM, bem como das funções de referência MLPR, LR, DTR e persistência. Os cálculos das métricas de avaliação foram feitos a partir das funções pré-existentes na biblioteca Scikit-learn do Python. Observa-se que a rede FTDNN&ADAM apresentou melhor desempenho em todas as métricas aplicadas.

Tabela 8 – Resultados obtidos para a FTDNN&BP, FTDNN&ADAM e modelos de referência

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&ADAM	18,419	13,506	79,899	14,206	10,416
FTDNN&BP	18,680	13,743	79,326	14,407	10,599
DTR	19,645	14,322	77,140	15,151	11,046
MLPR	19,891	15,330	76,565	15,341	11,823
LR	20,833	16,200	74,293	16,067	12,494
Persistência	27,609	19,685	54,669	21,293	15,182

Fonte: o próprio autor.

O tempo médio (Texec) e o tempo total de execução (Texect) de cada modelo é mostrado na Tabela 9. O tempo total de execução para os modelos que utilizam RNA é dado pela multiplicação do tempo médio pelo número total de redes criadas manualmente (NR1). Para os modelos de referência, Texec é igual a Texect, sendo o tempo de execução da técnica da busca em grade para cada modelo.

Tabela 9 – Tempo de execução para a FTDNN&BP, FTDNN&ADAM e modelos de referência

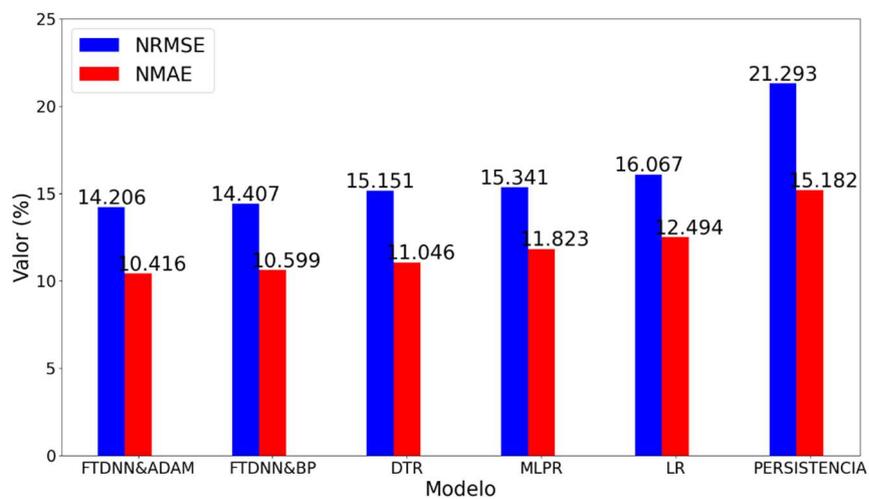
Modelo	Texec	Texect
FTDNN&ADAM	1,5min	15h
FTDNN&BP	1,5min	15h
DTR	<1min	<1min
MLPR	17min	17min
LR	<1min	<1min

Persistência	<1min	<1min
--------------	-------	-------

Fonte: o próprio autor.

A Figura 19 apresenta o gráfico de barras para as métricas de avaliação NRMSE e NMAE dos modelos executados no estudo de caso 1. Pela análise gráfica, percebe-se que a discrepância de valores entre o modelo de melhor performance (FTDNN&ADAM) e o modelo de pior performance (persistência), além de acentuada, é maior em termos de NRMSE do que de NMAE.

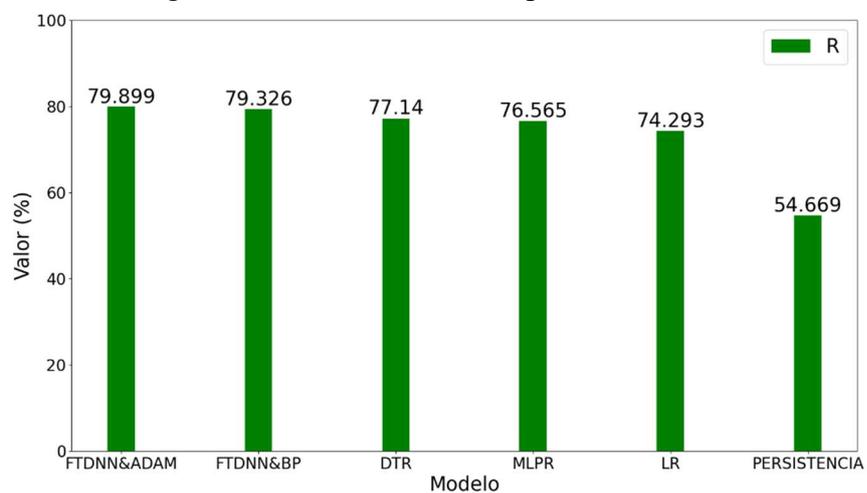
Figura 19 – Gráfico de barras para as métricas NRMSE e NMAE



Fonte: o próprio autor.

Já o gráfico de barras para a métrica de avaliação R dos modelos executados no estudo de caso 1 é mostrado na Figura 20, onde também se percebe uma diferença acentuada entre o melhor e o pior modelo.

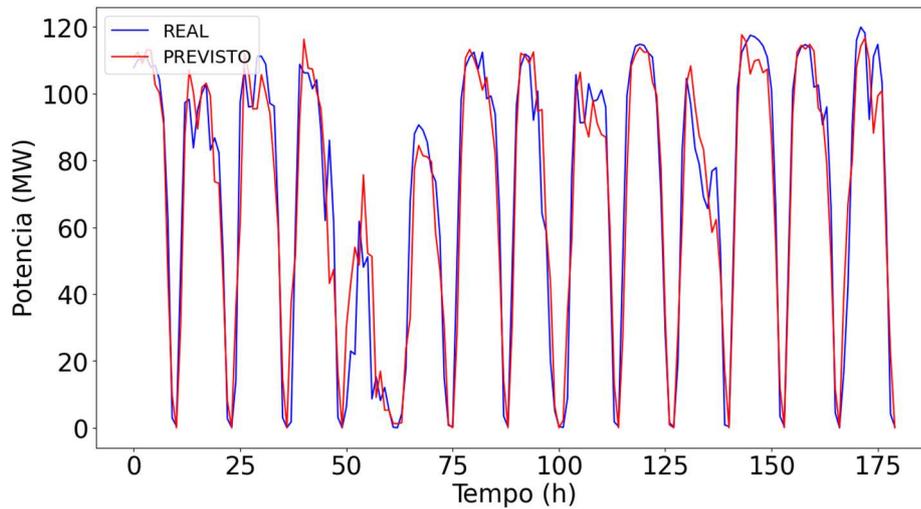
Figura 20 – Gráfico de barras para a métrica R



Fonte: o próprio autor.

O gráfico das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) com valores reais e previstos para o melhor resultado obtido no estudo de caso 1 (FTDNN&ADAM) é apresentado na Figura 21. O intervalo de 180h foi escolhido para facilitar a visualização.

Figura 21 – Dados reais e previstos para o modelo FTDNN&ADAM



Fonte: o próprio autor.

A partir do estudo de caso 1 percebe-se que a rede FTDNN&ADAM apresentou a melhor performance em termos de métricas de avaliação dentre todos os algoritmos avaliados nesta seção. Mesmo modelos já bastante consolidados na literatura, como o LR e o DTR apresentaram valores de RMSE, MAE e R maiores que os algoritmos que utilizam FTDNN neste estudo de caso.

5.2 Estudo de Caso 2: FTDNN com Técnica de Treinamento PSO, PSO-AWF e CPSO

Os algoritmos PSO, PSO-AWF e CPSO também foram testados como algoritmos para treinamento da rede para algumas faixas de dados, como número de neurônios da camada escondida (NNHL): 2; 5; 10; 25; 50; atraso (LAG): 1; 3; 6, 10 e valores limites dos pesos (W): [-1,5;1,5]; [-10;10]. Escolheu-se esses dois intervalos para os pesos tendo em vista abranger uma elevada faixa de valores. Os melhores parâmetros para os algoritmos PSO, PSO-AWF e CPSO quando usados na otimização dos pesos da rede FTDNN e o número total de redes criadas pela combinação dos parâmetros ajustados manualmente (NR2) são mostrados na Tabela 10.

Tabela 10 – Melhores parâmetros para a FTDNN&PSO, FTDNN&PSO-AWF e FTDNN&CPSO

MODELO	NNHL	LAG	W	NR2
FTDNN&PSO	5	3	[-1,5;1,5]	40
FTDNN&PSO-AWF	5	10	[-1,5;1,5]	40

FTDNN&CPSO	5	1	[-1,5;1,5]	40
------------	---	---	------------	----

Fonte: o próprio autor.

