**UNIVERSIDADE DE SÃO PAULO INSTITUTO DE ENERGIA E
AMBIENTE PROGRAMA DE PÓS-GRADUAÇÃO EM ENERGIA**

WELLINGTON WINÍCIUS FERREIRA FONSECA

**REDE NEURAL RECORRENTE PARA PREVISÃO DE CURTO PRAZO
DA USINA FOTOVOLTAICA DE 540 kWp DA USP**

**SÃO PAULO
2021**

WELLINGTON WINÍCIUS FERREIRA FONSECA

REDE NEURAL RECORRENTE PARA PREVISÃO DE CURTO PRAZO
DA USINA FOTOVOLTAICA DE 540 kW_p DA USP

Dissertação apresentada ao Programa de PósGraduação em Energia do Instituto de Energia e Ambiente da Universidade de São Paulo para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Roberto Zilles
Coorientador: Prof. Dr. Fernando Ramos Martins

Versão Corrigida

SÃO PAULO

2021

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

Ferreira Fonseca, Wellington Winícius.

Rede neural recorrente para previsão de curto da usina fotovoltaica de 540 kWp da USP. / Wellington Winícius Ferreira Fonseca; orientador: Roberto Zilles. – São Paulo, 2021.

106 f.: il; 30 cm.

Dissertação (Mestrado em Ciências) – Programa de Pós-Graduação em Energia – Instituto de Energia e Ambiente da Universidade de São Paulo.

1. Sistemas fotovoltaicos. 2. Energia solar. 3. Redes Neurais I.

Título.

Elaborado por Maria Penha da Silva Oliveira CRB-8/6961

**UNIVERSIDADE DE SÃO PAULO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENERGIA**

Nome: FERREIRA FONSECA, Wellington Winícius

Título: Rede neural recorrente para previsão de curto prazo da usina fotovoltaica de 540 kWp da USP.

Dissertação apresentada ao Programa de Pós-Graduação em Energia do Instituto de Energia e Ambiente da Universidade de São Paulo para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Roberto Zilles

Coorientador: Prof. Dr. Fernando Ramos Martins

Aprovado em: 09/08/2021

Banca Examinadora

Prof. Dr. Adnei Melges de Andrade

Instituição: IEE-USP

Profa. Dra. Rachel Ifanger Albrecht

Instituição: IAG-USP

Dr. Francisco José Lopes de Lima

Instituição: INPE

Prof. Dr. Roberto Zilles

Instituição: IEE-USP

Orientador e Presidente da Comissão

AGRADECIMENTOS

Agradeço primeiramente aos meus pais, Julio Cesar Fonseca e Eliane Mariza Ferreira Fonseca, e a minha parceira, Gabriela Sabra de Albuquerque Choueri, por todo o amor, por todo o apoio, e por sempre estarem ao meu lado.

Agradeço aos meus orientadores, Prof. Dr. Roberto Zilles e Prof. Dr. Fernando Ramos Martins, pelos ensinamentos e pelas experiências compartilhadas. Ao professor Zilles, por me acolher em seu grupo de pesquisa e me dar a oportunidade de desenvolver este trabalho. Ao professor Fernando, pelos anos de aprendizado e companheirismo que se estendem desde minha primeira graduação.

Aos pesquisadores membros do LSF também pelo acolhimento e pelas contribuições, em especial ao Dr. Marcelo Pinho Almeida e ao Eng. Teddy Arturo Melendez pela constante disposição em dar suporte com o sistema de aquisição de dados.

Aos amigos e companheiros do LABMAR/Unifesp pelos bons momentos e conhecimentos compartilhados, que contribuíram para o enriquecimento deste trabalho bem como minha formação profissional.

A CAPES pelo auxílio financeiro que permitiu a execução desta pesquisa.

Por fim, mas não menos importante, aos amigos do IEE e da Unifesp que contribuíram seja compartilhando ideias ou simplesmente tornando a trajetória mais leve e prazerosa.

RESUMO

FERREIRA FONSECA, Wellington Winícius. Rede neural recorrente para previsão de curto prazo da usina fotovoltaica de 540 kWp da USP. 106 f. Dissertação (Mestrado) – Instituto de Energia e Ambiente, Universidade de São Paulo, São Paulo, 2021.

O desenvolvimento de modelos de previsão de energia está entre as principais estratégias para manter um rápido crescimento de energias renováveis intermitentes, como a Solar Fotovoltaica (FV), ao mesmo tempo que auxilia na mitigação de desafios relacionados à gestão de sistemas conectados à rede. Neste trabalho, um método estatístico baseado em Rede Neural Recorrente (RNN) foi desenvolvido para prever uma, duas e três horas futuras da produção de energia de uma usina fotovoltaica conectada à rede de 540 kWp, instalada no Instituto de Energia e Meio Ambiente do Universidade de São Paulo (USP). As entradas padrão do modelo RNN são dados meteorológicos, dados ambientais e série temporal da produção de energia da usina FV. A série histórica foi dividida em meses secos (abril a setembro) e chuvosos (outubro a março). Este trabalho também visa avaliar a influência do aerossol atmosférico como preditor do modelo de previsão. Assim, dados de reanálise de Profundidade Óptica de Aerossol (AOD) interpolados de hora em hora a partir do modelo MERRA-2 foram incluídos como entrada adicional em um segundo modelo RNN. Um algoritmo de força bruta determinou a arquitetura de três camadas ocultas com trinta neurônios cada. A mesma arquitetura foi aplicada para todos os horizontes de tempo avaliados, bem como com ou sem dados AOD. A precisão dos modelos RNN de previsão foi avaliada por meio do erro quadrático médio (MSE), raiz do erro quadrático médio nominal (NRMSE) e Skill Score (SSMSE) em comparação com um modelo de persistência smart. Os modelos desenvolvidos apresentaram RMSE variando entre 4,53% e 7,24% para os meses secos e 5,93% e 8,1% para os meses chuvosos. O método aplicado neste trabalho superou de forma consistente o modelo de persistência para todos os cenários avaliados, especialmente durante os meses chuvosos, com valores de SSMSE observados variando entre 0,67 e 0,92. Finalmente, a inclusão de dados de AOD resultou no aumento do SSRMSE em 2,27% no desempenho do modelo de previsão durante os meses secos.

Palavras-chave: Previsão de geração; Aprendizado de máquinas; Rede Neural Artificial; Energia solar; Sistemas fotovoltaicos.

ABSTRACT

FERREIRA FONSECA, Wellington Winícius. Recurrent Neural Network for short time forecasting for the photovoltaic plant in USP. 106 p. Master of Science Dissertation – Instituto de Energia e Ambiente, Universidade de São Paulo, São Paulo, 2021.

The development of power forecast models is among the main strategies to maintain a rapid growth of intermittent renewable energy, such as Solar Photovoltaics (PV) while supporting to mitigate challenges related to the management of grid-connected systems. In this work, a statistical method based on Recurrent Neural Network (RNN) has been developed to predict one, two, and three hours ahead of the power production of a 540 kWp grid-connected PV plant installed at the Energy and Environment Institute of the University of São Paulo (USP) in Brazil. The default inputs of the RNN are meteorological data, environmental data, and the time series of the power output of the PV plants. The time series was separated into dry months (April to September) and rainy months (October to March). This work also aims to evaluate the influence of atmospheric aerosol as a predictor for the forecast model. Thus, reanalysis data of Aerosol Optic Depth (AOD) interpolated hourly from the MERRA-2 model was included as additional input in a second RNN model. A brute force algorithm determined the architecture of three hidden layers with thirty neurons each. We applied the same architecture for all time horizons evaluated, as well as with or without AOD data. We evaluated the accuracy of the forecast RNN models through Mean Squared Error (MSE), Nominal Root Mean Squared Error (NRMSE), and Skill Score (SS_{MSE}) comparing to a smart persistence model. Both developed models presented a RMSE varying between 4.53% and 7.24% for dry months and between 5.93% and 8.1% for rainy months. The method applied in this work consistently outperformed the reference persistence model for all evaluated scenarios, especially during rainy months. With SS_{MSE} values observed ranges from 0.67 to 0.92. Finally, the inclusion of AOD data improved the model performance SS_{RMSE} by 2.27% during dry months.

Keywords: PV power forecasting; machine learning; artificial neural networks; Solar energy; Photovoltaic systems;

LISTA DE FIGURAS

Figura 1.1 – (a) Potencial de geração solar fotovoltaica anual no Brasil, em kWh/kWp. (b) Potencial de geração solar fotovoltaica anual na Alemanha, em kWh/kWp. Observa-se que as regiões Norte e litoral da região Sul possuem potenciais FV mais baixos do Brasil, com valores próximos a 1.200 kWh/kWp comparáveis à região Alemã com maior potencial. Fonte: Adaptado de GLOBAL SOLAR ATLAS (2021) e PEREIRA et al. (2017)	19
Figura 1.2 – Evolução da capacidade instalada acumulada de geração distribuída solar fotovoltaica no Brasil entre os períodos de 2012 e 2020. Fonte: Adaptado ABSOLAR (2020)	19
Figura 2.1 – Principais componentes de um módulo solar fotovoltaico. Fonte: adaptado de Sunbank (2021)	28
Figura 2.2 – Limite teórico de Shockley-Queisser (linha contínua preta) de eficiência FV em função da banda proibida. As regiões azul, verde e vermelha representam as tecnologias que tenham atingido, respectivamente mais que 75%, entre 50 – 75% e menos que 50% do limite S-Q Fonte: Adaptado de Polman (2016)	29
Figura 2.3 – Ângulos notáveis em solarimetria. (a) ângulo zênite; ângulo solar; ângulo azimute. (b) declinação solar e ângulo horário. Fonte: Adaptado de Pereira et al (2017).	31
Figura 2.4 – Densidade de potência do espectro solar em função do comprimento de onda. Fonte: Zilles et al (2012).	32
Figura 2.5 – Representação gráfica de processo de espalhamento múltiplo envolvendo espalhamento de primeira (P), segunda (Q) e terceira (R) ordem. Fonte: Adaptado de Liou (2002).	33
Figura 2.6 – Componentes da radiação solar. Fonte: VIGNOLA et al (2012).	35
Figura 2.7 – Massa, concentração e raio médios de aerossóis atmosféricos presentes em diferentes ambientes. Fonte: Adaptado Ramachandran (2018).	36
Figura 2.8 – Médias mensais (2001 – 2017) dos valores de AOD nos comprimentos de onda 500nm, 675nm, 870nm e 1020nm para a cidade de São Paulo. Fonte: Adaptado de Bet (2021).	37
Figura 2.9 – Classificação de modelos com base nas resoluções espacial e temporal. Fonte: Adaptado de Diagne et al. (2013).	43

Figura 2.10 – Esquema genérico de uma Rede Neural Artificial que apresenta uma camada oculta. Fonte: Adaptado Srikrishnan (2015).....	44
Figura 2.11 – Modelo de um único neurônio. Fonte: Guarnieri (2006)	45
Figura 2.12 – Representação gráfica das principais funções de ativação aplicadas a modelagem do recurso solar em RNAs. (a) função linear; (b) função logística; e (c) função tangente hiperbólica. Adaptado de John et al (2010).....	45
Figura 2.13 – Modelo de um único neurônio em Rede Neural Recorrente.....	46
Figura 3.1 – Instituto de Energia e Ambiente. Estação solarimétrica (S) e unidades de geração solar fotovoltaica BIPV, CRPV, CTPV e BAPV. Fonte: Adaptado de Google Earth.....	50
Figura 3.2 – Sistema BIPV - Estacionamento IEE-USP.....	51
Figura 3.3 – Sistema BAPV - Biblioteca Brasileira	51
Figura 3.4 – Sistema CRPV - Administração IEE-USP	52
Figura 3.5 – Sistema CTPV - Estrutura fixada no terreno atrás da administração do IEE-USP	52
Figura 3.6 – Normais Climatológicas de Temperatura, Umidade Relativa, Pressão Atmosférica e Velocidade Média dos Ventos na cidade de São Paulo, conforme a base de dados do INMET compreendendo o período entre 1981 e 2010.....	53
Figura 3.7 – Estação solarimétrica disposta no LSF e diagrama esquemático destacando seus equipamentos e sensores. Fonte: Adaptado de Rosa (2003).	54
Figura 3.8 – Análise de Componente Principal (PCA) em solução varimax para os sistemas fotovoltaicos BAPV, BIPV, CTPV e CRPV considerando as variáveis Índice de Claridade (Kt), Temperatura ambiente (Tp), Velocidade dos ventos (V), Profundidade ótica dos aerossóis (AOD), Umidade relativa do ar (U), Ângulo Zenital (Zn) e Pressão atmosférica (Pr).....	58
Figura 3.9 – Diagrama de funcionamento do Algoritmo de Força Bruta.....	62
Figura 3.10 – NRMSE (kWh) do sistema referência CTPV para o Algoritmo de Força Bruta (AFB).....	63
Figura 3.11 – Diagrama da arquitetura escolhida para o modelo RNN de previsão	64
Figura 3.12 – Esquema de nomenclatura dos modelos de previsão de geração FV. O modelo em referência é a previsão de geração no horizonte temporal de 1h a frente na estação seca quando presença de aerossol é empregada para o sistema CTPV. A RNN referência possui três camadas ocultas de trinta neurônios cada.....	64

Figura 4.1 – Regressão de treinamento sistema CTPV em meses secos para previsão de geração 1h no futuro.....	68
Figura 4.2 – a Regressão de treinamento sistema CTPV em meses secos para previsão de geração 2h no futuro.....	68
Figura 4.3 – Regressão de treinamento sistema CTPV em meses secos para previsão de geração 3h no futuro.....	69
Figura 4.4 – Regressão de treinamento sistema BIPV em meses secos para previsão de geração 1h no futuro.....	70
Figura 4.5 – Regressão de treinamento sistema BIPV em meses secos para previsão de geração 2h no futuro.....	71
Figura 4.6 – Regressão de treinamento sistema BIPV em meses secos para previsão de geração 3h no futuro.....	71
Figura 4.7 – Regressão de treinamento sistema BAPV em meses secos para previsão de geração 1h no futuro.....	72
Figura 4.8 – Regressão de treinamento sistema BAPV em meses secos para previsão de geração 2h no futuro.....	73
Figura 4.9 – Regressão de treinamento sistema BAPV em meses secos para previsão de geração 3h no futuro.....	74
Figura 4.10 – Regressão de treinamento sistema CTPV em meses chuvosos para previsão de geração 1h no futuro.	77
Figura 4.11 – Regressão de treinamento sistema CTPV em meses chuvosos para previsão de geração 2h no futuro.	77
Figura 4.12 – Regressão de treinamento sistema CTPV em meses chuvosos para previsão de geração 3h no futuro.	78
Figura 4.13 – Regressão de treinamento sistema BIPV em meses chuvosos para previsão de geração 1h no futuro.	79
Figura 4.14 – Regressão de treinamento sistema BIPV em meses chuvosos para previsão de geração 2h no futuro.	80
Figura 4.15 – Regressão de treinamento sistema BIPV em meses chuvosos para previsão de geração 3h no futuro.	80
Figura 4.16 – Regressão de treinamento sistema BAPV em meses chuvosos para previsão de geração 1h no futuro	81

Figura 4.17 – Regressão de treinamento sistema BAPV em meses chuvosos para previsão de geração 2h no futuro	82
Figura 4.18 – Regressão de treinamento sistema BAPV em meses chuvosos para previsão de geração 3h no futuro	83
Figura 4.19 – Comparação da geração normalizada entre os sistemas de geração FV86	
Figura 4.20 – Diagrama da rotina de seleção da base de dados conjunto IEE.....	87
Figura 4.21 – Regressão de treinamento do modelo conjunto em meses secos para previsão de geração 1h no futuro	88
Figura 4.22 – Regressão de treinamento do modelo conjunto em meses secos para previsão de geração 2h no futuro	89
Figura 4.23 – Regressão de treinamento do modelo conjunto em meses secos para previsão de geração 3h no futuro	90
Figura 4.24 – Regressão de treinamento do modelo conjunto em meses chuvosos para previsão de geração 1h no futuro	91
Figura 4.25 – Regressão de treinamento do modelo conjunto em meses chuvosos para previsão de geração 2h no futuro	91
Figura 4.26 – Regressão de treinamento do modelo conjunto em meses chuvosos para previsão de geração 3h no futuro	92
Figura 4.27 – Validação modelo com dados BIPV inéditos para meses secos em ausência de AOD.....	93
Figura 4.28 – Validação modelo com dados BIPV inéditos para meses secos em presença de dados de AOD.....	93
Figura 4.29 – Validação modelo com dados BIPV inéditos para meses chuvosos em ausência de AOD.....	94
Figura 4.30 – Validação modelo com dados BIPV inéditos para meses chuvosos em presença de dados de AOD.....	94

LISTA DE TABELAS

Tabela 3.1 – Tempo computacional de execução do Algoritmo de Força Bruta	62
Tabela 4.1 – Resumo dos resultados de Erro Quadrático Médio (MSE) em ($\times 10^{-3}$) do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses secos nos horizontes temporais de uma, duas e três horas.	75
Tabela 4.2 – Resumo dos resultados de Raiz do Erro Quadrático Médio Nominal (NRMSE), em kWh, do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses secos nos horizontes temporais de uma, duas e três horas.	76
Tabela 4.3 – Resumo dos resultados de Skill Score (SS_{mse}) do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses secos nos horizontes temporais de uma, duas e três horas.	76
Tabela 4.4 – Resumo dos resultados de Erro Quadrático Médio (MSE) em, $\times 10^{-3}$, do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses chuvosos nos horizontes temporais de uma, duas e três horas.	84
Tabela 4.5 – Resumo dos resultados de Raiz do Erro Quadrático Médio Nominal (NRMSE), em kWh, do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses chuvosos nos horizontes temporais de uma, duas e três horas.	84
Tabela 4.6 – Resumo dos resultados de Skill Score (SS_{mse}) do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses chuvosos nos horizontes temporais de uma, duas e três horas.	85
Tabela 4.7 – Skill Score do desempenho do modelo conjunto IEE em referência ao modelo BIPV	95
Tabela 4.8 – Comparação de desempenho dos modelos de previsão de curto prazo propostos na literatura recente. O desempenho é comparado por meio do RMSE (%) para os horizontes temporais de 1h, 2h e 3h.	96

LISTA DE SIGLAS

AFB – Algoritmo de força bruta
ANFIS – *Adaptative Neurofuzzy Inference System*
AM – Massa atmosférica
ANEEL – Agência Nacional de Energia Elétrica
AOD – Profundidade ótica dos aerossóis
ARIMA – Autoregressive Integrated Moving Average
ARMA – Autoregressive Moving Average
BAPV – Sistema Fotovoltaico Aplicado a Edificação
BIPV – Sistema Fotovoltaico Integrado a Edificação
BP – *Backpropagation*
BPTT – *Backpropagation Through Time*
COP – *Conference of Parties*
CRPV – Sistema Fotovoltaico com Correção do Fator de Potência
CTPV – Sistema Fotovoltaico sobre Terreno
DHI – Irradiância Difusa Horizontal
DL – *Deep Learning*
DNI – Irrância Direta Normal
ELM – Aprendizado de Máquinas Extremo
EPE – Empresa de Pesquisas Energéticas
ETA/CPTEC – Modelo Eta desenvolvido no Centro de Previsão de Tempo e Estudos Climáticos
FV – Fotovoltaico
GD – Geração Distribuída
GEE – Gases do Efeito Estufa
GHI – Irradiância Global Horizontal
IBGE – Instituto Brasileiro de Geografia e Estatística
IEE – Instituto de Energia e Ambiente
INMET – Instituto Nacional de Meteorologia
IV – Intervalo do comprimento de onda Infra-Vermelho
KNN – K-Nearest Neighbor
KT – Índice de Claridade
LSF – Laboratório de Sistemas Fotovoltaicos

LSTM – Long-Short Term Memory

MERRA-2 – *Modern-Era Retrospective Analysis for Research and Application, Ver. 2*

MLP – *Multilayer Perceptron*

MRE – Erro Relativo Médio

MSE – Erro Quadrático Médio

NAN – *Not a Number*

NPU – *National Penghu University*

NRMSE – Raíz do Erro Quadrático Médio Normalizado

NWP – Modelos Numéricos Meteorológicos

PCA – Análise de Componente Principal

PCABT – Potência em corrente alternada em baixa tensão

PNT – Previsão Numérica do Tempo

REN – Resolução Normativa

RMSE – Raíz do Erro Quadrático Médio

RNA – Rede Neural Artificial

RNN – Rede Neural Recorrente

RP – *Resilient Backpropagation*

SARIMA – *Seasonal Autoregressive Integrated Moving Average*

SIN – Sistema Interligado Nacional

S-Q – *Shockly-Queisser Limit*

SS – *Skill Score*

SVM – Máquina de Suporte Vetorial

SVR – Regressão de Suporte Vetorial

TOA – Topo da Atmosfera

USP – Universidade de São Paulo

UV – Intervalo do comprimento de onda Ultra-Violeta

VIS – Intervalo do comprimento de onda da Luz Visível

WPD – *Wavelet Packet Decomposition*

WRF – Modelo numérico *Weather Research and Forecasting*

SUMÁRIO

Capítulo 1 : INTRODUÇÃO	17
1.2. OBJETIVOS	22
1.3. MOTIVAÇÃO E JUSTIFICATIVA	22
1.4. ESTRUTURA CAPITULAR	25
Capítulo 2 : FUNDAMENTAÇÃO TEÓRICA.....	26
2.1. APROVEITAMENTO DO RECURSO SOLAR FOTOVOLTAICO	26
2.2. TECNOLOGIA DE GERAÇÃO FOTOVOLTAICA.....	27
2.3. CONCEITOS FÍSICOS ASSOCIADOS AOS PROCESSOS RADIATIVOS NA ATMOSFERA	30
2.3.1. IRRADIÂNCIA GLOBAL, DIRETA E DIFUSA.....	34
2.4. AEROSSÓIS ATMOSFÉRICOS	35
2.4.1. PROPRIEDADES ÓPTICAS DOS AEROSSÓIS ATMOSFÉRICOS E SUA INFLUÊNCIA NA RADIAÇÃO SOLAR	38
2.5. ÍNDICE DE CLARIDADE (K_t).....	38
2.6. MODELOS DE PREVISÃO DE GERAÇÃO SOLAR FOTOVOLTAICO	40
2.6.1. REDES NEURAIS ARTIFICIAIS.....	43
2.6.2. RESUMO DE PUBLICAÇÕES NACIONAIS E INTERNACIONAIS	47
Capítulo 3 : METODOLOGIA.....	50
3.1. SISTEMAS DE GERAÇÃO SOLAR FOTOVOLTAICA.....	50
3.2. CLIMATOLOGIA LOCAL	52
3.3. BASE DE DADOS OBSERVADOS	53
3.4 DADOS DE PROPRIEDADE ÓPTICA DE AEROSSÓIS	55
3.5. MÉTODO DE TRATAMENTO DOS DADOS	56
3.5.1 TRATAMENTO DE DADOS SOLARIMÉTRICOS.....	56

3.5.2. TRATAMENTO DOS DADOS DOS SISTEMAS FOTOVOLTAICOS	57
3.5.3. NORMALIZAÇÃO DOS DADOS	57
3.5.4. ANÁLISE DE PRÉ-SELEÇÃO DOS PREDITORES	57
3.6. VALIDAÇÃO ESTATÍSTICA	59
3.7. REDES NEURAIS ARTIFICIAIS RECORRENTES	61
Capítulo 4 : RESULTADOS E DISCUSSÃO	66
4.1. MESES SECOS	67
4.1.1. CTPV	67
4.1.2. BIPV	69
4.1.3. BAPV	72
4.2. MESES CHUVOSOS	76
4.2.1. CTPV	76
4.2.2. BIPV	78
4.2.3. BAPV	81
4.5. INFLUÊNCIA DO ÍNDICE DE PROFUNDIDADE ÓTICA DO AEROSSOL ATMOSFÉRICO (AOD) NO DESEMPENHO DOS MODELOS PREDITIVOS	85
4.6. MODELO CONJUNTO IEE	86
4.4.1. DESENVOLVIMENTO E TREINAMENTO	86
4.4.1.1. MESES SECOS	87
4.4.1.2. MESES CHUVOSOS	90
4.4.2. VALIDAÇÃO DO MODELO COM DADOS DE GERAÇÃO INÉDITOS	92
4.7. COMPARAÇÃO DE DESEMPENHO COM MODELOS ESTATÍSTICOS APRESENTADOS NA LITERATURA RECENTE	95
Capítulo 5 : CONCLUSÕES E CONSIDERAÇÕES FINAIS	98
REFERÊNCIAS BIBLIOGRÁFICAS	102

Capítulo 1 : INTRODUÇÃO

1. INTRODUÇÃO

A questão energética é um tópico bastante relevante no momento atual em razão da sua relação intrínseca com a qualidade de vida da sociedade, com a infraestrutura necessária para dar suporte ao desenvolvimento socioeconômico do país e, finalmente, com a questão ambiental associada à variabilidade climática e poluição atmosférica (GOLDEMBERG; LUCON, 2007). Durante as últimas décadas vem sendo extensivamente discutido acerca das mudanças climáticas associadas à ação antrópica, sobretudo acerca das expressivas emissões de gases do efeito estufa (GEE). O mais recente compromisso global com a questão climática foi feito através do Acordo de Paris em 2016, que busca manter o aumento da temperatura média do planeta abaixo de 1,5 °C (COP, 2015). Nesse contexto, há um esforço em escala mundial para aumentar a participação de fontes renováveis de energia, sobretudo os recursos oriundos do fluxo energético solar, direta ou indiretamente. Dentre o aproveitamento das fontes renováveis, a geração de eletricidade solar fotovoltaica é a mais promissora dentre as tecnologias conhecidas atualmente, apresentando a maior densidade de potência e o modo mais apropriado à geração elétrica descentralizada (SMIL, 2019), com potencial de expressivo crescimento ao passo que haja desenvolvimento tecnológico que permita o aumento da eficiência das tecnologias de conversão.

No entanto, apesar da longa discussão acerca das mudanças climáticas e a importância de uma transição energética, aproximadamente 85,2% da matriz energética primária global é oriunda de recursos fósseis (BRITISH PETROLEUM, 2018). Além disso, a parcela renovável da matriz energética primária é fortemente concentrada no aproveitamento dos recursos hídrico e biomassa tradicional, representando 19,9% e 41,4%, respectivamente, do montante renovável de energia (REN21, 2018). Sobretudo em função de suas propriedades e processos de conversão, o aproveitamento de fontes renováveis de energia apresenta maior representatividade no setor elétrico. A matriz elétrica mundial possui uma participação de 26% por fontes renováveis de energia. Contudo, mais da metade desse montante ainda vem das denominadas “velhas renováveis” – hidroeletricidade e biomassa tradicional (REN21, 2018).

O Brasil apresenta uma matriz energética notoriamente renovável, sobretudo se comparada ao restante do mundo. A matriz energética primária brasileira conta com 48,4% de energia produzida a partir de fontes renováveis enquanto a média mundial encontra-se em 13,9% (BRITISH PETROLEUM, 2018; EPE, 2021). A participação de fontes renováveis é ainda mais notável no que tange a geração de eletricidade, representando 84,8% da matriz elétrica brasileira. No entanto, há uma forte concentração da geração hidráulica, responsável por 65,2% da oferta interna de energia elétrica nacional. A fins comparativos, o aproveitamento das energias eólica e solar, que são os recursos energéticos que apresentaram maior taxa de crescimento em termos de capacidade instalada ao longo da presente década, representam 8,8% e 1,7% da matriz elétrica, respectivamente (EPE, 2021). A dependência da geração hidráulica fragiliza a seguridade da matriz elétrica brasileira, tornando-a mais suscetível a períodos de estiagem, como foi observado neste ano (CPRM, 2021). Além disso, existem problemas socioambientais associados a implementação de grandes empreendimentos de geração por meio de hidroeletricidade. Este contexto aponta para a necessidade de diversificação da matriz energética brasileira a fim de garantir a segurança energética nacional. A Empresa de Pesquisa Energética (EPE) estima o aumento da participação de fontes renováveis diversas (PCH, Solar, Eólica e Biomassa) para 30% da matriz elétrica, em contraste com a queda para 52% da participação de empreendimentos hidrelétricos de grande porte no setor elétrico até o ano de 2026 (EPE, 2017).

O Brasil possui um enorme potencial para geração solar fotovoltaica, apresentando irradiação elevada com baixa variabilidade sazonal, se comparado com países como Alemanha, Espanha e Portugal, onde o mercado desta tecnologia está bem estabelecido. (PEREIRA et al., 2017). Mesmo as regiões com menor incidência de radiação solar no Brasil, como o litoral da região sul, apresentam potencial de geração solar fotovoltaica similar às regiões mais propícias ao aproveitamento solar fotovoltaico na Alemanha, conforme demonstra a Figura 1.1(a). Apesar do cenário favorável em termos de condições climatológicas, a energia solar fotovoltaica ainda é pouco explorada no Brasil. Os líderes mundiais em termos de capacidade instalada são atualmente China (~170 GW), Estados Unidos (60 GW), Japão (56 GW), Alemanha (~40 GW) e Índia (~30 GW). Juntos representam cerca de 70% da capacidade instalada mundial que encontra-se na ordem de 505 GW (REN21, 2018).

Figura 1.1 – (a) Potencial de geração solar fotovoltaica anual no Brasil, em kWh/kWp. (b) Potencial de geração solar fotovoltaica anual na Alemanha, em kWh/kWp. Observa-se que as regiões Norte e litoral da região Sul possuem potenciais FV mais baixos do Brasil, com valores próximos a 1.200 kWh/kWp comparáveis à região Alemã com maior potencial. Fonte: Adaptado de GLOBAL SOLAR ATLAS (2021) e PEREIRA et al. (2017)

De forma geral, a geração solar fotovoltaica é dividida em duas categorias principais: centralizada e distribuída. A geração centralizada consiste em projetos de maior porte, comumente implantados nas regiões Nordeste, Centro-Oeste e Sudeste do país, que apresentam os maiores valores médios de irradiação direta (PEREIRA et al., 2017). A geração distribuída (GD), por sua vez, é a geração elétrica de pequena escala realizada próximas aos locais de consumo, comumente em telhados e propriedades privadas de consumidores. Essa forma de geração está em acelerada expansão pelo território nacional, conforme demonstra a Figura 1.2. Em 2020, a capacidade instalada GD superou a centralizada, em termos cumulativos (ABSOLAR, 2020).

Figura 1.2 – Evolução da capacidade instalada acumulada de geração distribuída solar fotovoltaica no Brasil entre os períodos de 2012 e 2020. Fonte: Adaptado ABSOLAR (2020)

O sistema brasileiro de geração solar fotovoltaica centralizada registra até final de junho de 2021 com uma capacidade instalada outorgada de 3,3 GW e mais 2,5 GW encontram-se em construção (ANEEL, 2021). Por sua vez, a GD conta com uma capacidade instalada de aproximadamente 4 GW, por meio de 335.536 sistemas FV conectados à rede (ABSOLAR, 2020). O rápido crescimento observado a partir de 2012 é reflexo dos incentivos regulatórios implementados por meio da REN 482/2012 e REN 687/2015 que tornaram o investimento em sistemas fotovoltaicos mais atrativo para o consumidor. Essas resoluções normativas implementam o mecanismo de compensação de energia elétrica e a possibilidade de utilização de créditos de energia gerados em outras propriedades do consumidor, desde que se encontrem dentro da área de concessão da distribuidora. A continuidade do rápido crescimento no número de sistemas FV GD instalados, no futuro próximo, é atualmente incerta. Nos últimos anos houveram discussões acerca da revisão da REN 482/2012, propondo a redução de compensação para 38% da energia injetada na rede pelos prosumidores. Ainda, está em trâmite no Congresso Nacional o PL 5.829/2019, que visa a criação do Marco Legal da GD, que é bem visto por empreendedores de sistemas fotovoltaicos GD. A discussão aprofundada do ponto de vista legal e tributário foge do escopo deste trabalho, mas decidimos por ao menos mencionar a situação visto a relevância do assunto.

A viabilidade econômica de projetos fotovoltaicos, sobretudo a geração distribuída, é mais facilmente atingida com a redução do custo dos componentes dos módulos fotovoltaicos, bem como quanto maior for a tarifação de eletricidade convencional da distribuidora na localidade de implantação do sistema de geração (PEREIRA et al., 2017). A avaliação da viabilidade econômica da geração fotovoltaica requer o correto dimensionamento do projeto para atender a demanda energética do usuário da instalação.

O conhecimento da irradiação incidente é necessário para o dimensionamento do projeto do sistema fotovoltaico, bem como a capacidade de produção que o mesmo terá em um dado período de tempo. A disponibilidade de radiação está intimamente ligada às condições ambientais e atmosféricas do local de interesse. A avaliação da irradiação solar incidente na superfície pode ser feita em estações de coleta de dados em pontos isolados e através simulação numérica com uso de modelos de transferência radiativa para os demais pontos do território. A utilização de modelos numéricos como o BRASIL-SR, por exemplo, é fundamental para o mapeamento da irradiação pelo território

nacional. O Atlas Brasileiro de Energia Solar (PEREIRA et al., 2017) apresenta as estimativas para variáveis relevantes ao aproveitamento do recurso solar para todo o país. A disponibilidade deste tipo de informação diagnóstica dá suporte ao setor público e a iniciativa privada na elaboração de novos projetos, que visam atender à crescente demanda energética e garantir a segurança da matriz elétrica nacional. No entanto, o monitoramento das variáveis meteorológicas e radiométricas são essenciais no local de interesse para a instalação de uma usina fotovoltaica em razão da redução de incertezas numéricas associadas à modelagem numérica.

Nesse sentido, é interessante utilizar um modelo numérico que analise a variabilidade temporal do recurso solar para geração fotovoltaica em determinada localidade. Durante a etapa de projeto, ferramentas como Atlas Brasileiro de Energia Solar auxiliam no dimensionamento da usina de geração, permitindo calcular o rendimento em função do tempo de operação e, considerando os investimentos com linhas de transmissão e outros custos pertinentes a implementação e a viabilidade econômica do projeto. Após a instalação da planta FV, modelos de previsão de geração servem como suporte na operação da usina e, também, auxiliam o operador do sistema elétrico.

Na literatura existem alguns trabalhos que se debruçaram sobre a temática de previsão de geração fotovoltaica, dos quais alguns serão brevemente descritos e discutidos a seguir. Um emprego comum de modelos numéricos é prever a geração de um sistema fotovoltaico em determinada localidade em função da produção gerada em plantas de referência. Este tipo de modelo demonstra certas limitações em função das particularidades de cada sistema, como sombreamento local e posicionamento dos módulos fotovoltaicos (LORENZ et al., 2009). Nesse tocante, outros trabalhos, aprimoraram esse tipo de modelo ao considerar a orientação do sistema em função do azimute e fazendo uso de modelos de decomposição e transposição das componentes direta e difusa da irradiação solar incidente na superfície (KILLINGER et al., 2016). Outros trabalhos introduzem o conceito de índices de céu claro para estimar a geração nos demais dias, além de modelar os valores de irradiação em função desses parâmetros (ENGERER; MILLS, 2014). Diferentemente, o presente projeto propõe modelar a geração fotovoltaica em função de uma base de dados de relativamente longo prazo, consistindo de variáveis atmosféricas, meteorológicas e de irradiação solar. A

base de dados será usada para alimentar um modelo estatístico a ser desenvolvido que utilizará uma técnica de aprendizado de máquinas.

1.2. OBJETIVOS

Esta dissertação de Mestrado teve por finalidade desenvolver uma metodologia para previsão de geração fotovoltaica baseado em aprendizado de máquina para uma usina em operação em região metropolitana sujeita sazonalmente a elevada carga de aerossóis na atmosfera. Séries temporais de dados meteorológicos e dados de irradiação solar incidente na superfície coletados no local da usina, assim como a série pretérita de dados de geração da planta fotovoltaica serão utilizadas para o desenvolvimento da metodologia.

Os objetivos específicos dessa pesquisa são:

1. Identificar os preditores mais relevantes para o desenvolvimento do modelo dentre o conjunto de dados de irradiação solar incidente, espessura ótica de aerossóis atmosféricos e variáveis meteorológicas com uso de análise de componentes principais;
2. Caracterizar a topologia do método de aprendizado de máquina, estabelecendo o número de camadas ocultas e o número de neurônios aplicados na Rede Neural Artificial;
3. Executar o treinamento do modelo de previsão de geração fazendo uso do conjunto de dados selecionados como mais relevantes;
4. Identificar as incertezas e a confiabilidade do modelo de previsão de geração solar fotovoltaica. Avaliar as condições em que o modelo de previsão de geração apresenta os melhores resultados;
5. Automatizar o modelo de previsão de geração solar fotovoltaica a partir de dados desconhecidos pelo modelo.

1.3. MOTIVAÇÃO E JUSTIFICATIVA

O aproveitamento do recurso solar fotovoltaico está intimamente ligado a um conjunto de variáveis ambientais, atmosféricas, meteorológicas e da radiação solar incidente. Algumas dessas variáveis comportam-se ciclicamente ao longo do dia e do ano, de tal forma que seu comportamento pode ser conhecido temporalmente através da análise de uma série histórica ou dos movimentos de rotação e translação da Terra,

como é o caso da posição relativa do Sol. Em contrapartida, alguns parâmetros climatológicos apresentam uma variabilidade espaço-temporal que impõe maior dificuldade de previsão, como é o caso da cobertura de nuvens. Esses parâmetros podem influenciar consideravelmente o desempenho do sistema fotovoltaico, sobretudo quanto à quantidade de energia elétrica gerada, como também podem influenciar a demanda por energia.

Como já foi abordado anteriormente, o aproveitamento de fontes energéticas de caráter intermitente vem ganhando participação na matriz elétrica brasileira nos últimos anos. Assim, o conhecimento detalhado das variáveis supracitadas configura-se de interesse para o planejamento da matriz energética brasileira. Nesse sentido, o desenvolvimento de modelos de previsão de geração elétrica tem o potencial de posar como uma importante ferramenta para dar suporte ao crescimento da penetração de geração fotovoltaica no Sistema Interligado Nacional (SIN).

A quantidade de eletricidade gerada por um módulo solar fotovoltaico está diretamente relacionada a quantidade de radiação solar incidente no mesmo, composta por suas componentes direta e difusa. Os principais agentes moduladores da radiação solar incidente são a cobertura efetiva de nuvens e a presença de aerossóis quando em condição de céu sem nuvens. A todo momento o planeta Terra apresenta entre 30 e 50% de cobertura de nuvens, variando em espessura ótica e altitudes (MARTINS, 2001). A influência das nuvens na transmitância radiativa é dada sobretudo pelo processo de espalhamento da radiação, que por sua vez depende da espessura ótica das nuvens associada com a distribuição de tamanhos das gotículas, conteúdo e do estado físico da água (PEREIRA et al., 2017). A distribuição das nuvens é irregular, variando horizontalmente e verticalmente em função à circulação atmosférica, distribuição de oceanos e porção continental, bem como particularidades climatológicas regionais. Portanto, a nebulosidade é o principal fator na variabilidade e intermitência do recurso energético solar, dificultando assim sua previsibilidade.

O segundo principal agente modulador da radiação solar incidente corresponde aos aerossóis atmosféricos. Por definição, aerossóis correspondem um conjunto de partículas sólidas ou líquidas em suspensão em um gás (HORVATH, 2000; RAMACHANDRAN, 2018). As partículas em suspensão na atmosfera terrestre podem absorver ou espalhar a radiação solar incidente, a depender de suas propriedades

ópticas, contribuindo, assim, para a irregularidade da distribuição da radiação solar na superfície terrestre. Os aerossóis presentes na atmosfera podem ser originados por processos naturais (como arraste eólico de sedimentos erodidos, por exemplo) ou por ação antrópica, sobretudo através da queima de biomassa e combustíveis fósseis. Além disso, pode haver também o transporte atmosférico oriundo de outras regiões. No caso do Brasil isso ocorre durante estação seca, quando a frequência de ocorrência de queimadas em regiões agrícolas cresce. A falta de chuva reduz a retirada dos particulados da atmosfera por deposição úmida, o que viabiliza o transporte de aerossóis por distâncias maiores. Nesse sentido, é coerente ponderar que locais com maior concentração urbana apresentarão, também, maior quantidade de aerossóis presentes na atmosfera. É interessante apontar que as regiões Sudeste, Centro-Oeste e Nordeste do Brasil apresentam coincidentemente o maior número de cidades com mais de 100.000 habitantes, além de apresentar os maiores potenciais de geração solar fotovoltaica, conforme demonstra a Figura 1.1 (a). Portanto, a análise da presença de aerossóis deve ser incorporada no desenvolvimento de um modelo de previsão de geração solar fotovoltaica.

Uma das principais barreiras da implementação em larga escala de aproveitamento solar fotovoltaico é a intermitência da energia solar disponível (LUIZ et al., 2018a). A intermitência da irradiância provoca variações significativas de corrente elétrica nas linhas de transmissão, tornando-se um desafio para o planejamento e operação do sistema elétrico, visto que a rápida variação da carga nas linhas de transmissão acarreta em maiores gastos com manutenção (ALMEIDA, 2017; LUIZ et al., 2018a). Uma forma de atenuar significativamente as flutuações da geração fotovoltaica é através da dispersão geográfica de sistemas fotovoltaicos (ALMEIDA, 2017). É importante ressaltar, também, que um sistema elétrico robusto, como o SIN Brasileiro, suporta uma maior participação de fontes intermitentes de geração elétrica. No entanto, essa característica não reduz a importância de se adotar técnicas que atenuem os impactos negativos da intermitência no sistema elétrico, destacando-se métodos de armazenamento de energia e, também, técnicas de previsão de geração elétrica.

Nesse contexto, o desenvolvimento de modelos numéricos da previsão de geração solar fotovoltaica permite evitar situações que apresentem riscos ao sistema elétrico, bem como auxilia na redução de perdas e custos operacionais oriundos a

medidas de mitigação (ALMEIDA, 2017). Outra vantagem da adoção de métodos preditivos é que estes permitem avaliar a geração elétrica em função da distribuição espacial do recurso solar e, portanto, auxiliam tanto no planejamento energético, bem como nas tomadas de decisão do Operador Nacional do Sistema Elétrico e também dos produtores de energia.

1.4. ESTRUTURA CAPITULAR

O presente capítulo tem por finalidade apresentar o contexto, relevância, motivações e principais objetivos desta dissertação de mestrado. O segundo capítulo apresenta a fundamentação teórica, apresentando conceitos básicos acerca da tecnologia FV, interações entre recurso solar e atmosfera e o ferramental teórico sobre modelos de previsão de geração FV. O terceiro capítulo apresenta a área de estudo e a metodologia empregada no desenvolvimento do modelo de aprendizado de máquinas. O quarto capítulo apresenta os resultados e as discussões. Por fim, o quinto capítulo apresenta as principais conclusões e considerações finais, bem como apresenta sugestões para trabalhos futuros a fim de dar seguimento a este estudo.

Capítulo 2 : FUNDAMENTAÇÃO TEÓRICA

2.1. APROVEITAMENTO DO RECURSO SOLAR FOTOVOLTAICO

Nas últimas décadas, a comunidade científica e a sociedade têm discutido extensivamente sobre as mudanças climáticas associadas à ação antrópica, principalmente no que se refere ao aumento das emissões de gases de efeito estufa (GEE). O mais recente compromisso global com o clima visa limitar o aumento da temperatura média do planeta à 1,5 °C (COP, 2015). Nesse contexto, há um esforço mundial em curso para aumentar a participação de fontes renováveis de energia, principalmente de recursos intermitentes. Dentre estes, a geração de eletricidade solar fotovoltaica demonstra-se como a mais promissora dentre as tecnologias conhecidas atualmente, apresentando a maior densidade de potência e o modo mais apropriado à geração elétrica descentralizada (SMIL, 2019).

O Brasil possui um grande potencial solar fotovoltaico em todo seu território, com baixa variabilidade ao longo do ano (PEREIRA et al., 2017). Nos últimos anos, a geração de energia fotovoltaica distribuída conectada à rede tem crescido rapidamente no Brasil devido a uma redução considerável nos custos de implementação e ao interesse geral da população em fontes de energia renováveis (VILAÇA GOMES et al., 2018). No entanto, a implementação em larga escala e de forma rápida de fontes de energia intermitentes impõe desafios para os operadores e integradores de fornecimento de energia. Para que possamos discutir e propor soluções para o aumento da participação de tecnologia solar fotovoltaica na matriz elétrica brasileira e mundial, é preciso apresentar conceitos teóricos referentes a disponibilidade do recurso solar que incide na superfície terrestre e conceitos técnicos inerentes a implementação desta tecnologia.

O presente capítulo visa apresentar de forma sucinta o ferramental teórico básico a fim de embasar as discussões apresentadas ao longo desta dissertação. A literatura disponível é bastante rica e os conceitos teóricos são profundamente examinados nas obras referenciadas neste capítulo, sobretudo em relação aos processos de fluxo radiativo e processos físico-químicos da atmosfera terrestre. Para um entendimento aprofundado dos conceitos técnicos e teóricos dos sistemas fotovoltaicos, recomendamos sobretudo a consulta das obras publicadas por Zilles et al. (2012) e para

consulta aprofundada dos conceitos teóricos da interação radiação e atmosfera recomendamos as obras publicadas por Iqbal (1983), Liou (2002), Jacobson (2005), e Ramachandran (2018).

2.2. TECNOLOGIA DE GERAÇÃO FOTOVOLTAICA

O efeito fotovoltaico é um fenômeno físico que permite a conversão da energia solar em eletricidade. Este fenômeno ocorre quando a radiação eletromagnética do Sol incide em materiais semicondutores, como o Silício (VILLALVA, 2012). Átomos de silício possuem quatro elétrons na camada de valência e formam ligações covalentes com outros átomos de silício próximos. O resultado é uma rede cristalina, que uma estrutura bastante ordenada. As células solares de silício são constituídas com duas regiões em que uma delas contém átomos de boro e outra que contém átomos de fósforo. Essa configuração é denominada célula de junção P-N. A região constituída com adição de átomos de boro é chamada de região tipo-P. A região que contém átomos de fósforo é chamada de tipo-N. Átomos de silício têm 4 elétrons na camada de valência, ao passo que átomos de boro e de fósforo têm 3 e 5 elétrons, respectivamente. Na interface entre as regiões P e N ocorre migração de elétrons dos átomos de fósforo (valência 5) que vão completar as ligações dos átomos de boro (valência 3) com os átomos de silício, que se ligam compartilhando átomos com seus vizinhos em ligações químicas estáveis conhecidas como covalentes. O resultado é o estabelecimento de um campo elétrico interno na interface das regiões P e N. É esse campo que irá dar lugar à corrente elétrica quando os fótons da luz incidente têm sua energia absorvida pelo material semicondutor. O processo funciona porque no silício cristalino, a energia necessária para que elétrons do material possam fazer parte da corrente elétrica é da mesma ordem da energia contida nos fótons da luz visível.

A Figura 2.1 ilustra os principais componentes de um módulo fotovoltaico. As células solares são apenas parte do módulo FV. Os demais componentes servem para dar resistência mecânica e suporte estrutural para o módulo, como é caso da cobertura frontal (normalmente feita de vidro) e a moldura (normalmente de alumínio). O material encapsulante, por sua vez, promove a junção entre as camadas do módulo e auxilia no aumento da vida útil das células solares. A cobertura posterior, comumente chamada de *backsheet*, isola o módulo eletricamente e também o protege da umidade. A depender da tecnologia empregada o *backsheet* pode ser substituído por outra superfície de vidro,

como no caso dos módulos bifaciais que estão ganhando participação no mercado recentemente (DULLWEBER et al., 2015; HENG et al., 2015; HEZEL, 2003; JOGE et al., 2002; KHAN; ALAM, 2015; ROSA-CLOT; TINA, 2018). Por fim, a caixa de junção é onde são feitas as conexões elétricas, permitindo associar módulos em série ou em paralelo, além de que é nela que é instalado os diodos de proteção (ou diodos de *bypass*), que permitem o módulo FV continuar gerando corrente elétrica caso este seja parcialmente coberto (por sombreamento ou fisicamente por sujeira ou objetos).

Figura 2.1 – Principais componentes de um módulo solar fotovoltaico. Fonte: adaptado de Sunbank (2021)

As características elétricas das células solares podem variar de acordo com fatores intrínsecos e extrínsecos. Fatores relacionados ao processo de fabricação, material semicondutor, as resistências internas (série e paralelo) das células, seletividade de absorção do espectro solar estão dentre os fatores intrínsecos que dependem da tecnologia FV empregada. Dentre os fatores extrínsecos, destaca-se fatores ambientais como a irradiância solar e a temperatura das células (ZILLES et al., 2012).

O aproveitamento solar fotovoltaico, apesar da grande evidência nas últimas décadas, não é uma tecnologia propriamente recente. Evidentemente, desde seu desenvolvimento diferentes tecnologias foram desenvolvidas e consequentemente aprimoradas. Além dos módulos fotovoltaicos de silício mono e poli-cristalino, que juntos representam cerca de 90% do mercado FV atual (POLMAN et al., 2016), existem outras tecnologias já maduras e com aplicações em larga escala como CIGS (Cobre, Índio,

Gálio e Selênio) e CdTe (Cádmio e Telúrio), como também tecnologias emergentes e aplicadas em pequena escala como silício de filme fino, *dye-sensitized* TiO₂, e células solares orgânicas (POLMAN et al., 2016). A Figura 2.2 demonstra a eficiência recorde de diferentes tecnologias FV. Nela, a linha preta representa o limite teórico de eficiência em função da banda proibida (*band gap*), chamado de limite de Shockley-Queisser (S-Q). As duas linhas cinzas delimitam os limites de 75% e 50% do limite S-Q, respectivamente.