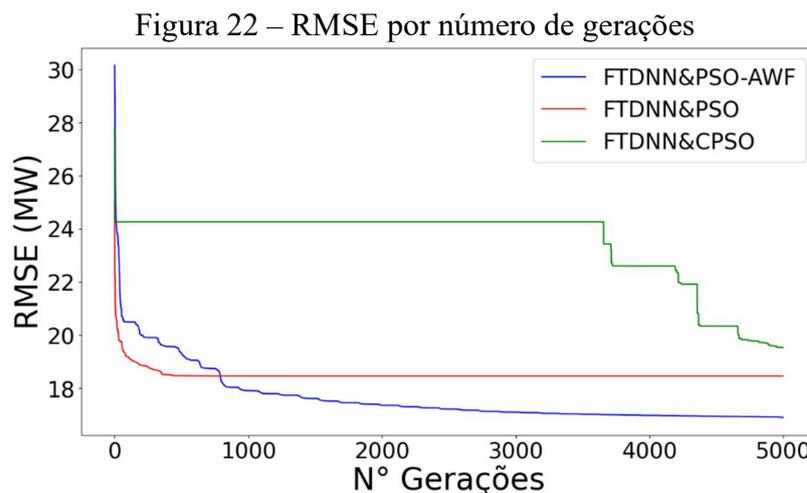
Utilizou-se 1 como valor para a constante social e cognitiva nas três técnicas de PSO utilizadas. Adotou-se como função de ativação a Sigmoid, número de gerações 5000 e população 500. A função Sigmoid foi escolhida por ter sido a que obteve melhor performance dentre as três utilizadas no estudo de caso 1, além de ter sido aplicada também nos trabalhos de Da Cunha (2021) e Bezerra (2022). As constantes social e cognitiva, o número de gerações e a população foram obtidos por tentativa e erro. Os valores adotados para os demais parâmetros do modelo PSO-AWF, como fator de envelhecimento (AF), fator de enfraquecimento (WF), expectativa de vida (LE) e número de partículas criadas após determinado número de gerações (NPartic), os quais também foram obtidos por tentativa e erro, são apresentados na Tabela 11.

Tabela 11 –Parâmetros para a rede FTDNN&PSO-AWF

MODELO	AF	WF	LE	NPartic
FTDNN&PSO-AWF	1	2	150	250

Fonte: o próprio autor.

Buscando-se eliminar a ocorrência de vieses, executou-se cada modelo três vezes. A Figura 22 apresenta o gráfico RMSE por número de gerações para a melhor dentre as três execuções dos modelos do estudo de caso 2, sendo o FTDNN&PSO-AWF o que apresentou melhor performance (erros menores para a mesma quantidade de épocas).



Fonte: o próprio autor.

A Tabela 12 resume os resultados médios obtidos após três execuções para os modelos FTDNN&PSO, FTDNN&PSO-AWF e FTDNN&CPSO. As métricas de avaliação são as

mesmas utilizadas no estudo de caso 1 e os valores dos modelos de referência foram repetidos para fins de comparação.

Tabela 12 – Resultados obtidos para a FTDNN&PSO, FTDNN&PSO-AWF, FTDNN&CPSO e modelos de referência

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&PSO-AWF	18,354	13,784	80,042	14,155	10,631
FTDNN&PSO	19,167	14,841	78,233	14,782	11,446
FTDNN&CPSO	20,453	15,864	75,211	15,774	12,235
DTR	19,645	14,322	77,140	15,151	11,046
MLPR	19,891	15,330	76,565	15,341	11,823
LR	20,833	16,200	74,293	16,067	12,494
Persistência	27,609	19,685	54,669	21,293	15,182

Fonte: o próprio autor.

O tempo médio (Texec) e o tempo total de execução (Texect) de cada modelo é mostrado na Tabela 13. O tempo total de execução é dado pela multiplicação do tempo médio pelo número total de redes criadas manualmente (NR2). Novamente os valores dos modelos de referência foram repetidos nesta tabela para fins de comparação.

Tabela 13 – Tempo de execução para a FTDNN&PSO, FTDNN&PSO-AWF, FTDNN&CPSO e modelos de referência

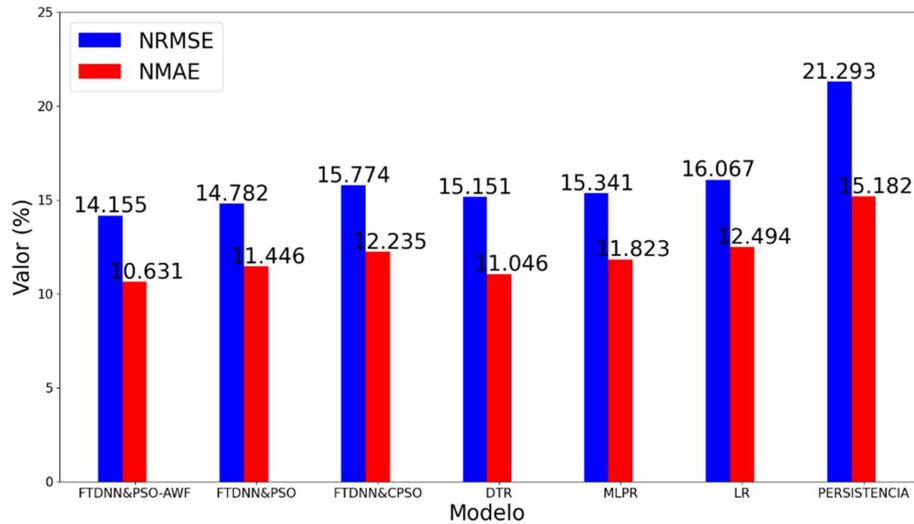
Modelo	Texec	Texect
FTDNN&PSO-AWF	1,25h	50h
FTDNN&PSO	5,98h	239,2h
FTDNN&CPSO	5,30h	212h
DTR	<1min	<1min
MLPR	17min	17min
LR	<1min	<1min
Persistência	<1min	<1min

Fonte: o próprio autor.

A Figura 23 apresenta o gráfico de barras para as métricas de avaliação NRMSE e NMAE dos modelos relacionados na Tabela 12. Pela análise gráfica, assim como ocorreu no estudo de caso 1, percebe-se que a discrepância de valores entre o modelo de melhor

performance (FTDNN&PSO-AWF) e o modelo de pior performance (persistência), além de acentuada, é maior em termos de NRMSE do que de NMAE.

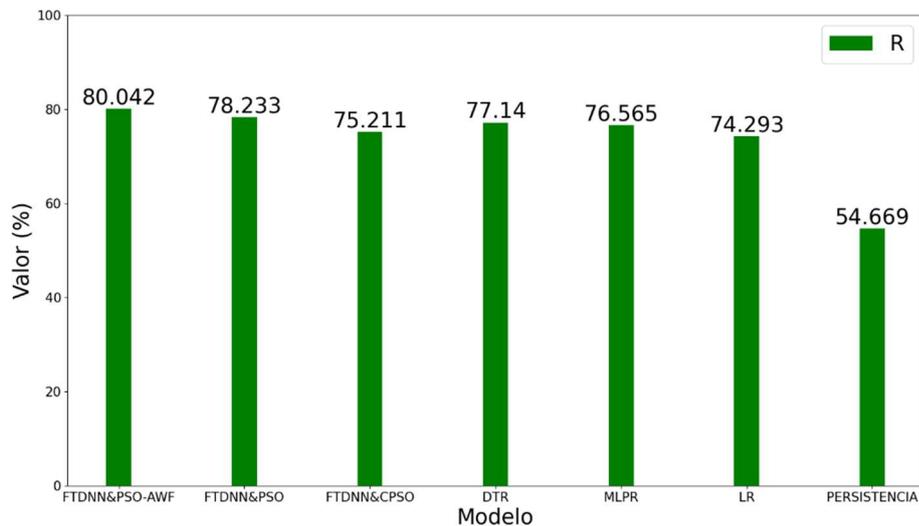
Figura 23 – Gráfico de barras para as métricas NRMSE e NMAE



Fonte: o próprio autor.

O gráfico de barras para a métrica de avaliação R dos modelos executados no estudo de caso 2 é mostrado na Figura 24, onde também se percebe uma diferença acentuada entre o melhor e o pior modelo.

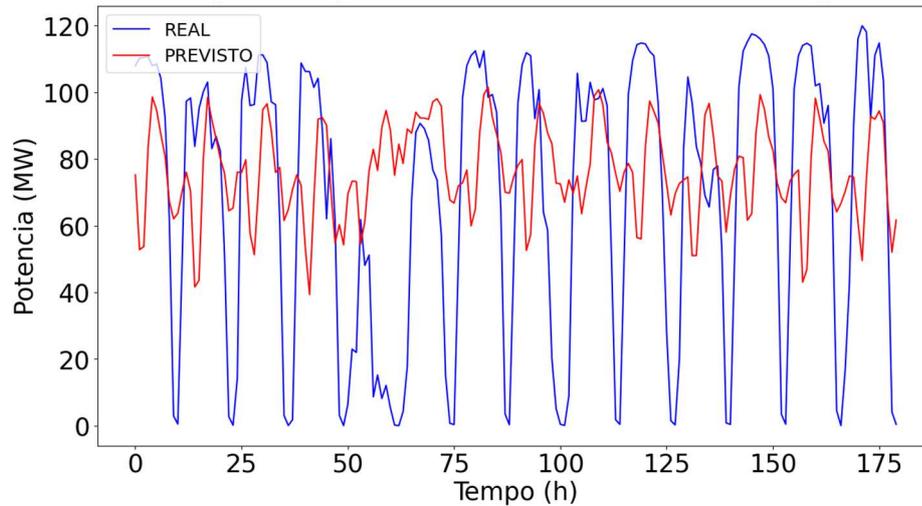
Figura 24 – Gráfico de barras para a métrica R



Fonte: o próprio autor.