Figura 2.2 – Limite teórico de Shockley-Queisser (linha contínua preta) de eficiência FV em função da banda proibida. As regiões azul, verde e vermelha representam as tecnologias que tenham atingido, respectivamente mais que 75%, entre 50 – 75% e menos que 50% do limite S-Q. Fonte: Adaptado de Polman (2016)

Os módulos fotovoltaicos sílico-cristalinos modernos apresentam um rendimento laboratorial entre 24 e 26%, ainda havendo espaço para aprimoramentos, visto que o limite teórico de eficiência é estimado em aproximadamente 33,7% para o *band gap* de operação destas tecnologias. No entanto, os módulos fotovoltaicos de silício mono e policristalino aplicados em projetos de larga escala e encontrados no mercado apresentam eficiências na ordem entre 18,5 e 22,4% (BLANKENSHIP et al., 2011; POLMAN et al., 2016). O desenvolvimento de novas tecnologias, bem como o aprimoramento das eficiências das já em aplicação, é de suma importância para o aumento da participação de energia solar fotovoltaica na matriz energética global. Mesmo pequenos avanços em termos de gerar mais energia elétrica fazendo uso de menos material e menor cobertura de solo podem contribuir em muito nos esforços de

uma possível transição energética em direção as fontes renováveis de energia, na qual a energia solar fotovoltaica é vista como protagonista. Esta discussão em particular é bastante extensa e foge do escopo deste trabalho, mas este é um ponto que vale a pena ser ao menos mencionado.

2.3. CONCEITOS FÍSICOS ASSOCIADOS AOS PROCESSOS RADIATIVOS NA ATMOSFERA

A energia irradiada pelo Sol é resultado de processos da fusão nuclear de átomos de Hidrogênio (H) formando átomos de Hélio (He), sendo propulsora de todos os processos térmicos, dinâmicos e químicos de nosso planeta. A disponibilidade do recurso solar que incide na Terra, bem como sua variação espaço-temporal, é fruto da posição relativa entre estes dois corpos. A órbita terrestre, elíptica, varia entre $1,47 \times 10^8$ e $1,52 \times 10^8$ km, fazendo com que a irradiância no topo da atmosfera varie entre 1.412 W/m^2 e 1.325 W/m^2 , respectivamente. A irradiância média no topo da atmosfera é de 1.366 W/m^2 , sendo esta denominada constante solar (JACOBSON, 2005; LIOU, 2002; VIGNOLA; MICHALSKY; STOFFEL, 2012)

Juntamente com movimento de translação, a rotação terrestre em torno de um eixo que apresenta inclinação de $23,45^\circ$ em relação ao plano da órbita terrestre configura a geometria Sol-Terra que irá determinar as estações do ano, duração dos dias e consequentemente a disponibilidade do recurso solar no topo da atmosfera (TOA). A radiação solar incidente em nosso planeta apresenta uma variação espaço-temporal diária e sazonal bem definida e conhecida, porém complexa. Para entendê-la é preciso utilizar conceitos geométricos importantes que iremos discutir brevemente a seguir e demonstrar o equacionamento nas próximas subseções deste capítulo. A Figura 2.3 ilustra os ângulos notáveis em solarimetria, descritos a seguir.

- **Declinação solar (δ):** é o ângulo formado entre a linha equatorial e a linha Sol-Terra variando entre $-23,45^\circ$ e $+23,45^\circ$ ao longo do ano. Os valores negativos são dados quando a linha Sol-Terra cruza o hemisfério sul, e o oposto ocorre para o hemisfério norte.
- **Ângulo horário (ω):** é o ângulo de deslocamento aparente formado pela rotação terrestre, variando entre -180° e $+180^\circ$ ao longo do dia. Os valores positivos são atribuídos pela manhã e os negativos pela tarde.

- **Ângulo zênite (θ_z):** é o ângulo entre a vertical e a linha Sol-Terra formado no ponto de observação.
- **Ângulo azimute:** é o ângulo entre a projeção da linha Sol-Terra no plano horizontal e o meridiano do observador, variando entre -180° e $+180^\circ$. Os valores positivos são dados no sentido horário a partir do norte geográfico.

Figura 2.3 – Ângulos notáveis em solarimetria. (a) ângulo zênite; ângulo solar; ângulo azimute. (b) declinação solar e ângulo horário. Fonte: Adaptado de Pereira et al (2017).

O Sol apresenta uma distribuição espectral parecida com a de um corpo negro a de um corpo negro a 5800K (IQBAL, 1983; JACOBSON, 2005). Em geral, o espectro de radiação solar é dividido em três faixas espectrais. Considerando dos comprimentos de onda mais curto ao mais longo, temos: região ultravioleta (UV) com comprimentos de onda inferiores a 400 nm, onde estão situados os raios UV; a região visível ao humano (VIS) entre 400nm e 750nm; e a região infravermelha (IV) com comprimentos de onda mais longos acima de 750 nm. O espectro de luz visível, apesar de uma pequena fatia da totalidade da distribuição espectral solar, apresenta os maiores valores de irradiância solar, concentrando aproximadamente 48% da energia incidente no topo da atmosfera (LIOU, 2002) .

A radiação solar incidente no topo da atmosfera terrestre é bastante semelhante à da distribuição espectral teórica de um corpo negro à temperatura próxima a do Sol. Sendo assim as perdas durante a propagação entre Sol-Terra são reduzidas. O mesmo não ocorre, no entanto, quando a radiação solar percorre a atmosfera terrestre até a superfície. Cerca de 30% da radiação incidente é refletida de volta ao espaço pelo TOA. Desse montante, nuvens e aerossóis são responsáveis por dois terços da radiação solar

refletida (RAMACHANDRAN, 2018). O restante da radiação incidente, com comprimento de onda curto, é absorvido pela superfície terrestre e pela atmosfera, sendo posteriormente reemitidos na forma de onda longa (PEREIRA et al., 2017; RAMACHANDRAN, 2018). A Figura 2.4 compara a densidade de potência ($\text{kW} \cdot \text{m}^{-2} \cdot \text{nm}^{-1}$) de um corpo negro a 6.000 K, no topo da atmosfera ($AM = 0$) e na superfície ($AM = 1,5$) em função dos diferentes comprimentos de onda. Onde AM é uma grandeza relacionada à massa óptica da atmosfera, representando o caminho ótico do fluxo radiativo em interação com a atmosfera. Sendo $AM = 1$ quando o ângulo zenital é igual a zero, que ocorre ao meio dia solar.

Figura 2.4 – Densidade de potência do espectro solar em função do comprimento de onda. Fonte: Zilles et al (2012).

Os processos de atenuação da radiação solar na atmosfera envolvem processos de absorção e espalhamento que ocorrem em função da composição gasosa da atmosfera, incluindo o conteúdo de água em nuvens e os aerossóis atmosféricos. A dispersão, ou espalhamento, é um processo físico fundamental associado à luz e sua interação com a matéria, ocorrendo em todos os comprimentos de onda ao longo do espectro eletromagnético. Este fenômeno é causado pela interação do fluxo radiativo com partículas com diâmetros variáveis que vão desde moléculas de gás ($10^{-4} \mu\text{m}$) a aerossóis ($1 \mu\text{m}$), gotículas de água ($10 \mu\text{m}$), cristais de gelo ($100 \mu\text{m}$) e gotas maiores d'água (1cm) presentes na atmosfera (LIOU, 2002).

Quando em presença de muitas partículas, cada uma destas está sujeita à interação com feixes de radiação direta ou mesmo feixes que tenham sido dispersos por outra partícula anteriormente. Uma partícula quando interage com a luz, espalha esta para todas as direções, provocando um espalhamento primário. Um destes feixes espalhados pode atingir uma segunda partícula, que por sua vez o espalha por todas as direções, provocando um espalhamento secundário. Este pode atingir uma terceira partícula, e assim por diante. Quando o processo de espalhamento de um feixe de luz envolve mais que uma partícula, este é denominado espalhamento múltiplo. Este fenômeno é um processo importante de transferência radiativa de energia na atmosfera terrestre, principalmente quando há presença de aerossóis e nuvens (LIOU, 2002). A Figura 2.5 ilustra o fenômeno da dispersão múltipla, no qual um feixe propagado na direção 'd' interage com partículas P, Q e R.

Figura 2.5 – Representação gráfica de processo de espalhamento múltiplo envolvendo espalhamento de primeira (P), segunda (Q) e terceira (R) ordem. Fonte: Adaptado de Liou (2002).

Outro processo importante no processo de atenuação do fluxo radiativo é o processo de absorção, que ocorre acompanhado da dispersão. A absorção, como o nome sugere, envolve a absorção da energia contida no fluxo radiativo, atenuando a quantidade de energia presente neste. Na atmosfera terrestre ocorre pouca absorção no espectro de luz visível. Ambas absorção e dispersão atenuam o fluxo radiativo, e esta atenuação é denominada extinção, que é resultado do conjunto destes dois processos. Estes fenômenos ocorrem ao longo de todo o espectro eletromagnético, mas como observado na Figura 2.4, são mais acentuados para comprimentos de onda compatíveis

aos principais gases do efeito estufa, que são: água, ozônio e dióxido de carbono, entre outros.

2.3.1. IRRADIÂNCIA GLOBAL, DIRETA E DIFUSA.

Como discutimos na seção anterior, a radiação solar ao propagar-se através atmosfera terrestre está sujeita a interagir com sua composição, e consequentemente ter sua energia atenuada. A quantidade de energia que chega a superfície terrestre é dinâmica e varia em função do tempo (posição do Sol) e da condição atmosférica (presença de nuvens e aerossóis).

Quando em condição de céu claro, ou seja, sem presença de nuvens e ao meio dia solar, cerca de 25% da radiação solar é absorvida e dispersa pela atmosfera, enquanto cerca de 1.000 W/m^2 da irradiância solar atinge a superfície terrestre sem sofrer uma atenuação considerável (VIGNOLA; MICHALSKY; STOFFEL, 2012). Este feixe de irradiação vindo diretamente do Sol é denominado Irradiância Direta Normal (DNI) . Do montante de radiação solar sofre interação com a atmosfera, parte é refletida de volta ao espaço enquanto o restante chega a superfície terrestre de forma atenuada em direções distintas daquela do feixe de radiação solar. Estes feixes dispersos formam a componente denominada Irradiância Difusa Horizontal (DHI). A DHI também pode ser refletida pela superfície, tendo sua intensidade definida pelo tipo de cobertura do solo. O somatório dessas duas componentes resulta no total de radiação solar que atinge a superfície no plano horizontal, sendo denominada Irradiância Global Horizontal (GHI). A Figura 2.6 ilustra as componentes da irradiância solar, demonstrando diferentes formas de espalhamento que a DHI pode apresentar.

A irradiância solar na superfície alcança valor máximo quando a superfície está perpendicular à direção de propagação do feixe de radiação solar. Quando o ângulo de incidência é diferente de 90° , a quantidade de energia é reduzida pelo produto do cosseno do ângulo formado pela normal e a inclinação de incidência da radiação solar.

Figura 2.6 – Componentes da radiação solar. Fonte: VIGNOLA et al (2012)

2.4. AEROSSÓIS ATMOSFÉRICOS

Por definição, aerossóis atmosféricos são uma mistura de partículas sólidas e líquidas suspensas em meio ao ar. Suas propriedades físicas (como tamanho e forma) e químicas são bastante variáveis e possuem grande influência em características atmosféricas como visibilidade, balanço de radiação, poluição do ar, entre outros (RAMACHANDRAN, 2018). O tamanho destas partículas varia entre $0,001\ \mu\text{m}$ até $100\ \mu\text{m}$, sendo divididos em três categorias principais: núcleos de Aitken ($0,001\ \mu\text{m} < r < 0,1\ \mu\text{m}$); partículas grandes ($0,1\ \mu\text{m} < r < 1\ \mu\text{m}$); e partículas gigantes ($r > 1\ \mu\text{m}$). Partículas com raio menor que $1\ \mu\text{m}$ são formadas principalmente por processos de conversão de gás para partículas. Por sua vez, partículas maiores que $1\ \mu\text{m}$ são formadas por processos mecânicos como arraste de poeira, sal marinho ou pólen. Podemos também agrupar os aerossóis atmosférico em moda fina ($< 1\ \mu\text{m}$) e moda grossa ($> 1\ \mu\text{m}$). A fonte da partícula irá determinar sua composição química e o mecanismo de formação irá determinar seu tamanho e formato (RAMACHANDRAN, 2018).

Nas seções anteriores discutimos brevemente os processos de atenuação do fluxo radiativo solar, nos quais a cobertura de nuvens tem um papel fundamental na modulação do recurso solar incidente. A presença de aerossóis atmosféricos é o segundo principal modulador da irradiância solar incidente na superfície, tendo sua importância destacada em condição atmosférica de céu claro (MARTINS, 2001). Levando em consideração que trabalhos recentes apontam para a influência de aerossóis atmosféricos na irradiância solar incidente na superfície, impactando inclusive

a geração solar FV (BET, 2021; CABALLERO et al., 2018; GUEYMARD, 2014; MARTINS, 1999; NOBRE et al., 2016; WILLOUGHBY; OSINOWO, 2018) dedicamos a presente seção para discutir sucintamente as propriedades dos aerossóis atmosféricos e sua interação com o recurso solar.

A fonte, ou origem, dos aerossóis pode ser classificada em primária e secundária. As fontes primárias são definidas como as formações diretas de aerossóis, sendo estas naturais em sua maioria, mas também envolvem ações antrópicas. Destacam-se oceanos (sal marinho), regiões áridas e semiáridas (poeira), biota terrestre (matéria orgânica de origem vegetal), queimadas/emissões (naturais ou ação antrópica direta), ação vulcânica e meteoritos. As fontes secundárias, por sua vez, são definidas como formações indiretas de aerossóis. Esta é a principal fonte de partículas menores que $1\ \mu\text{m}$, oriundas da conversão de gases precursores (naturais e antrópicos) em partículas sólidas e líquidas.

Os aerossóis atmosféricos presentes em grandes centros urbanos são misturas de fontes primárias e secundárias oriundos das indústrias, meios de transporte, geração energética e também por fontes naturais. A Figura 2.7 compara a massa, concentração e tamanho médios de aerossóis atmosféricos presentes em diferentes regiões.

Figura 2.7 – Massa, concentração e raio médios de aerossóis atmosféricos presentes em diferentes ambientes.
Fonte: Adaptado Ramachandran (2018).

Apesar das dimensões serem variáveis, a distribuição de aerossóis em centros urbanos é dominada por partículas de moda fina. Além disso, a concentração de

aerossóis nesses locais é pelo menos uma ordem de magnitude maior que em áreas rurais (RAMACHANDRAN, 2018). Considerando que as propriedades físico-químicas dos aerossóis influenciam em sua interação com a radiação solar e, também, o aumento de interesse em investir em instalação de sistemas fotovoltaicos em centros urbanos através da GD, estudos avaliando o impacto do aerossol no desempenho dos sistemas FV, como feito por Bet (2021), bem como em modelos preditivos é de grande relevância no âmbito dos esforços em direção ao aumento de fontes intermitentes de energia no sistema elétrico.

A cidade de São Paulo é a cidade mais populosa do Brasil, apresentando uma densidade demográfica superior a 8.100 habitantes/km² (IBGE, 2021). Assim, São Paulo apresenta uma forte atividade antrópica, principalmente devido à alta concentração de indústrias e circulação de veículos. Uma característica particular da condição atmosférica local em São Paulo é a elevada concentração de aerossóis durante os meses secos (DE ALMEIDA ALBUQUERQUE; DE FÁTIMA ANDRADE; YNOUE, 2012). A variação sazonal típica do aerossol atmosférico para a cidade de São Paulo, apresentada na Figura 2.8, demonstra o aumento da profundidade ótica dos aerossóis (AOD) durante os meses de Agosto, Setembro e Outubro.

Figura 2.8 – Médias mensais (2001 – 2017) dos valores de AOD nos comprimentos de onda 500nm, 675nm, 870nm e 1020nm para a cidade de São Paulo. Fonte: Adaptado de Bet (2021).

Aerossóis provenientes de poluição urbana e da queima de biomassa são caracterizados por altos valores de AOD para comprimentos de onda mais curtos (TOLEDANO et al., 2007). Analisando a Figura 2.8 é possível perceber que o período de aumento dos valores de AOD para a cidade de São Paulo coincide com o período

maior incidência de focos de queima de biomassa na região centro-oeste (RÉQUIA JÚNIOR; ARAVÉCHIA JÚNIOR, 2013).

2.4.1. PROPRIEDADES ÓPTICAS DOS AEROSSÓIS ATMOSFÉRICOS E SUA INFLUÊNCIA NA RADIAÇÃO SOLAR

Existem três propriedades ópticas principais que determinam os efeitos radiativos dos aerossóis atmosféricos e como estes interagem com a radiação solar incidente. São elas profundidade óptica (AOD – do inglês *Aerosol Optic Depth*), albedo simples (ω_0) e parâmetro de assimetria (g).

A profundidade óptica de aerossóis define a atenuação da radiação eletromagnética produzida pela interação da radiação solar com aerossóis atmosféricos ao passo que esta propaga na atmosfera. O albedo simples (ω) está relacionado com a fração de energia removida do feixe de radiação solar por espalhamento em relação à fração absorvida, com valores variando, portanto, entre 1 (somente espalhamento) e 0 (somente absorção). Esta última propriedade óptica está intrinsecamente ligada à distribuição de tamanho das partículas e é crucial para determinar a magnitude da fração espalhada (RAMACHANDRAN, 2018). Regiões continentais limpas apresentam albedo simples na ordem de 0,94, enquanto que aerossóis mais absorvedores com valores que podem atingir 0,67 são observados em regiões urbanas (RAMACHANDRAN, 2018).

Por fim, a distribuição angular da radiação solar espalhada por aerossóis é descrita pelo parâmetro de assimetria (g). Este parâmetro óptico é definido pela média do cosseno dos ângulos de espalhamento para o feixe de radiação disperso. Os valores de parâmetro de assimetria variam entre -1 e 1. Partículas com propriedades que promovem espalhamento isotrópico possuem um g próximo de zero. Para partículas que promovam espalhamento frontal possuem um g próximo de 1. Por fim, partículas que promovam um retro-espalhamento possuem um g próximo de -1. O parâmetro de assimetria é maior para distribuição de tamanho de aerossóis com moda em particulados de maior dimensão. Os valores de g para ambiente variam entre 0,61 – 0,63 para regiões continentais e aproximadamente 0,6 para regiões urbanas (RAMACHANDRAN, 2018).

2.5. ÍNDICE DE CLARIDADE (K_t)

O índice de claridade (K_t) é um parâmetro adimensional utilizado para estimar a atenuação do recurso solar em função apenas da condição atmosférica em um dado

momento. O cálculo do K_t é feito pela razão entre a Irradiação Global Horizontal medida em superfície e a calculada para o topo da atmosfera terrestre (IQBAL, 1983).

$$K_t = \frac{IG_{medida}}{IG_{topo\ atm}} \quad (\text{Eq. 2.2})$$

O cálculo da irradiação global horizontal no topo da atmosfera é dado pela equação

$$IG_{atm} = I_{sc} \cdot E_0 \cdot (\sin \delta \cdot \sin \phi + \cos \delta \cdot \cos \phi \cdot \cos \omega) \quad (\text{Eq. 2.3})$$

Onde I_{sc} é a constante solar, considerada 1367 W/m^2 ; E_0 é a excentricidade terrestre dada pela equação (2.3); δ é a declinação solar dada pela equação (2.5); ϕ é a latitude local considerada $-23,5589^\circ$; e ω é o ângulo solar dada pela equação (2.7).

O cálculo da declinação (Eq. 2.5) e da excentricidade solar (Eq. 2.6) utiliza o ângulo diário (Γ), expresso pela equação (2.4), na qual d_n é o dia juliano.

$$\Gamma = \frac{2\pi(d_n - 1)}{365} \quad (\text{Eq. 2.4})$$

Aplicando os valores de Γ calculados para cada dia, aplica-se o valor nas Equações (2.5) e (2.6), temos:

$$\delta = \left(0,006918 - 0,399912 \cos \Gamma + 0,070257 \sin \Gamma - 0,006758 \cos 2\Gamma + 0,000907 \sin 2\Gamma - 0,002697 \cos 3\Gamma + 0,00148 \sin 3\Gamma \right) \cdot (180/\pi) \quad (\text{Eq. 2.5})$$

$$E_0 = 1,000110 + 0,034221 \cos \Gamma + 0,001280 \sin \Gamma + 0,000719 \cos 2\Gamma + 0,000077 \sin 2\Gamma \quad (\text{Eq. 2.6})$$

O ângulo horário (ω) por sua vez é dado em função da hora solar real através da equação (2.7).

$$\omega = (12 - H_{solar\ real}) \cdot 15 \quad (\text{Eq. 2.7})$$

A hora solar real é calculada em função do horário local, considerando os ajustes da longitude padrão (no caso de São Paulo esse valor é -45°), longitude (Lon) local e levando em consideração a Equação do Tempo (E_t).

$$H_{solar\ real} = Hora\ Local + 4 (\phi_{padr\tilde{a}o} - Lon) + E_t \quad (\text{Eq. 2.8})$$

$$E_t = \left(\frac{0,000075 + 0,001868 \cos \Gamma - 0,032077 \sin \Gamma - 0,014615 \cos 2\Gamma}{0,04089 \sin 2\Gamma} \right) \cdot 229,18 \quad (\text{Eq. 2.9})$$

Aplicando as equações (2.3) à (2.9) a equação (2.2), obtemos o valor de K_t das medições horárias de Irradiação Global Horizontal incidente no sistema fotovoltaico.

O valor de K_t , está diretamente relacionado à disponibilidade do recurso solar e sua variabilidade temporal está intrinsecamente relacionado com a condição atmosférica. A presença de nebulosidade impacta seu valor uma vez que produz atenuação da radiação que incide na superfície. Quanto maior a nebulosidade, menor o valor de K_t . Em condições de céu sem presença de nuvens, quanto maior a espessura ótica de aerossóis (maior concentração de aerossóis na atmosfera), menor o valor de K_t . O Índice de claridade permite classificar as medições em função da condição de nebulosidade, sendo considerado condição de céu claro quando o K_t for da ordem de 0,7 ou superior (IQBAL, 1983).

Além da dinamicidade da composição da atmosfera terrestre para cada localidade, a obtenção de dados referentes aos dois principais modulantes do recurso solar, isto é, cobertura de nuvens e presença de aerossóis atmosféricos, é um desafio adicional na avaliação do recurso solar que incide na superfície terrestre (CRISOSTO; LUIZ; SECKMEYER, 2021; EBERHARDT et al., 2016; LUIZ et al., 2018b).

O índice de claridade K_t não apresenta a sazonalidade do ciclo diária associada à rotação da Terra em relação ao seu eixo uma vez que este fenômeno afeta igualmente a irradiância solar na superfície e no topo da atmosfera. Assim, o cálculo do índice de claridade foi utilizado neste trabalho com a função de normalizar a medição de radiação incidente no plano horizontal para ser posteriormente alimentado no modelo de aprendizado de máquinas.

2.6. MODELOS DE PREVISÃO DE GERAÇÃO SOLAR FOTOVOLTAICO

Juntamente dos sistemas de armazenamento, os modelos de previsão de energia estão entre as principais estratégias de apoio ao rápido crescimento da geração de energia solar fotovoltaica, mitigando as dificuldades de gerenciamento de sistemas

conectados à rede (DE GIORGI; CONGEDO; MALVONI, 2014). Recentemente, técnicas de previsão, especialmente aquelas que usam métodos de Inteligência Artificial, como as Redes Neurais Artificiais (RNA), vêm ganhando atenção devido à alta precisão e rápido desenvolvimento de novos métodos (SOBRI; KOOHI-KAMALI; RAHIM, 2018). O desenvolvimento de modelos de previsão de energia que considerem as condições atmosféricas e operacionais locais é de extrema relevância, especialmente em áreas urbanas densas. Os métodos de modelos preditivos encontram-se divididos em três categorias principais: métodos estatísticos; métodos físicos e métodos *ensemble*.

Os métodos estatísticos têm por principal característica serem relativamente de fácil implementação, possuírem habilidade de identificar padrões, mas também são bastante dependentes da série histórica de dados. Os modelos estatísticos têm a habilidade de reconstruir padrões em uma série temporal e aplicá-los na previsão sem que haja necessidade de equacionar matematicamente os processos dinâmicos inerentes à atmosfera (SOBRI; KOOHI-KAMALI; RAHIM, 2018). Este método é subdividido em cinco categorias diferentes: Máquina de Suporte Vetorial (SVM); Cadeias de Markov; Modelos Regressivos e Autoregressivos (ARMA, ARIMA, SARIMA); e Redes Neurais Artificiais (RNA), que iremos discutir mais profundamente na próxima seção.

Por sua vez, os métodos físicos, também conhecidos como métodos numéricos, utilizam dados medidos em satélites e medidos em solo para a simulação dos processos radiativos de absorção e espalhamento com os componentes atmosféricos. Modelos físicos são subdivididos em três métodos principais. O primeiro compreende-se Modelos Numéricos de Previsão do Tempo (PNT), que podem utilizar dados tanto em escala global ou mesoescala. Os modelos em escala global simulam as características atmosféricas a nível global, enquanto os em mesoescala simulam as características atmosféricas em uma área limitada, podendo ser a nível continental, nacional ou regional (MONTEIRO et al., 2013). Modelos PNT utilizam integração numérica acoplada a equações diferenciais que descrevem os mecanismos de fluxo de radiação solar, bem como os processos dinâmicos da atmosfera terrestre.

Em outra categoria dos modelos físicos, os modelos de transferência radiativa baseados em dados de cobertura de nuvens obtidos por imageamento de céu utilizando imagens do céu ou sensores em satélite. As imagens podem ser aplicadas para estimar, altitude de nuvens e determinação de movimento de nuvens. Este método analisa uma

sequência de imagens capturadas no local de interesse a fim de reconhecer a presença de nuvens, sua localização e seu movimento (LIMA et al., 2020; LUIZ et al., 2018b). Com isto, este método tem a capacidade de fazer previsões probabilísticas e determinísticas de parâmetros como cobertura de nuvens, irradiância solar e geração solar fotovoltaica.

Por fim, os métodos de modelo *ensemble*, também conhecidos como modelos híbridos. Estes modelos foram propostos a fim de resolver as fraquezas individuais dos modelos estatísticos e numéricos, melhorando assim a confiabilidade e a precisão da previsão do recurso solar. Os modelos *ensemble* podem ser divididos em dois métodos principais: cooperativo e competitivo.

O primeiro, cooperativo, no qual utiliza-se um segundo modelo para fazer um pré-processamento dos dados de entrada ou o pós-processamento dos dados de saída de um modelo preditivo pré-existente. Comumente modelos híbridos envolvem cooperação entre técnicas preditivas distintas, buscando o melhor de cada método. Em Guarnieri (2006) uma RNA-MLP (Rede Neural Artificial de *Perceptrons* Multi-camadas) foi empregada para refinar a previsão de irradiação solar prevista pelo modelo físico ETA/CPTC, apresentando ganhos entre 21 e 24% na Raiz do Erro Quadrático Médio (RMSE). Já em Lima et al. (2016) uma RNA foi utilizada para fazer o pós-processamento dos dados de saída do modelo WRF para dados de radiação no nordeste brasileiro, apresentando uma redução do RMSE em até 15% em comparação a utilização exclusiva do modelo físico.

No segundo método, o competitivo, são implementados modelos independentes utilizando técnicas distintas de tal forma que a previsão seja dada de maneira concorrente. Nos modelos híbridos competitivos são utilizados os mesmos preditores, porém estes são parametrizados diferentemente em função do processo de treinamento utilizado. Os valores previstos pelos modelos são, então, comparados e avaliados conjuntamente. Filipe et al. (2015) combinou modelos estatísticos e modelos físicos para prever a geração FV de uma planta de 16 kWp instalado no norte de Portugal em um horizonte temporal de 48h. Os autores utilizaram um modelo autoregressivo para analisar a série histórica de geração FV para horizonte intradiário e alimentaram a saída de um modelo NWP para alimentar uma RNA para prever a geração FV para horizontes temporais mais amplos. Os valores previstos por cada modelo foram combinados em

uma função de regressão, que os autores denominaram *regression pool*, resultando com um RMSE médio de 9,83% para as três primeiras horas e 13,37% para o restante do horizonte temporal.

A determinação da categoria de modelo preditivo indicado não é dada de maneira determinística. A escolha de técnica preditiva envolve avaliar fatores particulares do problema que procura-se solucionar. A Figura 2.9 classifica os modelos preditivos em função das resoluções espacial e temporal. Modelos estatísticos encontram aplicação para maiores resoluções espaciais e temporais, ou seja, para áreas mais limitadas (até 1 km) e curto intervalo de tempo (de minutos até 10 horas). Modelos físicos, por sua vez, abrangem menores resoluções espaciais, na mesoescala e escala global, e seus horizontes temporais abrangem desde o intra-diário até previsões de longo prazo.

Figura 2.9 – Classificação de modelos com base nas resoluções espacial e temporal. Fonte: Adaptado de Diagne et al. (2013).

2.6.1. REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNA) são um método de Aprendizado de Máquina que generaliza uma regressão logística ao alimentar variáveis em um algoritmo multicamadas dotadas de unidades de processamento que compõe um modelo de neurônio artificial (SRIKRISHNAN et al., 2015). As RNA são potentes ferramentas estatísticas compostas por unidades de processamento simples que computam funções matemáticas a partir de uma série de dados fornecidas a esta (GUARNIERI, 2006). Estas unidades de processamento são dispostas em camadas paralelas e interligadas

a seus pares por conexões associadas a pesos. Como o próprio nome sugere, o modelo é inspirado nas interligações neurológicas humanas. Assim, as unidades de processamento são denominadas neurônios e suas interconexões denominadas sinapses. Os pesos atribuídos relacionam os dados de entrada em cada neurônio a um dado de saída desejado. Assim, o desenvolvimento de uma RNA passa por uma fase de treinamento no qual a rede ajusta os pesos sinápticos à cada iteração aplicando o método do gradiente para minimização do erro de sua estimativa. Quando a rede está treinada e ajustada, os pesos são fixados e a RNA pode ser utilizada como um modelo capaz de estimar saídas a partir de um novo conjunto de dados de entrada.

A Figura 2.10 representa um esquema genérico de uma RNA convencional apresentando cinco entradas, uma camada oculta com três neurônios e uma camada de saída com um único neurônio. Cada neurônio apresenta um nível de atividade interna (v_j) representado pela combinação do produto das variáveis de entrada (x_i) a seus respectivos pesos sinápticos (w_i) juntamente do somatório com um valor ajustável (b_i), denominado viés, ou *bias*, e em seguida processada por uma função de ativação antes que esta informação seja transferida para os neurônios da próxima camada.

Figura 2.10 – Esquema genérico de uma Rede Neural Artificial que apresenta uma camada oculta. Fonte: Adaptado Srikrishnan (2015)

A Figura 2.11 ilustra de maneira mais detalhada os componentes de entrada e saída em um neurônio. A função de ativação processa a informação a ser transferida à próxima camada de acordo com o nível de atividade (Eq. 2.10).

Figura 2.11 – Modelo de um único neurônio. Fonte: Guarnieri (2006)

A função de ativação tem por finalidade processar o valor de saída do neurônio antes que esta informação seja transmitida para os neurônios da próxima camada. As três principais funções de ativação aplicadas a modelagem do recurso solar são a linear, a logística e a tangente hiperbólica.

Figura 2.12 – Representação gráfica das principais funções de ativação aplicadas a modelagem do recurso solar em RNAs. (a) função linear; (b) função logística; e (c) função tangente hiperbólica. Adaptado de John et al (2010).

A função linear (Figura 2.12.a) não altera a saída da função de atividade, não havendo um limite rígido para os valores de saída. Esta função, quando utilizada, comumente é aplicada na camada de saída da RNA. As funções logística (Figura 2.12.b) e tangente hiperbólica (Figura 2.12.c) são denominadas funções sigmóides. Estas funções têm por finalidade normalizar o valor de saída da função de atividade interna, limitando o valor de saída entre 0 e 1 para a função logística (Eq. 2.11) e -1 e 1 para a função tangente hiperbólica (Eq. 2.12). As funções de ativação utilizadas em camadas ocultas são por padrão do tipo sigmoide. Contudo, estas funções também podem ser utilizadas na camada de saída.

Atividade Interna:
$$v_j = \sum_{i=1}^n w_{ji}x_i + b_j \quad (2.10)$$

Função Logística:
$$\varphi(v_j) = \frac{1}{1+e^{-av_j}} \quad (2.11)$$

Função Tangente Hiperbólica: $\varphi(v_j) = \frac{1 - e^{-v_j}}{1 + e^{-v_j}}$ (2.12)

Os modelos de redes neurais artificiais são classificados em duas categorias principais quanto a sua arquitetura: redes neurais não-recorrentes (*feedforward*) e recorrentes (*feedback*). Ambas utilizam metodologia do gradiente para minimização do erro e operam conforme as equações apresentadas anteriormente. A Figura 2.12 ilustra os componentes de processamento de um único neurônio disposto em uma Rede Neural Recorrente (RNN), alimentado por m variáveis de entrada e um atraso de d iterações no processo de realimentação. A diferença fundamental é que as Redes Neurais Recorrentes possuem realimentação dos dados de saída como dado de entrada na iteração seguinte. Desta forma, as RNN são indicadas para identificar padrões sequenciais nos dados fornecidos para seu treinamento, considerando então a variação do tempo. Enquanto as RNA utilizam o método de *Back Propagation* (BP), as RNN utilizam o método de *Back Propagation Through Time* (BPTT) indicando justamente a contabilização da variabilidade através da percepção de uma sequência nos dados.

Uma RNA convencional *feedforward* tem a capacidade de resolver problemas não lineares complexos, mas não consegue reter informações de iterações anteriores. Nesse sentido, as RNN vêm ganhando atenção na solução de problemas onde estados anteriores de uma série temporal influenciam o estado atual (AL-SBOU; ALAWASA, 2017; MISHRA; PALANISAMY, 2018; PANG; NIU; O'NEILL, 2020), como é o caso da previsão do recurso solar para geração solar fotovoltaica.

Figura 2.13 – Modelo de um único neurônio em Rede Neural Recorrente.

2.6.2. RESUMO DE PUBLICAÇÕES NACIONAIS E INTERNACIONAIS

Mellit e Pavan (2010) utilizaram uma RNA com perceptrons multicamada (MLP) para prever a irradiância com 24h de antecedências para um sistema instalado na Itália. O modelo desenvolvido considerou como dados de entrada as variáveis de temperatura ambiente, irradiância total diária e o dia do mês. O estudo obteve coeficientes de correlação de 98% e 94% para dias de céu claro e encoberto, respectivamente. Em outro trabalho, Mellit et al. (2013) propuseram duas redes neurais artificiais para diferentes condições de cobertura de nuvens tomando como preditores dados de irradiância e temperatura ambiente a fim de prever a geração de um módulo de silício policristalino de 50 Wp. O modelo ANN-1 prevê geração FV para dias com céu encoberto, apresentando irradiação diária inferior a $400 \text{ W.m}^{-2}/\text{dia}$. O modelo ANN-2, por sua vez, prevê geração FV em dias ensolarados com irradiação superior a $400 \text{ W.m}^{-2}/\text{dia}$. Comparado a outros métodos preditivos, o modelo RNA desenvolvido apresenta melhor desempenho alcançando um Erro Relativo Médio (MRE) de 2,5% e 2,3% para dias de céu encoberto e ensolarado, respectivamente.

Pedro & Coimbra (2012) propuseram diversos métodos para realizar a previsão com 1 e 2 horas de antecedência para um sistema de 1 MWp instalado em Merced, California. Os modelos desenvolvidos compreendem os métodos de Persistência, ARIMA, kNN, RNA e RNA otimizada por meio de Algoritmo Genético (GA/RNA). Foram utilizados somente dados inerentes do sistema gerador. Dados exógenos ao sistema, como irradiância e variáveis ambientais, não foram considerados. A base de dados compreende o período de Novembro de 2009 a Agosto de 2011, sendo utilizado o período de Janeiro de 2011 em diante como conjunto de teste para os modelos. O conjunto teste foi dividido em três períodos de variabilidade do recurso: P1 de Janeiro à Abril; P2 para Maio e Junho; P3 para Julho e Agosto. O terceiro período apresentou os melhores resultados. O modelo RNA apresentou valores de RMSE de 7,16% e 12,96% para 1 e 2 horas, respectivamente. Por fim, o modelo GA/RNA apresentou valores de RMSE de 6,4% e 8,99% para 1 e 2 horas, respectivamente.

Kardakos et al. (2013) propuseram dois métodos de modelos estatísticos e os comparou com um modelo de persistência. O primeiro método foi um modelo *Seasonal Autoregressive Integrated Moving Average* (SARIMA) desenvolvido com dados de radiação solar modelados a partir de um modelo NWP. O segundo método foi uma RNA

com múltiplos dados de entrada. Ambos foram utilizados para prever a geração FV de forma intradiária e com um dia de antecedência, apresentando resultados em um passo horário. O trabalho apresentou um RMSE médio de 11,26–11,46%, 12,89% e 13,71% para os modelos SARIMA, RNA e Persistência, respectivamente.

Wu et al. (2014) utilizaram dados de geração FV, irradiação, temperatura ambiente e temperatura dos módulos a fim de prever a geração FV com uma hora de antecedência. Os dados são oriundos de três localidades distintas localizadas na Malásia. Os autores utilizaram cinco métodos estatísticos distintos, sendo eles ARIMA, SVM, RNA e ANFIS. O estudo propõe, também, o desenvolvimento de um modelo híbrido por meio da combinação dos modelos estatísticos utilizando um algoritmo genético. O modelo híbrido proposto pelos autores apresentou RMSE variando entre 3,43% e 6,57%.

Almonacid et al. (2014) propuseram um método para prever a geração FV 1h a frente utilizando um modelo RNA dinâmico. Os dados de entrada utilizados escolhidos foram irradiância e temperatura ambiente, em função de seu impacto no desempenho do sistema gerador. A comparação entre o valor previsto e real foi realizada por meio de uma análise de regressão linear. O estudo indicou uma correlação entre irradiância global e temperatura ambiente muito próxima a 1 e o RMSE observado foi de 3,38%.

De Giorgi et al. (2014) utilizaram um método estatístico de análise multi-regressiva e uma Rede Neural Artificial de Elmann a fim de prever a geração produzida por um sistema de 960 kWp, ligado à rede na Itália, para diferentes horizontes temporais intradiários. Foram considerados três vetores de entrada distintos combinando a série temporal de potência de saída (W), temperatura ambiente (°C), irradiância no plano inclinado de 15° (W/m²) e a temperatura dos módulos (°C). Os autores concluíram que o melhor desempenho foi obtido quando todas as variáveis ambientais foram inclusas, indicando a habilidade de um modelo de rede neural recorrente em reconhecer padrões em séries temporais de dados. O RMSE observado varia entre 10,91% e 12,6% para previsões com uma hora de antecedência e variando entre 15,61% e 18,55% para previsões com três horas de antecedência.

Lima et al. (2016) utilizaram uma combinação entre o modelo WRF e um método de pós-processamento utilizando uma RNA para estimar a radiação solar incidente no

Nordeste Brasileiro. Os dados de saída do modelo WRF foram utilizados como dados de entrada na RNA. Os resultados de saída do modelo foram comparados com dados de irradiação solar de superfície obtidos através de estações de coleta automatizadas. A metodologia proposta apresentou uma redução do erro do viés de aproximadamente 50 W/m² e o RMSE médio foi reduzido em aproximadamente 30% quando a RNA foi acoplada em detrimento de quando somente o modelo WRF foi utilizado.

Hossain et al. (2017) utilizaram um modelo de Aprendizado de Máquinas Extremo (ELM) para prever a geração elétrica de três sistemas FV conectados à rede instalados no telhado do laboratório PEARL da Universidade da Malásia com 1h e 1 dia de antecedência. Foram utilizados dados de irradiância solar, temperatura ambiente e das células, velocidade dos ventos e geração FV. As medições são feitas em um passo de 5min. Para previsão para o dia seguinte foram utilizados a média diária dos preditores. Para previsão com uma hora de antecedência foram utilizadas médias horárias. A metodologia ELM empregada foi comparada com outros métodos preditivos populares como RNA e SVR. O desempenho do modelo ELM apresentou RMSE variando entre 2,77% e 3,87% para previsão horária, enquanto os modelos SVR e RNA apresentaram 3,2 – 4,72% e 3,56 – 6,31%, respectivamente.

Em Li et al. (2020) um modelo híbrido baseado em LSTM-WPD foi proposto para prever a geração com 1h de antecedência de uma usina FV de 26,52 kWp instalada em Alice Springs, Austrália, utilizando dados de saída do sistema PV e dados meteorológicos medidos em superfície. O estudo propôs um método de combinação linear das subséries decompostas pelo método *Wavelet Packet Decomposition* (WPD) a fim de otimizar o desempenho do modelo proposto. O desempenho do modelo LSTM-WPD apresentou RMSE de 0,89% quando o método de combinação linear foi empregado, em detrimento de 2,78% quando este não foi utilizado. O desempenho foi comparado com demais métodos preditivos populares como GRU (3,9%), RNN (3,99%) MLP (4,09%) e LSTM (3,91%).

Capítulo 3 : METODOLOGIA

3. MATERIAIS E MÉTODOS

3.1. SISTEMAS DE GERAÇÃO SOLAR FOTOVOLTAICA

O enfoque do projeto é a usina de geração fotovoltaica localizada na cidade universitária da Universidade de São Paulo, que possui 540 kWp de potência, distribuídos pelos sistemas de geração BIPV, CTPV, CRPV e BAPV conforme ilustrado na Figura. 3.1. A estação de coleta de dados solarimétricos e ambientais encontra-se situada no Laboratório Solar Fotovoltaico (LSF-IEE) nas coordenadas 23°33'32" S e 46°44'06" W.

Figura 3.1 – Instituto de Energia e Ambiente. Estação solarimétrica (S) e unidades de geração solar fotovoltaica BIPV, CRPV, CTPV e BAPV. Fonte: Adaptado de Google Earth

O sistema de geração BIPV (*Building Integrated Photovoltaic*) consiste em um gerador de 150 kWp localizado na cobertura do estacionamento de veículos do IEE-USP. O sistema é composto por 24 *strings* de 24 módulos do modelo YL260C-30b de 260 Wp. Cada string está ligada a um inversor de 5 kW. Diferentemente dos demais sistemas, a estrutura de suporte do sistema BIPV possui uma configuração curvada, apresentando três inclinações diferentes de 30°, 20° e 10° em relação ao plano horizontal (PIN, 2017; RODRIGUEZ, 2016).

Figura 3.2 – Sistema BIPV - Estacionamento IEE-USP

Fonte: Instituto de Energia e Ambiente, 2016.

O sistema de geração BAPV (*Building Applied PV*) consiste em um gerador de 156 kWp localizado na cobertura do prédio da biblioteca Brasileira. O sistema conta com 10 inversores da fabricante SMA modelo Sunny Tripower 15000TL, de 15 kW de potência nominal, nos quais em cada um são conectadas 3 *strings* de 20 módulos do modelo YL260C-30b de 260 Wp. A estrutura fixa dispõe os módulos a uma inclinação de 18° em relação ao plano horizontal (PIN, 2017).

Figura 3.3 – Sistema BAPV - Biblioteca Brasileira

Fonte: (PIN, 2017)

O sistema de geração CRPV consiste em um gerador de 78 kWp localizado na cobertura do prédio da administração do IEE-USP. Este sistema é, também, do tipo *Building Applied Photovoltaic* (BAPV), mas para fins de diferenciação será denominado CRPV daqui em diante. O sistema é composto por 15 *strings* de 20 módulos do modelo YL260C-30b de 260 Wp. A cada três strings são ligadas a uma caixa de junção, que por

sua vez está ligada a um inversor, totalizando 5 inversores que estão localizados dentro do prédio da administração.

Figura 3.4 – Sistema CRPV - Administração IEE-USP

Fonte: Instituto de Energia e Ambiente, 2016.

O sistema de geração CTPV consiste em um gerador de 156 kWp instalado sobre estrutura fixa no terreno atrás do prédio da administração do IEE-USP. O sistema é composto por 30 *strings* de 20 módulos do modelo YL260C-30b de 260 Wp. O sistema possui um inversor c.c./c.a. trifásico central localizado no eletrocentro localizado ao lado do sistema fotovoltaico.

Figura 3.5 – Sistema CTPV - Estrutura fixada no terreno atrás da administração do IEE-USP

Fonte: Instituto de Energia e Ambiente, 2016.

3.2. CLIMATOLOGIA LOCAL

A região metropolitana de São Paulo está situada numa região com clima classificado como subtropical úmido. A período de chuvas inicia-se normalmente em outubro e termina em março (INMET, 2010). As normais climatológicas para variáveis

de temperatura, umidade relativa, pressão atmosférica e velocidade dos ventos para a cidade de São Paulo estão apresentadas na Figura 3.6 com base nos dados do INMET de 1981 a 2010. Esses valores serão adotados como critérios de normalização dos preditores baseados nas variáveis ambientais adotadas no desenvolvimento do modelo de aprendizado de máquinas.

Figura 3.6 – Normais Climatológicas de Temperatura, Umidade Relativa, Pressão Atmosférica e Velocidade Média dos Ventos na cidade de São Paulo, conforme a base de dados do INMET compreendendo o período entre 1981 e 2010.

3.3. BASE DE DADOS OBSERVADOS

Os dados de radiação solar e meteorológicos empregados neste trabalho foram coletados por uma estação solarimétrica instalada no LSF, próxima dos sistemas de geração solar fotovoltaica. As características dos equipamentos instalados são listadas abaixo conforme descrito por Rosa (2003):

- **Sensores de temperatura e umidade relativa:** sensores Campbell Scientific-Vaisala Model HMP-45C. As medidas de umidade relativa apresentam um erro de 2% para valores abaixo de 90% e 3% para valores acima de 90%. As medidas de temperatura, por sua vez, apresentam um erro máximo de 0,5%, operando em temperaturas entre -40°C e 60°C;
- **Anemômetros:** Modelo 05103 R.M. Young Wind Monitor, com erro máximo de 0,3 m/s e 3° para medidas de velocidade e direção do vento, respectivamente;

- **Piranômetros:** Modelo Kipp & Zonen CM 11, apresentando uma variação máxima anual de 0,5% e um erro direcional menor que 10 W/m^2 ;
- **Anel de sombreamento:** Kipp & Zonen CM 121;
- **Sistema automático de coleta de dados:** CR10X da Campbell Scientific, com erro máximo de 0,1%.

A Figura 3.7 mostra a estação solarimétrica disposta no LSF-IEE (à esquerda) e o diagrama esquemático de seus equipamentos.

Figura 3.7 – Estação solarimétrica disposta no LSF e diagrama esquemático destacando seus equipamentos e sensores. Fonte: Adaptado de Rosa (2003).

A base de dados utilizada compreende o período de janeiro de 2017 a novembro de 2020. As medições meteorológicas e solarimétricas são coletadas automaticamente em um passo horário. Os dados utilizados para as etapas de desenvolvimento, treinamento e validação do modelo de Redes Neurais Artificiais dividem-se em dados de aerossóis atmosféricos, meteorológicos e de irradiação solar descritos a seguir:

- Dados de irradiação solar – monitorados com sensores dispostos no Instituto de Energia e Meio Ambiente da Universidade de São Paulo. Nesses sensores são coletados automaticamente e armazenados diariamente os seguintes parâmetros:
 - Irradiação global horizontal (kWh/m^2)
 - Irradiação difusa global horizontal (kWh/m^2)
 - Irradiação global em plano inclinado de 30 graus (kWh/m^2)
- Dados ambientais – coletados pelos sensores mencionados anteriormente são armazenados com periodicidade horária.
 - Umidade relativa (%)

- Temperatura média (°C)
- Velocidade média do vento (m/s)
- Pressão atmosférica (mmHg)

c) Dados de geração fotovoltaica – medições de Potência em Corrente Alternada em Baixa Tensão (PcaBT), dada em kW, armazenadas no sistema de monitoramento e armazenamento de dados da usina FV. Os valores são armazenados em um passo minuto a minuto.

3.4 DADOS DE PROPRIEDADE ÓTICA DE AEROSSÓIS

Apesar de haver uma estação da rede AERONET em operação na Cidade Universitária, a disponibilidade dos dados não acontece instantaneamente no momento da coleta de forma similar aos dados de irradiação solar incidente na superfície ou dados meteorológicos. Isso acontece por dois motivos: i) as propriedades óticas dos aerossóis são realizadas apenas em momentos em que há radiação solar direta incidente na superfície; ii) os dados são publicados apenas após procedimentos de controle de qualidade realizados pela agência responsável pela rede AERONET. Esse aspecto inviabiliza o uso operacional de um modelo de previsão da geração solar que utilize os dados de propriedades óticas observados em tempo real.