O gráfico das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) com valores reais e previstos para o melhor modelo do estudo de caso 2 (FTDNN&PSO-AWF) após a primeira iteração é apresentado na Figura 25.

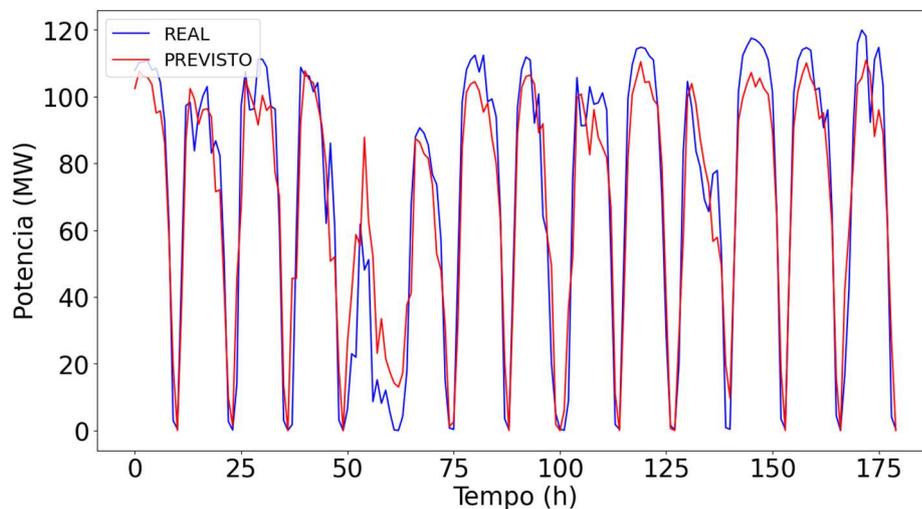
Figura 25 – Dados reais e previstos para o modelo FTDNN&PSO-AWF (primeira iteração)



Fonte: o próprio autor.

Já o gráfico das primeiras 180h do conjunto de teste com valores reais e previstos para o modelo FTDNN&PSO-AWF após a última iteração é apresentado na Figura 26. Comparando-se as Figuras 25 e 26 é possível se perceber a evolução do algoritmo FTDNN&PSO-AWF ao longo das 5000 iterações pré-definidas (número de gerações).

Figura 26 – Dados reais e previstos para o modelo FTDNN&PSO-AWF (última iteração)



Fonte: o próprio autor.

A partir do estudo de caso 2, percebe-se que a rede FTDNN&PSO-AWF apresentou a melhor performance em termos de métricas de avaliação e tempo de execução dentre todos os modelos que utilizam a técnica PSO. O gráfico de RMSE por época de treinamento evidencia que os erros relacionados à melhor técnica são menores, situação ocasionada pela diversidade gerada pelas partículas que são incluídas ou excluídas ao longo da etapa de treinamento.

5.3 Estudo de Caso 3: FTDNN com Ajuste Automático de Parâmetros

Neste estudo de caso o PSO-AWF é utilizado para otimizar os parâmetros NNHL e LAG e a técnica Adam para treinar a rede FTDNN. A este modelo atribuiu-se a sigla FTDNN&ADAM&PSO-AWF(L,N). Em uma segunda implementação, o PSO-AWF otimizou os parâmetros NNHL e LAG e realizou também o treinamento da rede, recebendo a sigla FTDNN&PSO-AWF(L,N,W). Assim sendo, os modelos implementados no estudo de caso 3 tratam-se de modelos híbridos desenvolvidos a partir dos melhores modelos dos estudos de caso 1 e 2, com a diferença dos parâmetros serem ajustados automaticamente e não manualmente conforme aconteceu até então. No primeiro caso, aplicou-se 5 combinações aleatórias de NNHL e LAG (5 redes) e 50 gerações para otimização destes parâmetros. No segundo modelo utilizou-se também 5 redes aleatórias, com número de gerações 50 para a otimização dos parâmetros NNHL e LAG e 500 para treinamento. Todos os valores citados foram definidos empiricamente. O parâmetro NNHL foi escolhido aleatoriamente dentro do intervalo [1,50] e o LAG no intervalo [1,10], não sendo excedido assim os limites aplicados nos estudos de caso com ajuste manual de parâmetros. Utilizou-se 1 como valor para as constantes social e cognitiva. A Tabela 14 resume os parâmetros iniciais e os melhores parâmetros obtidos automaticamente pelas duas implementações.

Tabela 14 – Parâmetros iniciais e melhores parâmetros definidos automaticamente

Parâmetros Iniciais		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N)	5	10
FTDNN&PSO-AWF(L,N,W)	41	2
Melhores Parâmetros		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N)	28	10
FTDNN&PSO-AWF(L,N,W)	32	4

Fonte: o próprio autor.

Os parâmetros AF, WF, LE e NPartic utilizados nos dois modelos com ajustes automáticos de NNHL e LAG encontram-se listados na Tabela 15. Para o algoritmo FTDNN&ADAM&PSO-AWF(L,N) usou-se taxa de aprendizagem 0,001, que corresponde à taxa padrão da função para treinamento de RNA pré-existente na biblioteca Scikit-learn do Python. O número de épocas de treinamento foi feito igual a 250 e o tamanho do lote 32,

mesmos parâmetros utilizados nos modelos com treinamento ADAM e ajuste manual de parâmetros (estudo de caso 1).

Tabela 15 – Parâmetros relacionados à técnica PSO-AWF nos modelos automatizados

MODELO	AF	WF	LE	NPartic
FTDNN&ADAM&PSO-AWF(L,N)	1	2	100	5
FTDNN&PSO-AWF(L,N,W) – Estruturação da RNA	1	2	100	5
FTDNN&PSO-AWF(L,N,W) – Treinamento da RNA	1	2	150	250

Fonte: o próprio autor.

A Tabela 16 apresenta os resultados médios obtidos após três execuções dos modelos automatizados, tendo o FTDNN&ADAM&PSO-AWF(L,N) superado o algoritmo FTDNN&PSO-AWF(L,N,W).

Tabela 16 – Resultados obtidos para os modelos automatizados

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&ADAM&PSO-AWF(L,N)	18,542	13,565	79,631	14,300	10,462
FTDNN&PSO-AWF(L, N, W)	19,003	14,518	78,605	14,656	11,197

Fonte: o próprio autor.

O tempo médio (Texec) que é igual ao tempo total de execução (Texect) de cada algoritmo com ajuste automático de parâmetros é mostrado na Tabela 17. Texec e Texect são iguais tendo em vista que as redes FTDNN&ADAM&PSO-AWF(L,N) e FTDNN&PSO-AWF(L,N,W) para previsão 1h à frente necessitaram ser rodadas apenas uma vez (a combinação dos melhores parâmetros aconteceu automaticamente, e não manualmente).

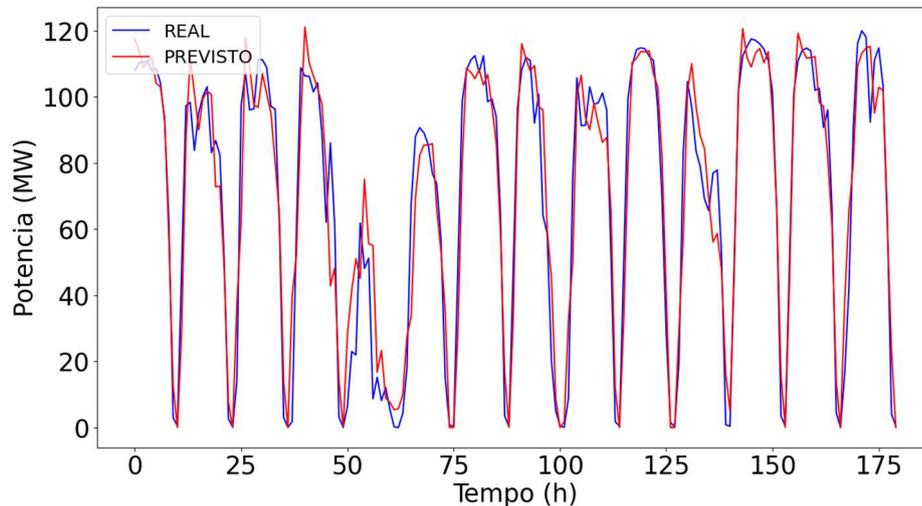
Tabela 17 – Tempo de execução para os modelos automatizados

Modelo	Texect
FTDNN&ADAM&PSO-AWF(L,N)	2,26h
FTDNN&PSO-AWF(L,N,W)	37,24h

Fonte: o próprio autor.

A Figura 27 apresenta o gráfico das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) para o melhor resultado do modelo FTDNN&ADAM&PSO-AWF(L,N).

Figura 27 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N)



Fonte: o próprio autor.

A partir do estudo de caso 3 pode-se concluir que o modelo FTDNN&ADAM&PSO-AWF(L,N) apresenta tempo de execução bastante reduzido em comparação com os modelos que utilizam FTDNN e ajuste manual de parâmetros (2,26h x 50h, respectivamente), além de possuir valores de RMSE, MAE e R próximos ao do melhor algoritmo implementado até então (FTDNN&PSO-AWF). Desta maneira, a implementação deste modelo é bastante satisfatória do ponto de vista de métricas de avaliação e, principalmente, de tempo de execução.

5.4 Estudo de Caso 4: FTDNN com Ajuste Automático de Parâmetros e Previsão de Potência para 3h e 6h à frente

Neste estudo de caso, o modelo com melhor performance do estudo de caso 3, ou seja, a melhor dentre as duas implementações com ajuste automático de parâmetros (FTDNN&ADAM&PSO-AWF(L,N)) é utilizado para previsão da potência gerada pela usina solar no horizonte de 3h à frente, 6h à frente, 3h à frente com resolução horária (de 1h até 3h) e 6h à frente com resolução horária (de 1h até 6h). Para tanto, utilizou-se como variáveis de entrada tanto os dados meteorológicos da estação meteorológica WS01 quanto dados de potência gerada, situação chamada neste trabalho de abordagem multivariada.

Semelhante ao que ocorreu no estudo de caso 3, aplicou-se 5 combinações aleatórias de NNHL e LAG (5 redes) e 50 gerações para otimização destes parâmetros, AF, WF, LE e NPartic iguais aos da Tabela 11, constantes social e cognitiva 1, taxa de aprendizagem 0,001, 250 épocas

de treinamento e tamanho do lote 32. A Tabela 18 resume os melhores parâmetros obtidos automaticamente para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de potência para 3h e 6h à frente.

Tabela 18 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente

Parâmetros Iniciais		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – 3h	47	9
FTDNN&ADAM&PSO-AWF(L,N) – 6h	34	10
Melhores Parâmetros		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – 3h	24	10
FTDNN&ADAM&PSO-AWF(L,N) – 6h	31	10

Fonte: o próprio autor.

A Tabela 19 apresenta os resultados médios obtidos após três execuções do modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente.

Tabela 19 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&ADAM&PSO-AWF(L, N) – 3h	25,478	18,628	61,538	19,650	14,367
FTDNN&ADAM&PSO-AWF(L, N) – 6h	26,508	19,478	58,359	20,444	15,022

Fonte: o próprio autor.

O tempo médio (Texec) que é igual ao tempo total de execução (Texect) do algoritmo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente é mostrado na Tabela 20. Texec e Texect são iguais tendo em vista que a rede FTDNN&ADAM&PSO-AWF(L,N) para previsão 3h e 6h à frente necessitou ser rodada apenas uma vez (a combinação dos melhores parâmetros aconteceu automaticamente, e não manualmente).

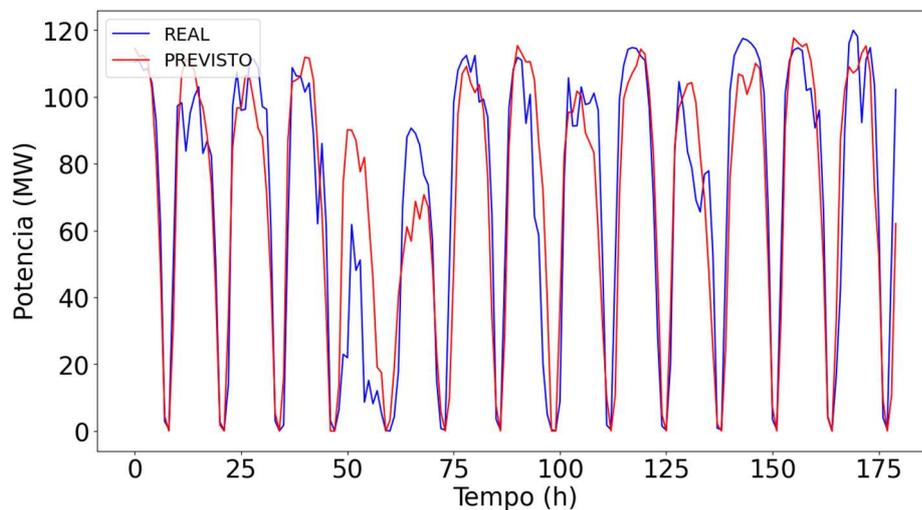
Tabela 20 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 3h e 6h à frente

Modelo	Texect
FTDNN&ADAM&PSO-AWF(L,N) – 3h	2,24h
FTDNN&ADAM&PSO-AWF(L,N) – 6h	2,09h

Fonte: o próprio autor.

A Figura 28 apresenta o gráfico das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) do melhor resultado para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de potência para 3h à frente.

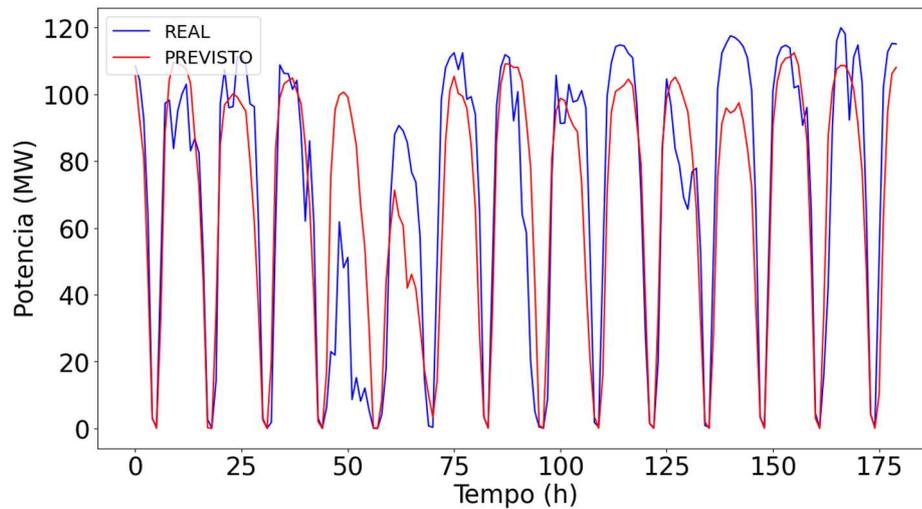
Figura 28 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão para 3h à frente



Fonte: o próprio autor.

Já o gráfico das primeiras 180h do conjunto de teste do melhor resultado para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de potência para 6h à frente é apresentado na Figura 29. As discrepâncias maiores existentes entre os picos de valores reais e previstos em comparação com o gráfico dos modelos de previsão para 1h à frente, explicam as piores métricas de avaliação obtidas para o modelo que prevê 3h e 6h à frente.

Figura 29 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão para 6h à frente



Fonte: o próprio autor.

Na busca de se ampliar ainda mais a investigação do impacto causado por previsões para um horizonte de tempo maior que 1h à frente, utilizou-se o modelo FTDNN&ADAM&PSO-AWF(L,N) para previsão da potência gerada pela usina FV 3h à frente com resolução horária (de 1h até 3h) e 6h à frente com resolução horária (de 1h até 6h), ou seja, levando-se em conta também intervalos intermediários.

A Tabela 21 resume os melhores parâmetros obtidos automaticamente para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de potência de 1h até 3h e de 1h até 6h à frente.

Tabela 21 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente

Parâmetros Iniciais		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – 1h - 3h	41	10
FTDNN&ADAM&PSO-AWF(L,N) – 1h - 6h	39	10
Melhores Parâmetros		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – 1h - 3h	50	10
FTDNN&ADAM&PSO-AWF(L,N) – 1h - 6h	45	10

Fonte: o próprio autor.

A Tabela 22 apresenta os resultados médios obtidos após três execuções do modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente. Apesar

da previsão ser realizada para vários pontos simultaneamente, calculou-se as métricas de avaliação para os intervalos de previsão 1h, 3h e 6h à frente.

Tabela 22 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente

Modelo	RMSE (MW) - 1h	MAE (MW) - 1h	R (%) - 1h	NRMSE (%) - 1h	NMAE (%) - 1h
FTDNN&ADAM&PSO-AWF(L, N) – 1h - 3h	18,426	13,431	79,884	14,211	10,359
Modelo	RMSE (MW) - 3h	MAE (MW) - 3h	R (%) - 3h	NRMSE (%) - 3h	NMAE (%) - 3h
FTDNN&ADAM&PSO-AWF(L, N) – 1h - 3h	25,577	19,081	61,236	19,726	14,716
Modelo	RMSE (MW) - 1h	MAE (MW) - 1h	R (%) - 1h	NRMSE (%) - 1h	NMAE (%) - 1h
FTDNN&ADAM&PSO-AWF(L, N) – 1h - 6h	18,409	13,457	79,924	14,198	10,378
Modelo	RMSE (MW) - 6h	MAE (MW) - 6h	R (%) - 6h	NRMSE (%) - 6h	NMAE (%) - 6h
FTDNN&ADAM&PSO-AWF(L, N) – 1h - 6h	26,584	19,525	58,120	20,503	15,059

Fonte: o próprio autor.