Com o intuito de investigar a contribuição do uso de dados de aerossóis atmosféricos na redução de desvios das previsões de geração, optou-se por utilizar dados de profundidade ótica de aerossóis estimados pelo modelo de reanálise meteorológica MERRA-2. O modelo em questão é a segunda versão do projeto “análise retrospectiva da era moderna para pesquisas e aplicações”, apresentando, entre outras variáveis, propriedades observáveis dos aerossóis com resolução espacial de $0,5^\circ \times 0,625^\circ$. A base de dados de reanálise em questão compreende o período entre 1980 até atualmente, apresentando registros para intervalos de 3 em 3 horas (OLAUSON; BERGKVIST, 2015). Foi utilizado o índice de profundidade ótica (AOD) a 550 nm, considerando que este comprimento de onda coincide com o pico de densidade de potência do espectro de radiação solar que chega a superfície terrestre (vide figura 2.4). Desta forma, o AOD 550 nm é eficaz em dividir o espalhamento em função da interação entre radiação e partículas presentes na atmosfera.

3.5. MÉTODO DE TRATAMENTO DOS DADOS

A base de dados foi tratada e organizada de maneira estabelecer sincronia entre as variáveis meteorológicas, solarimétricas e de geração elétrica para que as etapas seguintes de desenvolvimento do modelo de previsão possam ser executadas. Sobretudo para as variáveis meteorológicas (temperatura, umidade relativa, velocidade do vento e pressão barométrica), para tratamento de controle foram removidas medições sinalizadas com código de erro na base de dados, sendo desconsideradas todas as medições feitas para este mesmo horário. Foram consideradas somente medições que tenham ocorrido de maneira concorrente à disponibilidade de dados de geração dos sistemas FV. Ou seja, para as situações em que nenhum dos sistemas FV tenham apresentado dado de geração (por inatividade ou falha na coleta automática de dados) os dados solarimétricos e meteorológicos também foram descartados para este intervalo de tempo. O método de tratamento dos dados solarimétricos e de geração, bem como a análise estatística realizada para pré-seleção dos preditores do modelo serão descritas mais detalhadamente na seção seguinte.

A seguir, vamos descrever os procedimentos adotados com cada uma das bases de dados utilizados neste documento.

3.5.1 TRATAMENTO DE DADOS SOLARIMÉTRICOS

Para o desenvolvimento do modelo está sendo considerado o dado de Irradiação Global incidente em superfície horizontal (em kWh/m²) através do cálculo do K_t , denominado índice de claridade (IQBAL, 1983), que considera a razão entre o valor medido na estação solarimétrica e a irradiação no plano horizontal no topo da atmosfera terrestre.

$$K_t = \frac{IG_{medida}}{IG_{topo\ atm}} \quad (3.1)$$

O índice K_t fornece uma avaliação da atenuação da radiação solar em seu percurso na atmosfera e retira a sazonalidade diária relacionada com o trajetória do Sol. O valor de K_t possibilita classificar a condição atmosférica em termos de nebulosidade, sendo considerado que o K_t superior a 0,7 pode ser atribuído à condição de céu claro (sem nebulosidade) (IQBAL, 1983). Os valores de K_t compreendem o intervalo entre 0 e 1, não necessitando de normalização.

3.5.2. TRATAMENTO DOS DADOS DOS SISTEMAS FOTOVOLTAICOS

Para contabilizar a quantidade de energia produzida por cada um dos sistemas de geração solar fotovoltaica (BAPV, BIPV, CTPV e CRPV) foi considerada a potência em corrente alternada de baixa tensão (PcaBT), visto que esta contabiliza a potência instantânea total do sistema antes de passar pelo transformador. Nesse sentido, o cálculo da energia produzida em kWh é dado pela equação (3.2).

$$Energia (kWh) = (\sum_{i=1}^{60} PcaBT_i) \frac{1}{60} \quad (3.2)$$

3.5.3. NORMALIZAÇÃO DOS DADOS

A normalização é uma etapa importante do pré-tratamento dos dados para alimentá-los a etapa de treinamento de um modelo de aprendizado de máquinas. A normalização consiste em converter cada uma das variáveis a um limiar variando entre 0 e 1, proporcionalmente a um valor de referência inerente à variável. A normalização das variáveis foi feita através de um cálculo de mínimo e máximo, descrito na equação (3.3).

$$X_{iNorm} = \frac{(X_i - X_{min})}{(X_{max} - X_{min})} \quad (3.3)$$

As variáveis temperatura, velocidade dos ventos, umidade relativa e pressão atmosférica foram normalizadas com base no máximo absoluto anual observados para a normal climatológica para a cidade de São Paulo (INMET, 2010). Os dados de geração elétrica dos sistemas fotovoltaicos, por sua vez, foram normalizados de maneira proporcional às suas capacidades instaladas. Por fim, as demais variáveis, como ângulos solarimétricos notáveis e a profundidade ótica dos aerossóis foram normalizados através dos máximos e mínimos observados na série temporal.

3.5.4. ANÁLISE DE PRÉ-SELEÇÃO DOS PREDITORES

Considerando que as variáveis empregadas no desenvolvimento do modelo correspondem a variáveis ambientais e meteorológicas, é possível inferir que estas não só apresentam influência na geração dos sistemas fotovoltaicos, mas também

apresentam uma inter-relação. Nesse sentido, a análise da Correlação de Pearson de maneira individual pode não demonstrar a real influência destas na geração elétrica. Foi então escolhido o método de Análise de Componente Principal (PCA) das variáveis disponíveis nas bases de dados descritos acima. O método PCA analisa o conjunto de dados disponíveis buscando identificar as componentes representativas da base de dados por meio de combinações lineares entre as variáveis (JOHNSON, 1998). Neste estudo, a transformação ortogonal varimax foi adotada assumindo a ocorrência de duas componentes ortogonais entre si.

Figura 3.8 – Análise de Componente Principal (PCA) em solução varimax para os sistemas fotovoltaicos BAPV, BIPV, CTPV e CRPV considerando as variáveis Índice de Claridade (K_t), Temperatura ambiente (T_p), Velocidade dos ventos (V), Profundidade ótica dos aerossóis (AOD), Umidade relativa do ar (U), Ângulo Zenital (Z_n) e Pressão atmosférica (P_r).

A Figura 3.8 apresenta a solução ortogonal para os quatro sistemas fotovoltaicos BAPV, BIPV, CTPV e CRPV. No eixo das abcissas (x) encontra-se o fator de carga, ou *loadings*, para o primeiro fator e no eixo das ordenadas (y) os fatores de carga para o segundo fator. Foram consideradas as variáveis Índice de Claridade (K_t), Temperatura ambiente (T_p), Velocidade dos ventos (V), Profundidade ótica dos aerossóis (AOD), Umidade relativa do ar (U), Ângulo Zenital (Z_n) e Pressão atmosférica (P_r). A literatura aponta que quando uma carga superior ao módulo de 0,4 é observada a variável apresenta correlação relevante ao fator. Os quatro sistemas fotovoltaicos apresentaram resultados semelhantes na análise PCA. As variáveis temperatura ambiente e velocidade dos ventos foram consideradas relevantes para ambos os fatores latentes. As variáveis K_t e U apresentaram cargas superiores ao módulo 0,841 para o primeiro

fator, mas de baixa relevância para o segundo fator. Inversamente, o ângulo zenital apresentou uma forte relevância ao segundo fator com uma carga de 0,99, mas com relevância próxima de 0,121 para todos os sistemas FV. A profundidade ótica de aerossóis a 550nm apresentou cargas próximas de 0,26 em relação ao primeiro fator para todos os sistemas. Vale mencionar aqui que toda a base de dados foi utilizada, incluindo os períodos seco e chuvoso. A profundidade ótica dos aerossóis atmosféricos possui valores maiores durante a estação seca, como descrito anteriormente. Por fim, a pressão atmosférica apresentou cargas inferiores à 0,05 para os dois fatores latentes, sendo assim considerada negligenciável.

Observamos que em trabalhos recentes como os de Dolara et al. (2016) e Kashyap et al. (2015) e Lee & Kim (2019), além das variáveis solarimétricas e de geração elétrica, a hora do dia foi considerada como preditor da RNA, apresentando ganho na precisão do modelo preditivo. Desta forma, decidimos incluí-la juntamente das variáveis discutidas acima. Assim, o modelo de aprendizado de máquinas proposto nesta dissertação de mestrado conta com os seguintes preditores:

1. Índice de claridade;
2. Temperatura ambiente;
3. Velocidade dos ventos;
4. Umidade relativa
5. Ângulo zenital
6. Hora do dia;
7. Geração no instante zero;
8. AOD 550nm;

3.6. VALIDAÇÃO ESTATÍSTICA

A fim de avaliar o desempenho dos modelos preditivos propostos, fizemos a comparação da geração elétrica prevista com a produção de energia observada do sistema fotovoltaico. Os índices estatísticos usados neste estudo são: Erro Quadrático Médio (MSE) (Eq. 3.4), Raíz do Erro Quadrático Médio Nominal (NRMSE) (Eq. 3.5), *Skill Score* (SS) (Eq. 3.6).

$$MSE = \frac{1}{N} \sum_{i=1}^N (E_{prevista} - E_{saída})_i^2 \quad (3.4)$$

Onde $E_{saída}$ é a geração FV observada e a $E_{prevista}$ é a geração FV prevista pelo modelo de aprendizado de máquinas, e N é o número de medições observadas na série temporal. O MSE permite identificar a ocorrência de erro sistemático do modelo de previsão. Valores próximos a zero são o desejado para um modelo de previsão, e valores positivos do MSE indicam que o modelo superestima o valores observados em campo.

$$NRMSE = P_{kwp} \cdot 1h \cdot \sqrt{\frac{1}{N} \sum_{i=1}^N (E_{saída} - E_{prevista})_i^2} \quad (3.5)$$

Onde P_{kwp} é a capacidade instalada do sistema FV de referência. Este é integrado pelo período de uma hora, resultando em energia elétrica máxima produzida em kWh. O NRMSE permite avaliar a incerteza do modelo de aprendizado de máquinas em termos da unidade de geração elétrica.

$$SS_{MSE} = 1 - \frac{MSE_{RNN}}{MSE_{persistência}} \quad (3.6)$$

Onde o MSE_{RNN} é o MSE para a geração elétrica prevista pelo modelo de aprendizado de máquinas e o $MSE_{persistência}$ é o MSE para os valores previstos por um modelo de persistência. O índice de *skill-score* indica qual o ganho na redução de incerteza de um dado modelo quando comparado à um modelo de referência. Neste estudo o modelo de persistência inteligente (*smart*) (ANTONANZAS et al., 2016), descrito pela equação (3.7).

$$Persistência = Kt_i \cdot TOA_{i+h} \cdot \eta \cdot A_{mod} \cdot n_{mod} \quad (3.7)$$

Onde Kt_i é o índice de claridade no momento i , TOA_{i+h} é a irradiância horizontal no topo da atmosfera calculada para um horizonte temporal de h horas. η é a eficiência definida no *datasheet* do fabricante no módulo (15,92%). A_{mod} é a área de superfície de um módulo individual e n_{mod} é o número de módulos instalado no sistema FV de referência.

Valores de SS_{MSE} próximos de zero indicam que os dois modelos são equivalentes. Valores negativos mostram que o modelo de persistência possui menores desvios do que o modelo proposto. Valores positivos próximos a um são desejados e

demonstram que o modelo proposto reduziu significativamente as incertezas comparado ao modelo de referência (Persistência).

3.7. REDES NEURAIS ARTIFICIAIS RECORRENTES

Uma RNA convencional tem a capacidade de resolver problemas não lineares complexos, contudo não possui a habilidade de reter informações de iterações anteriores. Nesse sentido, as Redes Neurais Recorrentes vêm ganhando atenção na solução de problemas onde estados anteriores de uma série temporal influenciam o estado atual, como é o caso da previsão do recurso solar para geração solar fotovoltaica. Portanto, escolhemos uma RNN de múltiplas camadas (RNN) como a abordagem de aprendizado de máquina para prever a curto prazo a energia solar fotovoltaica no horizonte temporal com até três horas de antecedência. Os modelos RNN foram desenvolvidos para obter previsões precisas para diferentes condições atmosféricas: estação de clima seco (de abril a setembro) e estação chuvosa (de outubro a março). Portanto, a abordagem utilizada neste trabalho consiste no desenvolvimento de dois modelos de previsão diferentes, um para cada categoria de condição atmosférica. A mesma topologia foi empregada para os dois casos.

A fim de determinar a arquitetura ótima do modelo RNN utilizado neste trabalho foi implementado um Algoritmo de Força Bruta (AFB), que consiste em repetir o algoritmo de treinamento cinquenta vezes para cada das diversas topologias possíveis, escolhendo a melhor arquitetura com base na menor Raiz do Erro Quadrático Médio Nominal (NRMSE). O AFB foi desenvolvido levando em consideração que o mesmo número de neurônios deve ser empregado em todas as camadas ocultas, variando em um passo de cinco a cada iteração, a partir de 10 neurônios e parando em 50 neurônios. Foram considerados até quatro camadas ocultas e um atraso de até três iterações (1:3) em cada uma delas. A função Tangente Hiperbólica (Eq. 2.12) foi empregada em todas as camadas do modelo RNN. A função de treinamento utilizada foi o *Resilient Backpropagation* (RP) por ser rápida em reconhecer padrões no conjunto de dados e por apresentar menor requerimento de memória computacional (LOONKAR; MISHRA, 2016; MATHWORKS, 2021).

O computador utilizado para desenvolvimento dos modelos possui a seguinte configuração: Intel Core i7-4500U, 1,80GHz, 16GB de memória RAM e SSD 230GB. O

fluxograma da lógica de implementação do AFB é apresentado na Figura 3.9. A Tabela 3.1 elenca o tempo necessário, em segundos, para executar o AFB em função do número de camadas ocultas, permitindo também estimar a exigência computacional e o tempo de treinamento da arquitetura escolhida.

Tabela 3.1 – Tempo computacional de execução do Algoritmo de Força Bruta

Nº de camadas ocultas	H1	H2	H3	H4
função de treinamento	<i>trainlm</i>	<i>trainrp</i>	<i>trainrp</i>	<i>trainrp</i>
Tempo (s)	9370*	1817	1599	2884

trainlm = Levenberg-Marquardt

trainrp = Resilient Backpropagation

*Não completou o AFB. Abandonado em 9 iterações em 15 neurônios

Figura 3.9 – Diagrama de funcionamento do Algoritmo de Força Bruta

Os resultados obtidos através do AFB para o sistema CTPV são apresentados na Figura 3.10. Percebe-se que a partir de duas camadas ocultas o desempenho do modelo

é bastante semelhante, sobretudo entre 15 e 40 neurônios em cada camada. Contudo, o valor mínimo observado foi obtido na configuração de três camadas ocultas de 30 neurônios cada, representado pela linha tracejada RMSE mínimo. Contudo, observando que outras configurações também apresentaram desempenhos satisfatórios, foi considerado um limite superior considerando o percentil P25, representado pela linha pontilhada RMSE limite.

Figura 3.10 – NRMSE (kWh) do sistema referência CTPV para o Algoritmo de Força Bruta (AFB)

Desta forma, a arquitetura da rede neural recorrente escolhida para este trabalho foi de três camadas ocultas e trinta neurônios em cada uma destas, visto que esta configuração apresentou o melhor desempenho. Contudo, trabalhos subsequentes poderão empregar as outras arquiteturas situadas entre os dois limites de RMSE, caso o requerimento e tempo computacional seja um impeditivo. O diagrama da arquitetura escolhida para este trabalho está representado na Figura 3.11. A camada de entrada possui duas configurações possíveis. A primeira, que chamamos de convencional, é dada somente com os dados de índice de claridade, temperatura ambiente, velocidade do vento, umidade relativa, ângulo zenital, hora do dia e geração no instante zero. Na segunda, por sua vez, além das sete variáveis supramencionadas, estão inclusos dados modelados de aerossol. A mesma arquitetura foi utilizada para todos os sistemas de geração, bem como para as diferentes condições atmosféricas e horizontes temporais.

Figura 3.11 – Diagrama da arquitetura escolhida para o modelo RNN de previsão

Para fim de padronização, o sistema de nomenclatura dos modelos de previsão é composto pela sigla do sistema, a condição atmosférica (seco ou chuvoso), o horizonte temporal e a arquitetura utilizada (número de camadas ocultas e neurônios). A Figura 3.12 apresenta o sistema de nomenclatura utilizado. Nela os valores expressos em branco são variáveis para cada sistema e condição operacional. Nos modelos RNN em que a presença de aerossol seja utilizada como sinal de entrada, a sigla AOD é colocada ao final do código de referência.

Figura 3.12 – Esquema de nomenclatura dos modelos de previsão de geração FV. O modelo em referência é a previsão de geração no horizonte temporal de 1h a frente na estação seca quando presença de aerossol é empregada para o sistema CTPV. A RNN referência possui três camadas ocultas de trinta neurônios cada.

As matrizes das variáveis de entrada (índice de claridade, temperatura ambiente, velocidade do vento, umidade relativa, zênite, hora do dia e geração instantânea) e variável alvo (geração elétrica com deslocamento de uma, duas ou três horas no futuro) foram separadas em duas estações do ano: “Meses Secos” (período entre abril e setembro) e “Meses Chuvosos” (período entre outubro e março). Todas as RNN foram desenvolvidas conforme a arquitetura demonstrada na Figura 3.11, seguindo as mesmas configurações utilizadas no AFB anteriormente.

O conjunto de dados foi separado em três seções: treinamento, validação e teste. O conjunto de treinamento tem por finalidade ajustar os pesos e os vieses das unidades de processamento da RNN a cada iteração, também denominada *epoch*. O conjunto de validação, por sua vez, tem a finalidade de monitorar a melhora no desempenho do modelo, ou seja, a redução do MSE, à cada *epoch*. Quando a RNN não obtém êxito em reduzir o MSE por seis *epochs* consecutivas o algoritmo de treinamento é interrompido e os parâmetros do melhor desempenho mais recente são restaurados. Por fim, o conjunto de teste não impacta o ajuste dos parâmetros, tampouco serve como critério de parada. Este conjunto tem por finalidade avaliar o desempenho do modelo perante um conjunto de dados inédito (não utilizado no treinamento ou no critério de parada). A distribuição da base de dados foi feita da seguinte maneira: 60% para treinamento, 20% para validação e 20% para teste. O critério de seleção de cada seção foi feito de maneira aleatória, a fim de que cada etapa de desenvolvimento do modelo conte com dados que representem a diversidade dos eventos ao longo da estação do ano. Tendo em vista o caráter aleatório da seleção dos conjuntos de treinamento, validação e teste, o algoritmo de treinamento foi executado em uma rotina de 50 iterações buscando o melhor desempenho do modelo preditivo para o conjunto de teste.

Capítulo 4 : RESULTADOS E DISCUSSÃO

4. RESULTADOS E DISCUSSÃO

O presente capítulo visa apresentar os resultados obtidos pelos modelos RNN e a discussão destes será feita de forma separada para meses secos e chuvosos. Dentro de cada seção serão apresentados os resultados para todos os sistemas de geração solar fotovoltaica analisados, apresentando também a comparação da influência da presença de dados modelados de aerossol na capacidade preditiva do modelo.

As Seções 4.1 e 4.2 apresentam as regressões de treinamento, validação, teste e geral para diferentes horizontes temporais, comparando com o desempenho do modelo com a ausência (“a” à esquerda) ou presença (“b” à direita) de dados modelados de AOD. As representações gráficas são dispostas em quatro *subplots*. No primeiro quadrante, os resultados obtidos na etapa de validação. No segundo quadrante, os resultados obtidos na etapa de treinamento. No terceiro quadrante, os resultados obtidos na etapa de teste. Por fim, no quarto quadrante, os resultados para o conjunto geral de dados.

As representações gráficas das regressões de treinamento apresentam a dispersão entre os valores reais de geração (*target*) no eixo x e os valores de geração previstos (*output*) no eixo y. Em cada gráfico há uma linha pontilhada cruzando a diagonal principal, representando a condição de igualdade entre valores medidos e previstos. Há, também, uma análise de densidade probabilística de Kernel. Esta análise representa a densidade dos dados em um determinado ponto do eixo de coordenadas, sendo que densidades elevadas são representadas por cores quentes e densidades baixas são representadas por cores frias. Por fim, são apresentadas as regressões lineares de cada conjunto de dados do processo de treinamento, representadas por linhas cheias coloridas (azul para treinamento, verde para validação, vermelha para teste e preta para geral). As equações lineares destas regressões estão representadas no eixo y. Quando o coeficiente angular (*m*) é próximo a 1, significa proximidade entre os valores medidos e previstos. Valores superiores a 1, o modelo está sobrestimando a geração FV. Valores inferiores a 1, o modelo está subestimando a geração FV.

São considerados três conjuntos de dados: treinamento, que ajusta os pesos e vieses do modelo a cada iteração; validação, que tem a finalidade de definir o critério de parada do processo de treinamento do modelo; e teste, que tem por finalidade de avaliar o desempenho do modelo para dados inéditos. A união destes três conjuntos resulta no Geral. Ao final de cada seção estão resumidos os resultados obtidos nas Tabelas 4.1 – 4.6, que compreendem os parâmetros estatísticos das etapas de treinamento para todos os sistemas geradores FV para todas as condições atmosféricas e horizontes temporais de previsão.

4.1. MESES SECOS

4.1.1. CTPV

O conjunto de dados dos meses secos para o sistema CTPV conta com 4.493 pontos de medições horárias, após tratamento dos dados, considerando o período entre 2017 e 2020. A Figura 4.1 apresenta a regressão linear de treinamento do sistema CTPV para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de correlação (R) variando entre 0,947 – 0,951 e os valores de coeficientes de inclinação (m) das retas de ajuste variando entre 0,899 – 0,916. O $\text{NRMSE}_{156\text{kWh}}$ teste observado é de 7,15 kWh e o SS observado para o conjunto de teste é 0,887. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,948 – 0,953 e valores de m variando entre 0,908 – 0,924. O $\text{NRMSE}_{156\text{kWh}}$ teste observado é de 7,06 kWh e o SS observado para o conjunto de teste é 0,894.

Figura 4.1 – Regressão de treinamento sistema CTPV em meses secos para previsão de geração 1h no futuro.

A Figura 4.2 apresenta a regressão linear de treinamento do sistema CTPV para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,912 – 0,926 e valores m variando entre 0,828 – 0,856. O $NRMSE_{156kW}$ teste observado foi 9 kWh e o SS para o conjunto teste foi 0,817. Quando dados de aerossol foram incorporados como sinal de entrada do modelo, foram observados valores R variando entre 0,906 – 0,92 e valores de m variando entre 0,824 – 0,845. O $NRMSE_{156kW}$ teste observado foi 9,39 kWh e o SS para o conjunto teste foi 0,824.

Figura 4.2 – a Regressão de treinamento sistema CTPV em meses secos para previsão de geração 2h no futuro.

A Figura 4.3 apresenta a regressão linear de treinamento do sistema CTPV para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,869 – 0,88 e valores de m variando entre 0,766 – 0,785. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 11,34 kWh e o SS para o conjunto teste foi de 0,78. Quando dados de aerossol foram incorporados como sinal de entrada do modelo, foram observados valores de R variando entre 0,863 – 0,888 e valores de m variando entre 0,782 – 0,81. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 11,3 kWh e o SS para o conjunto teste foi de 0,776.

Figura 4.3 – Regressão de treinamento sistema CTPV em meses secos para previsão de geração 3h no futuro.

O modelo RNN apresentou um comportamento satisfatório para o sistema CTPV, com um $\text{RMSE}\%$ máximo de 7,39%. Mesmo considerando o horizonte temporal mais amplo de 3h o coeficiente de inclinação da reta de ajuste (m) com conjunto de dados permaneceu acima de 0,75, comportando-se de forma semelhante ao critério de igualdade entre valores previstos e medidos. Contudo, é observado que ao passo que se expande o horizonte temporal de predição, o modelo passa a subestimar a geração FV. Destaca-se a maior densidade para pontos dispostos ao longo da diagonal principal, apontando uma maior probabilidade de comportamento condizente com o valor de geração elétrica real medida.

4.1.2. BIPV

O conjunto de dados dos meses secos para o sistema BIPV conta com 2.895 pontos de medições horárias, após tratamento dos dados. A Figura 4.4 apresenta a

regressão linear de treinamento do sistema BIPV para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,925 – 0,958 e valores de m variando entre 0,912 – 0,935. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 7,79 kWh e o SS para o conjunto teste foi de 0,894. Quando dados de aerossol foram incorporados como sinal de entrada do modelo, foram observados valores R variando entre 0,95 – 0,963 e valores de m variando entre 0,909 – 0,938. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 7,08kWh e o SS para o conjunto teste foi de 0,904.

Figura 4.4 – Regressão de treinamento sistema BIPV em meses secos para previsão de geração 1h no futuro

A Figura 4.5 apresenta a regressão linear de treinamento do sistema BIPV para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,898 – 0,911 e valores de m variando entre 0,836 – 0,868. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 11,27 kWh e o SS para o conjunto teste foi 0,737. Quando dados de aerossol foram incorporados como sinal de entrada do modelo RNN, foram observados valores de R variando entre 0,913 – 0,926 e valores de m variando entre 0,829 – 0,862. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 9,77 kWh e o SS para o conjunto teste foi de 0,829.

Figura 4.5 – Regressão de treinamento sistema BIPV em meses secos para previsão de geração 2h no futuro

A Figura 4.6 apresenta a regressão linear de treinamento do sistema BIPV para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,853 – 0,886 e valores de m variando entre 0,741 – 0,794. O $\text{NRMSE}_{150\text{kW}}$ teste observado foi de 12,8 kWh e o SS para o conjunto teste foi de 0,77. Quando dados de aerossol foram incorporados como sinal de entrada do modelo, foram observados valores de R variando entre 0,879 – 0,885 e valores de m variando entre 0,778 – 0,81. O $\text{NRMSE}_{150\text{kW}}$ teste observado foi de 12,76 kWh e o SS para o conjunto teste foi de 0,736.

Figura 4.6 – Regressão de treinamento sistema BIPV em meses secos para previsão de geração 3h no futuro.

O modelo RNN apresentou um comportamento satisfatório para o sistema BIPV, com um RMSE% máximo de 8,48%. Mesmo considerando o horizonte temporal mais amplo de 3h os valores de coeficiente de inclinação (m) da reta de ajuste do conjunto de dados permaneceram acima de 0,74. O modelo apresentou um comportamento semelhante ao sistema CTPV, sendo observado também que ao passo que se expande o horizonte temporal de predição, o modelo passa a subestimar a geração FV. Desta forma, as observações apontadas na seção anterior também valem para o sistema BIPV, levando em consideração, contudo, as características estruturais de ambos os sistemas de geração.

4.1.3. BAPV

O conjunto de dados referente aos meses secos para o sistema BAPV conta com 1.767 pontos de medições horárias, após tratamento dos dados. A Figura 4.7 apresenta a regressão linear de treinamento do sistema BAPV para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,951 – 0,958 e valores de m variando entre 0,902 – 0,937. O NRMSE_{156kW} teste observado foi de 8,23 kWh e o SS para o conjunto teste foi de 0,918. Quando dados de aerossol foram incorporados como dados de entrada, foram observados valores de R variando entre 0,945 – 0,963 e valores de m variando entre 0,91 – 0,959. O NRMSE_{156kW} teste observado foi de 7,98 kWh e o SS para o conjunto teste foi de 0,918.

Figura 4.7 – Regressão de treinamento sistema BAPV em meses secos para previsão de geração 1h no futuro

A Figura 4.8 apresenta a regressão linear de treinamento do sistema BAPV para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,885 – 0,922 e valores de m variando entre 0,807 – 0,847. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 11,59 kWh e o SS para o conjunto teste foi de 0,822. Quando dados de aerossol foram incorporados como dados de entrada do modelo RNN, foram observados valores de R variando entre 0,887 – 0,913 e valores de m variando entre 0,826 – 0,872. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 11,55 kWh e o SS para o conjunto teste foi de 0,842.

Figura 4.8 – Regressão de treinamento sistema BAPV em meses secos para previsão de geração 2h no futuro.

A Figura 4.9 apresenta a regressão linear de treinamento do sistema BAPV para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,819 – 0,864 e valores de m variando entre 0,679 – 0,735. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 15,22 kWh e o SS para o conjunto teste foi de 0,724. Quando dados de aerossol foram incorporados como dados de entrada, foram observadas correlações variando entre 0,801 – 0,844 e as inclinações das retas de ajuste variando entre de 0,658 – 0,722. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 15,6 kWh e o SS para o conjunto teste foi de 0,691.

Figura 4.9 – Regressão de treinamento sistema BAPV em meses secos para previsão de geração 3h no futuro.

Assim como para os sistemas FV anteriores, o modelo proposto também apresentou um comportamento relativamente satisfatório para o sistema BAPV, com um RMSE% máximo de 10,6%. Contudo, comparada a seus pares, o desempenho para este sistema demonstrou-se ligeiramente inferior. Este resultado era esperado levando em consideração a quantidade de medições válidas (1.767) para o processo de treinamento, validação e teste do modelo. Como foi discutido na Seção 2.6, modelos estatísticos são fortemente dependentes da disponibilidade de longas séries históricas de dados. Outro fator que pode ter levado a uma redução no desempenho do modelo RNN para o sistema BAPV é a localização deste sistema gerador. O sistema BAPV encontra-se instalado na cobertura da biblioteca Brasileira, sendo o sistema mais afastado da estação de solarimétrica. Variáveis como velocidade dos ventos, umidade, temperatura ambiente, irradiância e mesmo o ângulo zenital apresentam variabilidade espacial que pode ser significativa a depender da distância. No caso do recurso solar, em especial, a variabilidade pode ser significativa, mesmo em curtas distâncias, em função da cobertura de nuvens, cobertura do solo, sombreamento e dentre outras variáveis.

Vale apontar, contudo, que mesmo levando em consideração os desafios mencionados acima, o modelo RNN apresentou desempenho compatível a outros estudos semelhantes encontrados na literatura (DE GIORGI; CONGEDO; MALVONI, 2014; FILIPE et al., 2015; PEDRO; COIMBRA, 2012), conforme será discutido mais profundamente na Seção 4.5. O comportamento do modelo para o sistema BAPV pode ser de grande ajuda para elaborar um modelo que seja capaz de prever a geração solar

para uma determinada região, não estando limitada a um único ponto. Esse ponto em específico será discutido mais profundamente na Seção 4.4.

As Tabelas 4.1 e 4.2 apresentam em maior detalhe os resultados de MSE e NRMSE para todos os sistemas FV estudados. Nelas são dispostos, também, o desempenho do modelo de persistência considerado como referência. Os modelos para os sistemas CTPV, BIPV e BAPV desenvolvidos apresentaram desempenhos consistentemente superiores aos apresentados no modelo de persistência para todos os horizontes temporais analisados.

Na Tabela 4.3 é apresentado o *Skill Score* (SS_{mse}), que indica a habilidade preditiva de um modelo em comparação ao modelo de referência. Valores positivos de SS indicam um ganho de desempenho do modelo proposto, enquanto valores negativos de SS indicam um desempenho inferior deste. O modelo RNN proposto apresentou SS positivos para todos os cenários analisados, apresentando um valor mínimo de 0,65 e um valor máximo de 0,91.

Tabela 4.1 – Resumo dos resultados de Erro Quadrático Médio (MSE) em ($\times 10^{-3}$) do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses secos nos horizontes temporais de uma, duas e três horas.

	CTPV					BIPV					BAPV					IEE				
	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All
Meses secos	CTPV-S1-H3-N30-AOD	2,33	2,42	2,05	2,29	BIPV-S1-H3-N30-AOD	3,00	2,76	2,23	2,80	BAPV-S1-H3-N30-AOD	2,77	3,72	2,61	2,93	IEE-S1-H3-N30-AOD	2,68	2,54	3,03	2,72
	CTPV-S1-H3-N30	2,41	2,29	2,10	2,32	BIPV-S1-H3-N30	2,93	2,78	2,70	2,85	BAPV-S1-H3-N30	3,31	3,07	2,78	3,15	IEE-S1-H3-N30	3,58	3,79	2,90	3,49
	CTPV-P-S1	19,87	17,59	19,34	19,31	BIPV-P-S1	26,39	24,98	23,22	25,48	BAPV-P-S1	30,81	29,41	31,93	30,76	IEE-P-S1	19,56	19,24	20,11	19,60
	CTPV-S2-H3-N30-AOD	3,57	4,14	3,62	3,69	BIPV-S2-H3-N30-AOD	5,35	5,00	4,24	5,06	BAPV-S2-H3-N30-AOD	6,65	7,85	5,48	6,66	IEE-S2-H3-N30-AOD	4,04	5,05	4,40	4,31
	CTPV-S2-H3-N30	3,85	3,79	3,33	3,73	BIPV-S2-H3-N30	6,32	5,36	5,64	5,99	BAPV-S2-H3-N30	7,38	7,21	5,52	6,97	IEE-S2-H3-N30	4,61	5,05	3,84	4,54
	CTPV-P-S2	19,83	22,11	20,58	20,43	BIPV-P-S2	26,94	25,09	24,86	26,15	BAPV-P-S2	29,67	31,46	34,70	31,03	IEE-P-S2	20,47	21,30	20,46	20,63
	CTPV-S3-H3-N30-AOD	4,94	5,74	5,25	5,16	BIPV-S3-H3-N30-AOD	6,85	7,20	7,24	7,00	BAPV-S3-H3-N30-AOD	11,24	12,50	10,00	11,25	IEE-S3-H3-N30-AOD	6,24	5,98	5,77	6,10
	CTPV-S3-H3-N30	5,63	5,15	5,28	5,46	BIPV-S3-H3-N30	6,62	8,85	7,29	7,20	BAPV-S3-H3-N30	11,51	10,88	9,52	10,98	IEE-S3-H3-N30	6,48	6,93	6,67	6,61
	CTPV-P-S3	23,85	23,82	23,45	23,77	BIPV-P-S3	29,17	25,53	27,44	28,10	BAPV-P-S3	34,78	31,71	32,37	33,68	IEE-P-S3	24,43	22,70	20,82	23,36

Tr = Conjunto de treinamento

Val = Conjunto de validação

Te = Conjunto de teste

All = Tr + Val + Te

Tabela 4.2 – Resumo dos resultados de Raiz do Erro Quadrático Médio Nominal (NRMSE), em kWh, do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses secos nos horizontes temporais de uma, duas e três horas.

	CTPV					BIPV					BAPV					IEE				
	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All
Meses secos	CTPV-S1-H3-N30-AOD	7,52	7,67	7,06	7,46	BIPV-S1-H3-N30-AOD	8,21	7,89	7,08	7,93	BAPV-S1-H3-N30-AOD	8,20	9,51	7,98	8,44	IEE-S1-H3-N30-AOD	8,08	7,86	8,58	8,14
	CTPV-S1-H3-N30	7,65	7,46	7,15	7,52	BIPV-S1-H3-N30	8,12	7,91	7,79	8,01	BAPV-S1-H3-N30	8,97	8,64	8,23	8,76	IEE-S1-H3-N30	9,34	9,60	8,40	9,21
	CTPV-P-S1	21,99	20,69	21,70	21,68	BIPV-P-S1	24,37	23,71	22,86	23,94	BAPV-P-S1	27,38	26,75	27,88	27,36	IEE-P-S1	21,82	21,64	22,12	21,84
	CTPV-S2-H3-N30-AOD	9,32	10,04	9,39	9,48	BIPV-S2-H3-N30-AOD	10,97	10,61	9,77	10,67	BAPV-S2-H3-N30-AOD	12,72	13,82	11,55	12,73	IEE-S2-H3-N30-AOD	9,91	11,09	10,34	10,24
	CTPV-S2-H3-N30	9,67	9,60	9,00	9,53	BIPV-S2-H3-N30	11,93	10,98	11,27	11,61	BAPV-S2-H3-N30	13,40	13,25	11,59	13,03	IEE-S2-H3-N30	10,59	11,08	9,66	10,52
	CTPV-P-S2	21,97	23,20	22,38	22,30	BIPV-P-S2	24,62	23,76	23,65	24,26	BAPV-P-S2	26,87	27,67	29,06	27,48	IEE-P-S2	22,32	22,77	22,31	22,41
	CTPV-S3-H3-N30-AOD	10,96	11,82	11,30	11,21	BIPV-S3-H3-N30-AOD	12,42	12,73	12,76	12,55	BAPV-S3-H3-N30-AOD	16,54	17,44	15,60	16,54	IEE-S3-H3-N30-AOD	12,33	12,07	11,85	12,18
	CTPV-S3-H3-N30	11,70	11,19	11,34	11,53	BIPV-S3-H3-N30	12,20	14,11	12,80	12,72	BAPV-S3-H3-N30	16,73	16,27	15,22	16,35	IEE-S3-H3-N30	12,56	12,99	12,74	12,68
	CTPV-P-S3	24,09	24,08	23,89	24,05	BIPV-P-S3	25,62	23,97	24,85	25,14	BAPV-P-S3	29,09	27,78	28,07	28,63	IEE-P-S3	24,38	23,50	22,51	23,84

NRMSE encontra-se em kWh em relação a cada sistema

Tr = Conjunto de treinamento

Val = Conjunto de validação

Te = Conjunto de teste

All = Tr + Val + Te

Tabela 4.3 – Resumo dos resultados de Skill Score (SS_{mse}) do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses secos nos horizontes temporais de uma, duas e três horas.

		CTPV				BIPV				BAPV				IEE						
		Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te
Meses secos	CTPV-S1-H3-N30-AOD	0,8829	0,8625	0,8941	0,8814	BIPV-S1-H3-N30-AOD	0,8864	0,8894	0,9040	0,8902	BAPV-S1-H3-N30-AOD	0,9102	0,8737	0,9181	0,9049	IEE-S1-H3-N30-AOD	0,8627	0,8681	0,8496	0,8611
	CTPV-S1-H3-N30	0,8739	0,8803	0,8873	0,8778	BIPV-S1-H3-N30	0,8884	0,8801	0,8937	0,8880	BAPV-S1-H3-N30	0,8923	0,8897	0,9178	0,8974	IEE-S1-H3-N30	0,8208	0,7929	0,8532	0,8221
	CTPV-S2-H3-N30-AOD	0,8199	0,8128	0,8240	0,8192	BIPV-S2-H3-N30-AOD	0,8013	0,8006	0,8292	0,8064	BAPV-S2-H3-N30-AOD	0,7758	0,7504	0,8419	0,7854	IEE-S2-H3-N30-AOD	0,8027	0,7627	0,7852	0,7910
	CTPV-S2-H3-N30	0,8194	0,8117	0,8170	0,8174	BIPV-S2-H3-N30	0,7646	0,8135	0,7372	0,7708	BAPV-S2-H3-N30	0,7508	0,7956	0,8222	0,7753	IEE-S2-H3-N30	0,7729	0,7711	0,8099	0,7797
	CTPV-S3-H3-N30-AOD	0,7930	0,7592	0,7762	0,7829	BIPV-S3-H3-N30-AOD	0,7651	0,7179	0,7362	0,7509	BAPV-S3-H3-N30-AOD	0,6767	0,6056	0,6910	0,6661	IEE-S3-H3-N30-AOD	0,7443	0,7364	0,7229	0,7390
	CTPV-S3-H3-N30	0,7578	0,7947	0,7802	0,7701	BIPV-S3-H3-N30	0,7598	0,6610	0,7707	0,7439	BAPV-S3-H3-N30	0,6552	0,6783	0,7242	0,6739	IEE-S3-H3-N30	0,7241	0,7023	0,7102	0,7170

Tr = Conjunto de treinamento

Val = Conjunto de validação $SS_{mse} = 1 - \frac{MSE_{modelo}}{MSE_{persist}}$

Te = Conjunto de teste

All = Tr + Val + Te

4.2. MESES CHUVOSOS

4.2.1. CTPV

O conjunto de dados referente aos meses chuvosos para o sistema CTPV conta com 4.325 pontos de medições horárias, após tratamento dos dados, considerando o período entre 2017 e 2020. A Figura 4.10 apresenta a regressão linear de treinamento do sistema CTPV com 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,905 – 0,92 e valores de coeficiente angular das retas de ajuste (**m**) variando entre 0,831 – 0,865. O $NRMSE_{156kW}$ teste observado foi de 9,24 kWh e o SS para o conjunto teste foi de 0,923. Quando dados de aerossol foram incorporados como sinais de entrada do modelo, foram observados valores de R variando entre 0,903 – 0,922 e valores de **m** variando entre 0,847 – 0,87. O $NRMSE_{156kW}$ teste observado foi de 9,12 kWh e o SS para o conjunto teste foi 0,924.

Figura 4.10 – Regressão de treinamento sistema CTPV em meses chuvosos para previsão de geração 1h no futuro.

A Figura 4.11 apresenta a regressão linear de treinamento do sistema CTPV para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,865 – 0,882 e valores de m variando entre 0,757 – 0,777. O NRMSE_{156kW} teste observado foi de 11,19 kWh e o SS para o conjunto teste foi de 0,883. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,857 – 0,888 e valores de m variando entre 0,756 – 0,796. O NRMSE_{156kW} teste observado foi de 10,94 kWh e o SS para o conjunto teste foi de 0,889.

Figura 4.11 – Regressão de treinamento sistema CTPV em meses chuvosos para previsão de geração 2h no futuro.

A Figura 4.12 apresenta a regressão linear de treinamento do sistema CTPV para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,835 – 0,84 e valores de m variando entre 0,689 – 0,731. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 12,61 kWh e o SS para o conjunto teste foi de 0,857. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,83 – 0,86 e valores de m variando entre 0,693 – 0,732. O $\text{NRMSE}_{156\text{kW}}$ teste observado foi de 12,04 kWh e o SS para o conjunto teste foi de 0,863.

Figura 4.12 – Regressão de treinamento sistema CTPV em meses chuvosos para previsão de geração 3h no futuro.

O modelo proposto apresentou um comportamento satisfatório para o sistema CTPV, com um $\text{RMSE}_{\%}$ máximo de 8,21%. Mesmo considerando o horizonte temporal mais amplo de 3h o coeficiente de inclinação m permaneceu acima de 0,68. Destaca-se a maior densidade de dados dispostos ao longo da diagonal principal, apontando uma maior probabilidade de comportamento condizente com o valor de geração elétrica real medida mesmo em condições atmosféricas mais sujeitas a cobertura de nuvens.

4.2.2. BIPV

O conjunto de dados referente aos meses chuvosos para o sistema BIPV conta com 4.577 pontos de medições horárias, após tratamento dos dados. A Figura 4.13 apresenta a regressão linear de treinamento do sistema BIPV para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando

entre 0,905 – 0,925 e valores de m variando entre 0,841 – 0,871. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 9,4 kWh e o SS para o conjunto teste foi de 0,907. Quando dados de aerossol foram incorporados como dados de entrada do modelo RNN, foram observados valores de R variando entre 0,914 – 0,922 e valores de m variando entre 0,841 – 0,868. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 9,7 kWh e o SS para o conjunto teste foi de 0,9.

Figura 4.13 – Regressão de treinamento sistema BIPV em meses chuvosos para previsão de geração 1h no futuro.

A Figura 4.14 apresenta a regressão linear de treinamento do sistema BIPV para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,874 – 0,881 e valores de m variando entre de 0,761 – 0,78. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 12,06 kWh e o SS para o conjunto teste foi de 0,86. Quando dados de aerossol foram incorporados como sinal de entrada, foram observados valores de R variando entre 0,851 – 0,883 e valores de m variando entre de 0,745 – 0,779. O $\text{NRMSE}_{150\text{kWh}}$ teste observado foi de 11,74 kWh e o SS para o conjunto teste foi de 0,854.

Figura 4.14 – Regressão de treinamento sistema BIPV em meses chuvosos para previsão de geração 2h no futuro.

A Figura 4.15 apresenta a regressão linear de treinamento do sistema BIPV para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,834 – 0,846 e valores de m variando entre de 0,701 – 0,73. O NRMSE_{150kW} teste observado foi de 13,75 kWh e o SS para o conjunto teste foi de 0,815. Quando dados de aerossol foram incorporados como sinal de entrada do modelo, foram observados valores de R variando entre 0,826 – 0,849 e valores de m variando entre de 0,67 – 0,718. O NRMSE_{150kW} teste observado foi de 13,39 kWh e o SS para o conjunto teste foi de 0,825.

Figura 4.15 – Regressão de treinamento sistema BIPV em meses chuvosos para previsão de geração 3h no futuro.

O modelo proposto apresentou um comportamento satisfatório para o sistema BIPV, com um RMSE% máximo de 9,18%. Mesmo considerando o horizonte temporal mais amplo de 3h os coeficientes de inclinação m apresentou valores entre 0,67 e 0,87. Assim como ocorreu para os meses secos, o modelo RNN apresentou um comportamento semelhante ao sistema CTPV.

4.2.3. BAPV

O conjunto de dados referente aos meses chuvosos para o sistema BAPV conta com 3.562 pontos de medições horárias, após tratamento dos dados. A Figura 4.16 apresenta a regressão linear de treinamento do sistema BAPV para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,928 – 0,932 e valores de m variando entre de 0,874 – 0,9. O NRMSE_{156kW} teste observado foi de 10,16 kWh e o SS para o conjunto teste foi de 0,93. Quando dados de aerossol foram incorporados como dados de entrada, foram observados valores de R variando entre 0,928 – 0,936 e valores de m variando entre de 0,859 – 0,9. O NRMSE_{156kW} teste observado foi de 10,69 kWh e o SS para o conjunto teste foi de 0,923.

Figura 4.16 – Regressão de treinamento sistema BAPV em meses chuvosos para previsão de geração 1h no futuro

A Figura 4.17 apresenta a regressão linear de treinamento do sistema BAPV para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,873 – 0,9 e valores de m variando entre de 0,784 – 0,825. O NRMSE_{156kW} teste observado foi de 12,44 kWh e o SS para o conjunto teste foi de 0,898. Quando dados de aerossol foram incorporados como sinal de entrada do modelo,

foram observados valores de R variando entre 0,885 – 0,895 e valores de m variando entre 0,803 – 0,839. O $\text{NRMSE}_{156\text{kWh}}$ teste observado foi de 12,63 kWh e o SS para o conjunto teste foi de 0,89.

Figura 4.17 – Regressão de treinamento sistema BAPV em meses chuvosos para previsão de geração 2h no futuro

A Figura 4.18 apresenta a regressão linear de treinamento do sistema BAPV para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,816 – 0,866 e valores de m variando entre de 0,709 – 0,748. O $\text{NRMSE}_{156\text{kWh}}$ teste observado foi de 14,07 kWh e o SS para o conjunto teste foi de 0,87. Quando dados de aerossol foram incorporados como dados de entrada, foram observados valores de R variando entre 0,834 – 0,863 e valores de m variando entre de 0,715 – 0,749. O $\text{NRMSE}_{156\text{kWh}}$ teste observado foi de 14,74 kWh e o SS para o conjunto teste foi de 0,857.

Figura 4.18 – Regressão de treinamento sistema BAPV em meses chuvosos para previsão de geração 3h no futuro

Assim como para os sistemas anteriores, o modelo proposto também apresentou um comportamento relativamente satisfatório para o sistema BAPV, com um RMSE% máximo de 9,81%. Diferentemente do que foi observado na análise referente aos meses secos, durante o período chuvoso o modelo RNN para o sistema BAPV apresentou desempenhos semelhantes aos seus demais sistemas FV. Neste período, o sistema BAPV contou com 3.562 medições válidas, enquanto os sistemas CTPV e BIPV contaram com 4.325 e 4.577 medições válidas, respectivamente. Este comportamento corrobora com a literatura no que tange ao comportamento de métodos estatísticos de previsão, dependentes da série histórica de dados para desempenhar bem. Além disso, o desempenho do modelo para o sistema BAPV, sendo este o mais afastado da estação solarimétrica e os demais sistemas FV, indica que a aplicação deste método não está limitada a um ponto específico, podendo ser usada para a previsão FV para uma área próxima.

As Tabelas 4.4 e 4.5 apresentam em maior detalhe os resultados de MSE e NRMSE discutidos anteriormente. Assim como ocorreu nos meses secos, durante o período chuvoso o modelo proposto neste trabalho apresentou desempenho consistentemente superior aos apresentados no modelo de persistência para todos os horizontes temporais analisados.

Na Tabela 4.6 é apresentado o *Skill Score* (SS_{mse}) para os sistemas geradores CTPV, BIPV e BAPV, assim como um modelo conjunto que será discutido na próxima

seção. O modelo RNN apresentou SS_{mse} positivos para todos os cenários analisados, apresentando um valor mínimo de 0,82 e um valor máximo de 0,93.

Comparando ambos os períodos secos e chuvosos, o segundo conjunto apresentou valores de MSE e NRMSE mais elevados, o que é esperado visto a maior presença de nebulosidade e consequentemente o aumento da incerteza do recurso solar que incide a superfície terrestre. Em contrapartida, o SS_{mse} demonstrou-se mais robusto e estável ao longo do horizonte temporal de previsão nos meses chuvosos. Isso indica o êxito do método neste trabalho em desempenhar previsões de geração FV satisfatoriamente em condições de operação na qual a previsão em condições atmosféricas mais desafiadoras.

Tabela 4.4 – Resumo dos resultados de Erro Quadrático Médio (MSE) em, $\times 10^{-3}$, do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses chuvosos nos horizontes temporais de uma, duas e três horas.