O tempo médio (Texec) que é igual ao tempo total de execução (Texec) do algoritmo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente é mostrado na Tabela 23.

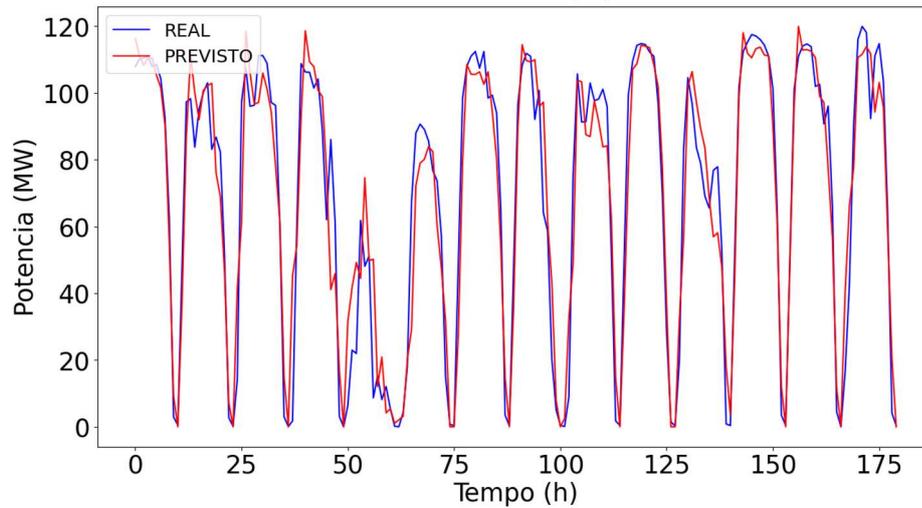
Tabela 23 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de 1h até 3h e de 1h até 6h à frente

Modelo	Texect
FTDNN&ADAM&PSO-AWF (L,N) – 1h - 3h	2,27h
FTDNN&ADAM&PSO-AWF (L,N) – 1h - 6h	2,33h

Fonte: o próprio autor.

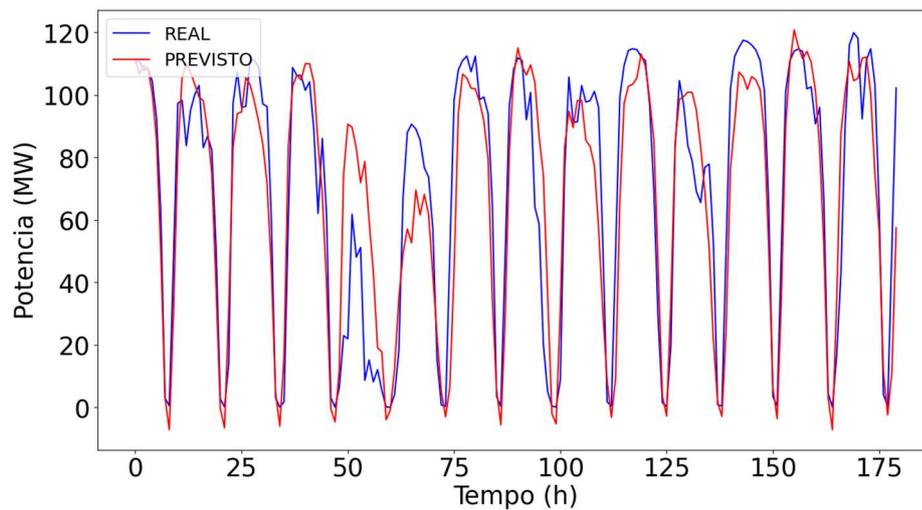
As Figuras 30 e 31 apresentam os gráficos das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) para o melhor resultado do modelo FTDNN&ADAM&PSO-AWF (L,N) com previsão de potência de 1h até 3h à frente.

Figura 30 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 3h à frente (1h)



Fonte: o próprio autor.

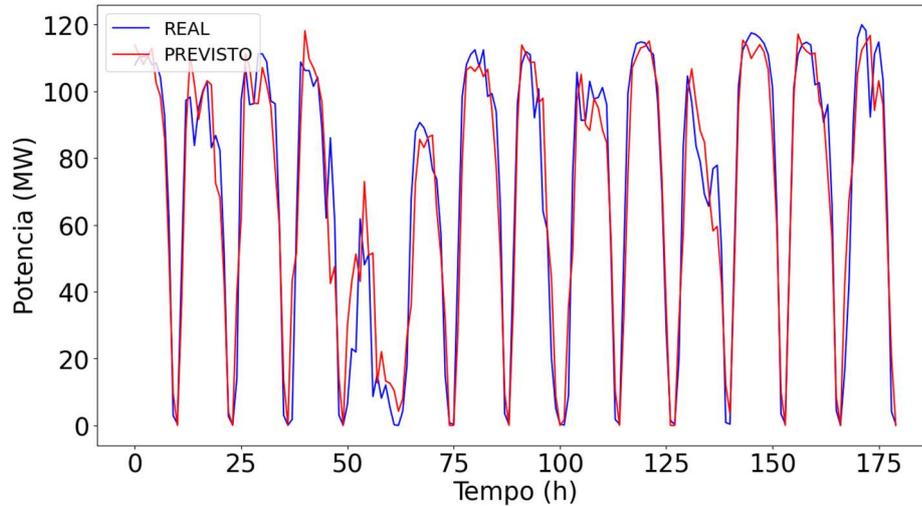
Figura 31 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 3h à frente (3h)



Fonte: o próprio autor.

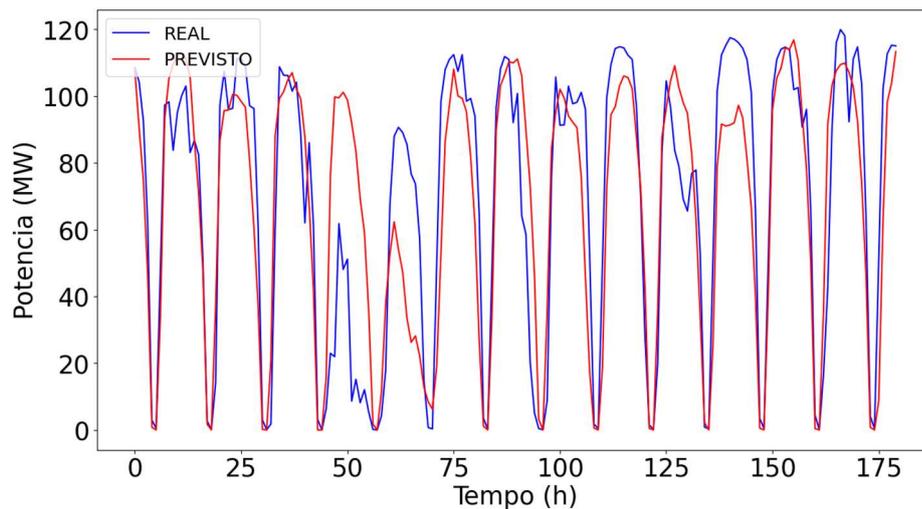
As Figuras 32 e 33 apresentam os gráficos das primeiras 180h do conjunto de teste para o melhor resultado do modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de potência de 1h até 6h à frente.

Figura 32 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 6h à frente (1h)



Fonte: o próprio autor.

Figura 33 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de 1h até 6h à frente (6h)



Fonte: o próprio autor.

O estudo de caso 4 demonstra que a previsão para 3h e 6h à frente com resolução horária (de 1h até 3h e de 1h até 6h) resultou em ganho com relação ao intervalo de 1h, tendo em vista que neste intervalo os valores das métricas de avaliação foram melhores que os obtidos pelo modelo FTDNN&ADAM&PSO-AWF(L,N) para previsão 1h à frente (estudo de caso 3).

5.5 Estudo de Caso 5: FTDNN com Ajuste Automático de Parâmetros e Abordagem Univariada, Multivariada Completa e com Validação Cruzada

Na primeira aplicação deste estudo de caso, o modelo com melhor performance do estudo de caso 3 (FTDNN&ADAM&PSO-AWF(L,N)) é utilizado para previsão horária da

potência gerada pela usina solar, utilizando-se como variável de entrada apenas dados de potência gerada, situação chamada neste trabalho de abordagem univariada.

Semelhante ao que ocorreu no estudo de caso 4, aplicou-se 5 combinações aleatórias de NNHL e LAG (5 redes) e 50 gerações para otimização destes parâmetros, AF, WF, LE e NPartic iguais aos da Tabela 11, constantes social e cognitiva 1, taxa de aprendizagem 0,001, 250 épocas de treinamento e tamanho do lote 32. A Tabela 24 resume os melhores parâmetros obtidos automaticamente para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada (Uni).

Tabela 24 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada

Parâmetros Iniciais		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – Uni	44	9
Melhores Parâmetros		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – Uni	21	10

Fonte: o próprio autor.