	CTPV					BIPV					BAPV					IEE				
	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All
Meses Chuvosos	CTPV-C1-H3-N30-AOD	3,76	4,05	3,42	3,75	BIPV-C1-H3-N30-AOD	4,77	4,07	4,18	4,51	BAPV-C1-H3-N30-AOD	4,05	4,27	4,70	4,22	IEE-C1-H3-N30-AOD	4,10	3,79	3,74	3,96
	CTPV-C1-H3-N30	3,50	4,12	3,51	3,63	BIPV-C1-H3-N30	4,39	5,16	3,92	4,45	BAPV-C1-H3-N30	4,31	4,79	4,24	4,39	IEE-C1-H3-N30	4,04	3,80	3,53	3,89
	CTPV-P-C1	45,75	44,55	45,27	45,41	BIPV-P-C1	44,17	46,18	42,71	44,28	BAPV-P-C1	61,16	55,65	61,17	60,06	IEE-P-C1	43,99	45,22	45,75	44,59
	CTPV-C2-H3-N30-AOD	5,40	6,03	4,91	5,43	BIPV-C2-H3-N30-AOD	7,23	7,68	6,12	7,10	BAPV-C2-H3-N30-AOD	6,57	7,23	6,56	6,70	IEE-C2-H3-N30-AOD	5,86	5,96	5,60	5,82
	CTPV-C2-H3-N30	5,39	5,80	5,14	5,42	BIPV-C2-H3-N30	6,69	6,30	6,47	6,57	BAPV-C2-H3-N30	7,99	6,15	6,36	7,29	IEE-C2-H3-N30	5,92	6,03	5,36	5,83
	CTPV-P-C2	46,15	44,38	44,50	45,47	BIPV-P-C2	45,28	44,63	42,03	44,50	BAPV-P-C2	61,84	63,37	59,72	61,72	IEE-P-C2	45,11	45,23	43,90	44,89
	CTPV-C3-H3-N30-AOD	6,82	7,27	5,96	6,74	BIPV-C3-H3-N30-AOD	8,21	9,57	7,96	8,43	BAPV-C3-H3-N30-AOD	9,72	10,05	8,92	9,63	IEE-C3-H3-N30-AOD	6,95	8,56	7,09	7,30
	CTPV-C3-H3-N30	6,87	6,52	6,54	6,73	BIPV-C3-H3-N30	8,06	8,54	8,40	8,22	BAPV-C3-H3-N30	8,31	11,05	8,13	8,82	IEE-C3-H3-N30	6,88	7,64	7,23	7,10
	CTPV-P-C3	46,66	46,65	43,65	46,06	BIPV-P-C3	46,35	45,01	45,45	45,90	BAPV-P-C3	65,16	62,95	62,31	64,15	IEE-P-C3	45,11	48,98	45,76	46,02

Tr = Conjunto de treinamento
Val = Conjunto de validação
Te = Conjunto de teste
All = Tr + Val + Te

Tabela 4.5 – Resumo dos resultados de Raiz do Erro Quadrático Médio Nominal (NRMSE), em kWh, do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses chuvosos nos horizontes temporais de uma, duas e três horas.

	CTPV					BIPV					BAPV					IEE				
	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All
Meses Chuvosos	CTPV-C1-H3-N30-AOD	9,57	9,93	9,12	9,55	BIPV-C1-H3-N30-AOD	10,36	9,56	9,70	10,08	BAPV-C1-H3-N30-AOD	9,92	10,19	10,69	10,13	IEE-C1-H3-N30-AOD	9,98	9,61	9,54	9,82
	CTPV-C1-H3-N30	9,22	10,02	9,24	9,39	BIPV-C1-H3-N30	9,94	10,77	9,40	10,01	BAPV-C1-H3-N30	10,24	10,80	10,16	10,34	IEE-C1-H3-N30	9,92	9,61	9,27	9,73
	CTPV-P-C1	33,37	32,93	33,19	33,24	BIPV-P-C1	31,52	32,23	31,00	31,56	BAPV-P-C1	38,58	36,80	38,58	38,23	IEE-P-C1	32,72	33,17	33,37	32,94
	CTPV-C2-H3-N30-AOD	11,46	12,11	10,94	11,49	BIPV-C2-H3-N30-AOD	12,75	13,15	11,74	12,64	BAPV-C2-H3-N30-AOD	12,65	13,27	12,63	12,77	IEE-C2-H3-N30-AOD	11,94	12,05	11,67	11,91
	CTPV-C2-H3-N30	11,46	11,88	11,19	11,49	BIPV-C2-H3-N30	12,27	11,90	12,06	12,16	BAPV-C2-H3-N30	13,94	12,23	12,44	13,32	IEE-C2-H3-N30	12,00	12,12	11,42	11,91
	CTPV-P-C2	33,51	32,86	32,91	33,26	BIPV-P-C2	31,92	31,69	30,75	31,64	BAPV-P-C2	38,79	39,27	38,12	38,76	IEE-P-C2	33,13	33,18	32,68	33,05
	CTPV-C3-H3-N30-AOD	12,89	13,30	12,04	12,81	BIPV-C3-H3-N30-AOD	13,59	14,68	13,39	13,77	BAPV-C3-H3-N30-AOD	15,38	15,64	14,74	15,31	IEE-C3-H3-N30-AOD	13,00	14,43	13,13	13,33
	CTPV-C3-H3-N30	12,93	12,60	12,61	12,80	BIPV-C3-H3-N30	13,46	13,86	13,75	13,60	BAPV-C3-H3-N30	14,22	16,40	14,07	14,65	IEE-C3-H3-N30	12,94	13,64	13,27	13,15
	CTPV-P-C3	33,70	33,69	32,59	33,48	BIPV-P-C3	32,29	31,82	31,98	32,14	BAPV-P-C3	39,82	39,14	38,94	39,51	IEE-P-C3	33,13	34,52	33,37	33,46

NRMSE encontra-se em kWh em relação a cada sistema
Tr = Conjunto de treinamento
Val = Conjunto de validação
Te = Conjunto de teste
All = Tr + Val + Te

Tabela 4.6 – Resumo dos resultados de Skill Score (SS_{mse}) do conjunto de treinamento dos sistemas CTPV, BIPV, BAPV e Conjunto IEE para meses chuvosos nos horizontes temporais de uma, duas e três horas.

	CTPV				BIPV				BAPV				IEE							
	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All	Modelo	Tr	Val	Te	All
Meses Chuvosos	CTPV-C1-H3-N30-AOD	0,9178	0,9090	0,9245	0,9174	BIPV-C1-H3-N30-AOD	0,8919	0,9120	0,9020	0,8981	BAPV-C1-H3-N30-AOD	0,9339	0,9233	0,9232	0,9297	IEE-C1-H3-N30-AOD	0,9069	0,9161	0,9182	0,9111
	CTPV-C1-H3-N30	0,9231	0,9086	0,9228	0,9202	BIPV-C1-H3-N30	0,8990	0,8944	0,9069	0,8995	BAPV-C1-H3-N30	0,9260	0,9256	0,9307	0,9269	IEE-C1-H3-N30	0,9114	0,9112	0,9185	0,9127
	CTPV-C2-H3-N30-AOD	0,8830	0,8642	0,8896	0,8806	BIPV-C2-H3-N30-AOD	0,8403	0,8278	0,8544	0,8405	BAPV-C2-H3-N30-AOD	0,8937	0,8859	0,8902	0,8914	IEE-C2-H3-N30-AOD	0,8702	0,8682	0,8725	0,8702
	CTPV-C2-H3-N30	0,8835	0,8694	0,8832	0,8807	BIPV-C2-H3-N30	0,8502	0,8503	0,8605	0,8524	BAPV-C2-H3-N30	0,8719	0,8962	0,8979	0,8818	IEE-C2-H3-N30	0,8716	0,8632	0,8724	0,8701
	CTPV-C3-H3-N30-AOD	0,8538	0,8442	0,8634	0,8537	BIPV-C3-H3-N30-AOD	0,8229	0,7873	0,8248	0,8163	BAPV-C3-H3-N30-AOD	0,8508	0,8403	0,8568	0,8499	IEE-C3-H3-N30-AOD	0,8460	0,8252	0,8451	0,8414
	CTPV-C3-H3-N30	0,8523	0,8552	0,8572	0,8538	BIPV-C3-H3-N30	0,8250	0,8148	0,8146	0,8209	BAPV-C3-H3-N30	0,8680	0,8403	0,8703	0,8625	IEE-C3-H3-N30	0,8510	0,8363	0,8387	0,8456

Tr = Conjunto de treinamento

Val = Conjunto de validação

Te = Conjunto de teste

All = Tr + Val + Te

$$SS_{mse} = 1 - \frac{MSE_{modelo}}{MSE_{persist}}$$

4.5. INFLUÊNCIA DO ÍNDICE DE PROFUNDIDADE ÓTICA DO AEROSSOL ATMOSFÉRICO (AOD) NO DESEMPENHO DOS MODELOS PREDITIVOS

As Figuras 4.1 – 4.18 apresentam as saídas do modelo preditivo quando o índice de profundidade ótica (AOD) em ausência de dados de aerossol atmosférico (a) e quando os dados de AOD 550nm são considerados como sinal de entrada do modelo (b). É possível perceber, tanto nas representações gráficas, quanto nas métricas estatísticas (Tabelas 4.1 – 4.6) que a influência do AOD na precisão do modelo de previsão é tímida, porém demonstra-se importante ao apresentar um padrão de comportamento bem definido. Para os meses secos, foi observado que a inclusão de dados de AOD 550 nm modelado como variável de entrada resultou em um ganho médio de $SS_{MSE\%}$ de 2,27% para até duas horas no futuro. Em contrapartida, o inverso ocorreu para a análise feita nos meses chuvosos. Neste caso, a inclusão do aerossol atmosférico como sinal de entrada do modelo RNN representou um ruído no modelo de previsão. Os valores de SS_{RMSE} variaram entre -4,13% e 4,51% independente do sistema FV e do horizonte temporal. Isso indica que em condições atmosféricas com maior presença de nuvens o desempenho do modelo independe da presença de dados de aerossóis. Ainda mais, nesta condição, a inclusão de AOD representa um ruído na camada de entrada da RNN, apresentando uma tendência de piora no desempenho.

Houveram, contudo, exceções ao comportamento descrito acima, sobretudo em relação ao modelo BAPV-S3-H3-N30. A substituição do AOD modelado por variáveis observadas localmente pode ajudar a elevar o ganho no modelo preditivo. Nesse sentido, uso de visibilidade atmosférica pode ser uma possibilidade a ser estudada.

4.6. MODELO CONJUNTO IEE

4.4.1. DESENVOLVIMENTO E TREINAMENTO

Os modelos preditivos de Redes Neurais Recorrentes multicamada apresentaram desempenhos satisfatórios quando analisados individualmente para cada sistema FV. Foi observado, contudo, que os modelos discutidos nas seções anteriores apresentaram comportamentos semelhantes para os sistemas CTPV, BIPV e BAPV em diferentes condições atmosféricas e operacionais. Além disso, observamos que a disponibilidade de uma base de dados robusta para treinamento influencia positivamente no desempenho do modelo preditivo. Nesse sentido, esta seção busca avaliar o desempenho de um modelo RNN que incorpore os dados de geração destes três sistemas no processo de treinamento da rede. A Figura 4.19 ilustra a comparação da geração real normalizada entre os sistemas FV. Da esquerda à direita, respectivamente, CTPV *versus* BIPV, BIPV *versus* BAPV e CTPV *versus* BAPV.

Figura 4.19 – Comparação da geração normalizada entre os sistemas de geração FV

Foram considerados somente os pontos de dados concorrentes entre os sistemas em pares. Ou seja, medições de geração que ocorreram no mesmo momento em ambos os sistemas em cada *subplot*. Novamente, ressalta-se a semelhança comportamental dos sistemas de geração sob mesma condição atmosférica, sobretudo no caso dos sistemas CTPV e BIPV. Houveram ocorrências, não representadas nos gráficos supramencionados, nas quais um sistema estava inoperante (geração nula) enquanto seu par apresentou geração elétrica, indicando que a saída nula não é fruto da falta de recurso solar, mas sim por outra ocorrência operacional (sombreamento, erro de armazenamento de dados, entre outros fatores).

Assim, as bases de dados desses três sistemas FV serão combinadas para formar uma base de dados mais robusta, visando analisar a capacidade preditiva do método de RNN para sistemas distintos dentro de uma área em proximidade. A Figura

4.20 ilustra o diagrama que delimita a rotina de seleção dos dados e a ordem de preferência dos dados disponíveis.

Figura 4.20 – Diagrama da rotina de seleção da base de dados conjunto IEE

A base de dados selecionada compreende um conjunto de dados no qual para o mesmo momento haja disponibilidade de dados para pelo menos um dos três sistemas, sendo diferente de zero e NaN (*Not a Number* – forma que o MATLAB identifica valores especiais que sejam complexos ou não-reais, utilizado aqui como um valor nulo não computável). A hierarquia de seleção obedece a ordem CTPV, BIPV e BAPV – dando origem ao conjunto de dados que será referido como conjunto IEE nas discussões apresentadas a seguir. O restante dos dados de geração do sistema BIPV não agregados no conjunto IEE será utilizado para validação do modelo preditivo e avaliar como este se comporta para dados desconhecidos.

4.4.1.1. MESES SECOS

O conjunto de dados dos meses secos para o conjunto IEE conta com 5.923 pontos de medições horárias, após tratamento dos dados, considerando o período entre 2017 e 2020. A Figura 4.21 apresenta a regressão linear de treinamento do conjunto IEE para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,928 – 0,945 e valores de m variando entre de

0,908 – 0,927. O $\text{NRMSE}_{150\text{kW}}$ teste observado foi de 8,40 kWh. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,942 – 0,953 e valores de m variando entre de 0,922 – 0,931. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 8,58 kWh.

Figura 4.21 – Regressão de treinamento do modelo conjunto em meses secos para previsão de geração 1h no futuro

A Figura 4.22 apresenta a regressão linear de treinamento do conjunto IEE para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,9 – 0,928 e os valores de m variando entre de 0,808 – 0,853. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 9,66 kWh. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,904 – 0,922 e valores de m variando entre de 0,815 – 0,847. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 10,34 kWh.

Figura 4.22 – Regressão de treinamento do modelo conjunto em meses secos para previsão de geração 2h no futuro

A Figura 4.23 apresenta a regressão linear de treinamento do conjunto IEE para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,862 – 0,873 e valores de m variando entre de 0,765 – 0,791. O NRMSE_{150kW} geral observado foi de 12,74 kWh. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,875 – 0,888 e valores de m variando entre de 0,761 – 0,793. O NRMSE_{150kW} geral observado foi de 11,85 kWh.

O modelo RNN para o conjunto IEE apresentou um comportamento satisfatório para os meses secos, com um RMSE% máximo de 8,13%.

Figura 4.23 – Regressão de treinamento do modelo conjunto em meses secos para previsão de geração 3h no futuro

4.4.1.2. MESES CHUVOSOS

O conjunto de dados dos meses chuvosos para o sistema conjunto IEE conta com 5.870 pontos de medições horárias. A Figura 4.24 apresenta a regressão linear de treinamento do IEE para 1h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,914 – 0,926 e valores de m variando entre de 0,84 – 0,875. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 9,27 kWh. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,913 – 0,922 e valores de m variando entre de 0,837 – 0,868. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 9,54 kWh.

Figura 4.24 – Regressão de treinamento do modelo conjunto em meses chuvosos para previsão de geração 1h no futuro

A Figura 4.25 apresenta a regressão linear de treinamento do conjunto IEE para 2h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,866 – 0,886 e valores de m variando entre de 0,757 – 0,789. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 11,42 kWh. Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,874 – 0,875 e valores de m variando entre de 0,757 – 0,791. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 11,67 kWh.

Figura 4.25 – Regressão de treinamento do modelo conjunto em meses chuvosos para previsão de geração 2h no futuro

A Figura 4.26 apresenta a regressão linear de treinamento do conjunto IEE para 3h de antecedência. Quando em ausência de dados de aerossol, foram observados valores de R variando entre 0,831 – 0,849 e valores de m variando entre de 0,706 – 0,732. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 13,27 kWh Quando dados de aerossol foram incorporados como dados de entrada foram observados valores de R variando entre 0,803 – 0,85 e valores de m variando entre de 0,697 – 0,728. O $\text{NRMSE}_{150\text{kW}}$ geral observado foi de 13,13 kWh.

Figura 4.26 – Regressão de treinamento do modelo conjunto em meses chuvosos para previsão de geração 3h no futuro

4.4.2. VALIDAÇÃO DO MODELO COM DADOS DE GERAÇÃO INÉDITOS

Os modelos conjuntos de previsão foram executados em diferentes condições atmosféricas (meses secos e chuvosos) ao longo de horizontes temporais variáveis de 1 a 3 horas de antecedência. As Figuras 4.27 – 4.30 apresentam uma comparação entre a geração real do sistema BIPV, expresso pela sigla NBIPV, e a prevista pelo modelo conjunto IEE, desenvolvido com os dados de geração dos sistemas CTPV, BIPV e BAPV.

As Figuras 4.27 – 4.28 apresentam as regressões lineares comparando os valores de geração real do sistema BIPV inéditos e os valores previstos pelo modelo conjunto IEE para os meses secos no horizonte tempo de uma, duas e três horas. Quando em ausência de dados de aerossol (Figura 4.27), foi observado um $\text{NRMSE}_{150\text{kW}}$ de 8,42; 10,99 e 13,43 kWh e um SS_{mse} de 0,876; 0,795 e 0,715 para uma, duas e três

horas de antecedência, respectivamente. As correlações observadas foram 0,928; 0,9 e 0,862, e as inclinações das retas de ajuste observadas foram 0,893; 0,787 e 0,702, respectivamente.

Figura 4.27 – Validação modelo com dados BIPV inéditos para meses secos em ausência de AOD

Quando em presença de dados de aerossol (Figura 4.28), foi observado um $\text{NRMSE}_{150\text{kWh}}$ de 8,01; 10,87 e 13,16 kWh e um SS_{mse} de 0,888; 0,799 e 0,726 para uma, duas e três horas de antecedência, respectivamente. As correlações observadas foram 0,953; 0,904 e 0,887. Já as inclinações das retas de ajuste observadas foram 0,903; 0,797 e 0,706.

Figura 4.28 – Validação modelo com dados BIPV inéditos para meses secos em presença de dados de AOD

Durante os meses chuvosos, quando em ausência dos dados de aerossol (Figura 4.29), foram observados um $\text{NRMSE}_{150\text{kWh}}$ de 9,93; 12,31 e 13,63 kWh e um SS_{mse} de 0,9; 0,848 e 0,82 para os horizontes temporais de uma, duas e três horas, respectivamente. As correlações observadas foram 0,921; 0,866 e 0,831. Já as inclinações das retas de ajuste observadas foram 0,836; 0,739 e 0,682.

Figura 4.29 – Validação modelo com dados BIPV inéditos para meses chuvosos em ausência de AOD

Quando os dados de aerossol foram incluídos (Figura 4.30) como entrada no modelo preditivo, foram observados um $\text{NRMSE}_{150\text{kW}}$ de 10,07; 12,32 e 13,88 kWh e um SS_{mse} de 0,896; 0,848 e 0,813 para os horizontes temporais de uma, duas e três horas, respectivamente. As correlações observadas foram 0,921; 0,874 e 0,803. Já as inclinações das retas de ajuste observadas foram 0,836; 0,737 e 0,683.

Figura 4.30 – Validação modelo com dados BIPV inéditos para meses chuvosos em presença de dados de AOD

O modelo conjunto desenvolvido com dados dos sistemas CTPV, BIPV e BAPV apresentou um desempenho semelhante aos modelos desenvolvidos individualmente para estes sistemas geradores FV. A Tabela 4.7 apresenta o *Skill Score* dos modelos conjuntos IEE tomando como referência o desempenho dos modelos específicos BIPV. A perda média de SS_{BIPV} é de 2,08% indicando que, de forma geral, o método conjunto apresenta um desempenho inferior em relação ao modelo especificamente treinado com os dados de geração do sistema BIPV. Contudo, o ônus em função da pequena diferença de capacidade preditiva pode ser mitigado ao utilizar deste método para previsão de geração em curto prazo de um sistema cuja série histórica é desconhecida. Por apresentar um comportamento coerente para diferentes sistemas operando sob uma condição atmosférica semelhante, o método proposto neste trabalho pode ser capaz de prever a geração de um conjunto de sistemas instalados dentro do perímetro de uma determinada área, desde que em condições operacionais semelhantes.

Tabela 4.7 – Skill Score do desempenho do modelo conjunto IEE em referência ao modelo BIPV

Estação		SECO						CHUVOSO					
Aerossol		Ausente			Presente			Ausente			Presente		
Horizonte		1h	2h	3h	1h	2h	3h	1h	2h	3h	1h	2h	3h
SS_{BIPV}		-0,1040	0,1043	-0,1140	-0,0193	-0,0375	-0,0995	0,0153	-0,0253	-0,0044	0,0014	0,0496	-0,0156

4.7. COMPARAÇÃO DE DESEMPENHO COM MODELOS ESTATÍSTICOS APRESENTADOS NA LITERATURA RECENTE

A Tabela 4.8 elenca o desempenho de modelos preditivos de geração solar fotovoltaica para horizontes temporais entre 1 e 3 horas de antecedência desenvolvidos nesta dissertação de mestrado e, também, por outros autores na literatura recente. Os trabalhos consultados utilizam exclusivamente métodos estatísticos em sua maioria, porém também comparamos resultados com modelos híbridos, como em Li et al. (2020), Wu et al. (2014) e Filipe et al. (2015) e Pedro & Coimbra (2012).

Para comparação, selecionamos nossos resultados obtidos para os sistemas CTPV e BIPV nas condições atmosféricas de clima seco e chuvoso, ambos com dados de aerossóis considerados como sinal de entrada. Também estão inclusos os modelos de persistência inteligente utilizado neste trabalho como modelo referência para cálculo de *Skill Score*. O modelo RNN apresentou melhores resultados para o sistema CTPV, apresentando RMSE% de 4,53%, 6,02% e 7,24% para 1 a 3h, respectivamente, para o período seco e 5,85%, 7,01% e 7,72% para 1 a 3h, respectivamente, para o período chuvoso. Estão destacados, em negrito, os melhores resultados obtidos neste trabalho e, também, os modelos desenvolvidos por outros trabalhos que apresentaram desempenhos superiores, comparado ao nosso desempenho em clima chuvoso.

Observamos que o desempenho dos modelos RNN desenvolvidos neste trabalho estão em concordância, e competitivos, com os modelos propostos na literatura recente. Como foi discutido anteriormente no Capítulo 2.6, métodos estatísticos são fortemente dependentes da disponibilidade e qualidade da série temporal de dados. Assim, a mesma técnica pode apresentar divergências de desempenho para diferentes localidades, como no caso do desempenho do modelo ANN proposto por Wu et al. (2014) para a planta FV instalada na National Penghu University (NPU).

Tabela 4.8 – Comparação de desempenho dos modelos de previsão de curto prazo propostos na literatura recente. O desempenho é comparado por meio do RMSE (%) para os horizontes temporais de 1h, 2h e 3h.

Referência	Planta/Local	modelo	1h	2h	3h
Autores	CTPV-S	RNN	4,53%	6,02%	7,24%
	BIPV-S	RNN	4,72%	6,52%	8,51%
	CTPV-C	RNN	5,93%	7,17%	8,09%
	BIPV-C	RNN	6,26%	8,04%	9,17%
	Persist-S	Smart-	13,91%	14,34%	15,31%
	Persist-C	Persist	21,28%	21,1%	20,89%
Wu et al. (2014)	GEO	ARIMA	9,52%	-	-
		ANN	7,78%	-	-
		Híbrido	5,64%	-	-
	TT	ARIMA	16,49%	-	-
		ANN	4,75%	-	-
		Híbrido	3,43%	-	-
	NPU	ARIMA	13,48%	-	-
		ANN	16,77%	-	-
		Híbrido	6,57%	-	-
Filipe et al. (2015)	Portugal	Híbrido	9,83% (médio 1-3h)		
Giorgi et al. (2013)	input I	Elman RNN	12,57%	-	18,55%
	input II	Elman RNN	12,6%	-	17,72%
	input III	Elman RNN	10,91%	-	15,61%
Hossain et al. (2017)	Plant1	ELM	2,77%	-	-
		ANN	3,56%	-	-
	Plant2	ELM	3,87%	-	-
		ANN	4,71%	-	-
	Plant3	ELM	3,32%	-	-
		ANN	6,31%	-	-
Li et al. (2020)	DKASC, Australia	LSTM	3,91%	-	-
		GRU	3,9%	-	-
		RNN	3,99%	-	-
		MLP	4,09%	-	-
		WPD-LSTM	0,89%	-	-
Pedro & Coimbra (2012)	Califonia	ARIMA	10,58%	14,21%	-
		ANN	7,16%	12,96%	-
		GA/ANN	6,4%	8,99%	-

De forma geral, o método de redes neurais recorrentes multicamadas proposto nesta dissertação de mestrado apresentou desempenhos superiores a métodos estatísticos como Redes Neurais Artificiais *feedforward* (ANN) e ARIMA. Este último que consistentemente apresentou desempenhos inferiores a métodos de aprendizado de máquinas de forma geral na literatura. Destaca-se, contudo, que nosso modelo de persistência inteligente apresenta um desempenho semelhante a modelos baseados em ARIMA, como alguns dos apresentados por Wu et al. (2014) e Pedro & Coimbra (2012).