A Tabela 25 apresenta os resultados médios obtidos após três execuções do modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada.

Tabela 25 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&ADAM&PSO-AWF(L, N) – Uni	19,891	14,852	76,557	15,341	11,455

Fonte: o próprio autor.

O tempo médio (Texec) que é igual ao tempo total de execução (Texec) do modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada é mostrado na Tabela 26.

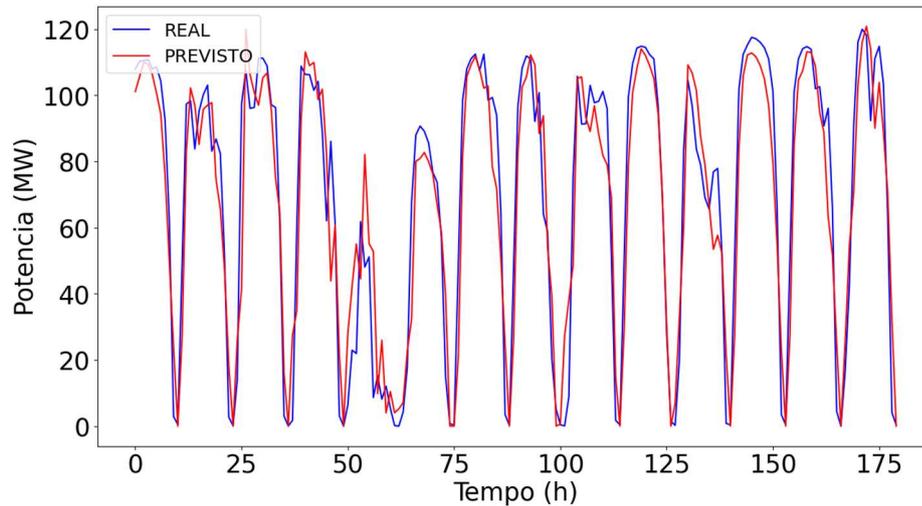
Tabela 26 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem univariada

Modelo	Texect
FTDNN&ADAM&PSO-AWF(L,N) – Uni	2,18h

Fonte: o próprio autor.

A Figura 34 apresenta o gráfico das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) do melhor resultado para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de potência para 1h à frente na abordagem univariada.

Figura 34 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão para 1h à frente na abordagem univariada



Fonte: o próprio autor.

Conforme informado no capítulo 2, o complexo FV em estudo neste trabalho possui um total de 13 estações meteorológicas espalhadas por toda a planta (WS01 a WS13). Até o presente momento utilizou-se os dados meteorológicos apenas da WS01, a mais completa do complexo, entretanto, buscando-se avaliar o efeito das variações meteorológicas ao longo de toda a planta, como segunda aplicação deste estudo de caso analisou-se também as métricas de avaliação de um modelo que utiliza tanto os dados meteorológicos da WS01 (irradiância horizontal, temperatura ambiente, umidade e velocidade do vento) quanto das WS02 a WS13 (irradiância horizontal, temperatura ambiente e velocidade do vento) para previsão de potência para 1h à frente, situação chamada neste trabalho de abordagem multivariada completa. As demais grandezas medidas pelas WS02 a WS13 (Tabela 01) não foram utilizadas devido a apresentarem muitas inconsistências (dados zerados, dados faltantes, dados negativos, etc.)

A Tabela 27 resume os melhores parâmetros obtidos automaticamente para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa (MC).

Tabela 27 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa

Parâmetros Iniciais		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – MC	9	10
Melhores Parâmetros		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – MC	47	10

Fonte: o próprio autor.

A Tabela 28 apresenta os resultados médios obtidos após três execuções do modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa.

Tabela 28 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&ADAM&PSO-AWF(L, N) – MC	18,911	13,307	78,791	14,585	10,263

Fonte: o próprio autor.

O tempo médio (Texec) que é igual ao tempo total de execução (Texect) do modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa é mostrado na Tabela 29. Este tempo é superior ao Texect para a abordagem multivariada, fato explicado pelo maior número de entradas existentes na abordagem multivariada completa.

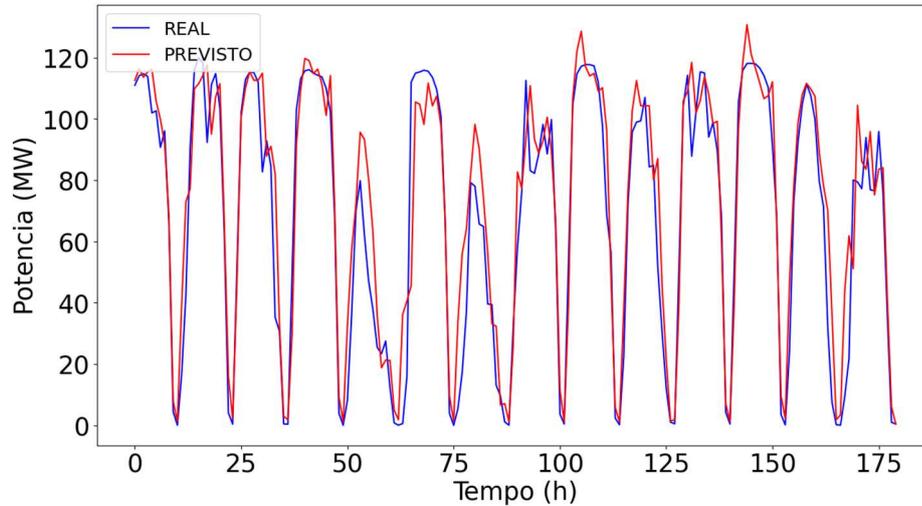
Tabela 29 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa

Modelo	Texect
FTDNN&ADAM&PSO-AWF(L,N) – MC	3,61h

Fonte: o próprio autor.

Além da remoção dos dados informada na seção 2.1 deste trabalho, removeu-se também para o caso da abordagem multivariada completa os dados de 30/06/2019 a 01/07/2019, tendo em vista que neste intervalo a WS08 apresentou dados de irradiância zerado. A Figura 35 apresenta o gráfico das primeiras 180h do conjunto de teste (de 14/07/2020 a 28/07/2020) do melhor resultado para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada completa.

Figura 35 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente na abordagem multivariada completa



Fonte: o próprio autor.

Buscando-se analisar a existência de vieses ao modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão de potência para 1h à frente na abordagem multivariada, aplicou-se ao mesmo a técnica de validação cruzada (validação por janela expansiva).

A Tabela 30 resume os melhores parâmetros obtidos automaticamente para o modelo FTDNN&ADAM&PSO-AWF(L,N) e previsão de potência para 1h à frente com validação cruzada (CV).

Tabela 30 – Parâmetros iniciais e melhores parâmetros obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada

Parâmetros Iniciais		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – CV	33	10
Melhores Parâmetros		
Modelo	NNHL	LAG
FTDNN&ADAM&PSO-AWF(L,N) – CV	37	10

Fonte: o próprio autor.

A Tabela 31 apresenta os resultados médios obtidos após três execuções do modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada.

Tabela 31 – Resultados obtidos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)
FTDNN&ADAM&PSO-AWF(L, N) – CV	18,780	13,826	79,094	14,484	10,663

Fonte: o próprio autor.

O tempo médio (Texec) que é igual ao tempo total de execução (Texec) do algoritmo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada é mostrado na Tabela 32.

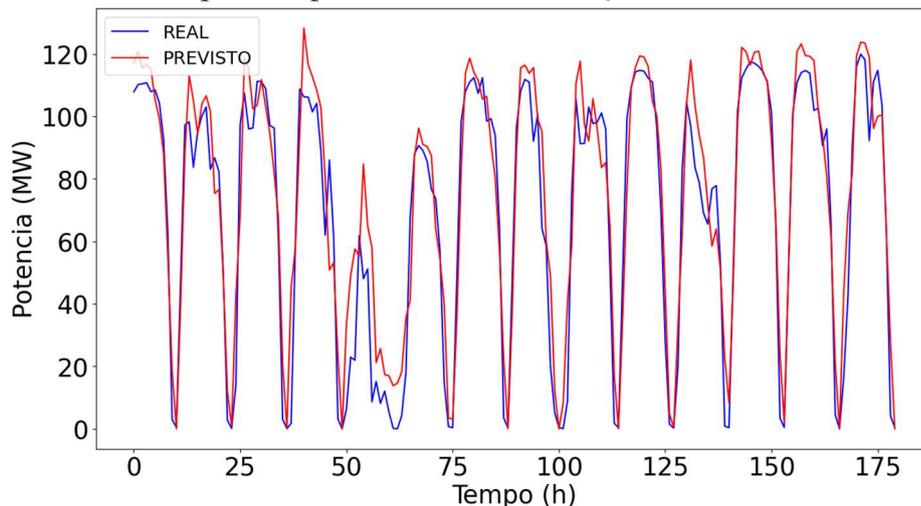
Tabela 32 – Tempo de execução para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada

Modelo	Texect
FTDNN&ADAM&PSO-AWF(L,N) – CV	6,78h

Fonte: o próprio autor.

A Figura 36 apresenta o gráfico das primeiras 180h do conjunto de teste (de 02/07/2020 a 16/07/2020) do melhor resultado para o modelo FTDNN&ADAM&PSO-AWF(L,N) na abordagem multivariada, com previsão de potência para 1h à frente e validação cruzada.

Figura 36 – Dados reais e previstos para o modelo FTDNN&ADAM&PSO-AWF(L,N) com previsão para 1h à frente e validação cruzada



Fonte: o próprio autor.

A partir do estudo de caso 5, percebe-se que não houve ganho com a implementação dos modelos de FTDNN para previsão 1h à frente com ajuste automático de parâmetros e abordagem univariada, multivariada completa e multivariada com validação cruzada. Os três

modelos experimentados não superaram as métricas de avaliação do algoritmo FTDNN&ADAM&PSO-AWF(L,N) com abordagem multivariada implementado no estudo de caso 3.

A Tabela 33 resume as métricas de avaliação de todos os dezoito modelos de previsão implementados neste trabalho. Apesar de apresentar resultados para apenas dezoito modelos, a referida tabela possui vinte linhas tendo em vista que para as implementações com ajuste automático de parâmetros e previsão para 3h e 6h à frente com resolução horária são apresentados dois resultados (dois horizontes de previsão). A partir da Tabela 33, percebe-se que o modelo FTDNN&PSO-AWF obteve a melhor performance dentre todas as implementações realizadas. O RMSE para todos os modelos foi maior que o MAE, o que faz sentido tendo em vista que, principalmente nos picos da geração, houve discrepâncias consideráveis entre os valores reais e previstos e o RMSE trata-se de uma métrica mais sensível às observações com desvios maiores. O coeficiente de correlação de Pearson acima de 80% para o melhor modelo indica uma correlação forte entre os valores reais e previstos pelo mesmo.

Tabela 33 – Resultados obtidos para todos os dezoito modelos implementados

Modelo	RMSE (MW)	MAE (MW)	R (%)	NRMSE (%)	NMAE (%)	Texect
FTDNN&ADAM	18,419	13,506	79,899	14,206	10,416	15h
FTDNN&BP	18,680	13,743	79,326	14,407	10,599	15h
DTR	19,645	14,322	77,140	15,151	11,046	<1min
MLPR	19,891	15,330	76,565	15,341	11,823	17min
LR	20,833	16,200	74,293	16,067	12,494	<1min
Persistência	27,609	19,685	54,669	21,293	15,182	<1min
FTDNN&PSO-AWF	18,354	13,784	80,042	14,155	10,631	50h
FTDNN&PSO	19,167	14,841	78,233	14,782	11,446	239,2h
FTDNN&CPSO	20,453	15,864	75,211	15,774	12,235	212h
FTDNN&ADAM&PSO-AWF (L, N)	18,542	13,565	79,631	14,300	10,462	2,26h
FTDNN&PSO-AWF(L, N, W)	19,003	14,518	78,605	14,656	11,197	37,24h
FTDNN&ADAM&PSO-AWF (L, N) – 3h	25,478	18,628	61,538	19,650	14,367	2,24h
FTDNN&ADAM&PSO-AWF (L, N) – 6h	26,508	19,478	58,359	20,444	15,022	2,09h

FTDNN&ADAM&PSO-AWF (L, N) – 1-3h (1h)	18,426	13,431	79,884	14,211	10,359	2,27h
FTDNN&ADAM&PSO-AWF (L, N) – 1-3h (3h)	25,577	19,081	61,236	19,726	14,716	2,27h
FTDNN&ADAM&PSO-AWF (L, N) – 1-6h (1h)	18,409	13,457	79,924	14,198	10,378	2,33h
FTDNN&ADAM&PSO-AWF (L, N) – 1-6h (6h)	26,584	19,525	58,120	20,503	15,059	2,33h
FTDNN&ADAM&PSO-AWF (L, N) – Uni	19,891	14,852	76,557	15,341	11,455	2,18h
FTDNN&ADAM&PSO-AWF (L, N) – MC	18,911	13,307	78,791	14,585	10,263	3,61h
FTDNN&ADAM&PSO-AWF (L, N) – CV	18,780	13,826	79,094	14,484	10,663	6,78h

Fonte: o próprio autor.

Pela análise do tempo de execução total (Texect), também presente na Tabela 33, é notório que a utilização dos modelos com ajuste automático de parâmetros trouxe ganhos com relação a este quesito. Para o horizonte de previsão de 1h, o melhor modelo com ajuste manual de parâmetros teve tempo de execução de 50h, enquanto no modelo de melhor performance com ajuste automático de parâmetros esse tempo foi de 2,26h. Além disso, a diferença entre as métricas de avaliação desses dois modelos não foi elevada, o que indica que os modelos que definem os parâmetros da rede FTDNN automaticamente são de grande utilidade dentro do contexto adotado nesse trabalho.

6 CONCLUSÃO

Neste trabalho foram avaliados modelos de previsão de geração de energia elétrica de uma usina solar fotovoltaica de 164 MWp, visando beneficiar a operação e integração à rede de usinas de potência de saída variável. Modelos de diferentes topologias foram implementados, os quais receberam como entrada uma série temporal de dados meteorológicos e de potência gerada e tiveram como variável alvo a potência gerada futura (instantes $t+1$, $t+3$ e $t+6$) da usina.

Conforme os resultados obtidos, a rede neural FTDNN com técnica de treinamento PSO-AWF (FTDNN&PSO-AWF) apresentou o melhor resultado dentre todos os modelos com ajuste manual de parâmetros, com RMSE 18,354 MW, MAE 13,784 MW, R 80,042%, NRMSE 14,155% e NMAE 10,631%, sendo superior inclusive a modelos já bastante consolidados na literatura como as funções de biblioteca em Python e o algoritmo com técnica de treinamento BP (FTDNN&BP). Já no caso dos modelos com ajuste automático de parâmetros e previsão horária, o modelo FTDNN&ADAM&PSO-AWF(L,N), parametrizado com PSO-AWF e treinado com o algoritmo Adam, foi o que obteve a melhor performance, com RMSE 18,542 MW, MAE 13,565 MW, R 79,631%, NRMSE 14,300% e NMAE 10,462%. Além disso, esse algoritmo apresentou tempo de execução bastante reduzido (2,26h) em comparação com os modelos que utilizam FTDNN e ajuste manual de parâmetros, além de possuir valores de RMSE, MAE e R próximos ao FTDNN&PSO-AWF.

Além da previsão para 1h à frente, implementou-se ainda modelos para previsão de 3h e 6h à frente com e sem resolução horária, tendo os modelos com resolução horária obtido melhores métricas de avaliação no intervalo 1h à frente em comparação com o modelo FTDNN &ADAM&PSO-AWF(L,N) implementado no estudo de caso 3. A alteração da quantidade de dados de entrada (abordagem univariada e multivariada completa), bem como a aplicação da técnica de validação cruzada na abordagem multivariada não trouxeram ganho para o melhor modelo automático com previsão horária implementado neste trabalho.

REFERÊNCIAS

- ADECE; FIEC; SEBRAE. **Atlas eólico e solar**: Ceará. 2019. 188p. Disponível em: <http://atlas.adece.ce.gov.br>. Acesso em: 28 ago. 2023.
- ALHMOUD, L.; AL-ZOUBI, A. M.; ALJARAH, I. Solar PV Power forecasting at Yarmouk University using machine learning techniques. **Open Engineering**, v. 12, p. 1078-1088, 2022.
- ANDRADE, F. D. A. *et al.* **Previsão de geração de energia fotovoltaica utilizando inteligência artificial em séries temporais**. In: XV Simpósio Brasileiro de Automação Inteligente – SBAI. Sociedade Brasileira de Automática (SBA), 2021.
- ARAIN, F. N. **Decision Tree Classification Algorithm**. 2021. Disponível em: <https://www.devops.ae/decision-tree-classification-algorithm/>. Acesso em: 28 ago. 2023.
- BEZERRA, E. C. **Abordagem auto-adaptativa baseada no conceito de expectativa de vida aplicada aos métodos Particle Swarm Optimization e máquinas kernel para previsão da velocidade do vento e geração eólica**. 2022. 135 f. Tese (Doutorado em Engenharia Elétrica) – Centro de Tecnologia, Universidade Federal do Ceará, Fortaleza, 2022.
- BEZERRA, E. C.; LEÃO, R. P. S.; BRAGA, A. D. S. A self-adaptive approach for particle swarm optimization applied to wind speed forecasting. **Journal of Control, Automation and Electrical Systems**, v. 28, p. 785-795, 2017.
- BOTCHKAREV, A. A new typology design of performance metrics to measure errors in machine learning regression algorithms. **Interdisciplinary Journal of Information, Knowledge, and Management**, v. 14, p. 45-79, 2019.
- BOUZON, M. F. **Estudo de algoritmos de otimização inspirados na natureza aplicados ao treinamento de redes neurais artificiais**. Dissertação (Mestrado) – Centro Universitário FEI, São Bernardo do Campo, 2021. 115f.
- BP. **BP statistical review of world energy 2022**. v. 71, 2022. Disponível em: <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html>. Acesso em: 28 ago. 2023.
- BROWNLEE, J. **How to control the stability of training neural networks with the batch size**. 2021. Disponível em: <https://machinelearningmastery.com/how-to-control-the-speed-and-stability-of-training-neural-networks-with-gradient-descent-batch-size>. Acesso em: 28 ago. 2023.
- BROWNLEE, J. **Linear Regression for Machine Learning**. 2020. Disponível em: <https://machinelearningmastery.com/linear-regression-for-machine-learning/>. Acesso em: 28 ago. 2023.
- CEPEL-CRESEB. **Manual de Engenharia para Sistemas Fotovoltaicos**. [S.l.], 2014. Disponível em: http://www.cresesb.cepel.br/publicacoes/download/Manual_de_Engenharia_FV_2014.pdf. Acesso em: 28 ago. 2023.