Os modelos propostos em Giorgi et al. (2013) são baseados no método de Redes Neurais Recorrentes de Elmann, atingindo RMSE% de 10,91% e 15,61% para 1 e 3 horas de antecedência, respectivamente. Esta arquitetura é a precursora das RNN modernas e caiu em desuso nos últimos anos. Nosso modelo RNN a superou consideravelmente para todos ambos os horizontes temporais. Ainda assim é interessante esta comparação a fim de verificar a evolução da técnica RNN observada ao longo da última década.

Quanto as técnicas preditivas em trabalhos mais recentes, destaca-se os métodos de *Deep Learning* (DL), como a rede neural recorrente de *Long Short Term Memory* (LSTM). Modelos baseados em LSTM, como proposto em Li et al. (2020), por apresentarem “portões de esquecimento”, possuem uma maior capacidade de armazenar informações mais relevantes (e descartar as menos relevantes) analisadas em uma série temporal, melhorando o desempenho preditivo nestes modelos. Em Li et al. (2020) um modelo híbrido baseado em LSTM-WPD foi proposto para prever a geração elétrica com 1h de antecedência de uma usina FV instalada em Alice Springs, Austrália. É importante ressaltar o clima subtropical desértico da região, apresentando baixa nebulosidade e precipitação ao longo do ano, com exceção do verão. Os desempenhos dos modelos individuais propostos superaram nossos modelos RNN, sobretudo os modelos GRU (3,9%) e LSTM (3,91%). O destaque especial encontra-se no modelo híbrido LSTM-WPD que apresentou um RMSE% baixíssimo de 0,89%.

Por fim, na literatura consultada, os modelos híbridos/*ensemble* apresentaram desempenhos superiores aos métodos estatísticos convencionais aplicados individualmente. Destacam-se os modelos híbridos propostos por Li et al. (2020), Wu et al. (2014) e Pedro & Coimbra (2012) que apresentaram comparação de seus modelos híbridos em relação a técnicas preditivas de maneira individual. O modelo RNN proposto neste trabalho é competitivo mesmo com modelos híbridos mais antigos, porém é superado em modelos híbridos mais recentes. Com isso, entendemos a aplicação de modelos de *Deep Learning* aliados a modelos *ensemble* ser uma tendência bem-vinda a fim de melhorar o desempenho de modelos de previsão de geração solar fotovoltaica.

Capítulo 5 : CONCLUSÕES E CONSIDERAÇÕES FINAIS

Este estudo propôs modelos baseados em aprendizado de máquinas para realizar a previsão de geração elétrica da planta FV de 540 kWp, instalada na Cidade Universitária da Universidade de São Paulo, para o horizonte temporal de 1 a 3 horas de antecedência. Para o desenvolvimento dos modelos de previsão foram utilizados dados meteorológicos, atmosféricos e ambientais coletados em superfície próximos aos sistemas fotovoltaicos. Em concordância com a literatura recente, que aponta para um melhor desempenho das Redes Neurais Recorrentes (RNN) em comparação com as Redes Neurais Artificiais convencionais (RNA) para a previsão de séries temporais, foram propostos neste trabalho modelos RNN multicamada desenvolvidos para diferentes condições atmosféricas (meses secos e meses chuvosos).

O clima da cidade de São Paulo tem por característica uma alta concentração de aerossóis atmosféricos durante os meses secos (abril a setembro). Estudos recentes buscaram avaliar a influência da presença de aerossóis atmosféricos na produtividade de sistemas FV. Dentre os estudos consultados, destaca-se o trabalho de Bet (2021) que analisou a influência da presença de aerossóis na produtividade da planta FV de 540 kWp da USP por meio da análise do fator espectral dos mesmos sistemas que são analisados por nós. Assim, um segundo objetivo deste estudo foi avaliar a influência no desempenho dos modelos preditivos frente à inclusão de dados de profundidade ótica de aerossóis como sinal de entrada nos modelos RNN. Para tanto, foram utilizados dados de AOD 550nm oriundos do modelo de reanálise MERRA-2.

Um algoritmo de força bruta (AFB) definiu a melhor arquitetura de nossas redes como três camadas ocultas com trinta neurônios cada. O vetor de entrada possui duas configurações distintas. A primeira, que chamamos de convencional, possui os dados de índice de claridade, temperatura ambiente, velocidade do vento, umidade relativa, ângulo zenital, hora do dia e geração elétrica no instante zero. A segunda acresce os dados de AOD 550nm ao primeiro vetor.

Como esperado, os modelos RNN apresentaram maior acurácia durante os meses secos em comparação com os meses chuvosos. Contudo, quando comparado ao modelo de persistência referência, durante o período chuvoso o modelo RNN

apresentou valores de SS entre 0,85 – 0,91, enquanto os valores de SS observados durante os meses secos estão entre 0,77 – 0,88. Desta forma, a técnica RNN apresentou robustez ao prever geração FV mesmo quando a condição atmosférica é mais desafiadora.

Os modelos desenvolvidos para o sistema CTPV apresentaram os melhores desempenhos. Para os meses secos, foram observados valores de RMSE% entre 4,53% e 7,24% para 1 e 3h de antecedência. Para os meses chuvosos, foram observados valores de RMSE% entre 5,93% e 8,1% para 1 e 3h de antecedência. O desempenho superior é dado em função do maior número de medições válidas disponíveis (8.818) para este sistema e também sua proximidade à estação solarimétrica. Os sistemas BIPV (7.472 medições) e BAPV (5.329 medições) apresentaram desempenhos semelhantes, com exceção do sistema BAPV para os meses secos devido a somente 1.767 medições válidas. Outro fator que pode ter levado ao sistema BAPV a apresentar desempenhos ligeiramente inferiores aos demais sistemas é em função deste ser o mais afastado da estação solarimétrica. Porém, comparando os desempenhos do modelo para o sistema BAPV em ambas as condições atmosféricas, entendemos a quantidade de dados ser a principal causa do desempenho inferior do modelo. Esta observação está de acordo com a natureza de modelos estatísticos, que são fortemente dependentes da base de dados disponível.

De forma geral, a inclusão dos dados modelados de aerossol resultou em um ganho médio de 2,27% baseado no $SS_{MSE\%}$ dos modelos RNN durante os meses secos, apresentando maior influência em previsões para 1 e 2h de antecedência. Em contrapartida, durante os meses chuvosos a inclusão do AOD como sinal de entrada no modelo não aparenta influenciar na capacidade preditiva. Observamos valores de $SS_{MSE\%}$ que variam entre -4,51% e 4,13% independentemente dos sistemas FV e horizontes temporais. Como discutido na Seção 2.4, quando há maior cobertura de nuvens a presença de aerossóis possui menor importância nos processos de atenuação do recurso solar. Desta forma, foi observado que a inclusão de dados de AOD para previsões durante os meses chuvosos representam um ruído no modelo RNN, piorando seu desempenho.

Neste trabalho um modelo de persistência *smart* foi utilizado como referência de comparação. Os modelos RNN apresentaram desempenhos significativamente

superiores aos modelos de persistência, com ganhos médios de 58,51% e 65,79% baseados no $SS_{RMSE\%}$ para os meses secos e chuvosos, respectivamente. Além disso, os modelos RNN desenvolvidos apresentaram desempenhos compatíveis, e inclusive competitivos, com a literatura recente acerca de modelos estatísticos de previsão de geração FV em curto prazo. O desempenho dos modelos RNN propostos neste trabalho foi consistentemente superior a métodos estatísticos convencionais consultados na literatura, apresentando ganhos médios variando entre 52,66% – 63,83% em relação a modelos ARIMA e 18,73% – 37,92% em relação a modelos RNA, para o horizonte temporal de 1h. Em contrapartida, modelos preditivos baseados em *Deep Learning*, como *Gated Recurrent Unit* (GRU), *Extreme Learning Machine* (ELM) e *Long-Short Term Memory* (LSTM) propostos em estudos recentes apresentam $RMSE_{\%}$ entre 3,44 – 3,91%, superando nossos modelos RNN. Contudo, técnicas de DL requerem maior capacidade computacional, representando um impeditivo a depender dos recursos técnicos disponíveis. Por fim, foi observada na literatura consultada um rápido desenvolvimento dos modelos estatísticos de previsão de curto prazo da geração solar fotovoltaica, sobretudo os modelos que utilizem técnicas de aprendizado de máquinas.

Os resultados obtidos neste estudo chamam à atenção para a influência da presença de aerossóis atmosféricos no desempenho de modelos de previsão solar FV. Como foi discutido anteriormente, em 2020 a capacidade instalada acumulada GD superou a centralizada, reforçando a importância da análise do recurso solar em ambientes urbanos, com maior concentração de aerossóis atmosféricos e que possuam alto potencial FV, como é o caso da cidade de São Paulo. É importante ressaltar, contudo, que os resultados obtidos neste trabalho são promissores, porém ainda inconclusivos, visto que foram utilizados dados de AOD modelados, apresentando assim erros do próprio modelo de reanálise MERRA-2. Portanto, recomenda-se novos estudos em diferentes regiões e também os que utilizem medições de superfície para uma avaliação mais aprofundada da influência da presença de aerossóis na precisão da previsão de energia. Para trabalhos futuros são sugeridos:

- Estudo de influência no desempenho do modelo RNN gerada por dados de profundidade ótica de aerossóis medidos em superfície, próximos aos sistemas de geração FV;

- Desenvolver uma rotina de renovação dos dados de treinamento do modelo RNN para os sistemas CTPV, BIPV e BAPV;
- Estudo de desempenho do modelo RNN para novos sistemas FV que venham a ser instalados na Universidade de São Paulo a fim de reafirmar a confiabilidade do modelo preditivo proposto;
- Investigar a influência da distância entre modelos FV e a estação solarimétrica no desempenho do modelo RNN proposto neste estudo, a fim de estimar o alcance que o modelo desempenha previsões satisfatórias;
- Desenvolvimento de modelos de aprendizado de máquinas baseados em técnicas de *Deep Learning* e, também, modelos híbridos. Além disso, comparar os métodos preditivos em termos de performance e tempo computacional.

REFERÊNCIAS BIBLIOGRÁFICAS

ABSOLAR. Energia Solar Fotovoltaica no Brasil: infográfico ABSOLAR. n. 25, p. 2, 2020.

AL-SBOU, Y. A.; ALAWASA, K. M. Nonlinear Autoregressive Recurrent Neural Network Model For Solar Radiation Prediction. **International Journal of Applied Engineering Research**, v. 12, n. 14, p. 4518–4527, 2017.

ALMEIDA, M. P. **Implicações técnicas da inserção em grande escala da geração solar fotovoltaica na matriz elétrica. 2017. 191p. Tese (Doutorado em Ciências) - Programa de Pós-Graduação em Energia da Universidade de São Paulo, São Paulo, 2017.** [s.l: s.n.].

ALMONACID, F. et al. A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator. **Energy Conversion and Management**, v. 85, p. 389–398, 2014.

ANEEL. **SIGA: Sistema de Informações de Geração.** Disponível em: <<https://bit.ly/2IGf4Q0>>.

ANTONANZAS, J. et al. Review of photovoltaic power forecasting. **Solar Energy**, v. 136, p. 78–111, 2016.

BET, L. G. **Impacto dos aerossóis atmosféricos no fator espectral de módulos fotovoltaicos em São Paulo.** [s.l.] Universidade Federal de São Paulo, 2021.

BLANKENSHIP, R. E. et al. Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. **Science**, v. 332, n. 6031, p. 805–809, 2011.

BRITISH PETROLEUM REPORT. 67 th edition Contents is one of the most widely respected. **Statistical Review of World Energy**, p. 40, 2018.

CABALLERO, J. A. et al. Spectral Corrections Based on Air Mass, Aerosol Optical Depth, and Precipitable Water for PV Performance Modeling. **IEEE Journal of Photovoltaics**, v. 8, n. 2, p. 552–558, 2018.

CPRM. **Crise hídrica no Brasil: o uso das águas subterrâneas como reforço no abastecimento público.** [s.l: s.n.].

CRISOSTO, C.; LUIZ, E. W.; SECKMEYER, G. Convolutional Neural Network for High-Resolution Cloud Motion Prediction from Hemispheric Sky Images. **Energies**, v. 14, n. 3, p. 753, 2021.

DE ALMEIDA ALBUQUERQUE, T. T.; DE FÁTIMA ANDRADE, M.; YNOUE, R. Y. Characterization of atmospheric aerosols in the city of São Paulo, Brazil: Comparisons between polluted and unpolluted periods. **Environmental Monitoring and Assessment**, v. 184, n. 2, p. 969–984, 2012.

DE GIORGI, M. G.; CONGEDO, P. M.; MALVONI, M. Photovoltaic power forecasting using statistical methods: Impact of weather data. **IET Science, Measurement and Technology**, v. 8, n. 3, p. 90–97, 2014.

DIAGNE, M. et al. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. **Renewable and Sustainable Energy Reviews**, v. 27, p. 65–76, 2013.

DOLARA, A. et al. PV hourly day-ahead power forecasting in a micro grid context. **EEEIC 2016 - International Conference on Environment and Electrical Engineering**, 2016.

- DULLWEBER, T. et al. PERC+: industrial PERC solar cells with rear Al grid enabling bifaciality and reduced Al paste consumption. **Progress in Photovoltaics: Research and Applications**, p. 12, 2015.
- EBERHARDT, I. D. R. et al. Cloud cover assessment for operational crop monitoring systems in tropical areas. **Remote Sensing**, v. 8, n. 3, p. 1–14, 2016.
- ENGERER, N. A.; MILLS, F. P. KPV: A clear-sky index for photovoltaics. **Solar Energy**, v. 105, p. 679–693, 2014.
- EPE. Plano Decenal de Expansão Energética 2026. p. 271, 2017.
- EPE. Balanço Energético Nacional Relatório Síntese 2021: ano base 2020. p. 77, 2021.
- FILIPE, J. M. et al. A hybrid short-term solar power forecasting tool. **2015 18th International Conference on Intelligent System Application to Power Systems, ISAP 2015**, 2015.
- GLOBAL SOLAR ATLAS. **Potencial fotovoltaico da Alemanha**. Disponível em: <<https://globalsolaratlas.info/download/germany>>.
- GOLDEMBERG, J.; LUCON, O. Energia e meio ambiente no Brasil. **Estudos Avançados**, v. 21, n. 59, p. 61–76, 2007.
- GUARNIERI, R. A. **Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do modelo ETA**. [s.l.: s.n.].
- GUEYMARD, C. A. A review of validation methodologies and statistical performance indicators for modeled solar radiation data: Towards a better bankability of solar projects. **Renewable and Sustainable Energy Reviews**, v. 39, p. 1024–1034, 2014.
- HENG, J. B. et al. >23% High-efficiency tunnel oxide junction bifacial solar cell with electroplated Cu gridlines. **IEEE Journal of Photovoltaics**, v. 5, n. 1, p. 82–86, 2015.
- HEZEL, R. Novel Applications of Bifacial Solar Cells. **Progress in Photovoltaics: Research and Applications**, v. 11, n. 8, p. 549–556, 2003.
- HORVATH, H. Aerosol - An introduction. **Journal of Environmental Radioactivity**, v. 51, n. 1, p. 5–25, 2000.
- HOSSAIN, M. et al. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. **Journal of Cleaner Production**, v. 167, p. 395–405, 2017.
- IBGE. **Instituto Brasileiro de Geografia e Estatística**. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/sp/sao-paulo.html>>. Acesso em: 2 mar. 2021.
- IEE. **Relatório da Produção de Energia Elétrica da Usina FV de 540 kWp**. [s.l.: s.n.].
- INMET. **Normais climatológicas do Brasil (1981 - 2010)**. Disponível em: <<https://portal.inmet.gov.br/normais>>.
- IQBAL, M. **An introduction to solar radiation**. [s.l.] ACADEMIC PRESS, 1983.
- JACOBSON, M. **Fundamentals of atmospheric modeling**. 2. ed. [s.l.] Cambridge University Press, 2005.

JOGE, T. et al. Applications and field tests of bifacial solar modules. **Conference Record of the IEEE Photovoltaic Specialists Conference**, p. 1549–1552, 2002.

JOHN, S.; KESAVAN, A.; HERSZBERG, I. The Use of Artificial Neural Networks in Damage Detection and Assessment in Polymeric Composite Structures. In: SAPUAN, S. M.; MUJTABA, I. M. (Eds.). . **Composite material technology: neural network application**. [s.l.] CRC Press, 2010. p. 372.

JOHNSON, R. A. **Applied multivariate statistical analysis**. Richard A. Johnson. Dean W. Wichers. - 4th ed. Prentice-Hall: EUA, 1998. [s.l.: s.n.].

KARDAKOS, E. G. et al. Application of time series and artificial neural network models in short-term forecasting of PV power generation. **Proceedings of the Universities Power Engineering Conference**, p. 2–7, 2013.

KASHYAP, Y.; BANSAL, A.; SAO, A. K. Solar radiation forecasting with multiple parameters neural networks. **Renewable and Sustainable Energy Reviews**, v. 49, p. 825–835, 2015.

KHAN, M. R.; ALAM, M. A. Thermodynamic limit of bifacial double-junction tandem solar cells. **Applied Physics Letters**, v. 107, n. 22, 2015.

KILLINGER, S. et al. Projection of power generation between differently-oriented PV systems. **Solar Energy**, v. 136, p. 153–165, 2016.

LEE, D.; KIM, K. Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information. **Energies**, v. 12, n. 2, 2019.

LI, P. et al. A hybrid deep learning model for short-term PV power forecasting. **Applied Energy**, v. 259, n. July, p. 114216, 2020.

LIMA, F. J. L. et al. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. **Renewable Energy**, v. 87, p. 807–818, 2016.

LIMA, F. J. L. et al. Previsão de curto prazo da radiação solar no brasil usando movimento das nuvens com base em dados de satélite. **Revista Brasileira de Energia Solar**, v. XI, p. 97–107, 2020.

LIU, K. **An introduction to atmospheric radiation**. 2. ed. [s.l.] Academic press, 2002.

LOONKAR, S.; MISHRA, D. Review of back-propagation algorithms for defect elimination with proposed DMASIC methodology. **2016 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2016**, p. 755–764, 2016.

LORENZ, E. et al. Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 2, n. 1, p. 2–10, 2009.

LUIZ, E. W. et al. Analysis of intra-day solar irradiance variability in different Brazilian climate zones. **Solar Energy**, v. 167, n. April, p. 210–219, 2018a.

LUIZ, E. W. et al. Comparison of methodologies for cloud cover estimation in Brazil - A case study. **Energy for Sustainable Development**, v. 43, p. 15–22, 2018b.

MARTINS, F. R. **Influência do processo de determinação da cobertura de nuvens e dos aerossóis de queimada no modelo físico de radiação BRASIL-SR**. São José dos Campos:

Instituto Nacional de Pesquisas Espaciais, 2001.

MARTINS, J. V. **O efeito de partículas de aerossol de queimadas da Amazônia no balanço radiativo da atmosfera.** Universidade de São Paulo - USP, 1999. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/43/43131/tde-20072012-121657/en.php>>

MATHWORKS. **Choose a multilayer neural network training function.** Disponível em: <<https://se.mathworks.com/help/deeplearning/ug/choose-a-multilayer-neural-network-training-function.html>>.

MELLIT, A.; PAVAN, A. M. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. **Solar Energy**, v. 84, n. 5, p. 807–821, 2010.

MELLIT, A.; SAĞLAM, S.; KALOGIROU, S. A. Artificial neural network-based model for estimating the produced power of a photovoltaic module. **Renewable Energy**, v. 60, p. 71–78, 2013.

MISHRA, S.; PALANISAMY, P. Multi-time-horizon Solar Forecasting Using Recurrent Neural Network. **2018 IEEE Energy Conversion Congress and Exposition, ECCE 2018**, p. 18–24, 2018.

MONTEIRO, C. et al. Short-term power forecasting model for photovoltaic plants based on historical similarity. **Energies**, v. 6, n. 5, p. 2624–2643, 2013.

NOBRE, A. M. et al. On the impact of haze on the yield of photovoltaic systems in Singapore. **Renewable Energy**, v. 89, p. 389–400, 2016.

OLAUSON, J.; BERGKVIST, M. Modelling the Swedish wind power production using MERRA reanalysis data. **Renewable Energy**, v. 76, p. 717–725, 2015.

PANG, Z.; NIU, F.; O'NEILL, Z. Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. **Renewable Energy**, v. 156, p. 279–289, 2020.

PEDRO, H. T. C.; COIMBRA, C. F. M. Assessment of forecasting techniques for solar power production with no exogenous inputs. **Solar Energy**, v. 86, n. 7, p. 2017–2028, 2012.

PEREIRA, E. B. et al. Atlas Brasileiro Energia Solar 2ª Edição. **Instituto Nacional de Pesquisas Espaciais**, p. 88, 2017.

PIN, M. **Levantamento do potencial de geração fotovoltaica com sistemas de microgeração aplicados à arquitetura no edifício do campus São Paulo da USP.** [s.l.: s.n.].

POLMAN, A. et al. Photovoltaic materials: Present efficiencies and future challenges. **Science**, v. 352, n. 6283, 2016.

RAMACHANDRAN, S. **Atmospheric aerosols: characteristics and radiative effects.** [s.l.] CRC Press, 2018.

REN21. **Global Status Report 2018.** [s.l.: s.n.].

RÉQUIA JÚNIOR, W. J.; ARAVÉCHIA JÚNIOR, J. C. O comportamento espacial e temporal das queimadas no Brasil: um estudo nas regiões mais críticas no período de 2000 a 2010 / The spatial and temporal behavior of burned in Brazil: a study in the most critical regions in the period

2000 to 2010. **Revista Ambiente**, v. 9, n. 2, p. 313–322, 2013.

RODRIGUEZ, M. **Avaliação operacional dos sistemas fotovoltaicos com seguimento solar: sistemas com e sem concentração**. [s.l.: s.n.].

ROSA-CLOT, M.; TINA, G. M. Photovoltaic Electricity. **Submerged and Floating Photovoltaic Systems**, p. 13–32, 2018.

ROSA, D. J. DE M. **Caracterização da radiação solar: o caso da cidade universitária/USP e da Ilha do Cardoso/Cananéia**. [s.l.] Universidade de São Paulo, 2003.

SMIL, V. Distributed generation and megacities: Are renewables the answer? **IEEE Power and Energy Magazine**, v. 17, n. 2, p. 37–41, 2019.

SOBRI, S.; KOOHI-KAMALI, S.; RAHIM, N. A. Solar photovoltaic generation forecasting methods: A review. **Energy Conversion and Management**, v. 156, n. November 2017, p. 459–497, 2018.

SRIKRISHNAN, V. et al. Using multi-pyranometer arrays and neural networks to estimate direct normal irradiance. **Solar Energy**, v. 119, p. 531–542, 2015.

SUNBANK. **Sunbank Solar Panel Breakdown**. Disponível em: <<https://mysunbank.com.au/pv-solar/solarpanel-howitsmade/>>.

TOLEDANO, C. et al. Aerosol optical depth and Ångström exponent climatology at El Arenosillo AERONET site (Huelva, Spain). **Quarterly Journal of Royal Meteorological Society**, n. 133, p. 795–807, 2007.

UNFCCC. CONFERENCE OF THE PARTIES (COP). **Paris Climate Change Conference-November 2015, COP 21. Adoption of the Paris Agreement**. Science. [s.l.: s.n.].

VIGNOLA, F.; MICHALSKY, J.; STOFFEL, T. **Solar and infrared radiation measurements**. [s.l.] CRC Press, 2012.

VILAÇA GOMES, P. et al. Technical-economic analysis for the integration of PV systems in Brazil considering policy and regulatory issues. **Energy Policy**, v. 115, n. January, p. 199–206, 2018.

VILLALVA, M. G. **Energia solar fotovoltaica: conceitos e aplicações**. 2º ed. São Paulo: Érica Editora Saraiva, 2012.

WILLOUGHBY, A. A.; OSINOWO, M. O. Development of an electronic load I-V curve tracer to investigate the impact of Harmattan aerosol loading on PV module performance in southwest Nigeria. **Solar Energy**, v. 166, n. December 2017, p. 171–180, 2018.

WU, Y. K.; CHEN, C. R.; ABDUL RAHMAN, H. A novel hybrid model for short-term forecasting in PV power generation. **International Journal of Photoenergy**, v. 2014, 2014.

ZILLES, R. et al. **Sistemas fotovoltaicos conectados à rede elétrica**. 1º ed. [s.l.] Editora Oficina de Textos, 2012.