CHU, Y. *et al.* Intra-hour irradiance forecasting techniques for solar power integration: a review. **iScience**, v. 24, n. 10, p. 103-136, 2021.

DA CUNHA, B. A. **Previsão intra-diária de geração fotovoltaica usando redes neurais recorrentes do tipo LSTM e dados históricos de energia.** Tese (doutorado) – Universidade Estadual Paulista (Unesp), Faculdade de Engenharia, Bauru, 2021. 80f.

DAMACENO, L. **Regressão Linear?**. 2020. Disponível em: <https://medium.com/@lauradamaceno/regress%C3%A3o-linear-6a7f247c3e29>. Acesso em 10. Out. 2023.

DANTAS, L. C. **Otimização de redes neurais artificiais de múltiplas camadas utilizando algoritmos genéticos e enxame de partículas.** Dissertação (Mestrado) – Universidade Estadual Paulista (Unesp), Instituto de Geociências e Ciências Exatas, Rio Claro, 2020. 95f.

DE BARROS SILVA, A. W. *et al.* Methodology Based on Artificial Neural Networks for Hourly Forecasting of PV Plants Generation. **IEEE Latin America Transactions**, v. 20, n. 4, p. 659-668, 2021.

ENGELBRECHT, A. P. **Computational intelligence: an introduction.** [S. l.]: John Wiley & Sons, 2007.

FACURE, M. **Funções de ativação: entendendo a importância da ativação correta nas redes neurais.** 2017. Disponível em: <https://matheusfacure.github.io/2017/07/12/activ-func/>. Acesso em: 28 ago. 2023.

FISCHER, D. R. **Desenvolvimento de uma ferramenta para previsão de curto prazo da geração de energia fotovoltaica.** Dissertação (Mestrado) – Universidade Federal de Santa Maria, Centro de Tecnologia, Programa de Pós-graduação em Engenharia Elétrica, 2021.

FRANCISCO, A. C. C. *et al.* Influência de parâmetros meteorológicos na geração de energia em painéis fotovoltaicos: um caso de estudo do Smart Campus Facens, SP, Brasil. **Revista Brasileira de Gestão Urbana**, [S. l.], 2019. Disponível em: <https://doi.org/10.1590/2175-3369.011.e20190027>. Acesso em: 28 ago. 2023.

GAMA, J. *et al.* **Concept drift in decision-tree learning for data streams.** In: Proceedings of the Fourth European Symposium on Intelligent Technologies and their Implementation on Smart Adaptive Systems; Verlag Mainz: Aachen, Germany; p. 218-225, 2004.

GBHAT. **Particle Swarm Optimization on Non-Convex Function.** 2021. Disponível em: https://gbhat.com/machine_learning/pso_nonconvex.html. Acesso em: 28 ago. 2023.

GLENN, H. **How Do Wind and Humidity Affect Solar Panel Efficiency?** 2018. Disponível em: <https://www.solar.com/learn/how-do-wind-and-humidity-affect-solar-panel-efficiency/>. Acesso em: 28 ago. 2023.

GOBBO, E. R.; SILVA, M. A. T. F.; BONE, R. B. **Do petróleo à energia fotovoltaica: a inserção do Brasil neste novo mercado.** In: VII Congresso Brasileiro De Energia Solar. Gramado, 2018.

GUO, W.; JIANG, X.; CHE, L. **Short-Term Photovoltaic Power Forecasting based on Machine Learning**. 2019. IEEE 3rd International Electrical and Energy Conference (CIEEC), p. 1276–1280, 2019.

HEFNY, H.; AZAB, S. Chaotic particle swarm optimization. **The 7th International Conference on Informatics and Systems (INFOS)**, p. 1-8, 2010.

INTERNATIONAL ENERGY AGENCY – IEA. **Key World Statistics 2021**. 2021. Disponível em: <https://www.iea.org/reports/key-world-energy-statistics-2021>. Acesso em: 28 ago. 2023.

ISAKSSON, E.; CONDE, M. K. **Solar Power Forecasting with Machine Learning Techniques**. 2018. Disponível em: <https://www.diva-portal.org/smash/get/diva2:1215661/FULLTEXT02>. Acesso em: 28 ago. 2023.

KENNEDY, J.; EBERHART, R. **Particle swarm optimization**. In: IEEE. PROCEEDINGS of ICNN'95-International Conference on Neural Networks. [S.l.: s.n.], v. 4, p. 1942–1948, 1995.

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

MALIK, F. **What Is Grid Search?** 2020. Disponível em: <https://medium.com/fintechexplained/what-is-grid-search-c01fe886ef0a>. Acesso em: 28 ago. 2023.

MARIO FILHO. **Validação de Séries Temporais em Machine Learning Usando Python**. 2023. Disponível em: <https://mariofilho.com/validacao-de-series-temporais-em-machine-learning-usando-python/>. Acesso em 28. Ago. 2023.

MAYER, M. J.; GRÓF, G. Extensive comparison of physical models for photovoltaic power forecasting. **Applied Energy**, 283, p. 116-239, 2021.

MOREIRA, J. R. S. **Energias renováveis, geração distribuída e eficiência energética**. 1. ed. Rio de Janeiro: LTC, 2019.

MUKILAN, K. *et al.* Prediction of Rooftop Photovoltaic Solar Potential Using Machine Learning. **International Journal of Photoenergy**, 2022. Article ID 1541938.

NOGUEIRA, W. B. **Avaliação de desempenho de uma usina solar fotovoltaica com base no agrupamento k-means**. 2022. 57f. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Elétrica) – Universidade Federal do Ceará, Fortaleza, 2022.

PAIXÃO, J. L. **Avaliação do impacto da microgeração fotovoltaica na rede de distribuição de energia elétrica**. Dissertação (Mestrado em Engenharia Elétrica), UFSM, Santa Maria, 2019.

PELISSON, Ângelo. **Aprendizado de máquina para previsão de geração de energia fotovoltaica em dados de estações solarimétricas**. Dissertação (Mestrado) – Universidade Federal de Santa Catarina, Campus Joinville, Programa de Pós-Graduação em Engenharia de Sistemas Eletrônicos, Joinville, 2021.

PERES, L. **Como detectar e tratar outliers com Python**. 2021. Disponível em: <https://medium.com/@lucapqg/como-detectar-e-tratar-outliers-com-python-ca2cf088c160>. Acesso em 28. Ago. 2023.

PRINCIPE, J. C.; EULIANO, N. R.; LEFEBVRE, W. C. **Neural and adaptive systems: fundamentals through simulations with CD-ROM**. [S. l.]: John Wiley & Sons, Inc., 1999.

QUINLAN, J. R. **Induction of decision trees**. *Machine Learning*, 1:81–106, 1986.

RUMELHART, D. E.; HINTON, G. E; WILLIAMS, R. J. Learning representations by back-propagating errors. *nature*, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986.

RUVIARO, S. R. *et al.* **Análise da variação da eficiência do módulo fotovoltaico em função da temperatura**. In: VII Congresso Brasileiro de Energia Solar, Gramado, 2018, p. 01-08.

SILVA, D. G. Segurança da informação por meio de redes neurais artificiais. *Revista Científica Multidisciplinar Núcleo do Conhecimento*, v. 6, n. 10, p. 7, p. 48-61, 2021.

SILVA, I. N. D.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. 2. ed. São Paulo: Artliber Editora, 2016.

SILVA, P. R. **Estudo do efeito da composição solar no desempenho de módulos FV caracterizados em condições reais de operação**. Dissertação (Mestrado) – Centro Federal de Educação Tecnológica de Minas Gerais, Programa de Pós-Graduação em Engenharia Elétrica em associação ampla com a Universidade Federal de São João Del Rei, Belo Horizonte, 2018.

SMITI, A. A critical overview of outlier detection methods. *Computer Science Review, Elsevier*, v. 38, p. 100-306, 2020.

SOLARPOWER EUROPE. **Global Market Outlook For Solar Power 2023 – 2027**. *Global Market Outlook*, [S. l.: s. n.], 2023.

SU, D.; BATZELIS, E.; PAL, B. **Machine learning algorithms in forecasting of photovoltaic power generation**. In: 2019 International Conference on Smart Energy Systems and Technologies (SEST). IEEE, 2019. p. 1-6.